Introduction to interferometry: Draper Labs, Sept. 16-18, 2019

Major Topic(s)

Fundamentals of Interferometry
- Interferometric radar
- Interferometric processing implementation
- Image registration
- Baseline determination

Decorrelation
- Scattering physics
- Baseline decorrelation
- Temporal decorrelation
- Examples

Measurement of surface topography
- Topographic algorithms
- Application examples

Velocity measurements
- Measurement of velocities
- Ocean current applications
- Glaciers and ice streams
- Airborne and spaceborne instrumentation

Surface deformation measurements
- Two- and three-pass algorithms
- Earthquakes and volcanoes

Motion compensation
- Azimuth integration
- Processing algorithms
- Examples

Backprojection processing
- Phase compensation
- Geocoding
- Processing algorithms
- Examples

REFERENCES

Books which may serve as useful references are listed below.

Bracewell, R. N., The Fourier Transform and Its Applications, McGraw-Hill, New York, 2nd edition, 1986.

Carrara, W.G., R.S. Goodman, and R.M. Majewski, Spotlight Synthetic Aperture Radar: Signal Processing Algorithms, Artech House, Norwood, MA, 1995.

Curlander, J.C. and R.N. McDonough, Synthetic Aperture Radar, Wiley Interscience, New York, 1991.

Elachi, C., Introduction to the physics and techniques of remote sensing, Wiley, New York, 1987.

Elachi, C., Spaceborne Radar Remote Sensing: Applications and Techniques, IEEE press, New York, 1988.

Goodman, J.W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968.

Kraus, J.D. Radio Astronomy, McGraw-Hill, New York, 1966. (Later editions good also)

Peebles, P.Z, Radar Principles, Wiley Interscience, New York, 1998.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes in C, the Art of Scientific Computing, Cambridge University Press, New York, 1988. (Any of the Numerical Recipes series will have useful algorithm information)

Sabins, F., Remote Sensing, 3rd ed., Freeman, New York, 1996. Soumekh, M., Fourier Array Imaging, Prentice Hall, Englewood Cliffs, New Jersey, 1994.

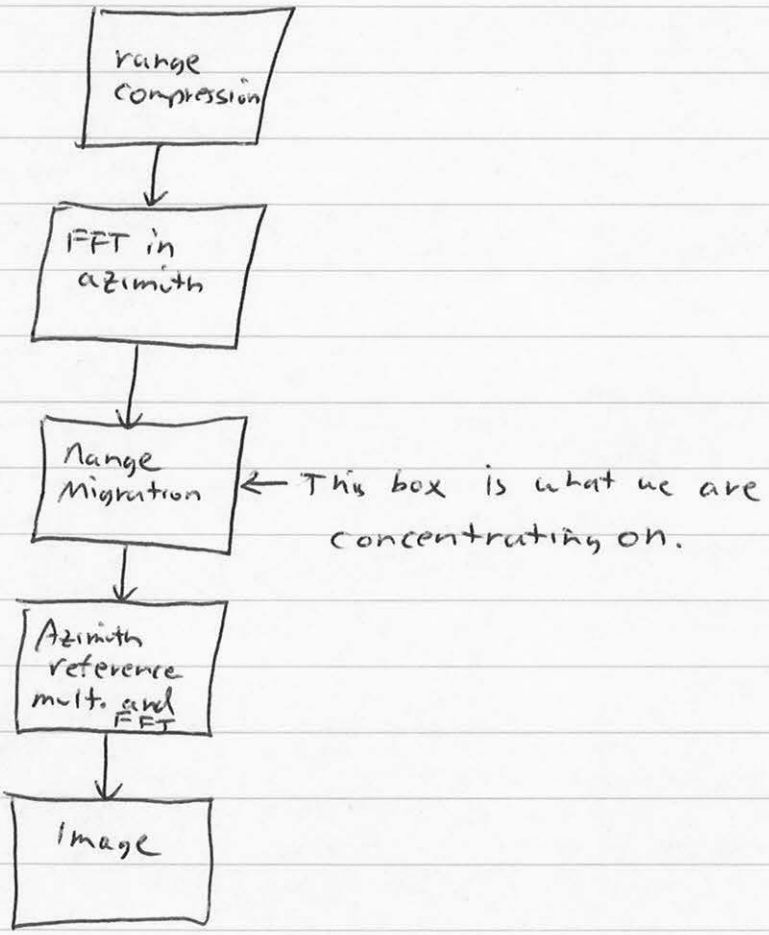
The following, though not on reserve, are also relevant and may be useful:

Cook, C.E., and M. Bernfeld, Radar Signals, Academic Press, New York, 1967.

Skolnik, M.I., Radar Handbook, McGraw-Hill, New York, 1970.

Implementation issues in range migration

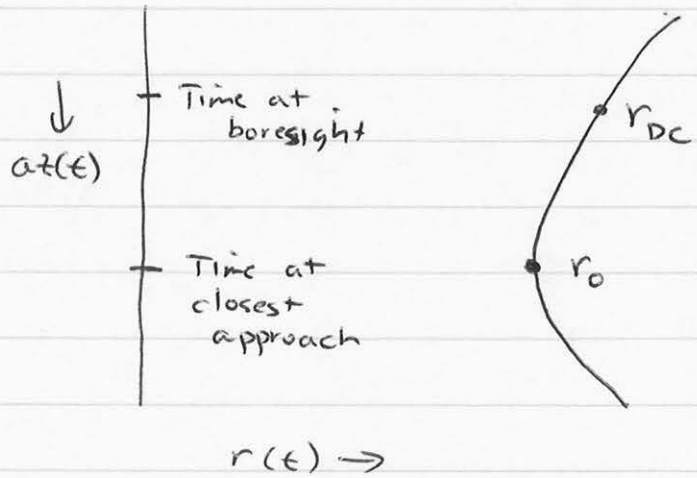
Let is assume we decide to process data using the interpolation method for range migration. We will go through design of a range migration routine for our SAR processor. Our flow diagram looks like



Here we have to decide a major issue for our processor: Just how we interpolate determines where the output pixels appear in the radar image. This means a great deal in terms of how we interpret our resulting image.

Deskewing

As we have found before, typically we process data at some centroid offset from the exact side-looking direction. Let's plot the range history first:



Two coordinate systems naturally present themselves: one referenced to the azimuth time and r_0 corresponding to closest approach, and one referenced to the azimuth time and r_{DC} for the antenna boresight direction. Since the (boresight time, r_{DC}) coordinate system is dependent on our choice of Doppler centroid while the (closest approach, r_0) system is independent of processing parameters, the former is called a <u>skewed</u> coordinate system while the latter is <u>deskewed</u>. It doesn't really matter which we use as long as we are consistent with our processing code and our analysis approach.

Some argue deskewing leads to more cartographically accurate images. This is largely a philosophical issue as a further processing step to convert from slant range to ground range is always necessary for mapping.

A range migration routine

So, what might a range migration interpolation routine look like? Assume we start off with a range compressed patch, which we have already transformed in the azimuth direction.

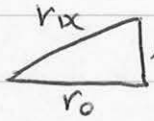
We want to trace out a different migration path for each range bin. Hence we begin by creating a loop over the range bins:

```
do i = 1, nbins
    r = r0 + (i-1) * δr        ← get range to line
```

(Note here as usual $\delta r = \frac{c}{2 f_s}$)

f_{DC} = <some function of range, perhaps a constant>

$r_{DC} = r\left(1 + \frac{f_{DC}^2}{8 v^2}\right)$ ← from construction



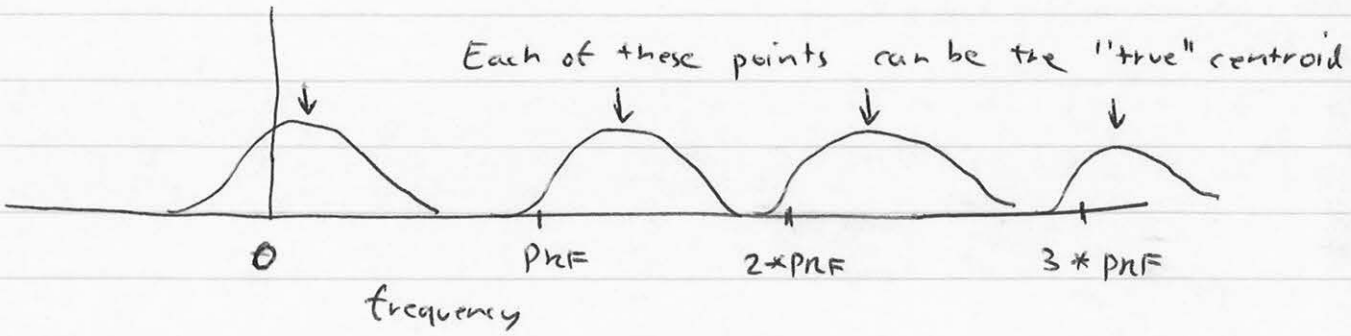
$f_R = \frac{-2 v^2}{\lambda r_{DC}}$

where we compute the range at Doppler centroid r_{DC} and the chirp rate f_R as a function of range.

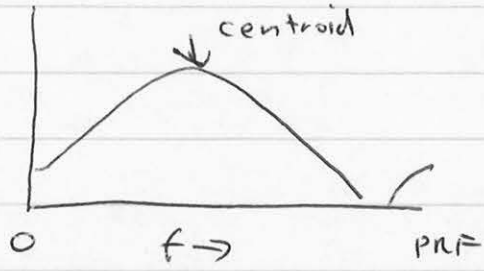
Now, we need to find the range to the object at each available Doppler bin, this is determine the range history as a function of azimuth frequency. Hence, we need to loop over frequency bins:

```
do j = 1, naz fft          ← j is azimuth bin # after transform
    freq = (j-1)/naz fft * prf
```

Here is a tricky part: we want to process on the correct "ambiguity". Recall that our azimuth spectrum is sampled at the prf and hence repeats in the frequency domain with that period:



The azimuth spectrum will be identical for each assumption of f_{dc}. Here we must go out on a limb and say while we don't know f_{DC} exactly, we probably do know it to the precision of a prf, and can thus properly calculate r_{DC} from the centroid measurement.



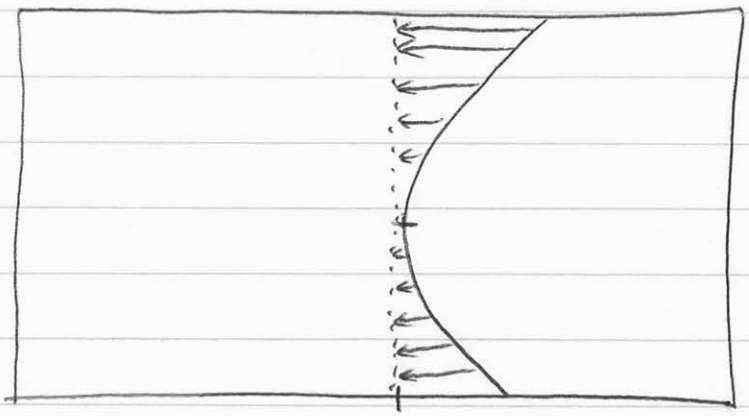
The centroid estimation gives f_{DC} to the "ambiguity" of a prf only!

The following code ensures that we ~~estim~~ use the proper frequency for each bin, and ensure that the frequency used is within $\frac{1}{2}$ prf of the true centroid.

$$n = \text{nint}\left(\frac{\text{freq} - f_{DC}}{\text{prf}}\right)$$ "nint" is nearest integer

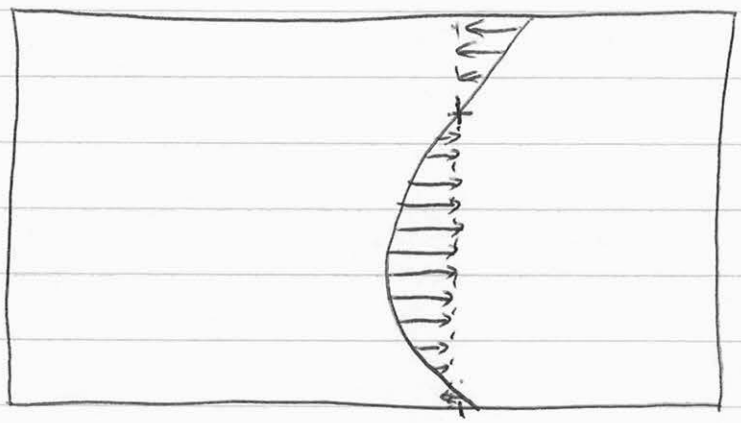
$$\text{freq} = \text{freq} - n * \text{prf}$$

Now, we need to decide about deskewing. Will we resample points to lie along the r_0 bin or the r_{DC} bin?



Deskewing

r_0



Not Deskewing

r_{DC}

We need to ~~train~~ shift the points by the difference of the true range history and the desired reference line in either case.

So, what is our range history as a function of frequency?

$$r^2(t) = r_0^2 + x^2$$

$$= r_0^2 + \left(\frac{f \cdot \lambda \cdot r(t)}{2v}\right)^2$$

~~$r^2(t) = r_0$~~

$$r^2(t)\left(1-\left(\frac{f\lambda}{2v}\right)^2\right) = r_0^2$$

$$r^2(t) = r_0^2\left(1+\left(\frac{f\lambda}{2v}\right)^2\right)$$

$$r(t) = r_0\left(1+\frac{f^2\lambda^2}{8v^2}\right)$$

So, for the deskew case, the offset $\Delta(f)$ is

$$\Delta(f) = \frac{f^2\lambda^2 r_0}{8v^2}$$

In the non-deskewed case

$$\Delta(f) = r_0\left(1+\frac{f^2\lambda^2}{8v^2}\right) - r_0\left(1+\frac{f_{DC}^2\lambda^2}{8v^2}\right)$$

$$= \left(f^2 - f_{DC}^2\right)\frac{\lambda^2 r_0}{8v^2}$$

We simply pick one of the above, and then interpolate:

$$V_{out} = \sum_{k=-3}^{4} v\left(i + \text{nint}\,\frac{\Delta(f)}{\Delta r} + k,\ j\right)\text{sinc}\left(k + \text{frac}\left(\frac{\Delta(f)}{\Delta r}\right)\right)$$

where frac( ) is the fractional part.

Also $v(i,j)$ is the transformed array.

Now that we have resampled the array to straighten out our range histories, we can multiply our columns by the conjugate of the transformed reference functions and create the images.

However, we still have the issue of deskewing in the azimuth direction (we have already corrected for the range part of the deskew correction). So, we must pay attention to the along-track position of the output pixel.

In other words, if we want to reference our output to the non-deskewed ("natural") position, we may need to apply a linear phase correction in the definition of the reference function. This simply means being careful in the matched filter definition: do we apply a reference with zero time at the point of closest approach or at the boresight time?

Some scattering issues

Now that we have learned to process the SAR data, it is time to begin analysis and interpretation of those data in terms of physical characteristics of the surface. While we won't spend a great deal of time on scattering and interaction topics, we will want a basic intuitive feel for the processes involved so we can understand our image data.

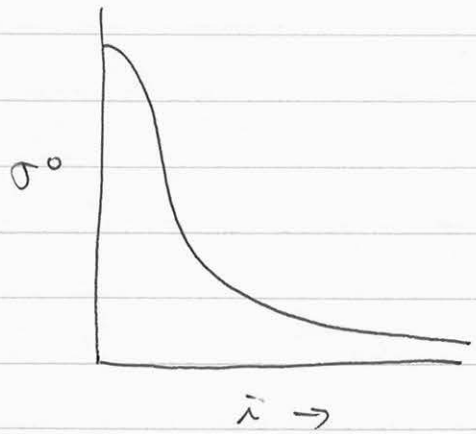
Scattering mechanisms

We have already expressed "brightness" in a radar image in terms of σ°, the normalized cross-section of a surface. Since σ° is proportional to the echo power in the radar equation, it embodies properties of the scattering surface itself and of the imaging geometry.

Decoupling the effects of geometry and surface parameters may be done by viewing the same surface at different incidence angles. But it is required if we want to interpret the apparent brightness in the images.

Consider viewing the same surface from two different angles:

If we view the scattering process as letting more energy escape in the near forward direction than near backward, we can picture surfaces as being brighter at near-normal incidence. In fact if we plot a typical curve of σ° vs i, we get something like

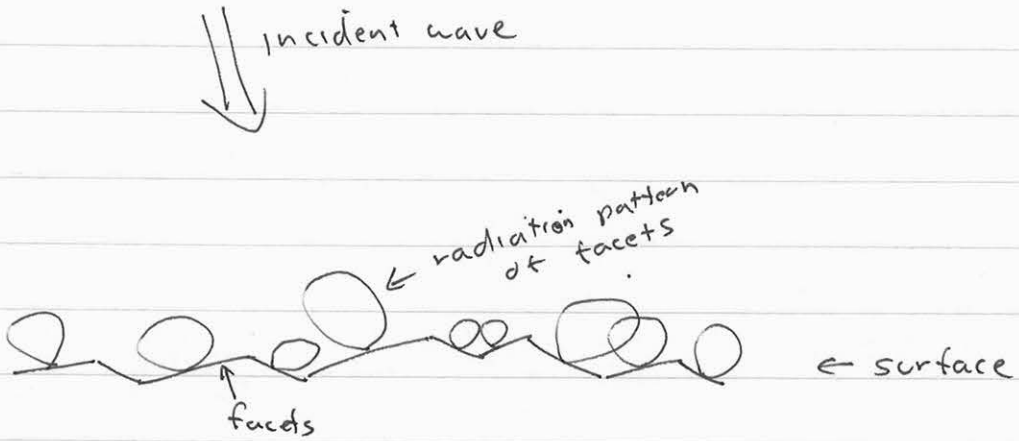


We have discussed already how the modulation of the local incidence angle in a radar image, coupled with a σ° curve like the one above, gives rise to the bright and dim areas associated with topography.

Now, the σ° response plotted above can be better understood if we try to relate its shape to surface parameters. It turns out to be easier to do this if we adopt simplified scattering models applicable to different regimes of i. This is equivalent to examining different scattering mechanisms for the different incidence angles.

Near-normal scattering - facet models

Consider first scattering for very small angles of incidence i. Then we can model the interaction as a series of mirrors, or facets, reflecting the incident wave:



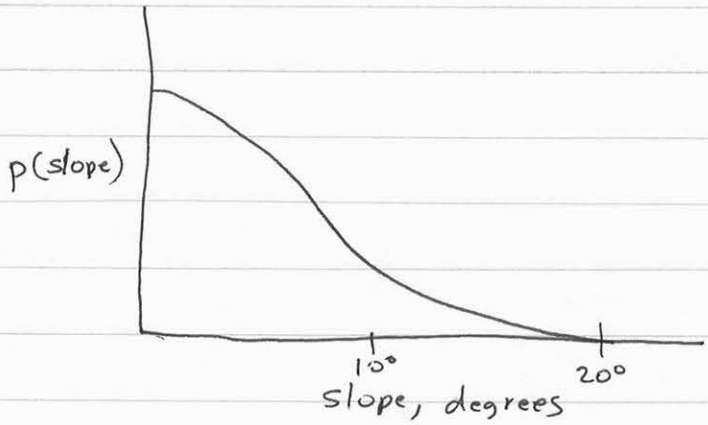
Each facet has an "equivalent radiation pattern" related to an equivalent antenna which reradiates incident energy. The return from each facet is greatest when illuminated normally, and fall off typically as a dipole (cosine) radiation pattern.

We can simplify matters still more by first assuming all facets are the same length, thus the reradiation patterns are the same size but of varying orientations. If all the returns add incoherently, which is the case if the facets are separated in range by a wavelength or more, Then the total backscattered power is simply related to the fraction of facets oriented at the radar, integrated over a range of angles equal to the width of each pattern.

If we have multiple facet sizes, we need simply to integrate over the distribution of sizes.

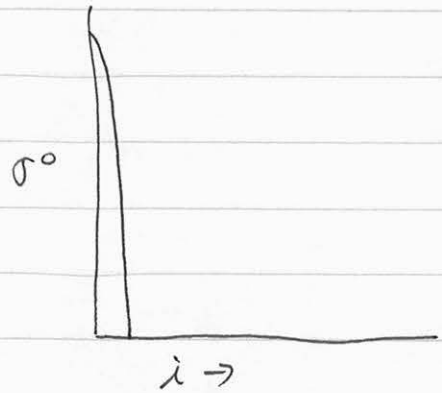
Hence the probability distribution of facets is the key parameter determining σ° for this model. A little thought shows this distribution is expressed as the probability distribution of surface slopes.

A typical slope distribution might be:

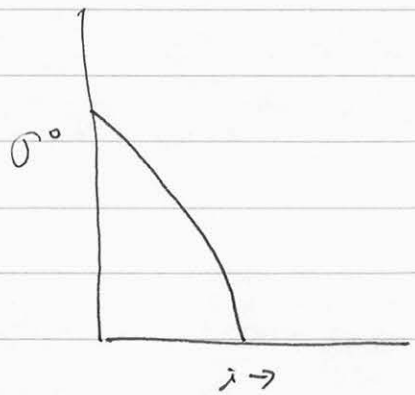


This surface, with a standard deviation of slopes ~ 10° is about as rough a surface as you will find. For most surfaces the angle of repose (slope) doesn't exceed about 20°.

Now, σ° will be related to a 2-D integral over slopes, hence we can ~~compte~~ compare σ° by computing the average power over the range of slopes for a given surface. For very smooth surfaces, such as water, the probability of a slope greater than 1-2° is rare. Thus we can infer surface slope or roughness by examining the near-nadir σ°:



Smooth surface



Rough surface

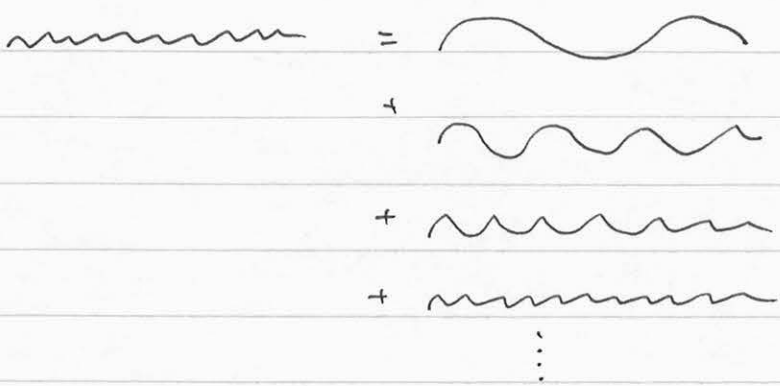
More detailed models incorporating curvature effects as well yield better approximations to the real σ°. These are called variously quasi-specular or Kirchoff scattering models and are dealt with in greater depth in EE354.

Greater Incidence Angles – Bragg scatter

The facet-type mechanism yields little energy for $i > \sim 20°$, but clearly we can image very much larger angles than this. One mechanism dominant at these angles is called Bragg scattering, and is valid for $i > \sim 20°$.

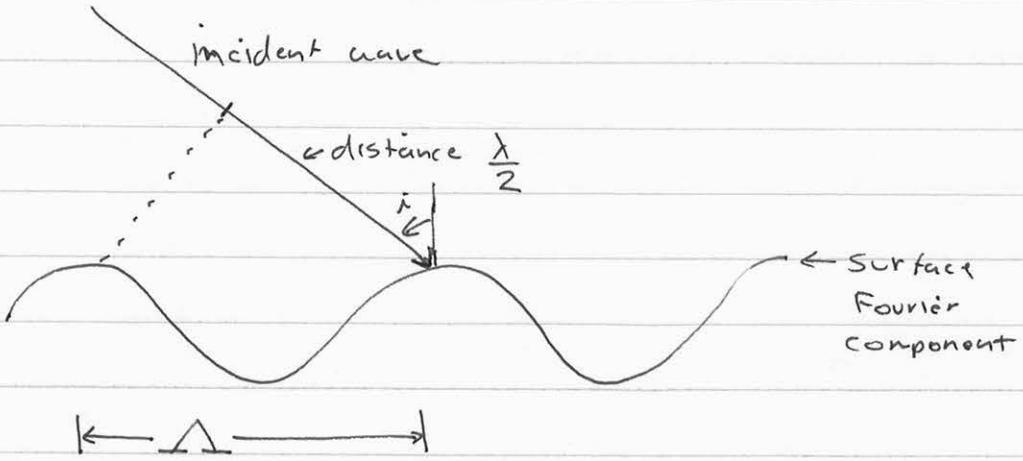
For the Bragg model, consider a surface divided into a large-scale, faceted component and a small-scale residual roughness component.

Putting the faceted part aside for the moment, consider the small-scale component alone. Break it up using Fourier decomposition:



Now, the amplitude of each of the Fourier components will be small. For the model to be easily evaluated we require the rms height of the surface to be less than about $\frac{\lambda}{8}$ in amplitude, which is why we removed the large-scale facet term.

Consider the Fourier component for which the distance Λ below is equal to $\frac{\lambda}{2}\sin i$:



Note that because of the round-trip distance of λ between like parts of the surface Fourier component, energy from everywhere on the surface adds up in phase, therefore this component is "brighter" than others. This component is the only one "matched" to the wavelength of the incident radiation.

If the surface component at this frequency has higher amplitude, the radar return will increase.

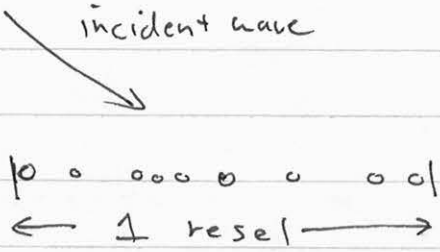
Therefore, in the Bragg model, we simply expand the surface to find the amplitude of the Fourier component resonant with our incident radiation. Eventually we could derive the following expression for σ^0 in the Bragg-dominated domain:

$$\sigma^0_{(pol)} = 8\left(\frac{2\pi}{\lambda}\right)^4 h^2 \cos^4(\theta)\ |\alpha_{pol}|\ W\left(\frac{4\pi}{\lambda}\sin\theta\right)$$

The subscript pol means that wave polarization affects the result through α_{pol}. Note that the cross-section is proportional to h^2, the rms surface height squared, and to the Fourier component of resonance $W\left(\frac{4\pi}{\lambda}\sin\theta\right)$.

Discrete scatterer model

Another model for larger incidence angles, which we'll use extensively when we consider interferometry, consists of a random distribution of discrete scatterers. Let each scatterer be an object with similar backscattering properties, and illuminate it with an incident wave:



The returns from each scatterer add coherently, but because the resel is >> than λ the phase is essentially random in the sum, and we can just add the powers to obtain σ^0.

Hence extending this argument we obtain

$$\sigma^0(\theta) = N \cdot \sigma_{1\,\text{scatterer}} \cdot \cos\theta$$

N: scatterers per unit ground area

$\cos\theta$: obliquity factor

$\sigma_{1\,\text{scatterer}}$: isotropic scatterers

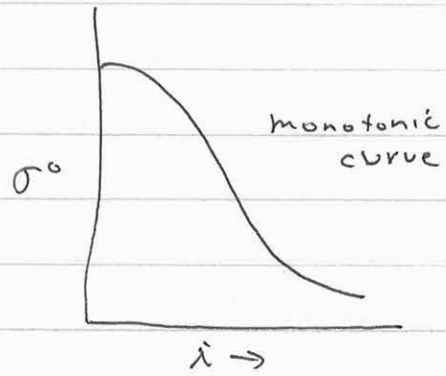
Or

$$\sigma^\circ(\theta) = N \cdot \sigma_{1\,\text{scatterer}} \cdot \cos^2\theta \quad \leftarrow \text{for } \cos(\theta) \text{ reradiation pattern}$$

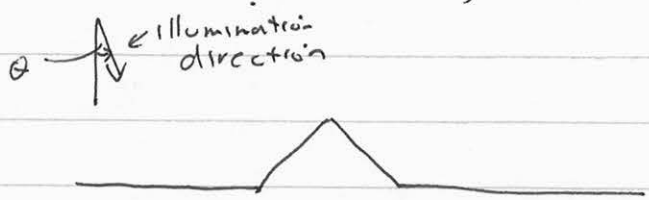
Both of these fall off more slowly than the Bragg model, which went as $\cos^4\theta$, so are applicable to even rougher surfaces than those ~~we~~ can model with Bragg scatter.

Slope modulation and topography

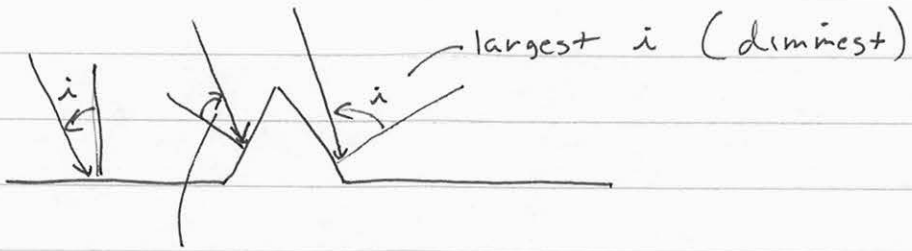
Consider our canonical σ°:



and image the following surface:



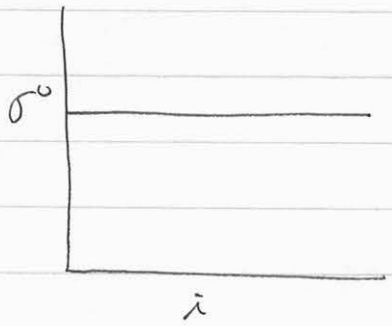
Let's determine the local incidence angles.



typically smaller i (brighter)

Thus this topographic feature would appear brighter on its near side and dimmer on its far side. This effect causes even subtle changes in topography to be visible in a radar image.

At very short wavelengths, where the surface is effectively extremely rough, σ^0 looks like



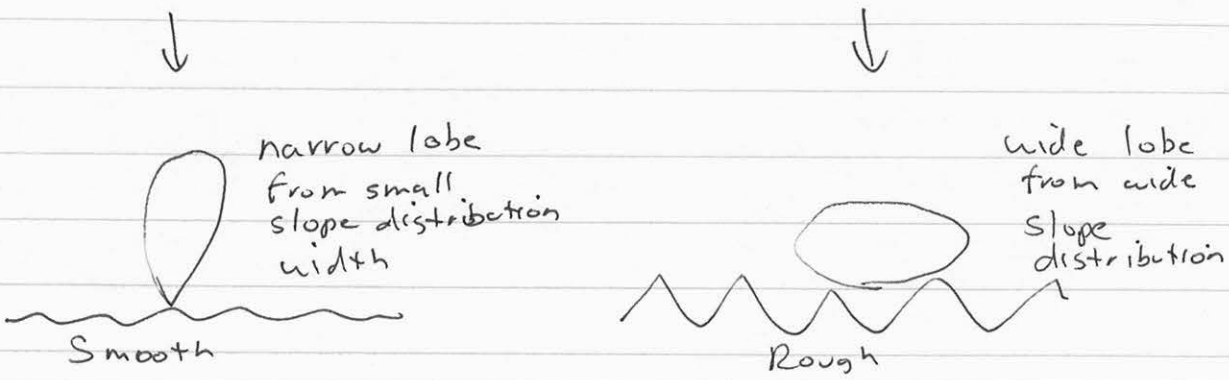
and sensitivity to topography decreases, or disappears. This happens often at optical wavelengths.

Surface roughness variations

Now we can understand two effects on radar echoes from roughness variations - increasing roughness decreases intensity from facet scattering and increases intensity from Bragg scattering?

Why is that?

Facet scatter (near normal)

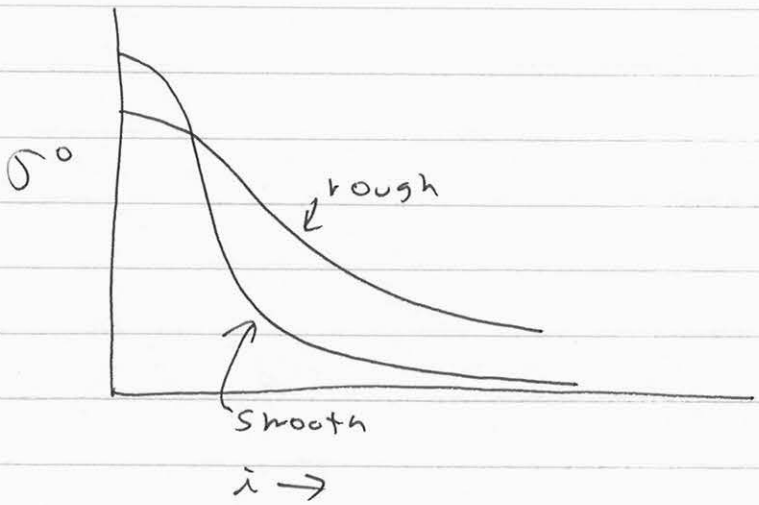


For Bragg, recall the Bragg law form:

$$\sigma^0_{pol} = 8\left(\frac{2\pi}{\lambda}\right)^4 h^2 \cos^4\theta \, \alpha_{pol} \, \omega\left(\frac{4\pi}{\lambda}\sin i\right)$$

σ^0 is proportional to h^2, so increased roughness leads to increased backscatter.

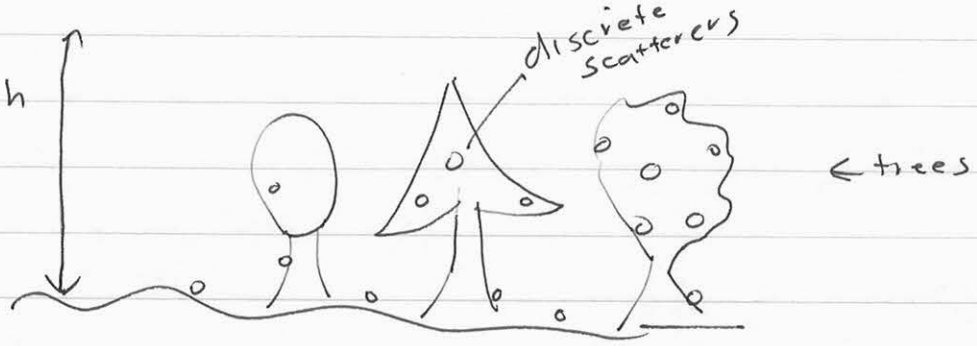
So, two surfaces might compare as follows:



Also, since topographic contrast depends on $\frac{\partial \sigma^\circ(\theta)}{\partial \theta}$, more shading is apparent in smoother areas than rougher areas.

Vegetation

Scattering from vegetation can be thought of as a form of scattering from an extremely rough surface:



These scatterers are widely distributed in a large volume, with equivalent h^2 very large. Hence we would expect little dependence on i in very dense canopies when ~~veg.~~ little or no energy penetrates to the ground. The sigma-zero curve is essentially flat.

Thus, one common method to calibrate radars without having to calculate incidence angle effects is to image a dense canopy such as the Amazon rain forest.

Wavelength Dependence of Backscatter

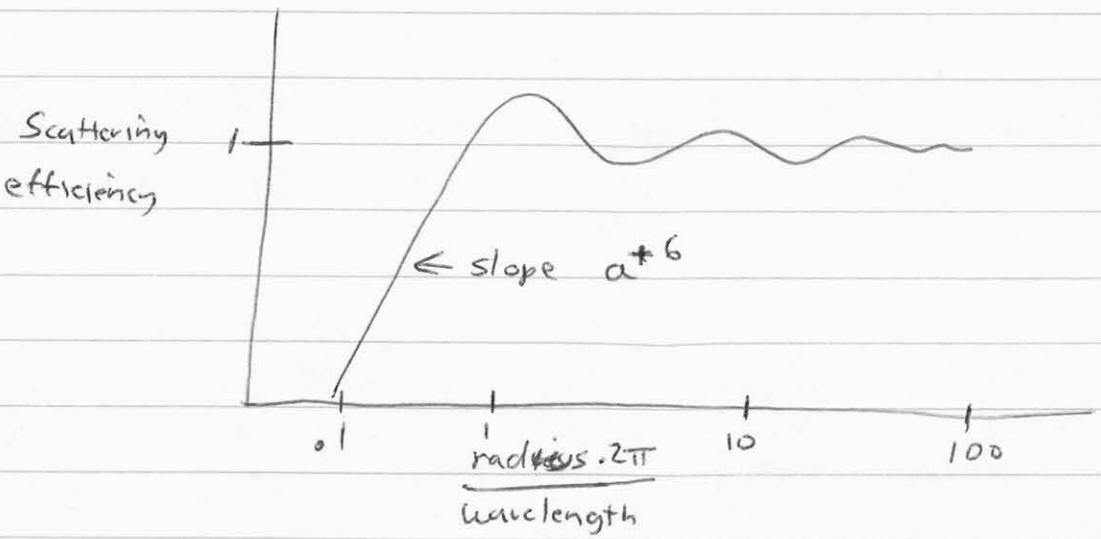
We often state that radar backscatter is dominated by structures similar in size to the radar wavelength. Thus use of multiple-wavelength radars permits investigation of a range of surface scatterer sizes.

The coupling between wavelength and scatterer size is self-evident for Bragg scatter, which matches the carrier λ to a particular Fourier component of the surface. But how does this relation evidence itself for other models and mechanisms?

Let's examine what happens in the case of discrete scatterers of various sizes.

Extinction / Scattering Cross-Section curves

The following curve relates scattering efficiency to particle size as a function of radius, in wavelengths:



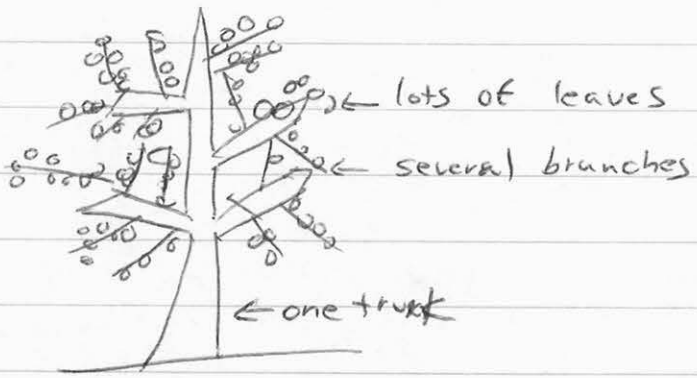
We interpret this curve as implying that the effective cross-section of a particle is equal to its physical cross-section

if the particle is comparable to a wavelength in size or larger, but falls off very steeply for smaller particles. In other words, the radar instrument is not very sensitive to sub-wavelength size particles.

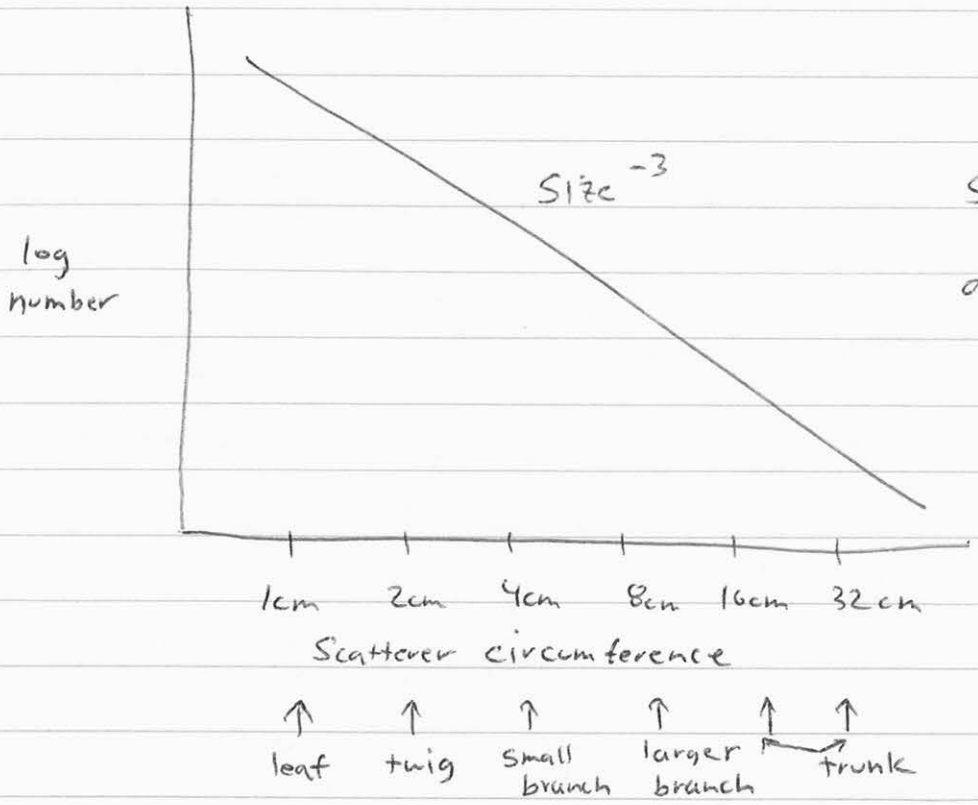
Incidentally, this is why the radar easily "sees" through clouds. The ice and water droplets in a cloud are much smaller than radar wavelengths, but are large compared to optical wavelengths. Thus the microwave radiation readily penetrates.

Scatterer size distribution

Now consider imaging a tree, which consists of a trunk, a few large limbs, many smaller branches and twigs, and a host of leaves.



Construct a table giving the number of each size scatterer and also the total cross-section of each, as well as the cross section normalized by scattering efficiency. Let's assume the number of scatterers of a given size obeys a power law distribution, falling off as size^{-3}, which is typical for some plants and many surface scatterer distributions.



Scatterer	Number on tree	Physical cross section	Effective cross-section $\lambda = 2$ cm	Effective cross-section, $\lambda = 8$ cm
Leaves	4096	$\frac{1024}{\pi}$	$\frac{32}{\pi}$	$\frac{1}{32\pi}$
Twigs	512	$\frac{512}{\pi}$	$\frac{512}{\pi}$ ← dominant	$\frac{1}{2\pi}$
Small branches	64	$\frac{256}{\pi}$	$\frac{256}{\pi}$	$\frac{8}{\pi}$
Larger branches	8	$\frac{128}{\pi}$	$\frac{128}{\pi}$	$\frac{128}{\pi}$ ← dominant
Trunk	1	$\frac{64}{\pi}$	$\frac{64}{\pi}$	$\frac{64}{\pi}$

Thus the return is dominated by λ-size scatterers!

GP265/EE255

Interferometric Techniques and Applications

Many applications of imaging radar involve interpretation of the radar amplitude data using photo-interpretive techniques and the same kinds of scattering models we have just discussed. We'll put these aside for now and concentrate instead on the measured phases of the radar echoes.

Since the phase of an individual radar echo is a statistical quantity that is uniformly distributed over the interval $(0, 2\pi)$ [for reasons we'll discuss below], it is only the phase difference between two measurements that we can readily measure. But measurement of phase differences is a well-developed concept in optics, denoted interferometry, and we can apply the same approaches to radar signals. This will allow us to measure surface shapes, or topography, as well as surface motions, useful for a great many geophysical investigations.

The radar distance measurement

When we measure a radar echo and form a radar image, we generate a high-resolution map of the 2-D brightness distribution of a surface. This distribution forms a "picture" of the terrain which we can interpret.

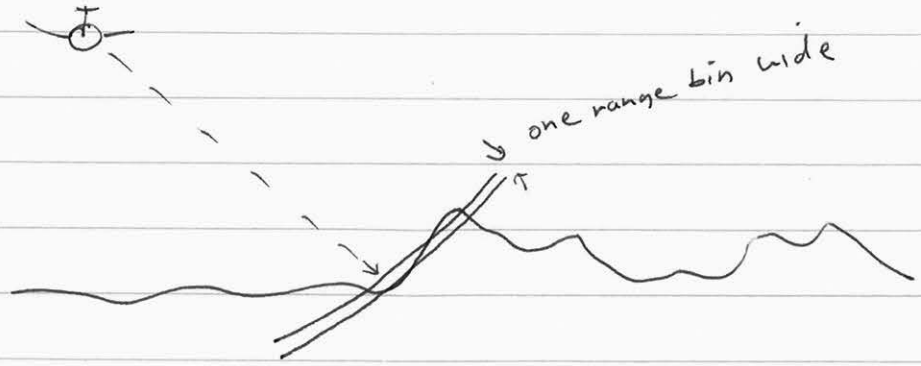
To form the image, say using the SAR algorithm, recall that we not only focus the image in the two dimensions but we remove any phase effects that follow from platform motion. Hence our image has a time delay associated with its distance (let's assume we use r_0 as a reference) but also the constant phase

$$\phi = \frac{-4\pi}{\lambda} r_0$$

Each pixel thus carries with it phase values that are related to the distance of that pixel from the radar:

Interferometry for topography

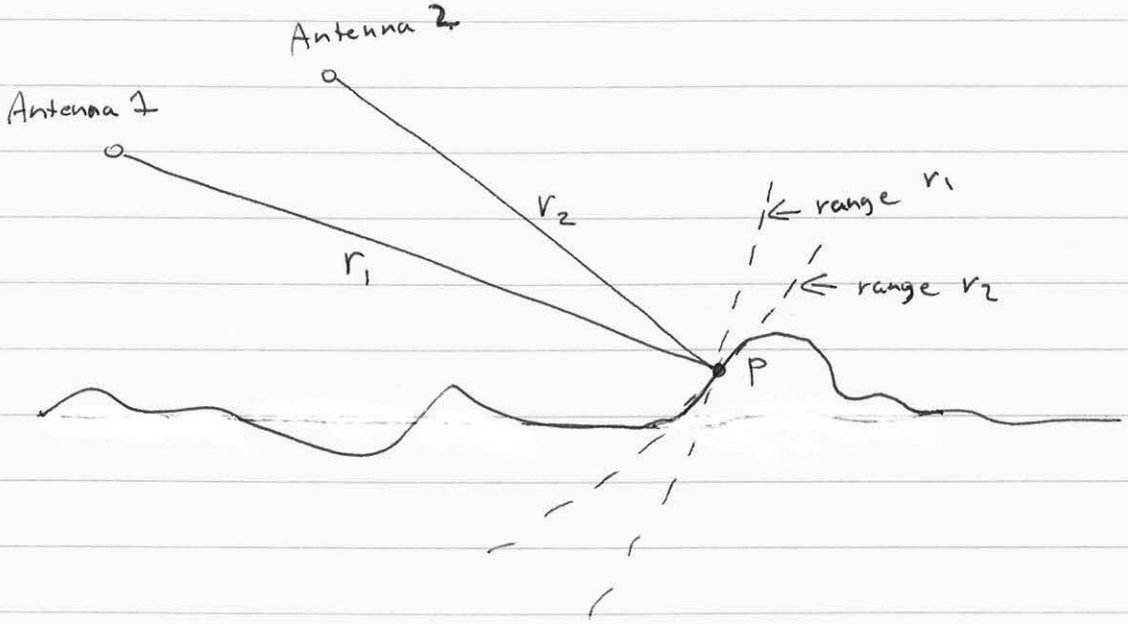
Consider imaging the following scene from a radar flying into the page:



Note that the return gives contributions from all points within a band the width of a range resolution element at the same time. Another way of thinking about this is that any scatterer located within that band would have the same time, Doppler, and phase history. In other words the radar is not at all sensitive to surface heights. Thus the radar image is two-dimensional in nature, and we can only infer topography from perceived changes in brightness.

Incidentally, ~~this~~ such a method of inferring topography from brightness is often used in radar image (or optical image) interpretation, and is called "shape from shading".

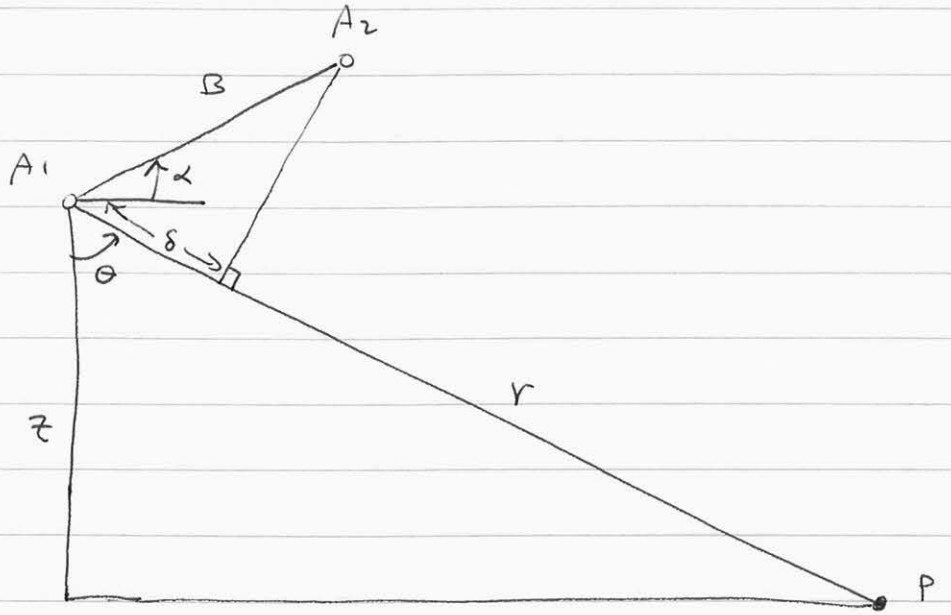
Now, let's change the geometry of the situation and add a second radar antenna. What are the distance measurements made by each to a point on the terrain?



The distance measurement r_1 from antenna 1 and r_2 from antenna 2 define two circles which intersect at the point P. If we could measure each then we could determine the 3-D coordinates of P.

If r_1 and r_2 are quite different, we can use this method of radar parallax (or radar stereo) for topographic mapping. But for $r_1 \approx r_2$ the method is not very sensitive to topography, and other considerations lead to operational difficulties.

Instead, consider the following implementation where we have made the assumption that $|r_1 - r_2| \ll |r_1|$, the so-called parallel-ray approximation:



Define a "baseline" B between the two antennas. Signals transmitted from A1 and received back at A1 travel a distance 2δ further than those transmitted and received at A2. (If we compare propagation lengths A1 – P – A1 and A1 – P – A2 the difference is only δ.)

Denoting the orientation of the ~~the~~ baseline with respect to horizontal as α, we have:

$$\delta = B \sin(\theta - \alpha)$$

and the height z:

$$z = r \cos\theta$$

Thus if we can measure δ we can infer the topographic height z. Now, how accurately can we do this? We need to evaluate the derivative

$$\frac{\partial z}{\partial \delta} = \frac{\partial z}{\partial \theta} \cdot \frac{\partial \theta}{\partial \delta}$$

We'll find it easier to evaluate the derivative in stages like this.

$$\frac{\partial z}{\partial \theta} = -r \sin\theta$$

$$\frac{\partial \delta}{\partial \theta} = B \cos(\theta - \alpha)$$

so $\frac{\partial z}{\partial \delta} = \frac{-r\sin\theta}{B\cos(\theta-\alpha)}$, and thus

$$\sigma_z = \frac{r\sin\theta}{B\cos(\theta-\alpha)}\,\sigma_\delta$$

Therefore, each error of one meter in δ leads to an error in z of about $\frac{r}{B}$ m, which can be large. Given the precision in range of one resolution element δ_r, the error in topography is about

$$\sigma_z = \frac{r}{B}\,\delta_r$$

which is why we would want to keep B very large.

But instead of comparing lengths by differencing ranges, let's use our knowledge of the phase of the returns. Then, if we were to compare phases we'd get the following additional relation:

$$\phi_1 - \phi_2 = \frac{-4\pi}{\lambda}\left(|r_1| - |r_2|\right)$$

or

$$\phi = \frac{-4\pi}{\lambda}\left(r_1 - r_2\right) = \frac{-4\pi}{\lambda}\,\delta$$

Now, we can typically measure phase to a few degrees or so, equivalent to measuring δ to about $\frac{1}{100}$ of a wavelength. Then

our relation for accuracy is about

$$\sigma_z = \frac{r}{B} \cdot \frac{\lambda}{100}$$

Even if $\frac{r}{B}$ is large, $\frac{\lambda}{100}$ is small so we get reasonable performance. Look at an ERS example:

$r = 800$ km

$B = 1000$ m

$\lambda = 6$ cm

$\delta_r = 9$ m

Stereo case: $\sigma_z \approx \frac{r}{B} \delta_r \approx 7200$ m ← Because of small B

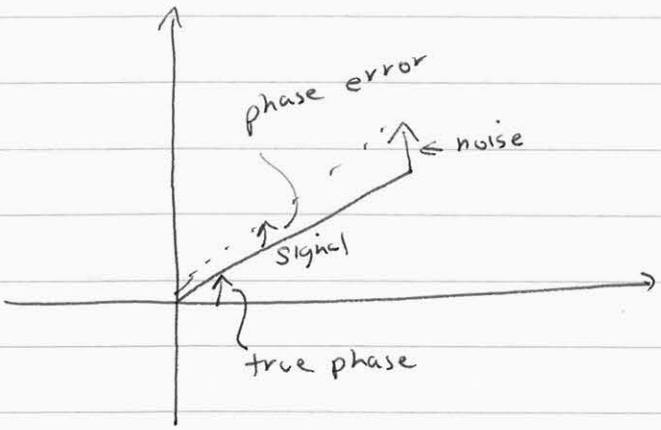
Interferometer case: $\sigma_z \approx \frac{r}{B} \frac{\lambda}{100} = \frac{1}{2}$ m

These are extreme examples but you see the power of the interferometric approach.

Accuracy of phase measurements

Clearly our interferometric performance depends critically on how well we can measure phase, which in turn depends on signal to noise ratios. We know how to estimate SNRs, so how does that map into phase accuracy?

Let's model an echo signal as consisting of a "signal" vector with unit length and a "noise" vector added to it: (These are both complex quantities)



Very approximately the phase error $\Delta\phi$ will be $\frac{|\text{noise}|}{|\text{signal}|}$ if the noise power is much less than the signal power. The variance of the noise σ_ϕ^2 will then be about

$$\sigma_\phi^2 \approx \frac{|\text{noise}|^2}{|\text{signal}|^2} = \frac{1}{SNR}$$

So, σ_ϕ in radians is about $\frac{1}{\sqrt{SNR}}$. If the SNR is 100, or 20 dB, $\sigma_\phi = 0.1$ radian $\approx 6°$.

Collecting much of the above together, we can summarize the above error analysis as:

$$\sigma_z = \frac{\lambda}{4\pi} \cdot \frac{r \sin\theta}{B\cos(\theta-\alpha)} \cdot \frac{1}{\sqrt{SNR}} \qquad \leftarrow \text{valid for high SNR}$$

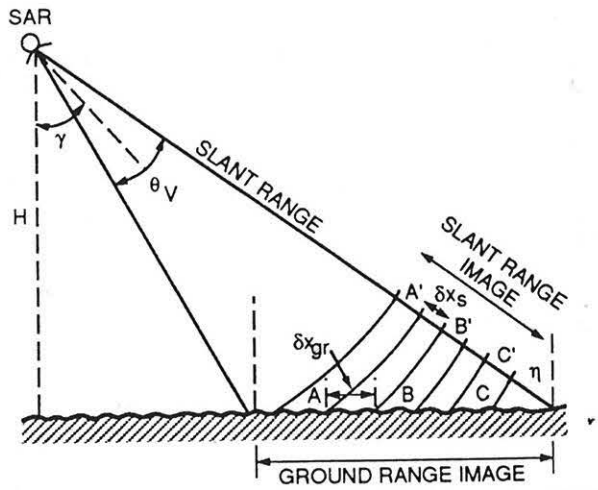
Cartographic applications

Once we have the interferometric technique available how can we use the data in cartographic studies? Let's briefly talk about geocoding and georeferencing, and then spend some time looking at several actual applications.

Geocoding and georeferencing (see chapter 8 of the text)

A typical SAR image is presented in a coordinate system where the range coordinate is slant range and azimuth is correct along-track distance. This leads to foreshortening or even layover in the resulting image. The issue is the nonlinear mapping of slant range into ground range, plus uncompensated topography.

Here are some illustrations of geometrical distortions: (from the text)



Definition of Slant-to-Ground Range Quantities

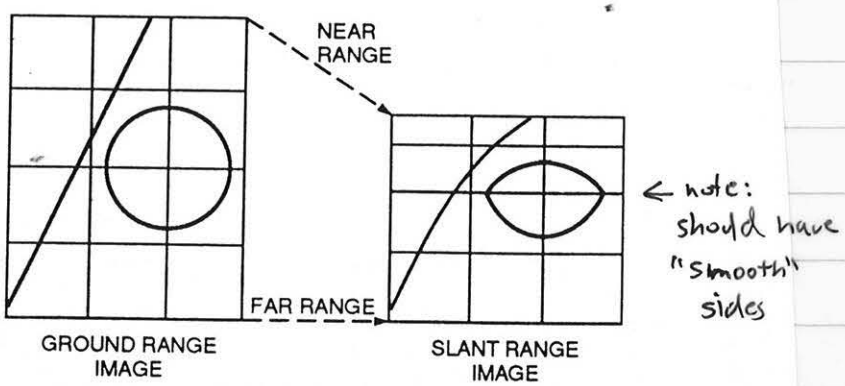
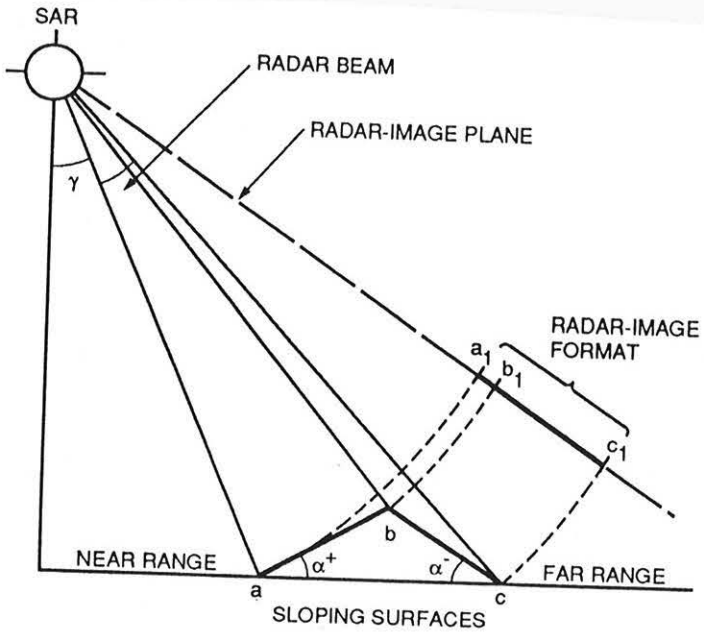


Figure 8.5 Relationship between slant range and ground range image presentation for a side looking radar.

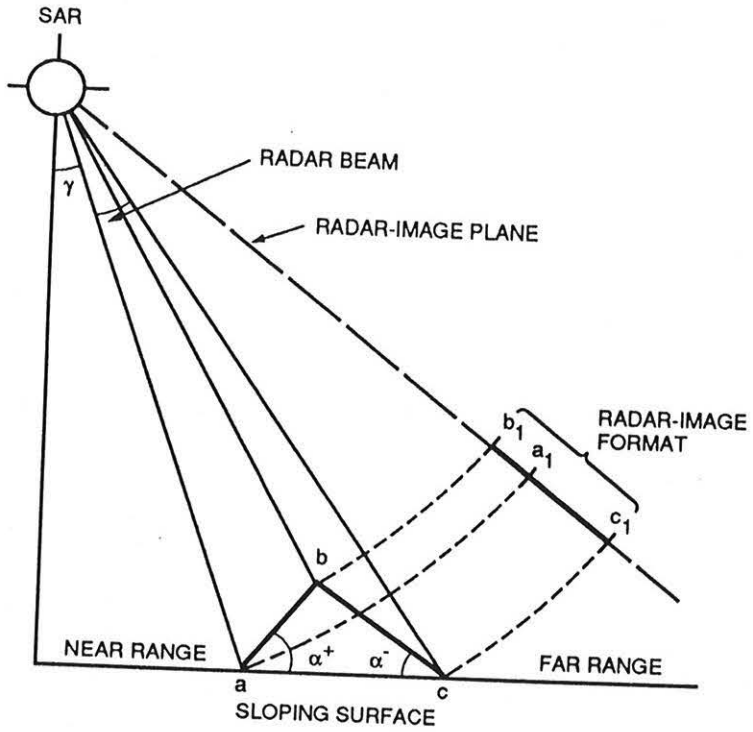
Foreshortening:



a

Extreme Foreshortening

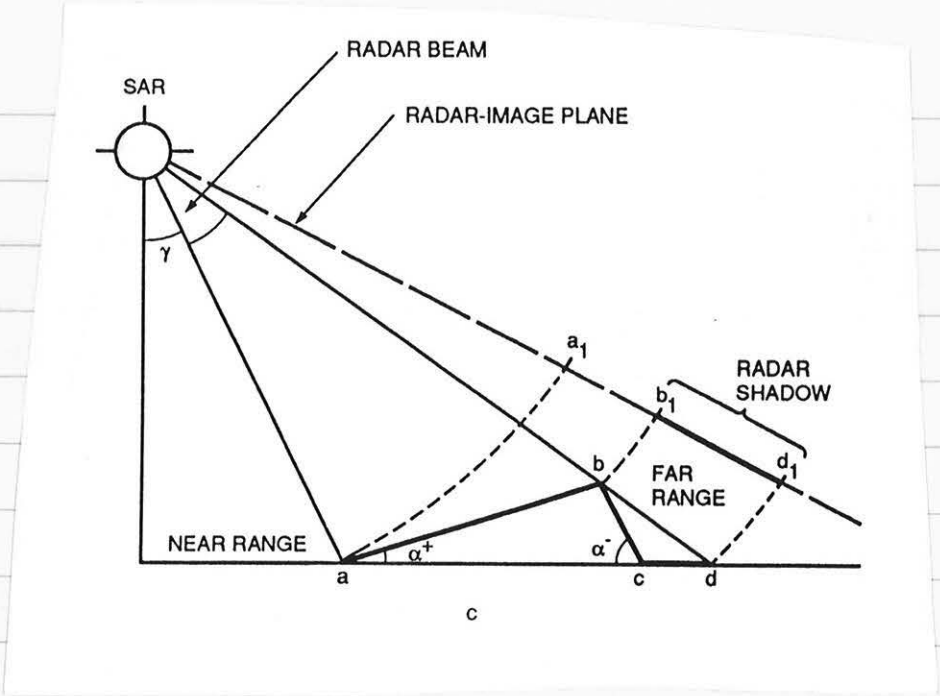
(Layover)



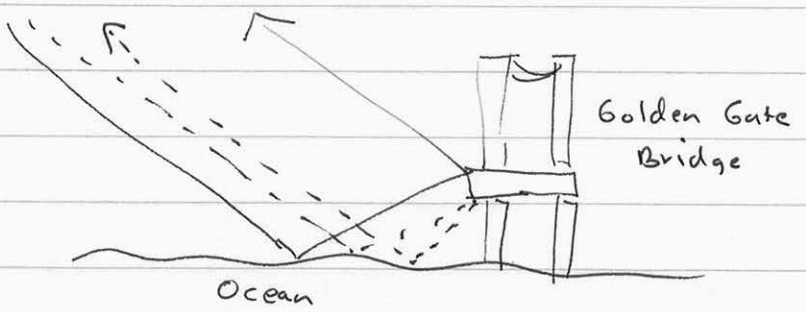
b

Figure 8.6 Geometric distortions in SAR imagery: (a) Foreshortening; (b) Layover; (c) Shadow; (d) A combination of imaging geometries illustrating secondary peak.

Shadowing



Multipath



Knowledge of the heights of the terrain allows us to correct for each of these effects. We know that

$$r_{ground}^2 + z^2 = r_{slant}^2$$

so if we resample according to

$$r_{ground} = \sqrt{r_{slant}^2 - z^2}$$

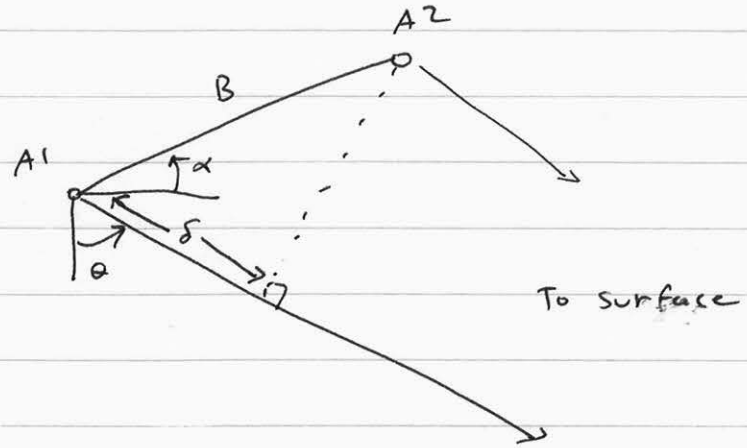
we'll get a "correct" image, at least as far as we can manage. Shadows and extreme foreshortening still provide distortion however.

GP 265 / EE 3.55

Forming the interferogram

We saw last time that if we implement an interferometer by using two antennas displaced across-track, then the phase of the difference signal from the two antennas contains information about the surface topography. Let's now examine in detail formation of such an interferogram.

From before, we had the following setup:



and we related the distance δ to the geometry by

$$\delta = B \sin(\theta - \alpha)$$

The path from A1 to the surface is δ longer than the corresponding path from A2, and δ depends on both the imaging geometry and the topography. This means that each pixel in the image from antenna A1 is displaced by 2δ in range from the corresponding pixel in the image from antenna A2. The total displacement is 2δ due to two-way propagation.

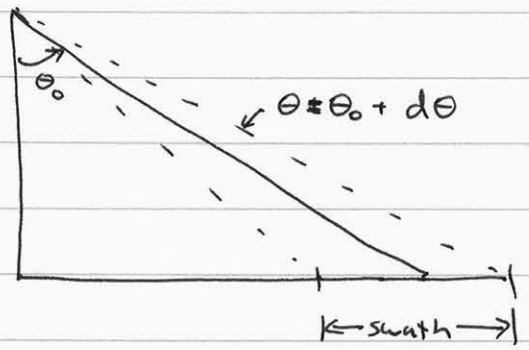
For the time being, let's neglect topography by assuming a flat surface, and investigate the form of the displacement across the swath.

Displacement vs. range

Starting with $\delta = B \sin(\theta - \alpha)$, we can relate θ to range and altitude through

$$z = r \cos\theta$$

We can solve for δ as a function of range, $\delta(r)$, by substituting the second expression into the former. But we will find it simpler to expand the result about a nominal look angle θ_0 first:



Hence

$$z = r \cos(\theta_0 + d\theta)$$

$$\approx r[\cos\theta_0 - \sin\theta_0 \, d\theta]$$

Also

$$\delta = B \sin(\theta_0 - \alpha + d\theta)$$

$$= B[\sin(\theta_0 - \alpha) + \cos(\theta_0 - \alpha) d\theta]$$

Eliminating $d\theta$:

$$z = r\cos\theta_0 - r\sin\theta_0\, d\theta$$

$$r\sin\theta_0\, d\theta = r\cos\theta_0 - z$$

$$d\theta = \frac{r\cos\theta_0 - z}{r\sin\theta_0}$$

and

$$\delta = B\sin(\theta_0-\alpha) + B\cos(\theta_0-\alpha)\,\frac{\cos\theta_0 - z/r}{\sin\theta_0}$$

$$= B\sin(\theta_0-\alpha) + \frac{B\cos(\theta_0-\alpha)}{\tan\theta_0} - \frac{B\cos(\theta_0-\alpha)}{\sin\theta_0}\,\frac{z}{r}$$

Now, expanding the last term about r_0 where $r = r_0 + dr$:

$$\frac{z}{r} = \frac{z}{r_0+dr}$$

$$= \frac{z}{r_0}\left(\frac{1}{1+\frac{dr}{r_0}}\right) \approx \frac{z}{r_0}\left(1-\frac{dr}{r_0}\right)$$

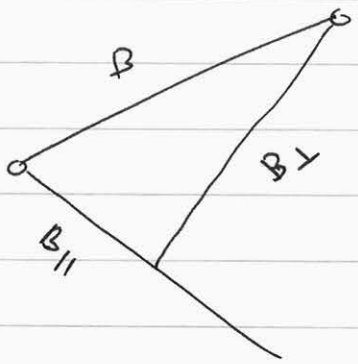
$$= \frac{z}{r_0} - \frac{z\,dr}{r_0^2}$$

So, for $\delta(dr)$ we get

$$\delta = B\sin(\theta_0-\alpha) + \frac{B\cos(\theta_0-\alpha)}{\tan\theta_0} - \frac{B\cos(\theta_0-\alpha)}{\sin\theta_0}\left(\frac{z}{r_0} - \frac{z}{r_0^2}dr\right)$$

$$= B\sin(\theta_0-\alpha) + \frac{B\cos(\theta_0-\alpha)}{\tan\theta_0}\,\frac{dr}{r_0}$$

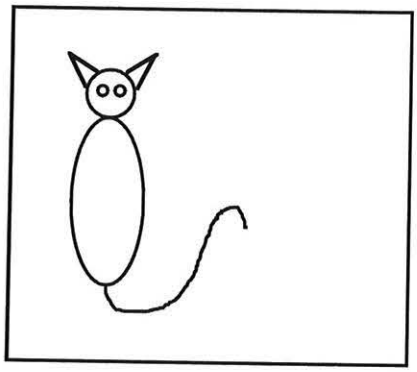
We can define the baseline at the nominal look angle θ_0 as having a parallel component $B_{\parallel} = B \sin(\theta_0 - \alpha)$ and a perpendicular compont $B_{\perp} = B \cos(\theta_0 - \alpha)$:



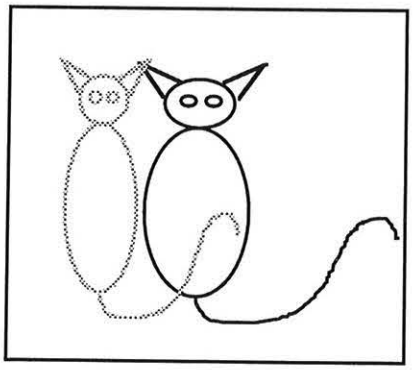
so $$\delta = B_{\parallel} + B_{\perp} \frac{dr}{r_0 \tan\theta_0}$$

All of this is a long way of saying that δ has a constant shift equal to the parallel component of the baseline, and a linearly increasing shift (stretch) dependent on $B_{\perp}$ and the geometrical factor $r_0 \tan\theta_0$.

A comparison of the images would then look like:



Antenna 1 image



Antenna 2 image - shifted and stretched

Now, forming the interferogram means relating the phase in a pixel in image 1 to the corresponding phase in ~~pixel~~ image 2. Thus, it is necessary to identify each pixel in image 2 in terms of its location in image 1.

We usually don't know the relationship for δ a priori, so we typically have to determine the function from our data. We do this by estimating the shift in image 2 by identifying corresponding points in the two images and solving for an equation for δ.

This is called <u>determining the interferometric offset field</u>.

<u>Offset determination</u>

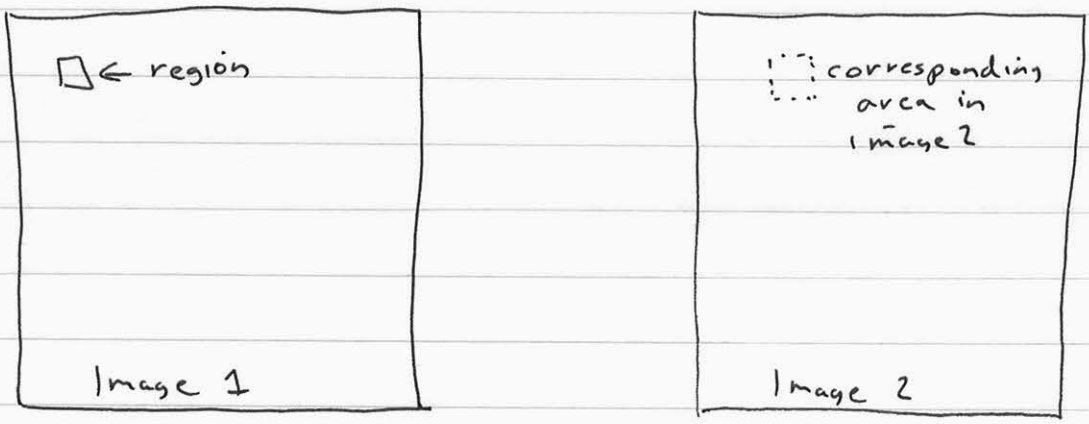
By comparing the positions of like scatterers in both interferometer images, we can derive empirically the relation

$$\delta = B_{\parallel} + B_{\perp} \frac{dr}{r_0 \tan\theta_0}$$

But since we almost always have noisy data this process is imperfect, and a noisy estimate of the offset results. We overcome this by estimating the offsets at a great many locations and solving for δ using least-squares approaches.

There are many ways to determine the offsets, and we'll discuss one here. Since in the general case we can have shifts both in range and azimuth we'll approach this as a 2-D problem.

Begin by identifying a small region within image 1:

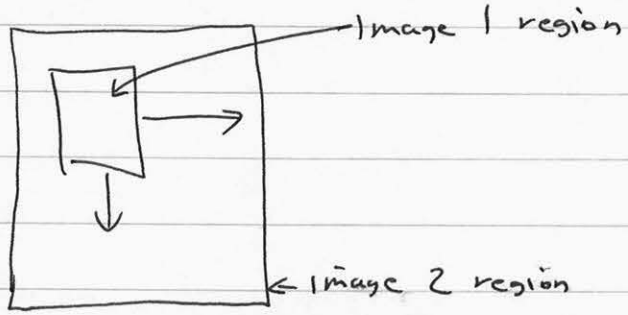


We know approximately, but not exactly, where the region would occur in Image 2. We can find the exact location by calculating the cross-correlation

$$\text{region}_{\text{Image 1}} \star\star \text{region}_{\text{Image 2}}$$

and determining the cross-correlation peak location. The location of the peak is the offset we seek.

Because we do not know the offsets exactly, we need to choose a larger region in Image 2 than in Image 1, and then cross-correlate:

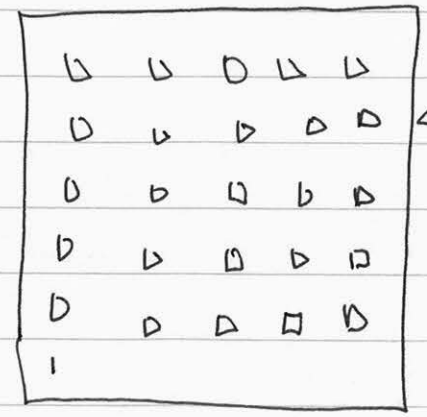


Now, the images are complex quantities when delivered from the processor. To find the peaks, however, we get better performance if we correlate the powers rather than the complex values. This eliminates additional noise that may be due to poor phase

fidelity in some regions.

In addition, because we are multiplying two detected images together we effectively have a cross product (in the correlation process) with twice the bandwidth of the original data. This could lead to aliasing if the original data were not oversampled by at least a factor of 2, which is almost never the case. Hence, for best results we would want to interpolate each region by a factor of 2 before the correlation.

Finally, we want to look over the entire image to get the best set of offset estimates. A good rule of thumb is to examine 5-10 locations in range and also in azimuth:



Once we determine all the correlations, we are ready to determine δ empirically. Say we plot the range offset as a function of range pixel number:

Note that the intercept is defined at the reference pixel, that is the one defined by θ_0.

From the measured slope of the line and its intercept, we solve for

$$\delta = B_{\parallel} + B_{\perp} \frac{dr}{r_0 \tan\theta_0}$$

effectively giving us the baseline components $B_{\parallel}$ and $B_{\perp}$.

We can later refine these estimates using known topographic corrections, but for now they are close enough.

The next step is to resample image 2 to the image 1 coordinates using the above relation for δ. In addition if there is an azimuth shift due to timing considerations we would remove that also at this time.

Finally, we form the interferogram itself by computing

$$i(x,y) = \text{image 1}(x,y) * \text{conjugate}(\text{image 2}(x+\delta, y + \text{az-offset}))$$

where the complex conjugate multiplication results in a phase differencing operation as desired.

Summary

Our interferogram formation is thus summarized as follows:

1) Process images from each antenna of about the same ground area.
2) Measure the offsets:
   a) Select a small region in image 1 and its

approximate neighbor in image 2.

c) detect complex values to obtain power images of small regions

b) interpolate by 2

d) cross-correlate and record offset peak location

e) repeat a-d for many locations

f) fit line in range dimension to offsets. Also determine azimuth shift, if any.

g) solve for $B_{\parallel}$ and $B_{\perp}$ from line parameters

3) Resample image 2 to image 1 coordinates

4) Cross multiply image 1 by conjugate of image 2.

Multilooking

Once we have the interferogram, we often want to multilook in order to improve our phase estimate accuracy at the expense of spatial resolution. This also reduces the required data volume for storage.

For the image data, recall that we average signal powers rather than complex values. But for the interferogram we need the average phase, so we average in this case the complex values.

In the image case, the phases were uniformly distributed and hence random from point to point, so we averaged power quantities. But for the interferogram the phase changes relatively slowly so coherent integration helps us.

Correlation

Another important measure of interferogram ~~quality~~ quality is the correlation of the two images, or how closely the phase in one image tracks that of the other. This quantity varies from 0 for images completely independent of each other to 1 for identical phase differences. We can calculate the correlation at many points in an image when we generate the multilooked image, according to the following definition

$$c = \frac{\sum \text{image1}_i \, \text{image2}_i^*}{\sqrt{\sum \text{image1}_i \, \text{image1}_i^*} \sqrt{\sum \text{image2}_i \, \text{image2}_i^*}}$$

where image1_i and image2_i are the complex values of a corresponding point in image 1 and image 2. Note that the numerator is simply the interferogram and the denominator is the product of the image amplitudes, not powers.

We can display correlations as a function of position and can evaluate how the correlation varies as a function of position. We'll discuss this more next time.

Decorrelation

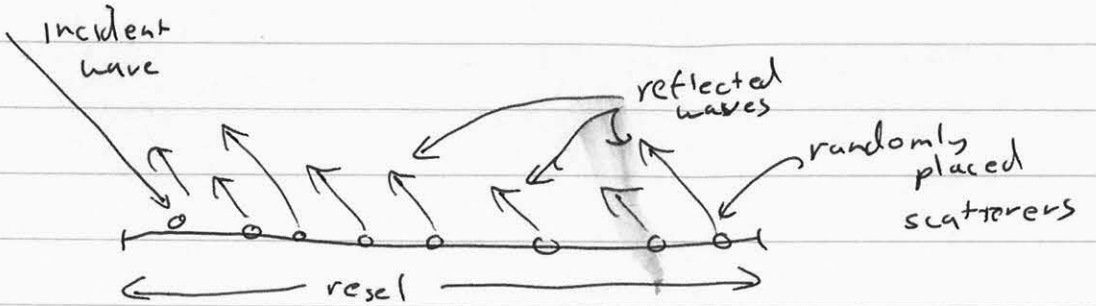
We mentioned, almost in passing, the correlation of a radar interferogram last time. In fact correlation is quite fundamental to understanding and interpreting interferometric measurements.

Correlation is closely coupled with our ability to estimate interferometric phase accurately, and hence use the data in practical applications. So, we will now take a closer look at it and its various sources in a radar system.

We will consider radar echoes to be "correlated" with each other if the measured phase and amplitudes coincide and hence represent the "same" interaction of radar signals and scattering terrain. In an imaging radar, this means that the observed "speckle" patterns are nearly the same. So first we need to understand speckle in a radar image.

Speckle

Speckle is the observable "graininess" in a radar image that is due to EM interactions rather than actual variations in σ° of a surface. Consider the echo from a single resel where we use the discrete scattering model from a handful of scatterers:



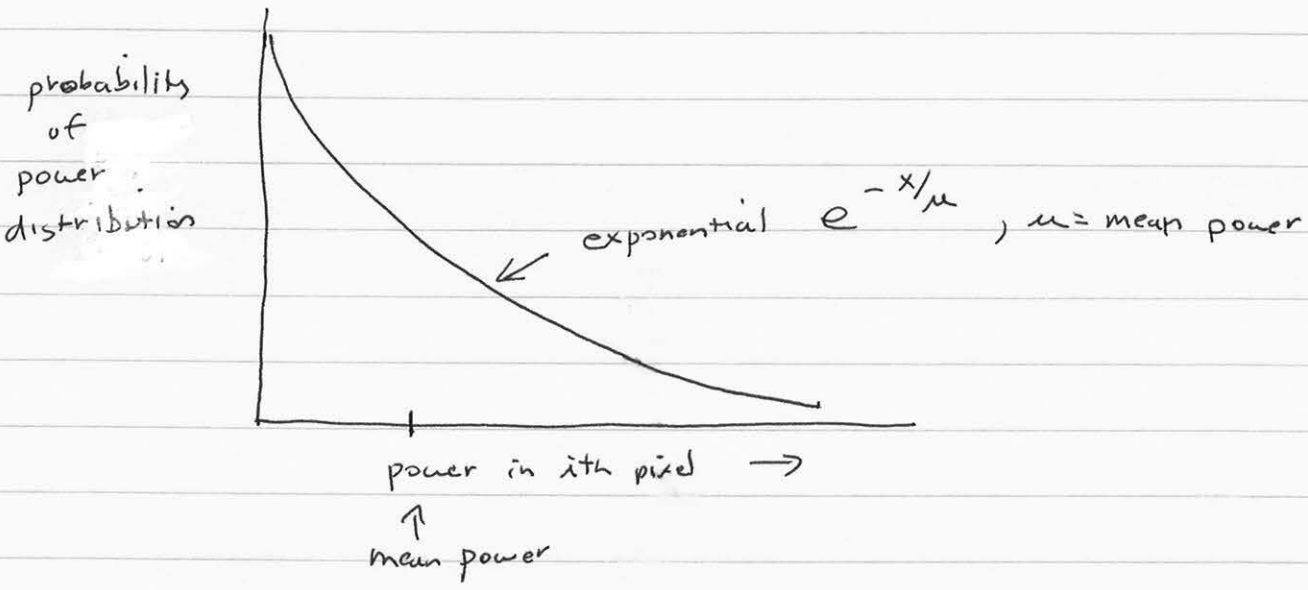
The received echo is the coherent sum of the reflected waves:

$$r(t) = \sum_{n \text{ scatterers}} a_i e^{-j \frac{4\pi}{\lambda} r_i}$$

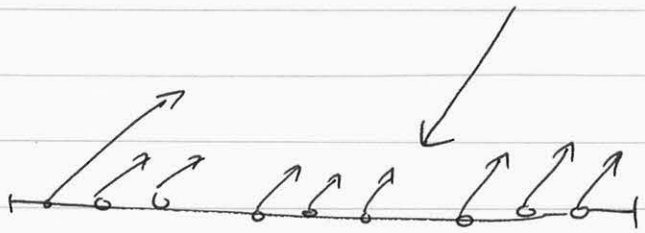
where a_i is the amplitude of the ith scatterer and r_i is the range to the ith scatterer.

If all the r_i are equal, the signals add in phase (coherently) for a large echo. In our previous figure, assuming a flat surface, that would be the case and we'd get a strong specular return.

But, since the resel size is typically much greater than a wavelength, the phase associated with r_i is essentially uniformly distributed over the interval $(0, 2\pi)$. Hence the waves will all interfere with each other, resulting in some pixels with a great degree of cancellation (near zero return) and some pixels where a lot of coherent addition occurs (bright return). The average power is well-related to the number of scatterers and the cross-section of each, but the realization of any scattering event is Rayleigh distributed in amplitude, or exponential in power.



Now, suppose we image the same set of scatterers from an entirely different viewing angle:



All of the distances r_i are different, hence each scatterer adds with a different phase than before. Thus, even though the same set of scatterers is illuminated, the phase will be entirely different. Thus changing viewing angle by too much makes the interferometric phase all noise, and no coherent analysis is possible.

Decorrelation theory

Consider two radar signals s_1 and s_2 acquired by two interferometric antennas, very near each other but going through separate receivers. We can model the signals as consisting of a common signal component c and individual noises n_1 and n_2:

$$s_1 = c + n_1$$

$$s_2 = c + n_2$$

Then the correlation by our previous formula is

$$\rho_{thermal} = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}}$$

where $\langle\ \rangle$ denotes ensemble averaging and the subscript "thermal" shows the noise is due to thermal properties.

Because c, n_1, and n_2 are uncorrelated random variables,

$$\rho_{thermal} = \frac{|c|^2}{|c|^2 + |n|^2}$$

Using $SNR = \frac{|c|^2}{|n|^2}$,

$$\rho_{thermal} = \frac{1}{1 + \frac{1}{SNR}}$$

which gives the relation of correlation to SNR.

Next, we introduce some additional decorrelation by changing the viewing angle slightly. As we discussed above, the individual echoes traveling distances r_i will add slightly differently, yielding a slightly different sum. We'll assume for now that the viewing angle changed only slightly, and we can model the two signals in this case as

$$s_1 = c + d_1 + n_1$$

$$s_2 = c + d_2 + n_2$$

where c again is the correlated part of the signal, n_i are the thermal noise components, and d_i represent additive "noise" due to the change in viewing direction. For reasons apparent later, we will call d_i the ~~spatial~~ spatial decorrelation terms, since they involve spatial movement of the antennas.

Using the same argument as above, consider the effective SNR from spatial decorrelation with no thermal noise term present:

$$\rho_{spatial} = \frac{|c|^2}{|c|^2 + |d|^2}$$

where we now have the subscript "spatial" instead of "thermal."

If we also include the thermal effects:

$$\rho_{spatial+thermal} = \frac{|c|^2}{|c|^2 + |d|^2 + |n|^2}$$

since c, d_i, and n_i are uncorrelated rvs. Let's rewrite as

$$\rho_{spatial+thermal} = \frac{|c|^2}{|c|^2 + |d|^2} \cdot \frac{|c|^2 + |d|^2}{|c|^2 + |d|^2 + |n|^2}$$

and, defining SNR as the ratio of all non-thermal to thermal powers, or $SNR = \frac{|c|^2 + |d|^2}{|n|^2}$

$$\rho_{spatial+thermal} = \rho_{spatial} \cdot \rho_{thermal}$$

$$= \frac{|c|^2}{|c|^2 + |d|^2} \cdot \frac{1}{1 + \frac{1}{SNR}}$$

Thus the total correlation is the product of the individual correlations.

Finally, we want to consider correlation when the two surfaces are imaged at different times. Even though the backscattered phase is a random quantity because it is the sum of the many echoes, each at distance r_i, if the scene area is unchanged between

observation times, and the radar is at the same location, each r_i is unchanged and the echoes will be completely correlated. If the surface has changed slightly, though, the echoes will be a little different.

How might a surface change between observations?

<DISCUSSION QUESTION>

We can model the decorrelation from slightly changed surfaces by introducing a term in the echo corresponding to surface change. A little thought and repetition of the previous argument leads to a formula of the form

$$\rho_{total} = \rho_{thermal} \cdot \rho_{spatial} \cdot \rho_{temporal}$$

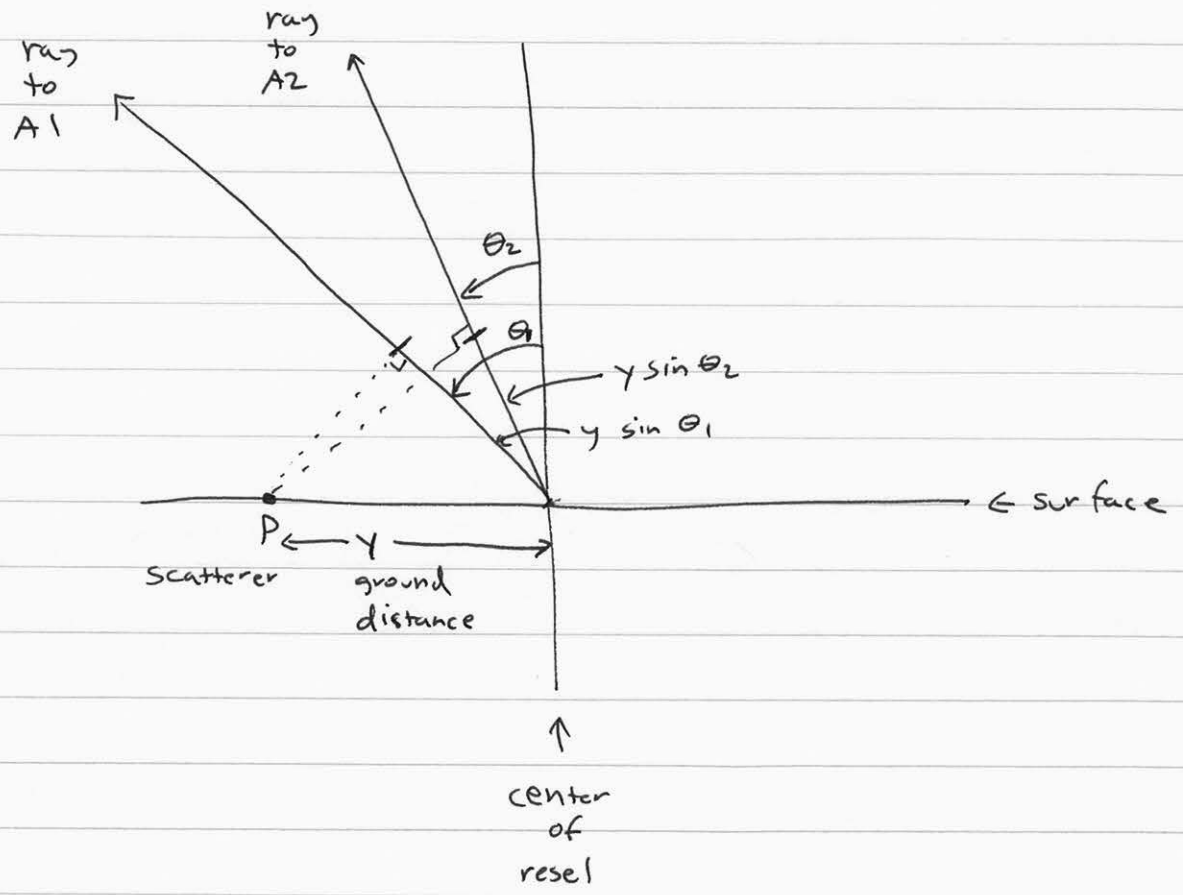
where $\rho_{temporal}$ represents decorrelation from a change in the surface between observations. $\rho_{spatial}$ referred to a change in sensor location, while $\rho_{thermal}$ modeled electronic changes inside the radar system.

Calculation of Spatial Baseline Decorrelation

Let's quantify the above discussion of $\rho_{spatial}$ and determine just how close the two interferometer antennas need to be ~~location~~ located for accurate phase estimation.

We begin by deriving a Fourier transform relation between the radar impulse response and the correlation as a function of difference in viewing angles.

Here's our imaging geometry:



Consider the phase of the reflection from P as compared to the reflection from the center of the resel. The phase from P is advanced by $y \sin\theta_i$ for the two antennas. If we model the total signal s_1 at antenna 1 as the integral over all of the resolution element, then

$$s_1 = \iint f(x-x_0, y-y_0) \exp\left(-j \frac{4\pi}{\lambda}(r + y\sin\theta_1)\right) \cdot w(x,y)\, dx\, dy + n_1$$

where $f(x,y)$ represents the complex backscatter from each point on the surface, and $w(x,y)$ is the system impulse response.

In other words, we simply add all the contributions from the points in the resel weighted by the impulse response - our usual linear system theory.

Similarly,

$$s_2 = \iint f(x-x_0, y-y_0) \exp\left(-j \frac{4\pi}{\lambda}(r + y\sin\theta_2)\right) w(x,y)\, dx\, dy + n_2$$

In both cases r is the range to the center of the pixel -- we'll let it be identical in s_1 and s_2. If this were not the actual configuration, only the mean phase of the return would change, but our conclusion below would remain unchanged.

Now, form the interferogram by computing $s_1 s_2^*$:

$$s_1 s_2^* = \iiiint f(x-x_0, y-x_0) f^*(x'-x_0, y'-y_0) \cdot \exp\left(-j \frac{4\pi}{\lambda} y(\sin\theta_1 - \sin\theta_2)\right)$$

$$\cdot w(x,y)\, w^*(x',y')\, dx\, dy\, dx'\, dy'$$

Now, if the surface scatterers are arranged randomly, uniformly distributed and uncorrelated,

$$\langle f(x,y) f^*(x',y') \rangle = \sigma^\circ \delta(x-x', y-y')$$

and we can reduce the above four-fold integral to

$$\langle s_1 s_2^* \rangle = \sigma^\circ \iint e^{-j \frac{4\pi}{\lambda} y \cos\theta\, d\theta} |w(x,y)|^2 dx\, dy$$

where θ is the average of θ_1 and θ_2 and $d\theta$ is their difference. Since the exponential kernel is linear in y, it can be viewed as a scaled Fourier transform relating the correlation function $\langle s_1 s_2^* \rangle$ to the power impulse response $|w(x,y)|^2$.

If the impulse response is approximated as a sinc function in both dimensions, usually a good approximation for high time bandwidth chirped systems, the transform may be evaluated and we find

$$\rho_{spatial} = 1 - \frac{2\cos\theta\,\delta_y\,d\theta}{\lambda}$$

where δ_y is the ground range resolution $\left(\frac{\delta_r}{\sin\theta}\right)$.

Since

$$d\theta = \frac{B_\perp}{r}$$

we can also use

$$\rho_{spatial} = 1 - \frac{2\cos\theta\,B_\perp\,\delta_y}{\lambda r}$$

Note that this is a linearly decreasing function of $B_\perp$, going from 1 at $B_\perp = 0$ (perfect correlation) to zero at a "critical" value

$$B_c = \frac{\lambda r}{2\cos\theta\,\delta_y}$$

If the perpendicular baseline $B_\perp$ approaches or exceeds B_c, no correlation is observed. This quantifies our observation that the interferometer antennas must be located near each other.

Rotation decorrelation

If the radar flight tracks are rotated with respect to each other between observations, a second source of decorrelation due to imaging geometry occurs, and can be quantified by

$$\rho_{rotation} = 1 - \frac{2 \sin\theta \, d\phi \, \delta_{az}}{\lambda}$$

(See Decorrelation paper for derivation)

Temporal Decorrelation

Finally, we can model the temporal decorrelation term as three-D motions of each of the individual scatterers, leading to the impressive-looking integral

$$s_1 s_2^* = \int\int\int\int\int\int f(x - x_0, y - y_0, z - z_0) f^*(x' - x_0, y' - y_0, z' - z_0)$$
$$\exp\{-j\frac{4\pi}{\lambda}(\delta y \sin\theta + \delta z \cos\theta)\} W(x,y) W^*(x',y') dx dy dz dx' dy' dz'$$

$$\langle s_1 s_2^* \rangle = \sigma^0 \int\int \exp\{-j\frac{4\pi}{\lambda}(\delta y \sin\theta + \delta z \cos\theta)\} p_y(\delta y) p_z(\delta z) d\delta y d\delta z$$

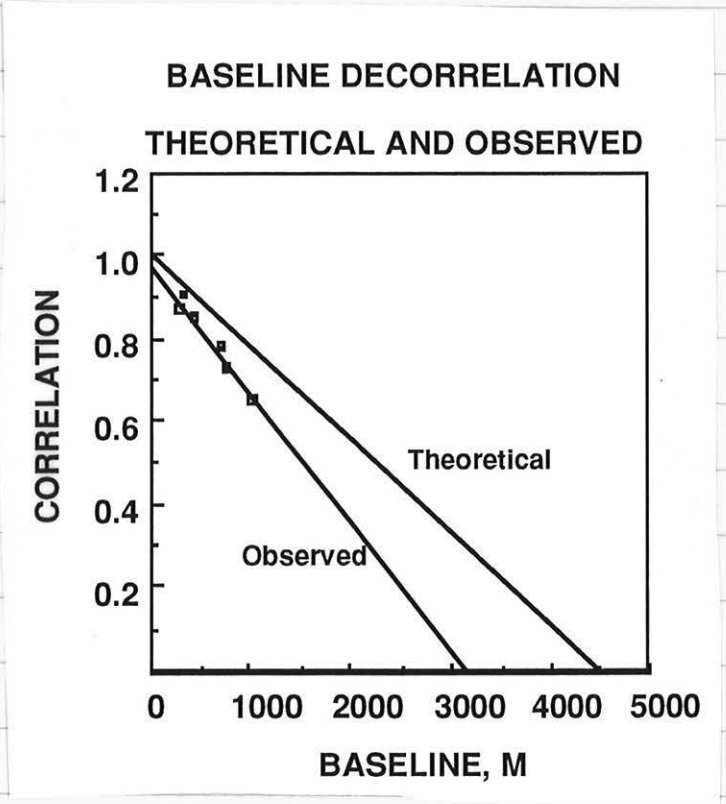
($p_y(\delta_y)$ and $p_z(\delta_z)$ are probability distributions of motions in y and z directions)

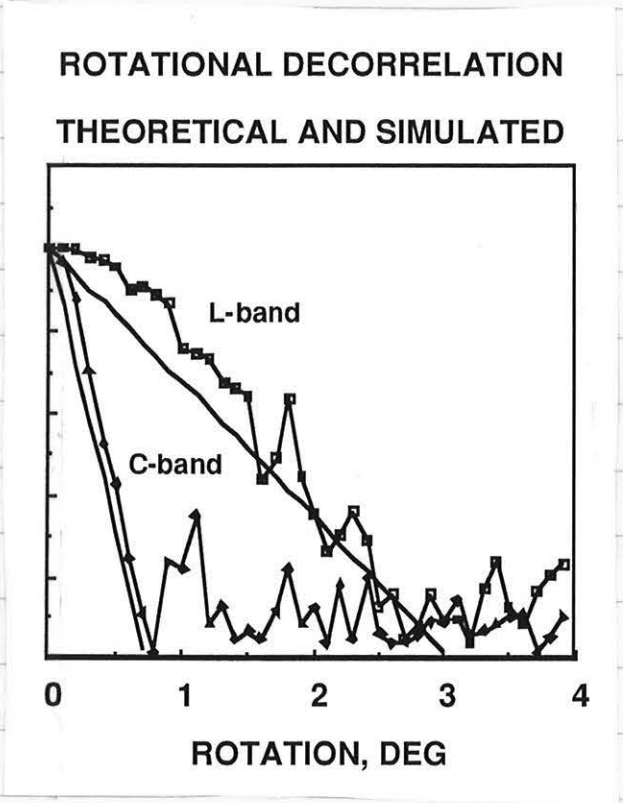
and its solution

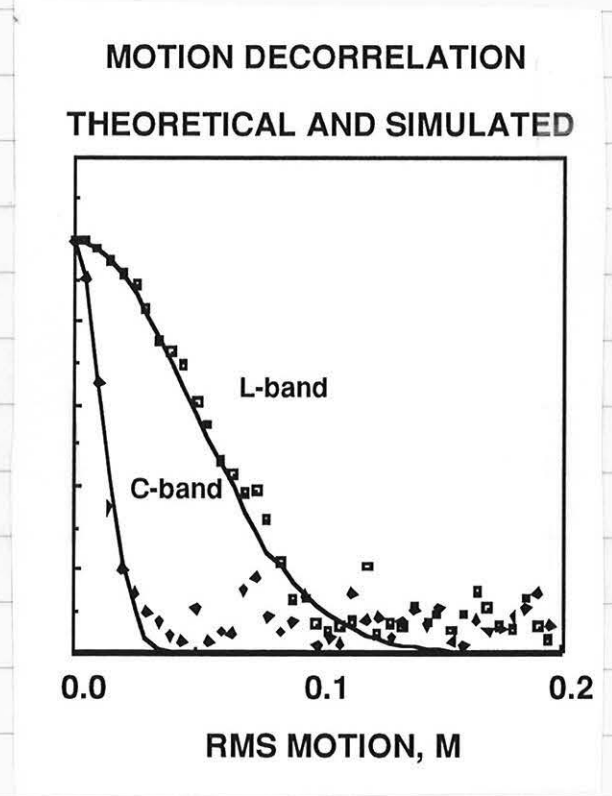
$$\rho_{temporal} = \exp\left(-\frac{1}{2}\left(\frac{4\pi}{\lambda}\right)^2 (\sigma_y^2 \sin^2\theta + \sigma_z^2 \cos^2\theta)\right)$$

where σ_y^2 and σ_z^2 are variances of motions in y and z.

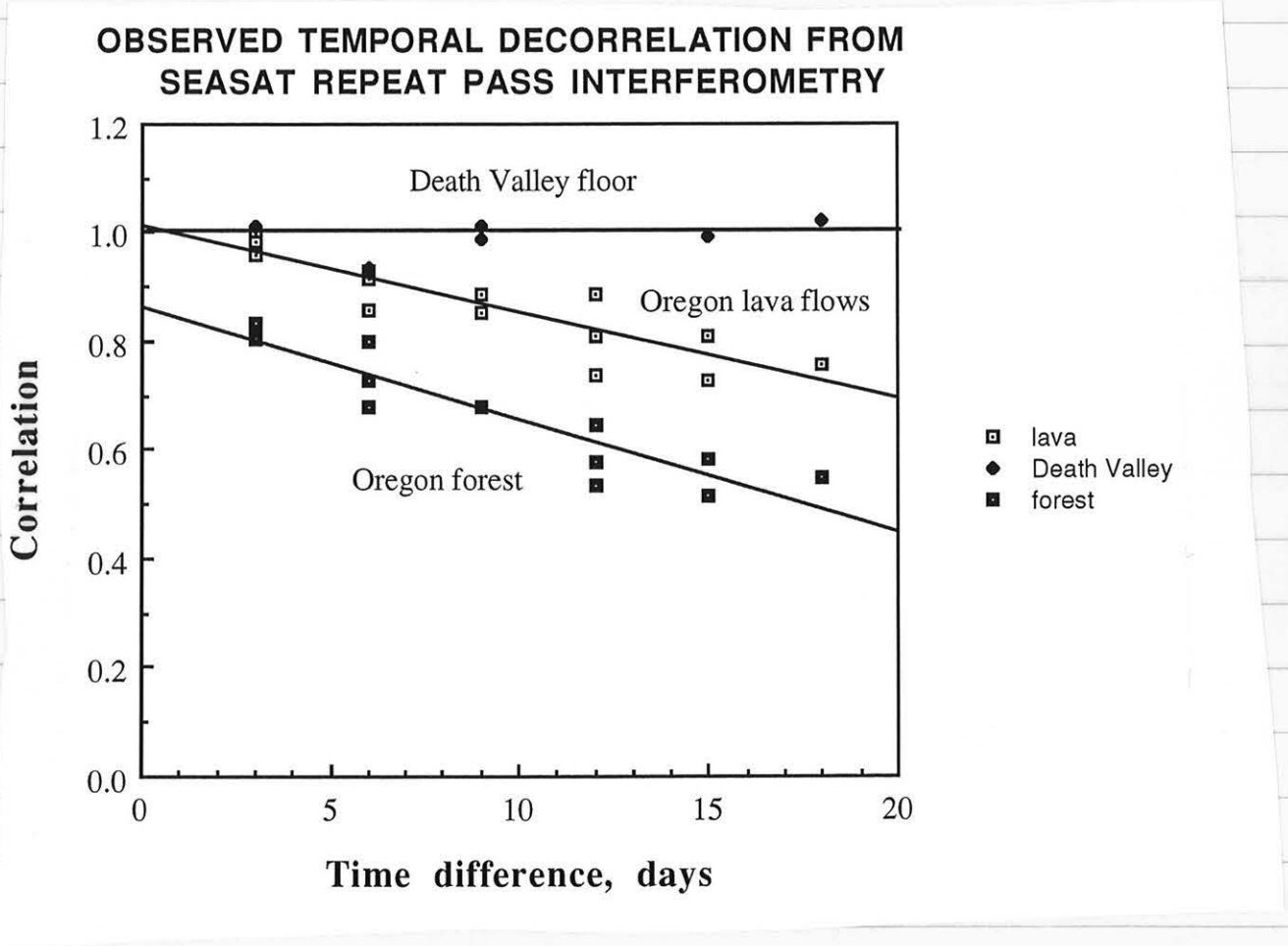
Sample data and simulations of these are given below:







Finally, here is an observed determination of temporal decorrelation at L-band:



Decorrelation in Interferometric Radar Echoes

Howard A. Zebker, *Senior Member, IEEE*, and John Villasenor, *Member, IEEE*

Abstract—**A radar interferometric technique for topographic mapping of surfaces promises a high resolution, globally consistent approach to generation of digital elevation models. One implementation approach, that of utilizing a single synthetic aperture radar (SAR) system in a nearly repeating orbit, is attractive not only for cost and complexity reasons but also in that it permits inference of changes in the surface over the orbit repeat cycle from the correlation properties of the radar echoes. Here we characterize the various sources contributing to the echo correlation statistics, and isolate the term which most closely describes surficial change. We then examine the application of this approach to topographic mapping of vegetated surfaces which may be expected to possess varying backscatter over time. We find that there is decorrelation increasing with time but that digital terrain model generation remains feasible. We present such a map of a forested area in Oregon which also includes some nearly unvegetated lava flows, and find that temporal decorrelation contributions to the height errors may be limited to 1.5 and 2.6 m for the forested and lava areas, respectively, if suitable attention is given to experiment design. Such a technique could provide a global digital terrain map.**

I. Introduction

INTERFEROMETRIC radar has been been proposed and successfully demonstrated as a topographic mapping technique by Graham [1], Zebker and Goldstein [2], and Gabriel and Goldstein [3]. A radar interferometer is formed by relating the signals from two spatially separated antennas; the separation of the two antennas is called the baseline. The spatial extent of the baseline is one of the major performance drivers in an interferometric radar system— if the baseline is too short the sensitivity to signal phase differences will be undetectable, while if the baseline is too long additional noise due to spatial decorrelation corrupts the signal. The theory of spatial baseline noise has previously been described by Li and Goldstein [4], and by Rodriguez and Martin [5], and Li and Goldstein have also shown some experimentally measured determinations of the spatial decorrelation noise level. In this paper we will review that work, develop Fourier transform relations between radar impulse response and the baseline and rotation-induced decorrelation functions, and utilize the results in separating the effects of temporally and spatially induced decorrelation. We then will produce a topographic map of a heavily forested area in Oregon, and assess its accuracy considering both spatial and temporal decorrelation. Finally, we will speculate on the utility of the correlation measurement itself as a remote sensing observable.

Manuscript received August 12, 1991; revised February 20, 1992. The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The authors are with the Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive Pasadena, CA 91109.

IEEE Log Number 9200760.

Two distinct implementation approaches have been discussed for topographic radar interferometers; they differ in how the interferometric baseline is formed. In the first case the baseline is formed by two physical antennas which illuminate a given area on the ground simultaneously— this is the usual approach for aircraft implementations where the physical mounting structures may be spaced for sufficient baseline. This is the approach used by Zebker and Goldstein [2] for the NASA CV-990 radar, and it is also currently used in the TOPSAR topographic mapping radar mounted on the NASA DC-8 aircraft [6]. This implementation has been suggested for spaceborne use by Rodriguez and Martin [5] and informally by others. In this case either the wavelength is chosen to be quite short (< 1 cm), or for longer wavelengths tethered satellites are required to generate a baseline of adequate length [7].

The second type of implementation, which we analyze here, is to utilize a single satellite antenna in a nearly-exact repeating orbit, forming the interferometer baseline by relating radar signals on repeat passes over the same site. Even though the antennas do not illuminate the same area at the same time, if the ground is completely undisturbed between viewings the signals will be highly correlated and a spatial baseline may be synthesized. Topographic maps using this technique have been demonstrated by Goldstein *et al.* [8], Gabriel *et al.* [9], and Gabriel and Goldstein [3].

The amount of decorrelation observed in these repeat-pass interferometers is important for two reasons. First, the amount of surface change over time describes processes occurring on time scales of the orbit repeat time and size scales on the order of a radar wavelength. Measurement of interferometer correlation thus provides a means to sense remotely a wide variety of surficial processes such as vegetation growth, glacier motion, permafrost freezing and thawing, and soil moisture induced effects.

The second area of interest in understanding temporal decorrelation is that it constitutes an important error source in the operation of a repeat pass geometry topographic mapping radar. The orbit selection will be driven by a combination of tolerable error levels, the attainable baseline, and the expected decorrelation with time of signals from the regions of interest to be mapped. Since this implementation approach may be employed using existing and planned general purpose radar satellites such as SEASAT, ERS-1, and RADARSAT, it is attractive in achieving the widest possible utilization of those systems.

We note here that there exists a class of radar interferometers specifically designed for measurement of radar echo phase differences on repeat images separated in time by less than a second, that is many times shorter than the temporal baselines we consider below. These are the "along-track"

or "front–back" interferometers implemented on aircraft for the measurement of ocean surface motion such as currents or swell wave spectra (see Goldstein and Zebker [10] or Goldstein *et al.* [11] for a description of such instruments and their application). Since the geophysical phenomena for decorrelation of ocean surfaces are quite different than those for land processes, we defer a description of these to a later work, and will not consider them further.

II. Background

Coherent radar echoes, that is, those with measurable phase and amplitude, will be correlated with each other if each represents nearly the same interaction with a scatterer or set of scatterers. For imaging radars, another way of stating this is that the observed "speckle" patterns are similar. Speckle, according to a widely used model originally developed for laser scattering, may be modeled by postulating that at least several scattering centers are present in each resolution cell of the radar image; the total scattered field is then the coherent sum of the individual fields from each scattering center. If the scatterers are randomly positioned within the cell, and the cell is assumed to have dimension many wavelengths in size, the phase of each will be random and the sum will be well characterized by a zero mean, complex Gaussian random number with variance proportional to the average radar cross section of the surface.

Even though the radar signal in this case possesses Gaussian statistics, if we duplicate the radar imaging experiment at a later time but do not alter the position or cross section of the subresolution scatterers, the received signal will be identical to the original signal. In this sense, the signal is a spatially random process, but slowly varying with time, and repeated echoes will be highly correlated because variation is slow compared with the repeat observation frequency.

This does not imply that all observations of the same resolution cell will be correlated, however, as altering the observation geometry leads to decorrelation as the apparent relative positions of the scatterers change. Thus additional constraints on the repeat incidence and aspect angles are required for observation of echo correlation, but careful data acquisition and processing can minimize these effects. We will quantify the decorrelation due to each of these effects below.

In this paper we are concerned with the measurement of radar echo correlation and its interpretation in terms of the above effects. In particular, we would like to separate decorrelation due to actual changes of the target from that dependent on sensor geometry. In this manner we may infer geophysical properties of the surface without being confused by instrumental effects. In addition, understanding the sensor effects permits a more effective and useful system design and performance analysis, resulting in a controlled and quantified error budget.

For the purposes of this paper we will refer to the sensor geometry effects as spatial in nature and those due to target change as temporal effects, as the dominant source of decorrelation for a well-designed system observing a truly stationary target is spatial baseline noise caused by viewing the surface with two antennas at slightly different aspect angles. This is the effect which has been described by Li and Goldstein [4] and by Rodriguez and Martin [5]. The change in the target surface with time, the temporal effect, then causes additional decorrelation which is related only to properties of the surface.

III. Theory

We consider here the three sources of decorrelation introduced above: spatial baseline decorrelation, decorrelation due to rotation of the target between observations, and decorrelation from surface motion of the individual scattering centers within each resolution element. Two derivations of baseline decorrelation have been presented previously by Li and Goldstein [4] and by Rodriguez and Martin [5]; here for clarity we rederive the main results in a slightly different form and in addition obtain a Fourier transform relation between the correlation function and the system impulse response. We will verify the results by observation in the next section, and present data indicating the dependence of phase error on system parameters. For rotation, we find a similar transform relation and also present a numerical calculation of the decorrelation as a function of angle. We then verify that it is not important for the data analysis described in the next section. Finally, for the temporal decorrelation we plot decorrelation as a function of the degree of motion of the individual scatterers. In each case, we show the dependence of the correlation function on parameters of either the sensor or the target, as appropriate.

A. Overview

Consider two radar signals s_1 and s_2 acquired by two antennas observing the same target at the same time, but with different receivers. If we model the signals as consisting of a correlated part c common to the signal at both antennas and also of thermal noise parts n_1 and n_2, such as

$$s_1 = c + n_1$$

$$s_2 = c + n_2 \tag{1}$$

then we may evaluate the correlation $\rho_{thermal}$ between them as a function of noise in the usual manner:

$$\rho_{thermal} = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}} \tag{2}$$

where $\langle \cdot \rangle$ denotes ensemble averaging. Since the noise and signal are uncorrelated, we obtain

$$\rho_{thermal} = \frac{|c|^2}{|c|^2 + |n|^2} \tag{3}$$

Noting that the thermal signal-to-noise ratio (SNR) is $\frac{|c|^2}{|n|^2}$, (3) may be equivalently written

$$\rho_{thermal} = \frac{1}{1 + SNR^{-1}} \tag{4}$$

Next, we generalize (1) by including a term representing that portion of the signal which is uncorrelated between

antennas due to, say, spatial baseline decorrelation— a result of nonidentical viewing directions (see below). Then

$$s_1 = c + d_1 + n_1$$

$$s_2 = c + d_2 + n_2 \tag{5}$$

where c is the correlated part of the return, d_i is the uncorrelated part exclusive of thermal noise, and n_i again represents thermal noise. We now can calculate the correlation $\rho_{spatial}$ in the infinite SNR case:

$$\rho_{spatial} = \frac{|c|^2}{|c|^2 + |d|^2} \tag{6}$$

and also the correlation if thermal noise is included (these follow from simple application of (2) and (5) above):

$$\rho_{spatial+thermal} = \frac{|c|^2}{|c|^2 + |d|^2 + |n|^2} \tag{7}$$

Since the signal itself consists of both the correlated and decorrelated components, the SNR is $\frac{|c|^2+|d|^2}{|n|^2}$, thus (7) may be written

$$\begin{aligned} \rho_{spatial+thermal} &= \frac{|c|^2}{|c|^2 + |d|^2} \cdot \frac{|c|^2 + |d|^2}{|c|^2 + |d|^2 + |n|^2} \\ &= \frac{|c|^2}{|c|^2 + |d|^2} \cdot \frac{1}{1 + SNR^{-1}} \\ &= \rho_{spatial} \cdot \rho_{thermal} \end{aligned} \tag{8}$$

A similar argument leads to a further, and final, generalization for a pair of signals consisting of a correlated part, a decorrelated part due to spatial decorrelation, and a decorrelated part due to temporal phenomena, yielding the following for the total observed correlation:

$$\rho_{total} = \rho_{temporal} \cdot \rho_{spatial} \cdot \rho_{thermal} \tag{9}$$

We note that this derivation incorporates an assumption that the thermal noise powers at each antenna are equal, and that it is a trivial extension of the above to account for the situation for differing noise levels.

In summary, if any three of the quantities in (9) are known, the fourth may be determined. For data analyzed in this paper, we know quite well our imaging geometry and signal to noise ratio and can measure the total correlation ρ_{total}, therefore, the temporal component, which contains the information about the target, may be inferred. In the rest of this section we will present theoretical bases useful in determining the various correlation parameters.

B. Spatial Baseline Decorrelation

In order to determine the spatial decorrelation $\rho_{spatial}$, we need to calculate, from knowledge of our imaging geometry, what fraction of the received radar echo should be correlated between antennas. In this case, we know the interferometer baseline and need to determine the correlation as a function of that baseline.

We first derive a Fourier transform relation between the radar impulse response and the baseline decorrelation function

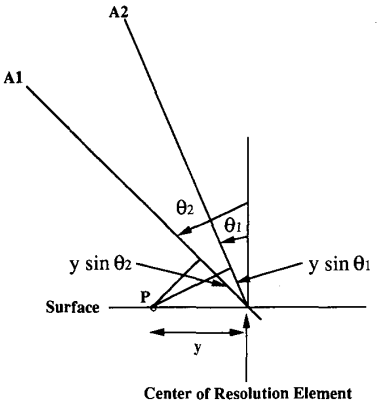


Fig. 1. Interferometer imaging geometry. Radar antennas A1 and A2 both illuminate the same patch of ground centered at $y = 0$. Incidence angles θ_1 and θ_2 result in phase offsets for all points P displaced by distance y of $y \sin\theta_1$ and $y \sin\theta_2$, respectively. Difference of these phases is measured interferometer phase.

as a function of the difference in viewing angles of the two interferometer antennas. Consider a radar interferometer operating with geometry as shown in Fig. 1 where two antennas A1 and A2 illuminate a patch on the surface at incidence angles θ_1 and θ_2, respectively. The along-track (azimuth) distance is x and the across-track (ground range) distance is y; the distance from the sensor itself to the center of a resolution element is r. Then the signal s_1, measured in the final processed image at position (x_0, y_0), from a radar antenna A1 may be represented as

$$\begin{aligned} s_1 = \int\int & f(x - x_0, y - y_0) \exp\{-j\frac{4\pi}{\lambda}(r + y\sin\theta_1)\} \\ & \cdot W(x, y)dxdy + n_1 \end{aligned} \tag{10}$$

where $f(x, y)$ represents the complex backscatter at each point on the surface, λ is the radar wavelength, $W(x, y)$ is the system impulse response, and n_1 is the noise associated with the receiver. Similarly, the signal from antenna A2 is

$$\begin{aligned} s_2 = \int\int & f(x - x_0, y - y_0) \exp\{-j\frac{4\pi}{\lambda}(r + y\sin\theta_2)\} \\ & \cdot W(x, y)dxdy + n_2 \end{aligned} \tag{11}$$

The cross-correlation of the two signals, from which we determine the interferometer phase, is thus

$$\begin{aligned} s_1 s_2^* = & \int\int\int\int f(x - x_0, y - y_0) f^*(x' - x_0, y' - y_0) \\ & \exp\{-j\frac{4\pi}{\lambda}y(\sin\theta_1 - \sin\theta_2)\} W(x, y) W^*(x', y')dxdydx'dy' \end{aligned} \tag{12}$$

If the interferometer is arranged such that the range r is unequal at the two antennas only the mean phase of the correlation changes, but not the correlation magnitude.

Now, if the surface is taken to consist of uniformly distributed and uncorrelated scattering centers, then

$$\langle f(x, y) f^*(x', y') \rangle = \sigma^0 \delta(x - x', y - y') \tag{13}$$

where σ^0 is the average radar cross section, (12) reduces to

$$\langle s_1 s_2^* \rangle = \sigma^0 \int\int \exp\{-j\frac{4\pi}{\lambda} y \cos\theta\ \delta\theta\} |W(x,y)|^2 dxdy \tag{14}$$

where θ is the average look angle and $\delta\theta = \theta_1 - \theta_2$. The exponential term, since it is linear in y, can be interpreted as a Fourier kernel and thus we have a transform relation between the correlation function and the radar system impulse response; the correlation function is simply the transform of the intensity impulse response.

For the typical radar model where the impulse response is approximately

$$W(x,y) = \text{sinc}(x/R_x)\text{sinc}(y/R_y) \tag{15}$$

where R_x and R_y are the azimuth and range resolutions, and the sinc function is taken as $\frac{\sin \pi x}{\pi x}$, evaluation of (14) followed by normalization leads to the spatial baseline decorrelation function

$$\rho_{spatial} = 1 - \frac{2\cos\theta\ |\delta\theta| R_y}{\lambda} \tag{16}$$

The correlation function, which from (14) is simply the Fourier transform of the impulse response intensity, falls off linearly as $\delta\theta$, the difference in look angle for the two antennas, increases. Equivalently, this effect can be described in terms of the antenna baseline separation B in meters (assumed to be in the horizontal direction only) by

$$\rho_{spatial} = 1 - \frac{2|B|R_y \cos^2\theta}{\lambda r} \tag{17}$$

The minimum value of B for which for which $\rho_{spatial}$ equals zero is the critical baseline B_c, and occurs when the change in look angle between the two passes is sufficient to cause backscatter from each pixel to become completely uncorrelated. Specifically,

$$B_c = \frac{\lambda r}{2R_y \cos^2\theta} \tag{18}$$

In practice, the impulse response may also be modified by the nonideal characteristics of various elements in the radar itself as well as by windowing used during the processing, altering the baseline decorrelation function given in (17).

C. Decorrelation Due to Rotation

Another geometrical sensor effect that leads to decorrelation is rotation of the target with respect to the radar look direction. In other words, we cannot illuminate the same patch of surface from two different aspect angles and expect the signals to be fully correlated. To understand this source of decorrelation noise, consider a resolution element as shown in Fig. 2. Each scattering center at polar location (δ, ϕ) rotates to position $(\delta, \phi + d\phi)$. Transformation to rectangular coordinates $x = \delta\cos\phi$, $y = \delta\sin\phi$ permits us to express the change in position on the surface as a change in range; if the distance to a point before rotation is $r + \delta\sin\theta\sin\phi_1$, the distance after a small rotation $d\phi = \phi_1 - \phi_2$ is $r + \delta\sin\theta\sin\phi_2$. As the patch is rotated slightly, the range to and hence phase of each scattering center changes slightly, and their coherent sum will vary.

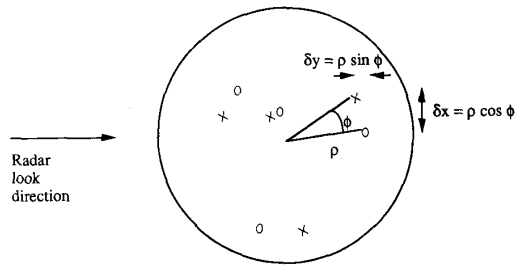


Fig. 2. Rotation of a resolution element by angle ϕ moves scattering centers from initial positions x to new positions o. Across-track component of displacement then yields slightly different phase shift for each scattering center, resulting in signal decorrelation.

We consider again two radar signals s_1 and s_2, representing the echo from a resolution element before and after rotation, respectively. By analogy with (12) above the cross-correlation of the two signals may be expressed as

$$\begin{aligned} s_1 s_2^* = \int\int\int\int & f(x - x_0, y - y_0) f^*(x' - x_0, y' - y_0) \\ & \exp\{-j\frac{4\pi}{\lambda}\delta\sin\theta(\sin\phi_1 - \sin\phi_2)\} W(x,y) \\ & W^*(x',y') dx dy dx' dy' \end{aligned} \tag{19}$$

and we obtain

$$\langle s_1 s_2^* \rangle = \sigma^0 \int\int \exp\{-j\frac{4\pi}{\lambda} x \sin\theta\ d\phi\} |W(x,y)|^2 dxdy \tag{20}$$

This second Fourier transform relation leads to the following expression for the rotation-induced decorrelation for $\frac{\sin x}{x}$ azimuth impulse response:

$$\rho_{rotation} = 1 - \frac{2\sin\theta |d\phi| R_x}{\lambda} \tag{21}$$

We have verified this result numerically by first determining a set of scattering centers randomly located within a resolution cell, and then altering the position of each according to a rotation of the entire cell. This process is repeated many times (1000) to obtain the ensemble average, which we present in Fig. 3. The relevant parameters here correspond to data acquired by the SEASAT satellite operating at L-band (λ = 24 cm, for a system description see observation section below), and also for a C-band (λ =5.66 cm) system in a similar orbit. Thus the C-band results will be approximately correct in assessing the performance of the ERS-1 radar satellite in interferometric applications.

The simulation results indicate that the signal decorrelates with angle, and nearly completely after rotation of about 2.8° at L-band and after about 0.7° at C-band, in agreement with (21). The functional dependence of the correlation depicted in Fig. 3 is not quite linear as we used a truncated impulse response for computational reasons, thus the transform of the azimuth response is not a triangle function. We have, however, preserved an "equivalent-width" response so that the critical rotation angle remains about the same.

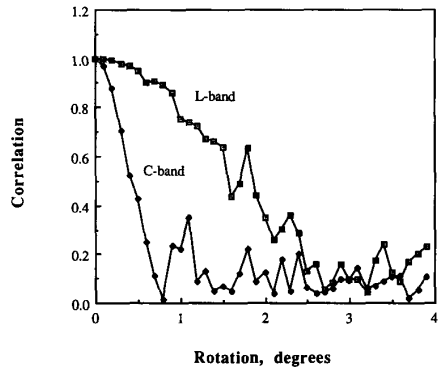


Fig. 3. Simulation results indicating dependence of correlation on rotation of resolution elements. Assumed L-band system parameters are those for SEASAT , while C-band parameters are similar but for wavelength of ERS-1 radar. Complete decorrelation results after a rotation of $2.8°$ at L-band, $0.7°$ at C-band.

In fact, the SEASAT data we analyze in this paper were acquired with orbits that are parallel within $0.02°$ and thus the decorrelation we observe from this effect is minimal.

D. Temporal Decorrelation

The final decorrelation source of interest here is the temporal effect, which follows from physical changes in the surface over the time period between observations. For the SEASAT case, the orbit repeat time was 3 days, so that temporal baselines of 3 days, 6 days, 9 days, and so forth are available. Since calculation of this effect depends on detailed changes of a given surface type, we present here only a sample calculation assuming Gaussian-statistic motion as a guide and leave predictive application-specific theories for later work. In the observation section of the paper which follows we will experimentally characterize temporal decorrelation from unvegetated, lightly vegetated, and heavily forested surfaces.

Once again, we begin with the expression for the correlation between two signals s_1 and s_2 (see (22) below) where we have generalized the backscatter function $f(z, y, z)$ to account for three-dimensional variability (volume scatter, important for vegetation models), and included terms related to change horizontal position δy and change in height δz of a scatterer in the exponential kernel. If we assume that changes in position of a scatterer are unrelated to the initial position, and are characterized by independent probability distributions $p_y(\delta y)$ and $p_z(\delta z)$, (22) reduces to (23) below.If the probability distributions are Gaussian, then after normalization the integral yields

$$\rho_{temporal} = \exp\{-\frac{1}{2}(\frac{4\pi}{\lambda})^2(\sigma_y^2 \sin^2\theta + \sigma_z^2 \cos^2\theta)\} \quad (24)$$

For the SEASAT geometry where the nominal incidence angle is $23°$, the contribution from displacements in z is greater than that for displacements in y as indicated by the geometrical factors in (24). In other words, we expect greater sensitivity to vertical changes than to horizontal changes for incidence angles less than $45°$, and thus surfaces with significant volume scattering, such as forests, should decorrelate most rapidly with time.

We again verify this result by simulation, where we restrict motion to the surface plane (δy only) for simplicity. For our sample calculation we alter the location of the scattering centers within a volume in a random direction by adding to each location a complex Gaussian distance of specified rms motion, thus the direction of each motion is uniformly random in angle. In Fig. 4 we plot temporal decorrelation at two wavelengths as a function of rms motion and also indicate the analytical result, where the wavelengths used are 0.24 m (L-band) and 0.0566 m (C-band). The remaining radar parameters are typical for SEASAT. We note that in this case 10 cm of rms motion is needed for complete decorrelation at L-band while only 2–3 cm rms motion decorrelates C-band signals.

We have here considered only random motions, that is, each scattering center moves independently of all others. If in fact the scatterers move together in one preferred direction, then instead of decorrelation a systematic phase shift would occur. This idea has been proposed and applied to measurement small surface changes (see, for example, [9]).

E. Implications for Topographic Mapping

Our final theoretical result is to assess the effect of decorrelation on the accuracy of inferred topographic maps. The major implication of decorrelation in an interferometer is that it adds noise to the radar echoes, increasing the standard deviation of inferred phase estimates and hence the derived height values. This topic has been previously addressed by Zebker and Goldstein [2], Li and Goldstein [4], and Rodriguez and Martin [5], although they did not consider the additional noise due to temporal decorrelation. Using the approximate formula given by Rodriguez and Martin [5, eq. 31], we can relate the phase standard deviations to height errors as follows:

$$\sigma_h = \frac{\lambda \rho \tan\theta}{4\pi B}\sigma_\phi \quad (25)$$

$$s_1 s_2^* = \int\int\int\int\int\int f(x - x_0, y - y_0, z - z_0) f^*(x' - x_0, y' - y_0, z' - z_0)$$
$$\exp\{-j\frac{4\pi}{\lambda}(\delta y \sin\theta + \delta z \cos\theta)\} W(x, y) W^*(x', y') dx dy dz dx' dy' dz' \quad (22)$$

$$\langle s_1 s_2^* \rangle = \sigma^0 \int\int \exp\{-j\frac{4\pi}{\lambda}(\delta y \sin\theta + \delta z \cos\theta)\} p_y(\delta y) p_z(\delta z) d\delta y d\delta z \quad (23)$$

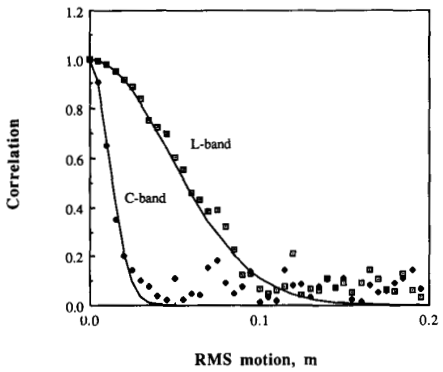


Fig. 4. Simulation results indicating dependence of correlation on random motion of scattering centers within resolution elements (points), with theoretical predictions (solid lines). As in previous figure, assumed L-band system parameters are those for SEASAT, while C-band parameters are those for wavelength of ERS-1 radar. Complete decorrelation results after rms motion of 10 cm at L-band, 2–3 cm at C-band.

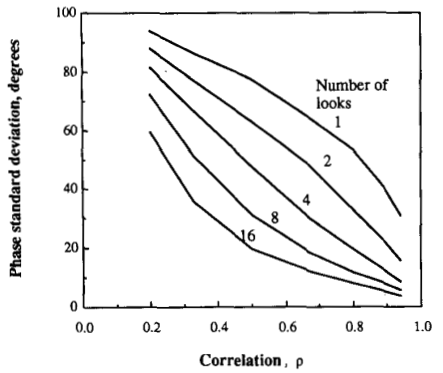


Fig. 5. Sensitivity of phase standard deviation to correlation and number of looks in processor. Increasing number of looks is an effective means to reduce statistical variation, especially for the first eight looks or so.

where ρ is range, θ is the look angle, B is the interferometer baseline, λ is the radar wavelength, and σ_h and σ_ϕ are the standard deviations of height and phase, respectively.

We have calculated, and present in Fig. 5, the expected phase standard deviations as a function of the interferometric radar system parameters of correlation and number of looks, where by number of looks we refer to the number of resolution elements averaged spatially in the complex interferogram to reduce statistical variations. For example, if the correlation ρ_{total} is 0.8 and we average four resolution elements in the interferogram, the resulting phases are determined with an uncertainty of 21°. Equation (25) may then be used to infer the resulting height precision, which would be 5 m for SEASAT operating with a baseline of 500 m. Note that "taking looks" is a particularly effective means to reduce errors when the number of looks is less than eight or so; therefore, systems should be designed with this in mind.

F. Observations

In this section, we present observations of decorrelation using SEASAT data, and derive a topographic map of a region in Oregon containing both lightly vegetated and heavily forested areas. First, we will describe the relevant parameters of the radar system and satellite orbit geometry. Next, we show measured spatial baseline decorrelation from data acquired over Death Valley, and temporal decorrelation from that site and several areas in the Oregon image. Finally, we derive the topographic map and estimate its accuracy over the several surface types included in the image.

TABLE I
SEASAT ORBIT PARAMETERS

Orbit Number	Date	Long. Asc. Node	Inclination
1226	Sept. 20	255.9989	108.0202
1269	Sept. 23	255.9930	108.0076
1312	Sept. 26	255.9885	108.0176
1355	Sept. 29	255.9863	108.0294
1398	Oct. 2	255.9868	108.0196
1441	Oct. 5	255.9896	108.0071
1484	Oct. 8	255.9950	108.0166

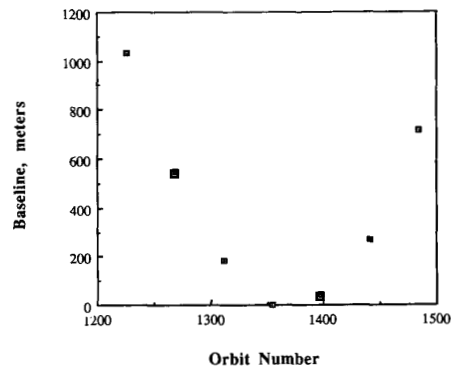


Fig. 6. Relative offsets of SEASAT orbits. Interferometer baseline for any pair of orbits may be found by difference of offsets for that pair. These values approximately correct for western U.S.

1) SEASAT Parameters: The data shown here were acquired by the SEASAT synthetic aperture radar satellite over a period from September 20–October 8, 1978. The corresponding orbit numbers range from 1226 to 1484. The SEASAT orbit altitude of 800 km provided for a nearly exact repeat track every 43 orbits (3 days). Orbital data are given in Table I.

A plot of approximate orbit offsets in meters over the western U.S. is shown in Fig. 6, where the independent variable is orbit number and the dependent variable is the relative offset of each orbit to orbit 1355 (chosen arbitrarily). Thus to find the interferometer baseline for any given pair of orbits, the relative locations from Fig. 6 must be differenced. The orbit position varies approximately quadratically with time. This should be kept in mind during the analyses presented below, when we are isolating observed temporal effects from spatial effects—the available time and space baselines are uncorrelated with each other.

The SEASAT radar consisted of a nominal 1-kW transmitter, operating at 1275 MHz and transmitting 33 μs pulses at a rate of 1647 pulses per second. The transmit waveform was range coded by a linear FM signal for 19-MHz bandwidth. Data

were transmitted to Earth using an analog downlink, and offset video signals were digitized on the ground to 5-bit accuracy at a sample rate of about 45.5 MHz. These samples were then processed on a general purpose computer using a conventional range/Doppler algorithm. Complex, single look high resolution pixels were generated, with a typical scene size of 1024 pixels in range by 4096 pixels in azimuth and corresponding to approximately 16 km by 16 km on the ground. As stated above, the nominal radar incidence angle was 23°.

2) Baseline Decorrelation: The arid, unvegetated floor of Death Valley in California serves as an ideal target for measurements of baseline decorrelation. The relatively strong backscatter from rough portions of the valley floor results in a high SNR and therefore minimizes the effect of the $\rho_{thermal}$ term in (9). More importantly, changes in the surface on a time scale of days or weeks are negligible, minimizing $\rho_{temporal}$ as a possible source of decorrelation. As a result, the observed correlation ρ_{total} is in effect a direct measure of $\rho_{spatial}$ and should fall off nearly linearly as the baseline B is increased. Interferograms obtained using small baselines are relatively free from degradations caused by baseline decorrelation, but are characterized by broad fringes and reduced accuracy in the resulting height maps. The lengths of the 21 baselines which can be synthesized using the seven SEASAT orbits given in Table I range from approximately 50 to 1100 m.

We estimated the critical baseline for our SEASAT data at about 4500 m by first estimating the system impulse response and then computing its Fourier transform as indicated above. Since our scenes of Death Valley, CA and Bend, OR did not contain any known point reflectors, we used the impulse response determined by the JPL SEASAT project which was documented in a JPL internal report [12]. They found that data from a calibration site at Goldstone Dry Lake in California were well modeled by an unweighted $\sin(x)/x$ function with intensity half-power width of 25 m. The transform of this response, as discussed above, is a linearly decreasing function which equals zero for a baseline of approximately 4500 m.

In order to compare observed baseline decorrelation with this theoretical estimate, we first formed six interferograms using images of Cottonball Basin in Death Valley acquired during orbits 1226, 1355, 1441, and 1484. For each pair of images we used a statistical correlation technique to estimate the relative offset and then resampled the data to coregister the images. Next, we selected regions of Cottonball Basin characterized by flat or smoothly sloping topography and therefore by straight, evenly spaced fringes in the interferograms. Using an iterative procedure, we identified and then removed from one of the images the phase ramp best corresponding to the observed fringes. Finally, for each region we calculated the correlation ρ_{total} ($= \rho_{spatial}$) between the images using the pixels within the area of interest. The resulting correlations, plotted in Fig. 7, show the near linear dependence on baseline expressed in (17). A critical baseline value of $B_c = 3200$ m, obtained by fitting a linear function to these data, is in very rough agreement with the value of 4500 m calculated using (17).

Note from Fig. 7 that our observed values of correlation fall below the theoretical expectation, which leads to a low estimate of the critical baseline. What this implies is that there

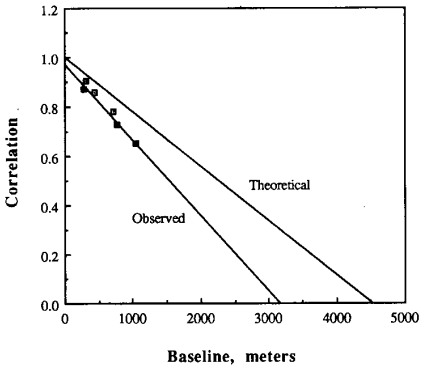


Fig. 7. Theoretical and empirically determined spatial baseline decorrelation functions. Ideal impulse–response analysis indicates a critical baseline, that is the baseline for which correlation equals zero, of 4500 m, while data fits to a value of 3200 m. The discrepancy is due to unmodeled decorrelation sources in the radar system. We thus use the empirically derived model for later analysis in order to compensate for these unkown error sources.

are additional unmodeled sources of decorrelation in our data. These sources can be, for example, interpolation noise in the processor or analysis routines, or that the impulse response we assumed is narrower than the true impulse response. Therefore, in the remainder of this paper we will model the baseline decorrelation by the empirically-derived function with a critical baseline of 3200 m rather than the theoretically ideal model. This approach allows us to isolate the temporal phenomena from any unknown processor-induced effects.

3) Temporal Decorrelation: We next considered an area in central Oregon characterized by diverse topography and containing both heavily forested areas and partially vegetated and bare lava flows. In contrast with Death Valley, we would expect more physical changes in the surface itself over the 18 days spanned by the seven SEASAT orbits. Another difference involved the topography and therefore the size of the areas over which we could consider decorrelation. The technique of removing fringes by applying a phase ramp to one of the images works only when the surface topography can be approximated by a plane. Many regions of Death Valley are indeed quite flat and are therefore well-suited to this approach. For data collected over the Oregon site, however, the selected areas must be large enough to produce statistically reliable results, but small enough so that the terrain can be approximated as flat. An alternative but less practical approach could involve utilizing arbitrarily large areas over which the fringes are removed using detailed knowledge of the true topography. Even if sufficiently detailed digital elevation maps were available, the increase in computational complexity would probably not justify the improvement in the results. We found it difficult to obtain good correlations using areas larger than about 20 × 20 pixels (a pixel corresponds to about 17 m) and used areas measuring 10 pixels on a side for most of this work.

Using a procedure identical to that described in the previous section, we processed data for six SEASAT passes (central Oregon was not imaged during orbit 1355) and formed 15 interferograms. We then selected both forested and unvege-

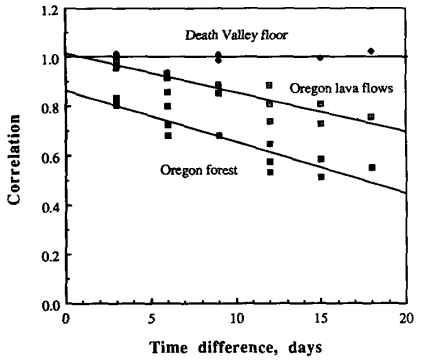


Fig. 8. Temporal decorrelation as a function of time for three surfaces. The floor of Death Valley exhibits no significant decorrelation over our 18-day observation period. The lightly vegetated and unvegetated lavas in Oregon show some temporal effects, and the heavily forested regions show the most temporal decorrelation. Even after 18 days, however, the correlation associated with the forested areas is still 0.5, enough for reasonably reliable topographic maps to be generated.

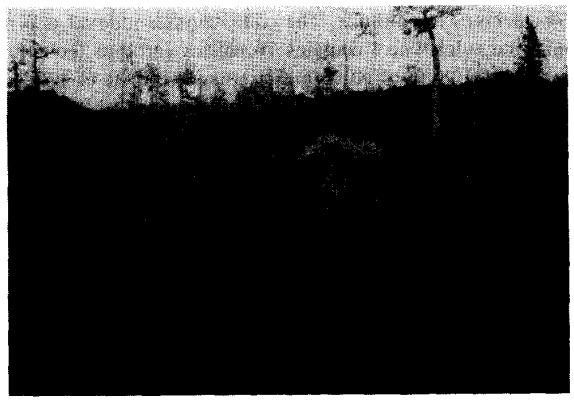


Fig. 9. Photograph of lava areas from which correlation values were measured. Note that the surface is rather devoid of vegetation, thus we do not understand the physical mechanism for the observed, albeit minor, decorrelation with time.

tated lava areas for analysis. After determining and removing from each interferogram the appropriate phase ramp for each small area, we calculated ρ_{total} and removed the contributions due to $\rho_{thermal}$ and $\rho_{spatial}$. The value which remained was $\rho_{temporal}$, which we interpret as an indication of the degree to which the area had changed in the time between the two images. A value of $\rho_{temporal} = 1$ indicates no change, while surface changes alter the exact complex backscatter and cause decorrelation, reducing $\rho_{temporal}$. By plotting $\rho_{temporal}$ as a function of time difference between images, an indication of the presence and degree of surface change results.

We present our temporal decorrelation results in Fig. 8. To verify that we had succesfully eliminated system errors which could give erroneous results suggesting gradual surface change where none existed, we first examined an area in Death Valley. Given the absence of precipitation and other factors which could change the nature of the surface on a time scale of several weeks, we expected to find no significant temporal decorrelation. As the plot of Fig. 8 shows, the surface of Death Valley remains fully correlated over our 18 days of observations, indicating minimal residual influence of system errors and demonstrating that the surface does remain unchanged. We next implemented a similar analysis on data acquired over forests and lava flows at the Oregon site. The forest decorrelates in what appears to be a linear fashion, reaching $\rho_{temporal} \approx 0.5$ for a time difference of 18 days. This is plausible given the volume scattering occuring for vegetated targets (see [13], or [14], for example, radar scattering models incorporating volume scattering from canopies), in which wavelength-order changes in the positions of branches significantly alter the speckle and therefore the correlation.

The temporal decorrelation results for the lava data, also shown in Fig. 8, are more difficult to explain. The lava appears to decorrelate at approximately the same rate as the forest, but with a higher initial value. Although the results plotted in Fig. 8 represent only one forested area and one lava area, correlations calculated in other areas produce similar results — signals from both the forest and the lava appear to decorrelate at about the same rate, but with the lava echoes consistently exhibiting higher correlation than the forest.

We first suspected that the decorrelation of the lava might be due to a system error. However, the same SEASAT orbits were used to acquire data over both Death Valley and Oregon (overflight times for the two sites are separated by only 3 min), and the intereferograms of Death Valley show no sign of temporal decorrelation. Many other possible error sources, such as those involving estimation of the baseline or SNR, are unrelated to the time difference between pairs of images and would not produce the results of Fig. 8. We also considered the possibility that the lava surface was vegetated and that we were in fact seeing changes in this vegatation rather than in the lava itself. The lava flows in question are approximately 5000 years old, and in certain areas pockets of soil have collected, permitting growth of brush and some trees. A visit to the area, however, showed that much of the lava has remained completely bare, as shown in the photograph of Fig. 9. Weather is also unlikely to provide an explanation. Throughout the entire period spanning these orbits (late September to early October 1978) the weather was dry with temperatures reaching 25–30 °C during the day and dropping to 10–15° at night. In addition, the data were acquired at approximately the same local time for each pass (about 10:30 AM) rendering dew and thermal expansion unlikely sources of decorrelation. We are left with the possibilities that 1) there is some progressive error source unique to the Oregon images which we have not eliminated, or 2) that there is true change occurring in the lava. It is true that the Oregon lava is blockier than the smooth floor of Death Valley, but why that would effect the observed decorrelation is unclear.

4) Topographic Map: In Fig. 10 we present a conventional radar image and also the interferometrically-produced topographic map of the study area in Oregon derived from a single pair of SEASAT passes (orbits 1226 and 1269). This pair was chosen to maximize spatial baseline (484 m) and minimize temporal baseline (3 days). In the figure, the brightness of each point is related to the magnitude of radar backscatter while the

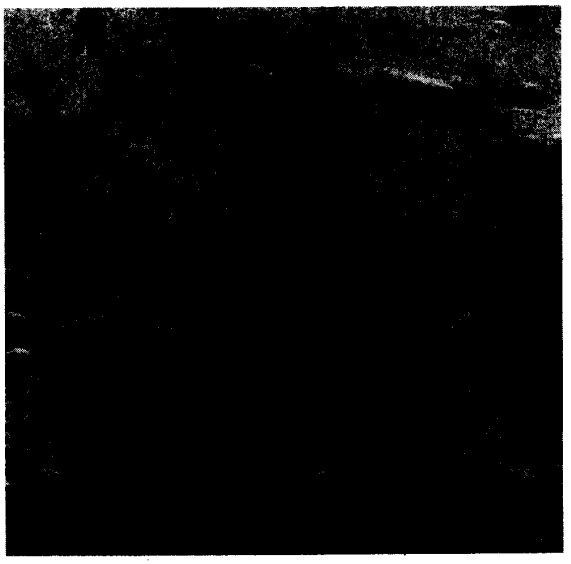
color denotes the altitude. The color contour interval is 6 m, while the color wheel contains 16 entries and thus the colors repeat every 96 m of altitude. The image consists of 1024 by 1024 points each with ground spatial dimensions 17 by 17 m, thus the image is slightly greater than 17 km on a side.

We note that relatively noise-free topography is available everywhere in the image, even over the most heavily forested area. These data were averaged to 16 looks. Given the 484-m baseline we estimate $\rho_{spatial}$ from (17) at approximately 0.85, and since the time separation was 3 days, from Fig. 8 we would expect approximately to observe $\rho_{temporal}$ of 0.97 and 0.8 over the lightly vegetated lavas and heavily forested areas, respectively. Thermal contributions are negligible for this scene. The total correlations ρ_{total} for the two types of targets are then 0.82 and 0.68. Examination of Fig. 5 yields phase standard deviations of 7 and 12°, respectively, for the lightly vegetated and heavily forested areas. Finally, using (25) we estimate the statistical variation contribution to the error in height to be 1.5 m over the lightly vegetated lava regions and 2.6 m over the forest. Of course, the actual accuracy is several times worse than this as the error budget is dominated by systematic errors such as uncertainty in baseline knowledge. Our main conclusion here is that if the temporal baseline is constrained at a few days or less, the additional height error due to temporal changes on the surface are not significant contributors to the overall error.

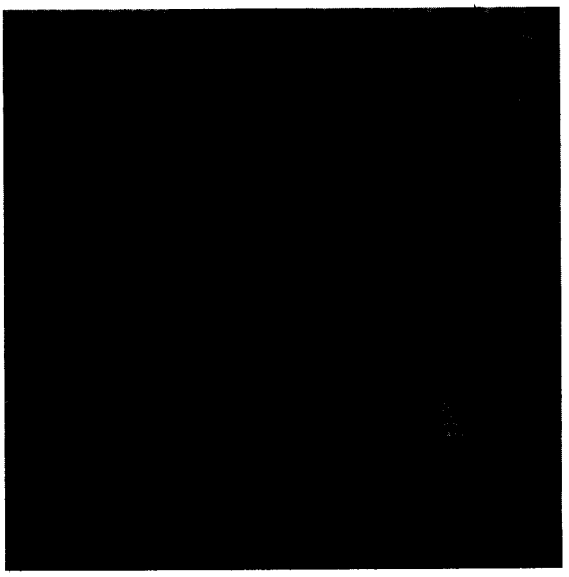
Of course additional optimization may be applied to the data of Fig. 10. For example, if additional height acuity is needed, we could average spatially to obtain more "looks," if maximum spatial resolution is not required.

IV. Summary

Correlation in pass-to-pass, interferometric radar can be degraded by thermal noise, lack of parallelism between the radar flight tracks, spatial baseline noise, and surficial change. The effects of decorrelation due to thermal noise can be easily evaluated and removed, while those due slight angular changes between flight tracks are negligible for data acquired using near-repeat orbits. Spatial baseline and rotation-induced decorrelation can be derived using the Fourier transform of the impulse response intensity, and increase linearly with baseline or rotation in an ideal system. Empirical results obtained using images of Death Valley confirm that, as the baseline increases, the overall correlation decreases due to spatial baseline noise. As the effects of these three sources of decorrelation can be quantified, their contributions to the observed overall correlation can removed, yielding a measure of the temporal decorrelation due to change in the target itself. We have shown that areas of Cottonball Basin in Death Valley remained unchanged over the 3-week period for which we have data, while a heavily forested area in Oregon exhibited significant temporal decorrelation. Lava surfaces in central Oregon also appeared to decorrelate, although the reasons for this are uncertain. We generated a topographic map from the images of central Oregon and achieve statistical contributions to height accuracy of 1.5 m over unvegetated areas and 2.6 m over forest. Our results demonstrate that generation of height maps of heavily vegetated areas using pass-to-pass interferometry is practical provided that the time between passes is at most several weeks.



(a)



(b)

Fig. 10. Radar image (a) and interferometrically derived topographic map (b) of Oregon forested area. Height contour levels are 6 m/color, or 96 m for one complete color cycle. The topography is clearly visible even in the most heavily forested regions. The layed over cones in the radar image are seen to be rectified in the topographic map, demonstrating that with three-dimensional data cartographically correct maps may be generated. The irregular border at the bottom of the topographic map is a result of the nonlinear stretch applied to rectify the image.

Acknowledgment

We would like to thank Soren Madsen and Dick Goldstein for many helpful suggestions and discussions about interfer-

ometer data analysis. We would also like to thank Soren for a careful reading of the manuscript.

REFERENCES

[1] L. C. Graham, "Synthetic interferometer radar for topographic mapping," *Proc. IEEE*, vol. 62, pp. 763–768, 1974.
[2] H. Zebker and R. Goldstein, "Topographic mapping from interfermetric SAR observations," *J. Geophys. Res.*, vol. 91, no. B5, pp. 4993–4999, 1986.
[3] A. K. Gabriel and R. M. Goldstein, "Crossed orbit interferometry: Theory and experimental results from SIR-B," *Int. J. Remote Sensing*, vol. 9., no. 8, pp. 857–872, 1988.
[4] F. Li and R. M. Goldstein, "Studies of multi-baseline spaceborne interferometric synthetic aperture radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 88–97, Jan. 1990.
[5] E. Rodriguez and J. Martin, "Theory and design of interferometric SAR's," submitted to *Proceeding of the IEE*, 1991.
[6] H. A. Zebker, S. Vetrella, J. Martin, *et al.*, TOPSAR: The NASA/ASI airborne topographic radar mapper, in preparation.
[7] A. Moccia and S. Vetrella, "A tethered interferometric SAR for a topographic mission," in press, *IEEE Trans. Geosci. Remote Sensing*, 1991.
[8] R. M. Goldstein, H. A. Zebker, and C. Werner, "Satellite radar interferometry: two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, no. 4, pp 713–720, July-Aug. 1988.
[9] A. K. Gabriel, R. M. Goldstein, and H. A. Zebker, "Mapping small elevation changes over large areas: differential radar interferometry," *J. Geophys. Res.* vol. 94, no. B7, 9183–9191, July 10, 1989.
[10] R. M. Goldstein and H. A. Zebker, "Interferometric radar map of ocean surface currents," *Nature*, vol. 328, no. 6132, pp. 707–709, Aug. 20, 1987.
[11] R. M. Goldstein, T. P. Barnett, and H. A. Zebker, "Remote sensing of ocean currents," *Science*, vol. 246, pp. 1282–1285, Dec. 8, 1989.
[12] S. H. Pravdo, B. Huneycutt, B. Holt, and D. Held, *SEASAT Synthetic Aperture Radar Data User's Manual*, JPL Publication 82–90, Jet Propulsion Laboratory, Pasadena, CA, Mar. 1, 1983.
[13] J. A. Richards, G. Sun, and D. Simonett, "*L*-band radar backscatter modeling of forest stands," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-25, pp. 487–498, 1987.
[14] S. L. Durden, H. A. Zebker, and J. J. van Zyl, "Modeling and observation of forest radar polarization signatures," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 290–301, Mar. 1989.

Howard A. Zebker (M'87–SM'89), for a photograph and biography, please see page 940 of this issue of the TRANSACTIONS.

John Villasenor (S'83–S'88–M'89), photograph and biography not available at the time of publication of this issue.

Accurate Estimation of Correlation in InSAR Observations

Howard A. Zebker, *Fellow, IEEE,* and Katherine Chen

***Abstract*—Interferometric synthetic aperture radar (InSAR) correlation, a measure of the similarity of two radar echoes, provides a quantitative measure of surface and subsurface scattering properties and hence surface composition and structure. Correlation is observed by comparing the radar return across several nearby radar image pixels, but estimates of correlation are biased by finite data sample size and any underlying interferometer fringe pattern. We present a method for correcting bias in InSAR correlation measurements resulting in significantly more accurate estimates, so that inverse models of surface properties are more useful. We demonstrate the value of the approach using data collected over Antarctica by the Radarsat spacecraft.**

***Index Terms*—Correlation, cryosphere, interferometry, interferometric synthetic aperture radar (InSAR), radar.**

I. Introduction

CORRELATION of interferometric synthetic aperture radar (InSAR) echoes is a fundamental observable in interferometric radar measurements. Correlation, a measure of the similarity of two radar echoes, is governed by several system parameters and by properties of the scatterers reflecting the radar signal [1]. Correlation has been used for qualitative identification or classification of surface features [2]–[6]. It has been used to guide phase unwrapping algorithms, for example [7]. Several applications rely on the quantification of correlation estimates, such as in the calculation of uncertainties in InSAR-derived topographic maps [8], extraction of forest parameters [9], estimation of penetration depths in snow [10], and more recently, in modeling of ice accumulation rates [11].

These quantitative studies rely on accurate measurements of correlation. Correlation is usually measured by comparing the radar return across several nearby radar image pixels, but properties of the signal and the imaging system result in biased estimates of the underlying correlation. In one example below, we show that blind calculation of correlation can overestimate correlation in some parts of an image by a factor of several, while underestimating it in other regions of the same image by a similar factor. We present here a method for correction of bias in InSAR data resulting in significantly more accurate correlation measurements. The approach is demonstrated using data collected over Antarctica during the Modified Antarctic Mapping Mission, the second Radarsat Antarctic mapping mission [12].

The need for quantitative evaluation of correlation has been pointed out by Werner *et al.* [13], who raised it in the context of surface change studies. They recognized the need to compensate for fringe patterns if one is to relate correlation to surface properties. Here, we extend that work by deriving the underlying mathematical model relating correlation to fringe rate, and develop an inverse method to correct estimation biases for the presence of interferometer fringes.

Manuscript received September 30, 2004; revised December 16, 2004.

The authors are with the Departments of Electrical Engineering and Geophysics, Stanford University, Stanford, CA 94305-2215 USA (e-mail: zebker@stanford.edu).

Digital Object Identifier 10.1109/LGRS.2004.842375

A. Sources of Bias

It is necessary to correct any estimation biases if we are to be able use correlation observations quantitatively. Two principal biases affect correlation estimates in InSAR, an underestimation resulting from the presence of interferometer fringes, and an overestimation due to limited sample size in the numerical computation of the estimate. Consider first the correlation definition

$$\rho = \frac{|\langle s_1 s_2^* \rangle|}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}} \tag{1}$$

where ρ is the correlation, s_1 and s_2 are complex signals received at the two radar antennas forming the interferometer, and $\langle \cdot \rangle$ denotes an ensemble average. In other words, the correlation is the average power of the cross multiplication of the two signals normalized by the square root of the products of the individual signal powers. Since in practice we are unable to obtain ensemble averages directly, we approximate them with local spatial averages and use an estimator of the form

$$\rho = \frac{\left| \sum_{i=1}^{n} s_{1,i} s_{2,i}^* \right|}{\sqrt{\sum_{i=1}^{n} s_{1,i} s_{1,i}^* \sum_{i=1}^{n} s_{2,i} s_{2,i}^*}} \tag{2}$$

where subscript i denotes the ith pixel of n neighboring pixels averaged. Note that approximating ensemble averages with spatial means implicitly assumes that the pixels have stationary statistics.

Using the correlation model of Zebker and Villsenor [1], we model the radar echoes as consisting of a correlated part c common to the signal at both antennas and also of noise parts n_1 and n_2

$$s_1 = c + n_1 \quad s_2 = c + n_2 \tag{3}$$

where the noise terms may be due to baseline, temporal, or thermal noise decorrelation, or any other artifacts present in the signals. Substituting into the definition [(1)] the correlation of these two signals becomes

$$\rho = \frac{|c|^2}{|c|^2 + |n|^2} \tag{4}$$

where $|n|^2$ is the average power of the "noise."

Now, suppose that the region on the surface over which we estimate correlation contains interferometer fringes. In that case, the correlated part of the signal c is not equal in both images, rather their relative phase varies according to

$$c_{2,i} = c_{1,i}e^{-j2\pi fi} \tag{5}$$

where for simplicity we have modeled a one-dimensional signal, i is the pixel index, and f is the fringe rate in fringes per pixel.

Note that when we evaluate the correlation according to (2), for the numerator we obtain

$$\begin{aligned}\rho_{(\text{num})} &= \left|\sum_{i=1}^{n} s_{1,i}s_{2,i}^*\right| \\ &= \left|\sum_{i=1}^{n}\left(c_{1,i}c_{1,i}^* e^{j2\pi fi} + c_{1,i}n_{2,i}^* \right.\right. \\ &\qquad \left.\left. + c_{1,i}^* e^{j2\pi fi}n_{1,i} + n_{1,i}n_{2,i}^*\right)\right|. \end{aligned} \tag{6}$$

Equation (6) shows both of the estimation bias effects discussed above. The first term will be less than the desired $|cc^*|$ because of the sinusoidal factor, and the difference in magnitude will be dependent on the fringe rate f. This will cause estimated correlations to be less than the true correlation of the signals. Each of the remaining three terms is zero in ensemble expectation, given independence of the noise and signal in each channel, but for any actual set of measurements the pixel summation of each term will not be zero. The presence of these terms will cause an overestimation of the correlation. While averaging many pixels, or looks, reduces the value of the overestimation bias, practical computation of the correlation always yields a correlation estimate that is higher than the true value.

There is in addition another source of overestimation bias that will result from the correction algorithm we describe in the next section. We will address this error after introducing our method.

Because both the underlying fringe rate and the number of looks used in an interferogram vary, correlation associated with a given pixel may be estimated as too high or too low, but rarely correctly.

B. Correction Algorithms

Our estimation method compensates for both the over- and underestimation biases described above in a two-step algorithm. Since most interferograms contain spatially variable fringes caused by viewing geometry, topography, and motion, in the first step we remove those fringes from the interferogram. Fringe removal comprises three calculations. First, we divide the interferogram phase data into small regions of reasonably uniform fringe rates by selecting a size for a block of pixels for which the observed fringes tend to be of constant spacing and direction. Typical box sizes are 8×8 or 16×16 pixels. Second, we estimate the initial phase and the average fringe rate inside each box using a Fourier filter. Finally, we subtract the estimated phase pattern from each box in the interferogram and reassemble the boxes to form a complete defringed image. These steps are illustrated in Fig. 1, in which we apply the algorithm to a 48×48 pixel image extracted from an interferogram of a part of Antarctica. The subset size for fringe removal in this example is 8×8 pixels. The defringed image no longer shows dominating phase fringes due to topography and ice motion.

(a) (b) (c)

Fig. 1. Sample measured interferogram consists of a 48×48 pixel matrix extracted from a larger interferogram acquired over Antarctica. The image matrix (a) is divided into 8×8 pixel boxes, the average fringe rate is estimated (b), and subtracted from the fringe pattern to result in a "flatter" interferogram (c). Each fringe represents one cycle of phase in the interferogram.

This approach is not without its flaws. Application of the phase-flattening algorithm forces detection of fringes in noisy areas, as discussed above, since even pure noise has a dominant term in its Fourier expansion. Hence, the algorithm overestimates correlation in noise-only areas by overcorrecting for fringes if none exist. This adds to the overestimation terms described in the previous section.

Another issue is the assumption of a uniform fringe rate over the small boxes used for the local fringe removal. If the rate is not constant, but rather a gradient in fringe rate is present, the filter will only measure a component of the topography or motion pattern and underestimate the correlation. This suggests that smaller box sizes are preferred. But smaller box sizes mean that the dominant Fourier coefficient of the noise portion of the data will be larger in a relative sense, so its subtraction will lead to an overestimation of correlation by overcorrecting the fringe pattern.

We can illustrate the uniformity problem by applying the defringing algorithm to simulated data with a fringe rate gradient. Fig. 2 shows a simulated 208×208 block of phase fringes that vary with increasing rate to the lower right corner of the block. Here, we applied the defringing algorithm to this fringe pattern using 8×8 and 16×16 box sizes. Note that the smaller box size produces less residual phase, as the bandwidth of the fringe signal is less over smaller areas. We correct only one dominant fringe rate in each box; hence, a smaller fringe bandwidth in an 8×8 box as compared to a 16×16 box leaves smaller residual phase.

Every interferogram has different features, and the defringing algorithm must be adapted accordingly. The filter box size must be small enough so that the fringe rate is fairly constant. At the same time, as we now show, it must be large enough so that noise is not mistaken for fringes and processed as such. To illustrate this error, consider the example of the correction algorithm shown in Fig. 1. Note misestimation in several cases in the upper right corner of the image, where the original matrix consists mainly of phase noise, yet the algorithm nonetheless identifies a dominant fringe pattern. This pattern is subtracted from the noise-only data, and as a result, correlation is overestimated in these areas. We will remove this additional component of algorithmic bias in the second algorithm step.

The presence of the unwanted terms in (6) and the overcorrection resulting from subtracting a component of the noise lead to

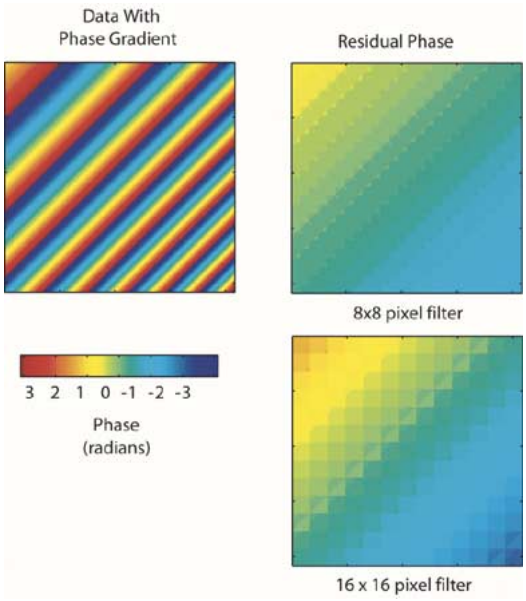


Fig. 2. Examples of applying defringing algorithm to data with nonconstant fringe rate. (Left) Fringe rate increases to lower right corner of input images, and defringed data are not uniformly zero phase. (Top right) Results from 8×8 filter boxes. (Bottom right) Results for for 16×16 box size.

overestimation bias. We model this bias by calculating correlation estimates of simulated data as a function of true correlation and the number of looks in the estimate. As described above, this results in overestimation of the correlation, and we record the difference between our simulated signal estimates and the underlying correlation. We then fit a polynomial correction curve to the bias observed in the simulated, which we then apply to the signal correlations derived from the flattened interferograms in order to infer their unbiased correlation. An example bias correction curve is shown in Fig. 3. Note that since the correction is essentially flat for correlation less than about 0.05, data in this region are often corrected to zero or less. This is a significant improvement over the previous lower limit of 0.2–0.25 where the observed correlations diverge in uncorrected data. We used a numerical inversion of this curve in our data processing, to which an approximate fourth-order polynomial is

$$-1.6185+23.3193\cdot\rho-112.4498\cdot\rho^2+244.5408\cdot\rho^3-193.906\cdot\rho^4$$

valid over the range $0.14 < \rho < 0.43$. We would like to point out that analytical expressions for correlation bias as a function of the number of looks have been derived [14], but these too are most easily inverted numerically as the expressions are quite complex.

In summary, our estimation method consists of two steps: adjusting the phase of the interferogram to remove fringe patterns that lead to underestimating the correlation and then applying a curve to remove the overestimation bias.

C. Application to a Large Interferogram

In this section, we apply the two-step correlation estimation algorithm to data acquired over a small region of Antarctica. The interferogram from this area possesses fringes from both topography and ice stream flow. Fig. 4 displays a portion of the original interferogram [Fig. 4(a)] and the same part of the defringed

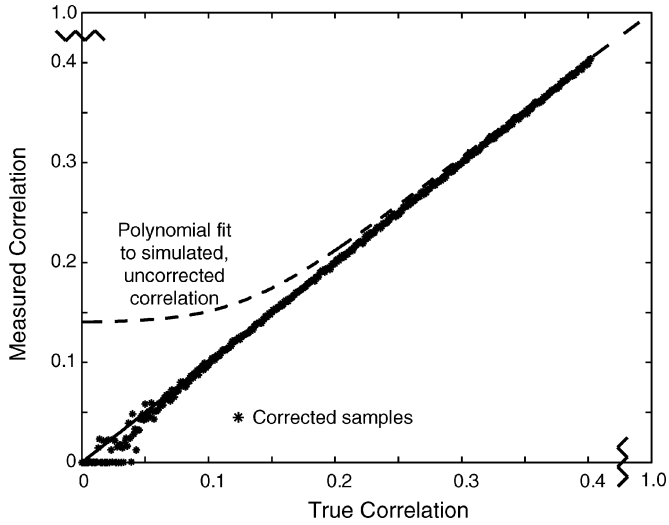


Fig. 3. Measured correlations (dashed line) exceed true correlations (solid line) due to bias from extra terms in estimator and from the reduction in noise variance after defringing algorithm is applied. The dashed line is a polynomial fit to the simulated data; its difference from "true" correlation is inverted to form the correction algorithm. Estimates approximate true correlations closely for $\rho_{\text{true}} > 0.4$, but exceed significantly for lower values. Dashed line derived from defringed measurements using an 8×8 pixel box, assuming 35 equivalent looks over the 64 pixels. Stars are correlations corrected for the overestimation bias. Dashed curve is essentially the same as the theoretical estimate derived by Hoen [10].

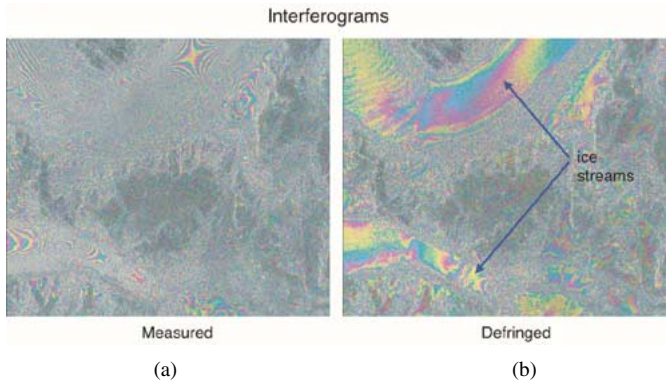


Fig. 4. (a) Measured interferogram exhibits dense fringe patterns as a result of topography and ice flow on the surface. The fringe rate is so high in most areas that aliasing in the printing process prevents their identification. In step 1) of our method, the image is divided into 8×8 pixel boxes, the average fringe rate is estimated, and subsequently subtracted from the fringe pattern. (b) Resulting interferogram, which shows fringes that vary much more slowly so that the local spatial averages needed for estimating correlation are closer to the underlying values, independent of fringe rate.

interferogram [Fig. 4(b)] after the first step of our method. The difference between the original and defringed interferograms is clearly visible. Dense fringes cover much of the original interferogram, whereas large areas of slowly varying phase are seen in most of the defringed interferogram. We note that some regions can have such dense fringes due to steep topography or ice motion that aliasing prevents accurate measurement of the fringe rate given the resolution of the radar, and we would expect reduced performance in these areas.

Correlation estimates over the full dataset, plus those after the overestimation bias correction (Fig. 3), are applied, are shown in Fig. 5. The correlations of the original, defringed, and bias-corrected data are shown in image form. We see that removing

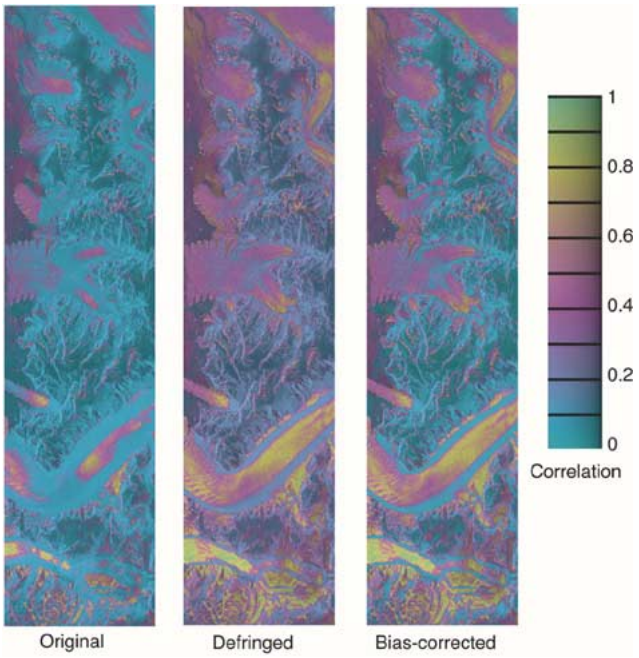


Fig. 5. Correlations derived from (a) original, (b) defringed, and (c) bias-corrected data are shown in image form. Fringe removal results in higher correlation values as underestimation bias is corrected, and this is quite evident in the ice streams. Correcting for overestimation biases yields high correlation values in some regions yet gives areas of very low correlation in the mountainous areas.

the phase fringes results in higher correlation values; this is especially visible in the ice streams. The lack of areas with low or zero correlation within the defringed correlation estimate image reveals the influence of the algorithmic bias. After correcting for this bias, the resulting correlation image [Fig. 5(c)] now expresses high correlation values in some regions but also contains areas of very low correlation in the mountainous areas.

This bias-corrected correlation image is the final product of the defringing and bias-correction steps and represents a better estimate of interferometric correlation than simple calculation of correlation. For problems requiring accurate quantification of correlation, inferred parameters such as accumulation rates will be that much more accurate.

ACKNOWLEDGMENT

The authors would like to thank the NASA Polar Processes Program and the Alaska SAR facility for related support for this work. The authors also note that all four referees returned useful comments that have improved the original submission significantly. In particular, the authors would like to thank C. Werner for several very good suggestions for completeness of this manuscript.

REFERENCES

[1] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 950–959, Sep. 1992.

[2] S. Lyons and D. Sandwell, "Fault creep along the southern San Andreas from interferometric synthetic aperture radar, permanent scatterers, and stacking," *J. Geophys. Res.*, vol. 108, no. B1, pp. ETG11–1–24, Jan. 2003.

[3] J. P. Matthews, H. Kamata, S. Okuyama, Y. Yusa, and H. Shimizu, "Surface height adjustments in pyroclastic-flow deposits observed at Unzen Volcano by JERS-1 SAR interferometry," *J. Volcanol. Geotherm. Res.*, vol. 125, no. 3–4, pp. 247–270, Jul. 2003.

[4] L. C. Smith, "Emerging applications of interferometric synthetic aperture radar (InSAR) in geomorphology and hydrology," *Can. Geotech. J.*, vol. 40, no. 5, pp. 385–398, Oct. 2003.

[5] C. L. Werner, S. Hensley, P. A. Rosen, and U. Wegmueller, *Comparison of Repeat Track Interferometric Correlation Signatures from ERS-1, ERS Tandem, SIR-C, and JERS-1 in ERS SAR Interferometry, Volume 2*. Paris, France: European Space Agency, 1997, pt. 1, vol. 406, European Space Agency Special Publiction.

[6] H. A. Zebker, P. Rosen, S. Hensley, and P. J. Mouginis-Mark, "Analysis of active lava flows on Kilauea volcano, Hawaii, using SIR-C radar correlation measurements," *Geology*, vol. 24, no. 6, pp. 495–498, 1996.

[7] C. W. Chen and H. A. Zebker, "Two -dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization," *J. Opt. Soc. Amer., A*, vol. 18, no. 2, pp. 338–51, Feb. 2001.

[8] H. A. Zebker, C. L. Werner, P. Rosen, and S. Hensley, "Accuracy of topographic maps derived from ERS-1 radar interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 823–836, Jul. 1994.

[9] R. N. Treuhaft and M. Moghaddam, "A unified analysis of radar interferometry and polarimetry for the estimation of forest parameters," in *Proc. IGARSS*, vol. 3, Seattle, WA, Jul. 6–10, 1998, pp. 1373–5.

[10] E. W. Hoen, "A correlation-based approach to modeling interferometric radar observations of the Greenland ice sheet," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2001.

[11] S. Oveisgharan and H. A. Zebker, "A snow accumulation map of part of greenland derived from InSAR correlation observations, Proc. Amer. Geophysical Union 2003 Fall Meeting," *EOS Trans. AGU*, vol. 84, no. 46, Fall 2003. Abstract C31B-0407.

[12] K. C. Jezek, "RADARSAT-1 antarctic mapping project; change-detection and surface velocity campaign," in *Papers from the 4th Int. Symp. Remote Sensing in Glaciology*, vol. 34, Annals of Glaciology, J. Winther, Ed., 2002, pp. 263–268.

[13] C. L. Werner, S. Hensley, and P. Rosen, "Application of the interferometric correlation coefficient for measurement of surface change," in *Proc. 1996 AGU Fall Meeting*, San Francisco, CA, 1996.

[14] R. F. Hanssen, *Radar Interferometry, Data Interpretation and Error Analysis*. Dordretch, The Netherlands: Kluwer, 2001, pp. 96–97.

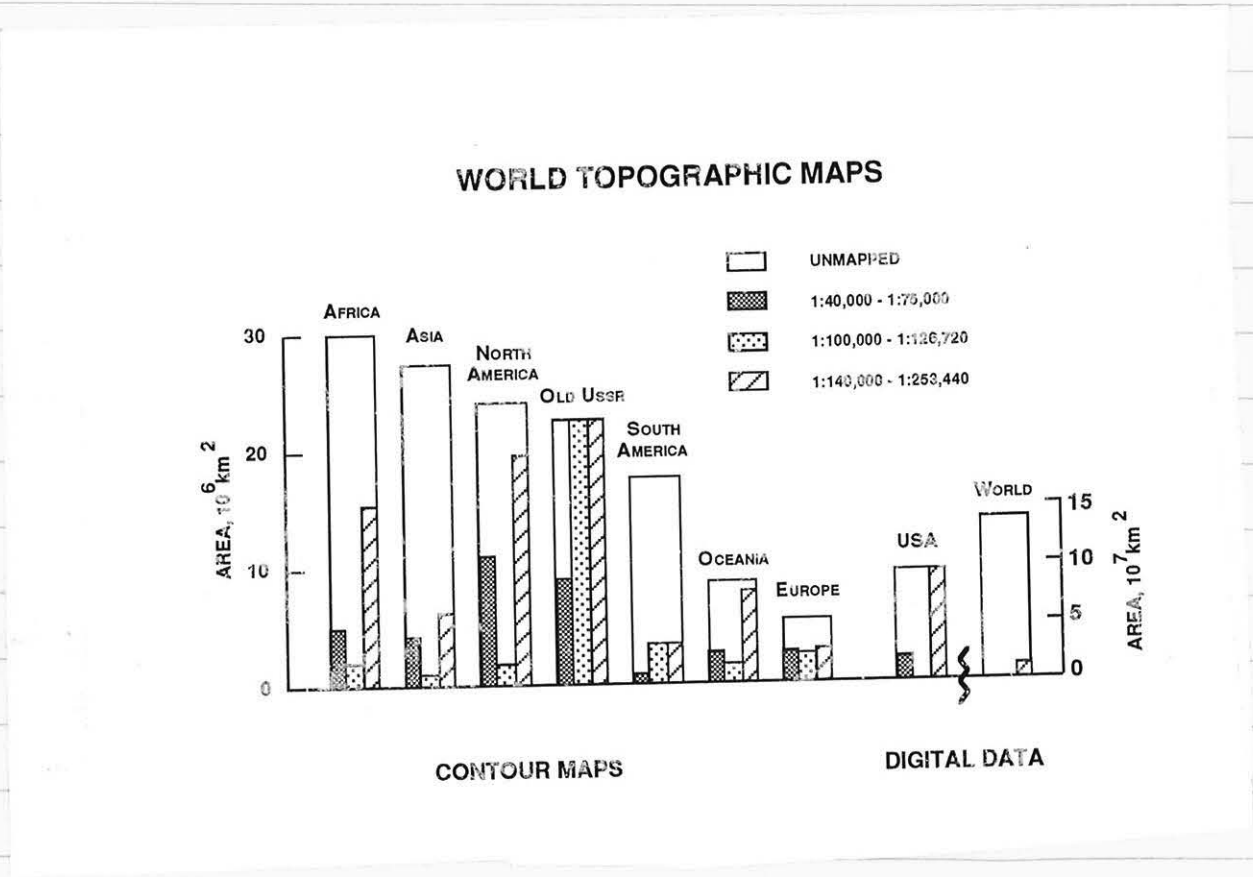
GP265
EE355

Handout #37

Measurement of Topography

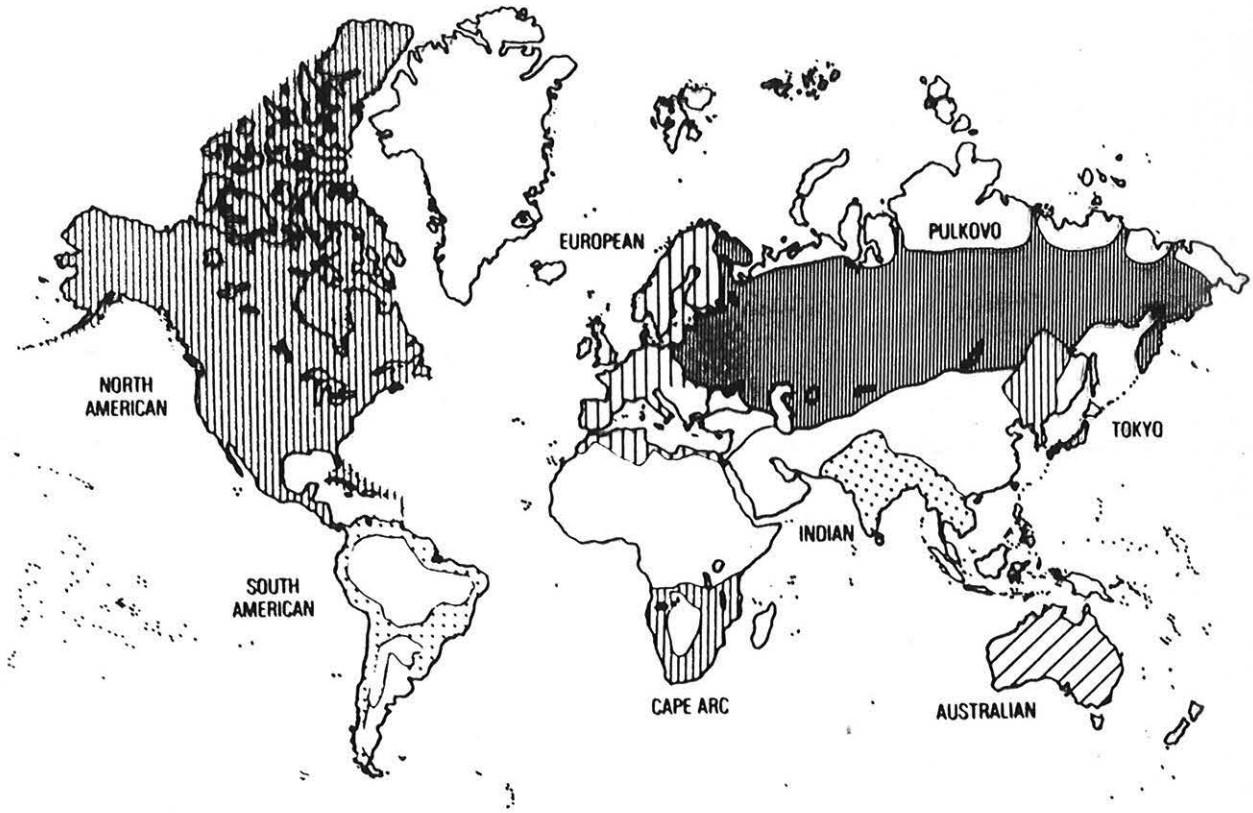
Now that we have demonstrated our ability to measure topography, let's look more in detail at the actual measurement of it. First, some rationale - why do we want to measure topography?

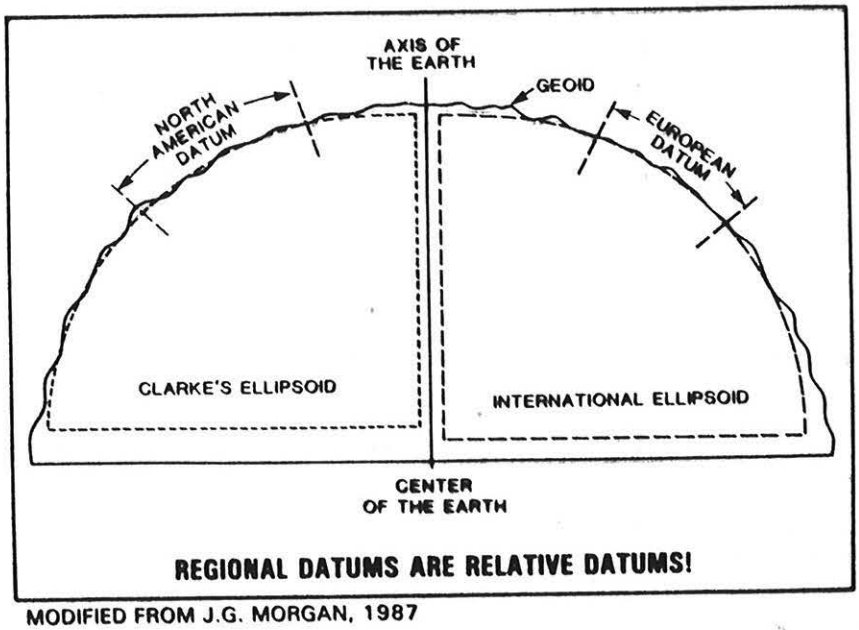
1. Not much of the world is well mapped, especially in digital form:



2. Different parts of the world are mapped on different "datums", or coordinate systems, making it difficult to implement global studies where values from one section of the Earth must be compared to those from other sections.

MAJOR GEODETIC DATUM BLOCKS





MODIFIED FROM J.G. MORGAN, 1987

Finally, let's take a look at what kind of accuracies we'd need for a variety of scientific studies:

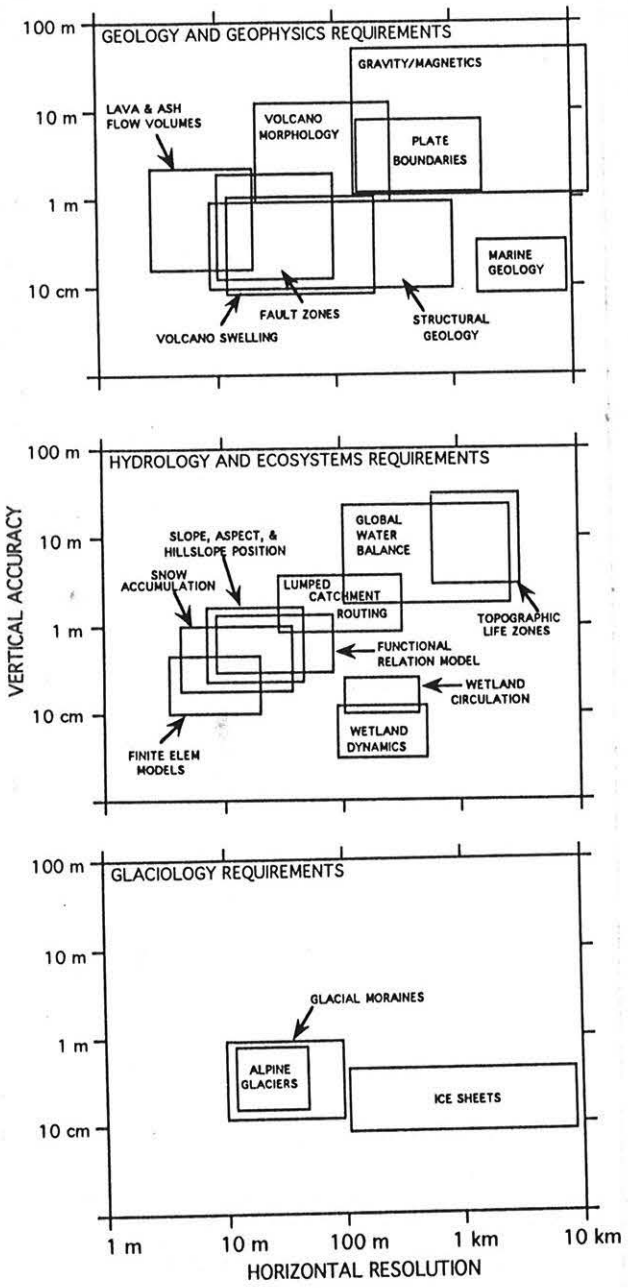
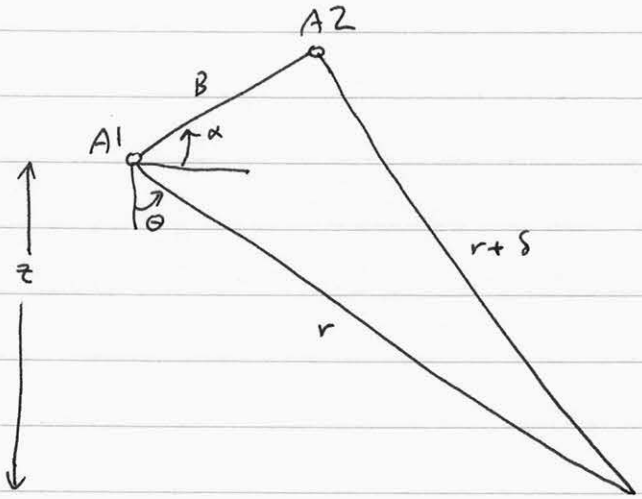


Fig. 1. Graphical depiction of horizontal and vertical topographic data accuracies required for several discipline studies. Each box represents a range of requirements for differing aspects of the studies.

Solution for Topography from Interferometric Phase

For our actual implementation of the topographic interferometer, we'll will rederive the phase vs. surface height equation without making the parallel-ray approximation, as the exact solution will be called for in precision applications. Again consider our construction:



We'll use previous definitions for all of the above quantities. From the law of cosines,

$$(r+\delta)^2 = r^2 + B^2 - 2rB\sin(\theta-\alpha)$$

from which

$$\sin(\theta-\alpha) = \frac{r^2 + B^2 - (r+\delta)^2}{2rB}$$

Also,

$$\delta = -\frac{\lambda}{4\pi}\phi$$

and

$$z = r\cos\theta$$

Hence given B, α, and r we can solve for θ, which gives us z. As before, we can estimate our phase error from the derivatives. We want

$$\frac{\partial z}{\partial \phi} = \frac{\partial z}{\partial \theta}\frac{\partial \theta}{\partial \delta}\frac{\partial \delta}{\partial \phi}$$

$$\frac{\partial z}{\partial \theta} = \frac{\partial}{\partial \theta} r\cos\theta = -r\sin\theta$$

For $\frac{\partial \theta}{\partial \delta}$:

$$\sin(\theta-\alpha) = \frac{r^2 + B^2 - (r+\delta)^2}{2rB}$$

$$\cos(\theta-\alpha)\,\partial\theta = -\frac{2(r+\delta)}{2rB}\partial\delta$$

$$\approx -\frac{\partial\delta}{B}$$

$$\therefore \frac{\partial\theta}{\partial\delta} = \frac{-1}{B\cos(\theta-\alpha)}$$

and

$$\frac{\partial\delta}{\partial\phi} = \frac{-\lambda}{4\pi}$$

so $$\frac{\partial z}{\partial\phi} = (-r\sin\theta)\left(\frac{-1}{B\cos(\theta-\alpha)}\right)\left(\frac{-\lambda}{4\pi}\right)$$

$$= -\frac{\lambda}{4\pi}\,\frac{r\sin\theta}{B\cos(\theta-\alpha)}$$

Which is the same result as we had in the approximate case. In other words, the error estimate for the parallel-ray approximation was good.

Now, There is a second error source in addition to the phase error. The baseline may not be known well either in terms of length B, or more likely in terms of position α. These contributions to height error are

$$\frac{\partial z}{\partial B} = \frac{\partial z}{\partial \theta} \cdot \frac{\partial \theta}{\partial B}$$

$$\frac{\partial \theta}{\partial B} \cdot$$

$$\cos(\theta-\alpha)\,\delta\theta = \frac{2}{\partial B} \frac{r^2+B^2-(r+\delta)^2}{2rB} \partial B$$

$$= \frac{2r\left(r^2+B^2-(r+\delta)^2\right) - 2rB \cdot 2B}{4r^2B^2}$$

$$= \frac{2r\left(B^2-2r\delta-\delta^2\right) - 4rB^2}{4r^2B^2} \partial B$$

$$\approx -\frac{\delta}{B^2} \partial B$$

$$\frac{\partial \theta}{\partial B} = \frac{-\delta}{B^2\cos(\theta-\alpha)}$$

$$\frac{\partial z}{\partial B} = \frac{r\sin\theta \sin(\theta-\alpha)}{B\cos(\theta-\alpha)}$$

This term usually is quite small in practical systems. But another baseline error term, from uncertainty in orientation, can be quite large.

$$\frac{\partial z}{\partial \alpha} = \frac{\partial z}{\partial \theta} \cdot \frac{\partial \theta}{\partial \delta} \cdot \frac{\partial \delta}{\partial \alpha}$$

We already found $\frac{\partial z}{\partial \theta} = -r\sin\theta$, and $\frac{\partial \theta}{\partial \delta} = \frac{-1}{B\cos(\theta-\alpha)}$.

So, we need $\frac{\partial \delta}{\partial \alpha}$

$$\sin(\theta-\alpha) = \frac{r^2 + B^2 - (r+\delta)^2}{2rB}$$

$$-\cos(\theta-\alpha)\,\partial\alpha = \frac{-2(r+\delta)}{2rB}\,\partial\delta$$

$$\approx -\frac{\partial\delta}{B}$$

$$\frac{\partial\delta}{\partial\alpha} = B\cos(\theta-\alpha)$$

Thus

$$\frac{\partial z}{\partial \alpha} = (-r\sin\theta)\left(\frac{-1}{B\cos(\theta-\alpha)}\right)(B\cos(\theta-\alpha))$$

$$= r\sin\theta$$

This term may be understood as an unknown "tilt" across the image. In other words we cannot distinguish a tilt in the interferometer itself from a slope on the surface.

Let's plug a few numbers into these to understand the contributions to error from each. We'll use the NASA TOPSAR airborne system as an example. Its parameters are

$r \approx 10$ km

$\theta \approx 30°$

$B = 1.5$ m

$\alpha = 63°$

$\lambda = 6$ cm

Using our formulas from above,

$$\text{Phase error} \Rightarrow \sigma_z = \frac{\lambda}{4\pi} \frac{r\sin\theta}{B\cos(\theta-\alpha)} \sigma_\phi$$

$$\text{Baseline length} \Rightarrow \sigma_z = \frac{r\sin\theta \tan(\theta-\alpha)}{B} \sigma_B$$

$$\text{Baseline angle} \Rightarrow \sigma_z = r\sin\theta \, \sigma_\alpha$$

So, we need to evaluate (or measure) the system uncertainties on the right hand sides above. Beginning with σ_ϕ, suppose the system has an SNR of 20 dB.

$$\sigma_{\phi,1\,look} = \frac{1}{\sqrt{SNR}} = 0.1 \text{ radians}$$

Now, this radar calculates about 10 looks in the interferogram before we estimate the height. We'll just use the following formula which is found in several papers (and by reference in today's other handout)

$$\sigma_{\phi,\,multilook} \approx \frac{\sigma_{\phi,\,1\,look}}{\sqrt{2 \cdot N}}$$

where N is the number of looks. Thus

$$\sigma_\phi \approx 1\tfrac{1}{4}^\circ \text{ or } 0.022 \text{ radians}$$

thus $$\sigma_z = \frac{\lambda}{4\pi} \frac{r\sin\theta}{B\cos(\theta-\alpha)} \sigma_\phi = 0.42\text{m}$$

For baseline length, aircraft deformation limits σ_B to about 10^{-4} m,

$$\sigma_z = \frac{r \sin\theta \tan(\theta-\alpha)}{B} \sigma_B = 0.216 \text{ m} \quad \leftarrow \text{half of phase error}$$

And for orientation, $\sigma_\alpha \approx 0.01°$, so

$$\sigma_z = r \sin\theta \, \sigma_\alpha = 0.88 \text{ m} \quad \leftarrow \text{twice the phase error}$$

So, we see that uncertainty in orientation is the major source of error in a typical system.

Solution for height

We obtained a set of equations above to allow us to determine heights from the phase measurements. This means if we know our imaging geometry and the absolute phase at each point we can solve for the topography. This requires, though, that we have "unwrapped" the phase at each point in our image.

We are going to defer discussion of phase unwrapping for the time being and discuss an approximate method of solution for now.

Begin with our equation relating surface topography to δ; through the look angle θ:

$$\sin(\theta-\alpha) = \frac{r^2 + B^2 - (r+\delta)^2}{2rB}$$

$$= \frac{r^2 + B^2 - r^2 - 2r\delta - \delta^2}{2rB}$$

$$\approx \frac{B}{2r} - \frac{\delta}{B} \qquad \text{(neglecting term of } \delta^2\text{)}$$

Note that if we neglect also the term $\frac{B}{2r}$ we set the plane parallel relation:

$$B \sin(\theta - \alpha) = -\delta \qquad \leftarrow \text{minus sign in our current coordinates}$$

Then since $z = r\cos\theta$,

$$\delta = -B \sin\left(\cos^{-1}\frac{z}{r} - \alpha\right)$$

$$= -B\left[\sin\left(\cos^{-1}\frac{z}{r}\right)\cos\alpha - \cos\left(\cos^{-1}\frac{z}{r}\right)\sin\alpha\right]$$

$$= -B\left[\sqrt{1-\frac{z^2}{r^2}}\cos\alpha - \frac{z}{r}\sin\alpha\right]$$

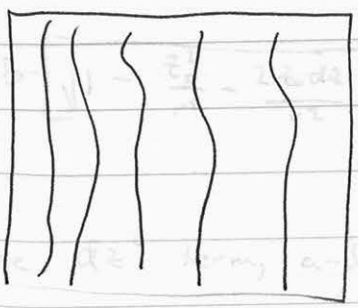
The phase we measure is then

$$\phi = \frac{4\pi}{\lambda} B\left[\sqrt{1-\frac{z^2}{r^2}}\cos\alpha - \frac{z}{r}\sin\alpha\right]$$

~~Expand the~~

This total phase contains the basic interferometer fringe pattern modulated by topographic variation. The "flat earth" pattern is simply that corresponding to $z = z_0$ everywhere.

This complicated-looking expression is simpler when viewed as an image!



← fringe pattern decreasing in spatial frequency with range, and modulated by topographic features.

Now, let's simplify and evaluate ϕ

$$\phi = \frac{4\pi B}{\lambda r}\left[\sqrt{r^2 - z^2}\cos\alpha - z\sin\alpha\right]$$

Expand about z_0 : $z = z_0 + dz$

(z_0: average height; dz: topography)

$$\phi = \frac{4\pi B}{\lambda r}\left[\sqrt{r^2 - z_0^2 - 2z_0 dz + dz^2}\cos\alpha - (z_0 + dz)\sin\alpha\right]$$

Neglecting dz^2 and making our usual approximation, (for $\sqrt{\ }$)

$$\phi = \frac{4\pi B}{\lambda r}\left[\sqrt{r^2 - z_0^2}\sqrt{1 - \frac{2z_0 dz}{r_0^2 + z_0^2}}\cos\alpha - z\sin\alpha - dz\sin\alpha\right]$$

$$= \frac{4\pi B}{\lambda r}\left[\sqrt{r^2 - z_0^2}\left(1 - \frac{z_0 dz}{r_0^2 - z_0^2}\right)\cos\alpha - z\sin\alpha - dz\sin\alpha\right]$$

Since we know the flat earth ($z = z_0$) pattern, subtract it off to be left with the topographic term:

$$\phi_{topo} = -\frac{4\pi B}{\lambda r}\left[\frac{z_0}{\sqrt{r^2 - z_0^2}}\cos\alpha + \sin\alpha\right]dz$$

$$= -\frac{4\pi B}{\lambda r}\left[\frac{1}{\tan\theta_0}\cos\alpha + \sin\alpha\right]dz$$

Hence the topography is linearly related to altitude, with the constant as shown above.

What does this imply? That constant-phase contours will appear at constant heights on the imaged topography. The phase will repeat every 2π radians, so therefore each "fringe" corresponds to height difference (the ambiguity height) defined by

$$2\pi = \frac{-4\pi B}{\lambda r}\left[\frac{1}{\tan\theta_0 \cos\alpha + \sin\alpha}\right] dz$$

$$dz = \frac{2\pi}{\frac{-4\pi B}{\lambda r}\left[\frac{1}{\tan\theta_0 \cos\alpha + \sin\alpha}\right]}$$

$$= \frac{-\lambda r}{2B\left[\frac{1}{\tan\theta_0 \cos\alpha + \sin\alpha}\right]}$$

For our TOPSAR system, the ambiguity height is

$dz \approx$ [redacted] ≈ 120 m

so contours in phase repeat every [redacted] ≈ 120 m, approximately. Note that the relation strictly holds only for a single value of r, and hence θ_0. Complete topographic reduction requires inclusion of the range dependence.

Relating error to correlation

Our equation for height uncertainty gives us the height error in terms of the standard deviation of the phase σ_ϕ

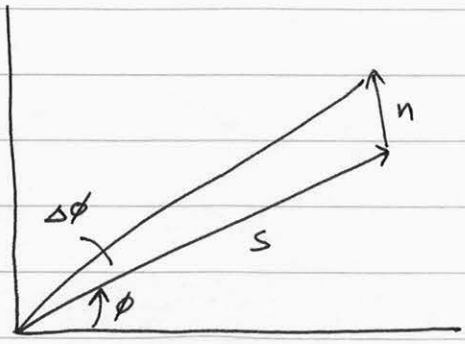
$$\sigma_z = [\text{redacted}]\ \frac{\lambda r}{4\pi B}\frac{\sin\theta}{\cos(\theta-\alpha)}\sigma_\phi$$

In order to use this expression we have to evaluate σ_ϕ either

from a very good system model, or else by direct measurement from the data. The latter is often dictated because we have only the data and no accurate system model. In addition, even if we know the system well we don't a priori know the σ^0 and hence the SNR at all points in the image.

But note that we already have a measure of system performance related to phase variance - the correlation. Thus, we can express the height uncertainty in terms of the correlation we estimate as we form the interferogram. This approach allows us to easily model variable height accuracy in a scene.

Let's more carefully rederive the relationship of σ_ϕ to SNR. As before let's look at the following construction, valid for high SNR, essentially the only case of interest.



This vector plot shows that $\Delta\phi$ is

$$\Delta\phi = \frac{\text{Component of } n \text{ perpendicular to } s}{|s|}$$

Since $SNR = \dfrac{\langle |s|^2 \rangle}{\langle |n_\perp|^2 \rangle + \langle |n_\parallel|^2 \rangle}$,

$$\sigma_\phi = \sqrt{\frac{1}{2 \cdot SNR}}$$ ← holds for large number of looks most accurately

and

$$\rho = \frac{1}{1+\frac{1}{SNR}}$$

thus

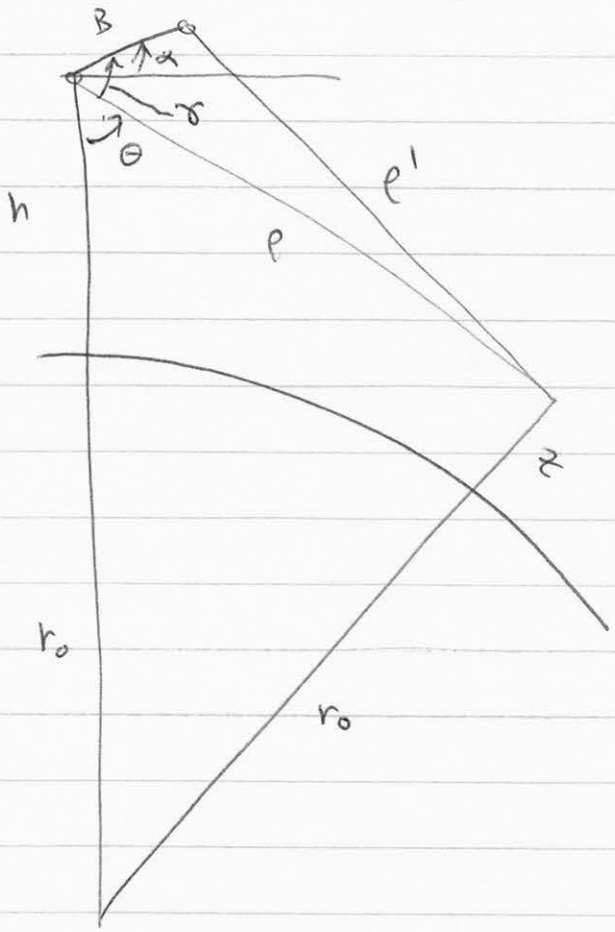
$$\sigma_\phi = \sqrt{\frac{1-\rho}{2\rho}}$$

so that

$$\sigma_z = \frac{\lambda r}{4\pi B} \frac{\sin\theta}{\cos(\theta - \alpha)} \sqrt{\frac{1-\rho}{2\rho}}$$

"Curved-Earth" phase pattern

What does the background phase from a curved planet look like in an interferogram?



The phase we measure is the difference $\rho' - \rho$. By the law of cosines

$$\rho'^2 = \rho^2 + B^2 - 2\rho B \cos\gamma$$

But $\gamma = \left(\frac{\pi}{2} - \theta\right) + \alpha$, so

$$\cos\left(\frac{\pi}{2} - \theta + \alpha\right) = \sin(\theta - \alpha)$$

so

$$\rho'^2 = \rho^2 + B^2 - 2\rho B \sin(\theta - \alpha)$$

$$\rho'^2 - \rho^2 = B^2 - 2\rho B \sin(\theta - \alpha)$$

$$\rho' - \rho \approx \frac{B^2 - 2\rho B \sin(\theta - \alpha)}{2\rho}$$

$$= \frac{B^2}{2\rho} - B \sin(\theta - \alpha) \qquad (\text{for } B \ll \rho)$$

Thus the phase as a function of θ and ρ is

$$\phi(\theta) = -\frac{4\pi}{\lambda}\left[\frac{B^2}{2\rho} - B\sin(\theta - \alpha)\right]$$

So, how does ϕ depend on ρ?

$$(z + r_0)^2 = (h + r_0)^2 + \rho^2 - 2\rho(h + r_0)\cos\theta$$

$$2\rho(h + r_0)\cos\theta = (h + r_0)^2 + \rho^2 - (z + r_0)^2$$

$$\cos\theta = \frac{(h + r_0)^2 + \rho^2 - (z + r_0)^2}{2\rho(h + r_0)}$$

or

$$\theta = \arccos\left(\frac{(h + r_0)^2 + \rho^2 - (z + r_0)^2}{2\rho(h + r_0)}\right)$$

and the curved earth pattern for no topography is this equation with $z = 0$.

Mapping the World's Topography Using Radar Interferometry: The TOPSAT Mission

HOWARD A. ZEBKER, SENIOR MEMBER, IEEE THOMAS G. FARR, RONALD P. SALAZAR, AND TIMOTHY H. DIXON

Invited Paper

Global-scale topographic data are of fundamental importance to many Earth science studies, and obtaining these data is a priority for the Earth science community. Several groups have considered the requirements for such a data set, and a consensus assessment is that many critical studies would be enabled by the availability of a digital global topographic model with accuracies of 2 and 30 m in the vertical and horizontal directions, respectively. Radar interferometric techniques have been used to produce digital elevation models at these accuracies and are technologically feasible as the centerpiece of a spaceborne satellite mission designed to map the world's land masses, which we denote TOPSAT. A radar interferometer is formed by combining the radar echoes received at a pair of antennas displaced across-track, and specialized data processing results in the elevation data. Two alternative implementations, one using a 2 cm-λ radar, and one using a 24 cm-λ radar, are technologically feasible. The former requires an interferometer baseline length of about 15 m to achieve the required accuracy, and thus could be built on a single spacecraft with a long extendable boom. The latter necessitates a kilometers-long baseline, and would thus be best implemented using two spacecraft flying in formation. Measurement errors are dominated by phase noise, due largely to signal-to-noise ratio considerations, and attitude errors in determining the baseline orientation. For the 2-m accuracy required by TOPSAT, the orientation must be known to 1 arc-second. For the single-spacecraft approach, where attitude would be determined by star tracking systems, this performance is just beyond the several arc-second range of existing instruments. For the dual-spacecraft systems, though, differential global positioning satellite measurements possess sufficient accuracy. Studies indicate that similar performance can be realized with either system.

I. INTRODUCTION

An accurate description of the surface elevation of the Earth is of fundamental importance to many branches of Earth science, as has been detailed in the reports of several working groups [1]–[4]. These studies have considered the characteristics of the available topographic database as well as existing and possible new scientific applications of high-resolution topographic data. Key findings of these groups are that there are significant deficiencies in available topographic data, that existing and potential new scientific applications are severely limited by these deficiencies, and that ready availability of a worldwide digital elevation model is a priority for these studies. The required characteristics of the data set are that i) the data cover the majority of the Earth's land surface, ii) all data be expressed in a common coordinate system, iii) the data exhibit uniform accuracy, and iv) the vertical accuracy and spatial resolution of the model be approximately 2 and 30 m, respectively.

The development of interferometric radar systems for the measurement of highly accurate digital elevation models (DEM's) has by now been well-documented in the literature [5]–[12]. The precision of the technique is now commensurate with conventional optical-stereo photogrammetric procedures [13], while data reduction time is a small fraction of that required by the optical-stereo instrument systems, which follows mainly from the automated nature of the radar data processing. In addition, because the radar signals easily penetrate even dense clouds in the atmosphere and are independent of solar illumination, there are also far fewer restrictions on data acquisition times and geometries from either airborne or spaceborne platforms.

Therefore, it is feasible to design a space mission dedicated to the acquisition of a global digital topographic database at accuracies superior to existing elevation models, including those described only in paper contour map format. In fact, NASA has been studying such a mission for several years to support many ongoing research programs and the general science community. Consideration of accuracy, facility of data reduction, and cost has led to the selection of the interferometric radar technique as the leading candidate for this mission, which we denote TOPSAT. The principal TOPSAT goal is to map the topography of the entire Earth in less than a year at an accuracy similar to that exhibited by 1:50000 maps.

Manuscript received December 7, 1993; revised March 22, 1994. The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under Contract with the National Aeronautics and Space Administration.

H. A. Zebker, T. G. Farr, and R. P. Salazar are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA.

T. H. Dixon is with the Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL 33149-1098 USA.

IEEE Log Number 9405553.

In this paper, we review the scientific need for global digital topographic data, including the requirements for the data set and a brief comparison of some data acquisition techniques. We then give a brief tutorial on the radar interferometric technique as applied to topographic mapping. This is followed by examples of the performance and application of such a system using data generated by TOPSAR, an airborne prototype of the space instrument, and also by repeat track analysis of data acquired by the ERS-1 radar satellite. Next, we present two possible designs of a spaceborne topographic radar system, one in which the interferometer is formed by using two antennas on a single spacecraft and one in which two platforms orbiting in formation are utilized. Finally, we present conceptual designs of the complete spacecraft systems needed to realize the TOPSAT mission.

II. Science Rationale

Digital topographic data are valuable among a wide community using maps on different scales for a variety of political, social, and scientific applications. In this section we will discuss quantifiable scientific studies enabled by large-scale, high-resolution elevation data. We then will present summaries of required accuracies for several discipline investigations. Finally, we will discuss several technologies used for generation of topographic data and show why the radar interferometric technique is the leading candidate for a spaceborne topographic mission.

Among the studies requiring continental topographic data are hydrology, ecology, glaciology, geomorphology, and atmospheric circulation. For example, in hydrologic and terrestrial ecosystem studies, topography exerts significant control on intercepted solar radiation, water runoff and subsurface water inventory, microclimate, vegetation type and distribution, soil development, and a host of additional interdependent parameters. The topography of the polar ice caps and mountain glaciers is important because it directly reflects ice-flow dynamics and is closely linked to global climate and sea-level change. Monitoring the amplitude of seasonal advance and retreat of mountain glaciers on a global basis and longer term trends of the polar ice sheets can give important information on the rate of global warming. Accurate mapping of the forms and slopes of young geomorphic features such as glacial moraines and feature offsets and scarps due to recent geological faulting can provide new information not only on the formative tectonic processes but also on the climatic and paleoclimatic processes contributing to their present form. Finally, models of the present and past general circulation of the atmosphere require topography as a fundamental input.

NASA has sponsored two working groups, the Topographic Science Working Group [3] and the Joint Topographic Science Working Group, to review previous panel reports and recommend a strategy for generating high-resolution, accurate digital elevation measurements of the entire land and ice surface of the Earth in a single consistent reference frame. The Joint Topographic Working Group, in particular, addressed horizontal and vertical resolution requirements for various disciplines. The working group report is still unpublished in its entirety, but we present a summary of their findings in Fig. 1 (private notes, authors Dixon and Zebker, who were participants in the group). This figure emphasizes the wide range (over several orders of magnitude) of requirements. Nevertheless, some common features stand out. First, several disciplines require very-high-resolution topographic data with horizontal resolution of a few tens of meters (approximately the resolution of current high-resolution space-based imaging systems such as Landsat TM and SPOT) and vertical precision of several meters or better. Acquisition of high horizontal resolution data with high vertical accuracy automatically satisfies all other lower resolution and accuracy requirements and thus is highly desirable. Vertical accuracy should not be significantly worse than vertical precision (we define the latter informally as the relative height uncertainty for adjacent pixels) to facilitate regional comparisons and comparisons of data taken at different times. High vertical precision (a few 10's of centimeters) over the polar ice sheets is particularly important to enable mass balance studies. Here, high horizontal resolution is less critical because slopes are generally lower, so widely separated measurements or averages over a few hundred meters do not, in general, cause large height biases.

Second, while high-resolution data are generally required only in specific regions, these regions may be located anywhere on the globe, and hence the data should be obtainable anywhere. This is virtually the same thing as a global requirement and demonstrates the desirability of space-based acquisition. However, if sensor power requirements, data rate, or ground processing time become significant cost drivers in a space-based mission, a compromise strategy for data acquisition and processing could be adopted whereby data are acquired or processed in high-priority regions first, building up a global data set more slowly. This approach must be traded off with the need in some applications to acquire a near-synoptic data set (see below).

The third requirement is in the area of multitemporal coverage for change detection and the related issues of synoptic coverage and accuracy. These are most critical for applications involving ice change and vegetation monitoring. While tree heights change slowly and are not usually the subject of topographic maps, the availability of temporally varying global height maps would permit global change studies such as forest destruction and regrowth. These ecological studies are not supported well by traditional mapping activities, but represent some of the new possibilities enabled by rapid global mapping technology. It is thus desirable to acquire data relatively quickly, ideally over a 1–2 year period or less, as opposed to building up a database more slowly, for example over a 5–10 year period as might be feasible with stereo-optical systems. It is feasible to acquire "near-synoptic" global data in 6 months with a radar interferometer. Seasonal or other shorter period effects will still have to be accounted for by modeling or other measurement. Obviously, if a global set could be acquired in 6 months, and the mission

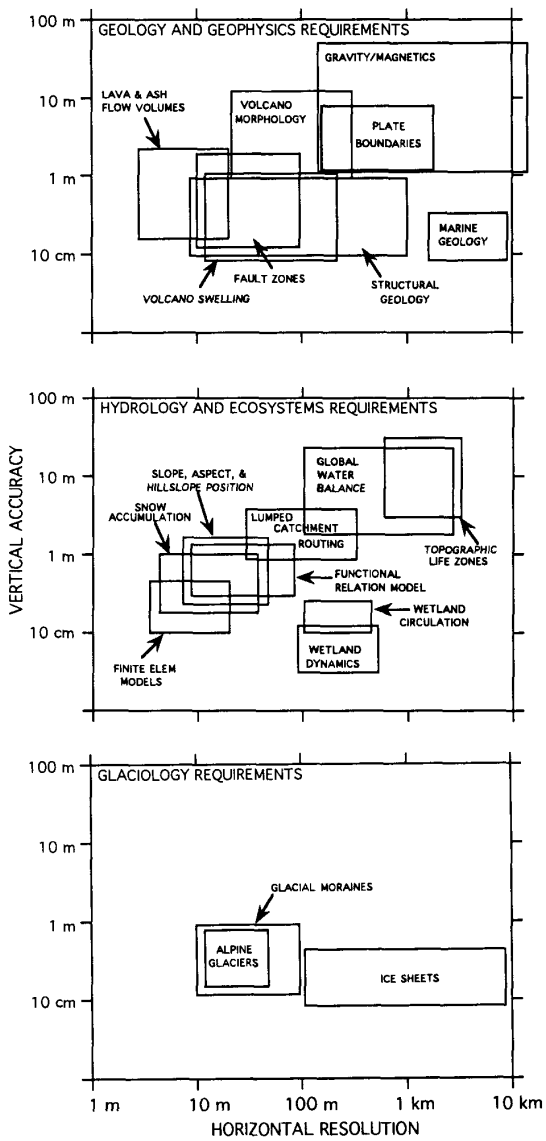


Fig. 1. *Graphical depiction of horizontal and vertical topographic data accuracies required for several discipline studies. Each box represents a range of requirements for differing aspects of the studies.*

continued for 3 years, changes over this period could be detected. Even if data acquisition ended after 1 year, future missions would benefit from a near-synoptic database for comparison purposes, assuming sufficient accuracy. The ability to compare with future data sets may be the most important constraint on the resolution and accuracy requirements of a topographic mission, as we cannot predict all possible future applications of a global, fine resolution, high-accuracy data set.

Examination of the figure above shows that many of the investigations would be enabled if elevation data at a vertical accuracy of 2–5 m and a spatial resolution of 30 m were available, requirements satisfied by the technological implementations discussed below. The principle exceptions here are the polar ice studies, which do not require fine spatial resolution but do need very precise (10-cm scale) vertical accuracies for adequate calculation of ice volumes. These values, along with coverage needs, form the mission requirements.

The existing inventory of topographic data has been produced from a variety of regional and local data sets representing a potpourri of horizontal and vertical datums, accuracies, styles, map projections, and resolutions making it nearly impossible to produce a uniform data set or assess the accuracy of the resulting derived product. The best existing digital database is the Digital Topographic Elevation Data (DTED) produced by the United States Defense Mapping Agency. DTED have been produced for about 70% of the Earth's land surface, mostly in the northern hemisphere. Completion of a global data set, exclusive of the Greenland and Antarctic ice sheets, is anticipated by year 2005. Current production of DTED is accomplished by automated photogrammetry using classified orbital stereoscopic images. Previously, DTED were produced by traditional analog photogrammetry and by digitization from contour maps. DTED, although not classified, are limited in distribution to members of the Department of Defense and their contractors. This distribution policy is currently under review. The quality characteristics of DTED, taken from public-domain DMA product specifications for photogrammetrically derived data, are about 90-m spatial resolution and tens of meters accuracy in the vertical direction.

A comparison of data requirements versus measurement performance demonstrates that DTED quality does not meet the needs of most scientific disciplines. The relatively coarse DTED horizontal grid size is insufficient for disciplines requiring local digital topographic data. Even with spatial averaging, the poor vertical accuracy of DTED, which is due mostly to large systematic errors, also precludes its suitability for most regional and global scientific disciplines.

There are at least three possible technologies for generation of future topographic data on a global scale: i) optical-stereo instrumentation, ii) laser profiling instruments, and iii) radar interferometry. Of these, the optical-stereo approach has the advantage that it utilizes existing or planned satellite systems justified by a broad spectrum of applications. Currently these include SPOT (Systeme Probatoire d'Observation de la Terre), JERS-1 (Japanese Earth Resources Satellite) OPS (Optical System), AVNIR (Advanced Visible and Near-Infrared Radiometer) on the Japanese ADEOS (Advanced Earth Observing System), HRMSI (High Resolution Multispectral Stereo Imager) on Landsat 7, and ASTER (Advanced Spaceborne Thermal Emission and Reflection) on EOS. Depending on the exact system involved, spatial resolutions ranging from 20 to 40 m and vertical accuracies of 10–60 m may be achieved [14], [15].

It is important to note that these accuracies cannot be achieved without suitable ground control point knowledge, clearly an undesirable need for a global system. In addition,

truly global coverage is unlikely, even with a space mission, due to orbital limitations and the requirement for two cloud-free scenes with compatible imaging geometry. For these reasons, stereo-optical data would likely be acquired in a piecemeal fashion, slowly building up coverage from a variety of missions with different orbits, illumination conditions, and accuracies. Thus space-based stereo-optical data would suffer from one of the most vexing problems with existing digital topographic databases, namely the lack of consistency.

Perhaps the major constraint on any stereo-optical approach is the existence of clouds in the Earth's atmosphere. Many areas of the globe are cloud-covered much of the time (especially high-relief or tropical areas) and have never been photographed from space. This is not to say that such areas are cloud-covered all of the time. However, any sun-synchronous orbital platform is constrained to fly near local noon (± 2 h), in order to minimize shadows and to ensure adequate solar illumination for passive optical sensors. Especially in tropical areas, cumulus clouds formed by solar heating of the ground and resultant convection generally start to form by mid-morning, severely limiting optical detection from sun-synchronous orbital platforms in certain locations.

The second approach is that of laser profiling, where one or more laser beams illuminate the Earth in a near-nadir direction to collect data directly beneath the satellite ground track. This approach has the advantage of very high vertical accuracy (≈ 0.1–1 m), but the disadvantage that for practical implementations only a very narrow swath may be acquired at one time. For example, if a 30-beam laser were employed with each beam separated by 30 m, the swath would be less than 1 km and complete orbital coverage, with overlaps, would take over 4 years, stressing the performnace in terms of lifetime and efficiency of laser transmitters. It also requires that the orbit be controlled to about 50 m, a challenge in itself. Finally, although only a single pass is required over each region of the Earth's surface, the same atmospheric limitations noted for the stereo imaging affect laser performance.

We note here that there are certain studies, such as the polar ice volume measurements, for which the laser altimeter's high vertical precision and low spatial coverage are ideal, and an overall global topographic study would benefit from the inclusion of a laser instrument to permit the polar study.

The final approach, radar interferometry, achieves the required resolutions and accuracies in a reasonable mission lifetime without interference from clouds in the atmosphere. If a very short radar wavelength is employed, there remains the possibility of interference from severe storms, which, fortunately, are much more rare than clouds in the sky. We describe this approach in detail in the next section, concluding that interferometric radar promises the highest quality product in the shortest time. The remaining discriminator is of course cost, and implementation studies are now investigating the cost issue in detail. For the remainder of this paper, we will assume that there is no significant cost advantage in selecting one of the optical approaches and will discuss only the radar implementation.

III. Radar Interferometry

A radar interferometer is formed by relating the signals from two spatially separated antennas; the separation of the two antennas is called the baseline. Radar interferometers have been realized in two ways. First, the two antennas may be mounted on a single platform. This is the usual implementation for aircraft systems [5], [6], [9], having the advantage of simultaneous observation (see below) but suffering from the disadvantage that the size of the airframe limits the achievable baseline. However, choosing a high operating frequency permits the baseline, measured in wavelengths, to be of sufficient length for meter-scale vertical accuracies. Second, synthetic interferometers have been formed by utilizing a single antenna on a satellite in a nearly exact-repeating orbit—the interferometer baseline is formed by relating radar signals on passes over the same site [7], [8], [12]. Even though the antennas do not illuminate the same area at the same time, if the ground is completely undisturbed between viewings the signals will be highly correlated and a spatial baseline may be synthesized. Here the choice of a baseline is limited only by orbit navigation accuracy, but the surface decorrelation properties must be considered. Topographic maps using this technique have been demonstrated [7], [16]–[18]. A third implementation, proposed for one possible implementation of the global spaceborne mission, is to utilize two spacecraft flying in formation. This has the advantage of obtaining arbitrary baselines while avoiding the temporal decorrelation phenomenon.

The performance of the radar interferometer depends on the radar instrument parameters, the orbit or aircraft attitude parameters, and the errors induced by the data processing and post-processing operations. For the repeat-pass implementation only, temporal decorrelation constitutes an important and in many cases the limiting error source in the operation of a topographic mapping radar. Zebker and Villasenor [18] investigated temporal decorrelation phenomena for the SEASAT 24 cm-λ (L-band) radar and were able to determine rates of decorrelation for several types of surfaces. A similar analysis of ERS-1 6 cm-λ (C-band) data [12] found that the decorrelation rates are often so much higher, and unpredictable, that the utility of the topographic maps derived from the radar measurements is limited. Therefore, the repeat pass implementation is a much less desirable candidate for global studies where complete coverage at uniform accuracy is required, and we will not consider it further here.

The theory of topographic mapping using radar interferometry has already been presented in some detail [6], [19], [20]—here we summarize the main results and establish notation. We note that for repeat pass imaging geometries, on each pass the radar acts as both a transmitter and receiver, therefore the path difference from each to a given point on the surface is twice what would be expected if a single spacecraft or aircraft with two physical antennas

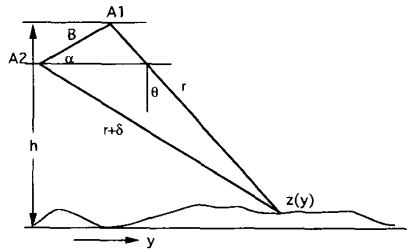


Fig. 2. Interferometer imaging geometry. Radar antennas $A1$ and $A2$ both illuminate the same patch of ground centered at $y = 0$. Incidence angles θ_1 and θ_2 result in phase offsets for all points P displaced by distance y of $y \sin\theta_1$ and $y \sin\theta_2$, respectively. Difference of these phases is measured interferometer phase.

were used. Thus some of the equations listed here differ from those in the references by a factor of two.

Given two antennas $A1$ and $A2$ as shown in Fig. 2, surface topography $z(y)$, the spacecraft altitude h above a tangent plane at the point of interest, the baseline distance B, the range to a point on the ground r, the look angle θ, and the angle of the baseline with respect to horizontal α, a radar signal transmitted from antenna $A1$ and received at both $A1$ and $A2$ will form an interferogram where the phase at each point is equal to the difference in path lengths δ. The measured phase of the interferometer is directly proportional to this distance, with the constant of proportionality $2\pi/\lambda$. Using the law of cosines we can determine the following equations for height as a function of these parameters:

$$\delta = \frac{\lambda\phi}{4\pi} \tag{1}$$

$$\sin(\alpha - \theta) = \frac{(r+\delta)^2 - r^2 - B^2}{2rB} \tag{2}$$

$$z(y) = h - r\cos\theta \tag{3}$$

where ϕ is the measured phase, and λ is the wavelength.

The above constitute a recipe for measuring topography with the interferometer. The two principal errors associated with the measurements arise from uncertainties in the measured phase and in the knowledge of baseline attitude. First, differentiation of (1)–(3) with respect to ϕ yields the error in height estimate as a function of the error in phase estimate

$$\sigma_z = \frac{\lambda r}{4\pi B}\tan\theta\sigma_\phi \tag{4}$$

where σ_z and σ_ϕ are the standard deviations of height and phase, respectively.

The second significant error source results from inaccuracies in knowledge of the interferometer baseline alignment. That is, it is impossible to distinguish a baseline angle knowledge error from a slope on the surface topography, and therefore extremely precise knowledge of the baseline geometry is required if absolute height estimation is needed. Again, differentiation but with respect to α yields

$$\sigma_z = r\sin\theta\sigma_\alpha. \tag{5}$$

Note that the error is independent of baseline length and depends only on attitude and range. This is a stringent constraint for spaceborne geometries where the range from the radar to the image swath can be many hundreds of kilometers. For TOPSAT's goal of 2-m accuracy the baseline orientation must be known to about 1 second of arc. We note that this requirement is for absolute accuracy only—relative height measurements corrected with ground control points to determine the absolute values do not require this accuracy. However, the additional costs associated with acquiring and integrating a worldwide ground control point data set probably outweigh the cost of the spacecraft systems needed to achieve arc second pointing knowledge.

Phase noise (4) in interferometric radar signals arises from several sources, including thermal noise, sampling and processing artifacts, and statistical correlation of the individual radar echoes before they are combined to form the interferogram. Thermal noise is of course minimized by using the greatest possible transmitter power and lowest noise receivers. Sampling and processing artifacts are a tradeoff between data system complexity and cost. All of these factors are well known in radar system design. However, for interferometric systems, the correlation property of the echoes represent a new factor limiting performance.

Decorrelation noise arises mainly from three sources, rotational, temporal, and baseline effects [18]. Rotation of the viewing angle between passes is important particularly when interferometric techniques are applied to satellites in crossing orbits [17], but these systems would never be practical for global mapping applications, and we will ignore this term here. Temporal decorrelation is important when the two radar echoes are not acquired simultaneously, as in the repeat pass technique, but as TOPSAT will likely be implemented using contemporaneous observations we can ignore that effect as well.

Baseline decorrelation results from viewing the surface at two slightly different angles and increases with increasing angle (or baseline). The correlation between echoes varies approximately linearly, decreasing from unity at zero baseline to zero at a critical baseline

$$B_c = \frac{\lambda r}{2R_y\cos^2\theta} \tag{6}$$

where R_y is the ground range resolution and a nearly horizontal baseline, appropriate for orbital implementations, is assumed [18]. Thus a tradeoff is involved in any interferometer design—the baseline must be large enough to give sufficient phase sensitivity to height (1)–(3) yet small enough as not to introduce too much decorrelation noise.

The baseline length may be optimized by expressing the uncertainty in phase σ_ϕ as a function of thermal signal-to-noise ratio, processing parameters, and baseline parameters, as [19]

$$\sigma_\phi = \frac{\sqrt{1-\gamma^2}}{\gamma\sqrt{2N_L}} \tag{7}$$

Table 1 TOPSAR and ERS-1 Radar System Parameters

Parameter	TOPSAR	ERS-1
Wavelength, m	0.0566	0.0566
Peak power, W	1000	4800
Pulse rate, Hz	600 nominal	1679 nominal
Pulse length, μs	5.0	37.1
Antenna length, m	1.6	10
Antenna width, m	0.11	1
Antenna gain, dB	25	43.2
Range bandwidth, MHz	40	15.55
Receiver noise temperature, K	2100	3700
Antenna baseline, m	2.58	variable
Baseline angle (α), deg	62.77	variable
Slant range resolution, m	3.75	9.6
Azimuth resolution, m	1.2	6.5
Platform altitude, km	8	790
Look angle, deg	20–65	23
Repeat interval, days	N/A	3, 35, 165

where N_L is the number of radar "looks"

$$\gamma = \frac{1 - (B/B_c)}{1 + (1/\mathrm{SNR})} \tag{8}$$

and SNR is the thermal signal-to-noise ratio.

It is worth noting that an active area of research in interferometric techniques involves the minimizing of baseline decorrelation at the expense of a loss of range resolution [8]. While in theory this is valid for flat surfaces, practical problems appear to limit its usefulness for practical systems. However, should the studies indicate that alternative processing could eliminate a major noise source, it would certainly be included in the data system design.

IV. Existing Radar Systems and Data Examples

In this section we illustrate interferometric radar principles by presenting data acquired by the TOPSAR airborne prototype topographic radar and by the ERS-1 radar satellite operating in a repeat pass mode. While TOPSAT would not be implemented by an ERS-1 repeat pass approach, the large areal coverage and global availability permitted by spacecraft systems is hard to appreciate by examining only aircraft strip maps, thus we include an example here.

The TOPSAR interferometric synthetic aperture radar system is implemented on the NASA DC-8 aircraft, where NASA /JPL also operates a multifrequency (P, L, and C bands), multipolarimetric radar (AIRSAR). The TOPSAR implementation uses much of the existing AIRSAR hardware. When in use, TOPSAR effectively replaces the C-band polarimeter instrument, but the remaining L- and P-band systems are undisturbed and operate together with the topographic mapper, producing simultaneous L- and P-band fully polarimetric, plus C-band VV polarization backscatter images in addition to the topographic product.

There are considerations specific to the TOPSAR environment which led to the existing design (Tables 1 and 2). The DC-8 airframe fuselage can support only a 2- to 3-m baseline without requiring significant modifications and this limits performance. The intrinsic range resolution of the AIRSAR is 3.75 m, thus the critical baseline at C-band from (6) is 150 m. The optimal baseline, balancing height sensitivity in the phase measurements and processing

Table 2 Design Control Tables

Parameter	TOPSAR (dB/dBW)	ERS-1 (dB/dBW)
Peak power	30	36.8
Antenna directional gain	28	45.9
Antenna efficiency	−5	−3
$1/(4\pi)$	−11	−11
$1/R^2$	−80	−118.6
Illuminated area	53	78.4
$sigma^0$	−15	−15
$1/(4\pi)$	−11	−11
$1/R^2$	−80	−118.6
Antenna area	−8	10
Antenna efficiency	−5	−3
System losses	−8	−3
Oversampling gain	5	1.8
Total	−106	−110.3
Thermal noise (kTB)	−119	−120.0
Signal-to-noise ratio	13	9.7

feasibility, is about one tenth to one fifth of the critical baseline, or about 15–30 m. Clearly the airframe will not support an interferometer at the optimum baseline and we must settle for the largest baseline attainable. We therefore chose to mount one antenna below the existing P-band antenna fairing and the second at window level, yielding a 2.58-m baseline. Although this is a factor of ten less than optimum, reasonable performance is achieved.

The ERS-1 satellite contains several instruments, including the synthetic aperture radar. Since only one antenna is used on the spacecraft, we must rely here on repeat pass analysis to form the interferometer. This radar also operates at C-band and has somewhat lower performance in terms of signal-to-noise ratio and resolution as compared to the aircraft system. However, the swath width is over 100 km in contrast to the 10-km swath imaged by TOPSAR.

The main factors affecting topographic mapping performance for the ERS-1 repeat pass case are baseline length, baseline alignment, and temporal decorrelation of the surface. For ERS-1 the critical baseline from (6) is 1100 m, and best performance is realized for a baseline length near 200 m. ERS-1 has been operating in a 35-day repeat cycle for nearly 2 years and fortunately many revisits to a site have been possible. Derivation of a DEM requires selecting a pair from the set of available data with a usable baseline. The knowledge of the baseline orientation is about 3 mrad using the best available orbit reconstructions, yielding an absolute height error from (5) of 50 m. This is not useful for global studies, but given a set of ground control points a digital elevation model may be derived.

For a repeat pass implementation, it is important to minimize temporal decorrelation, and areas of the world with little surface change, such as deserts, are the best candidates for application of ERS-1 as a topographic measurement tool. Forests and other vegetated areas as well as areas subject to freezing and thawing will return radar echoes that are less well correlated due to changes during the 35-day revisit interval [12]. Again, since this radar was

Fig. 3. TOPSAR image of Walnut Gulch/Tombstone, AZ. Perspective from elevation measurements, brightness represents radar backscatter coefficient.

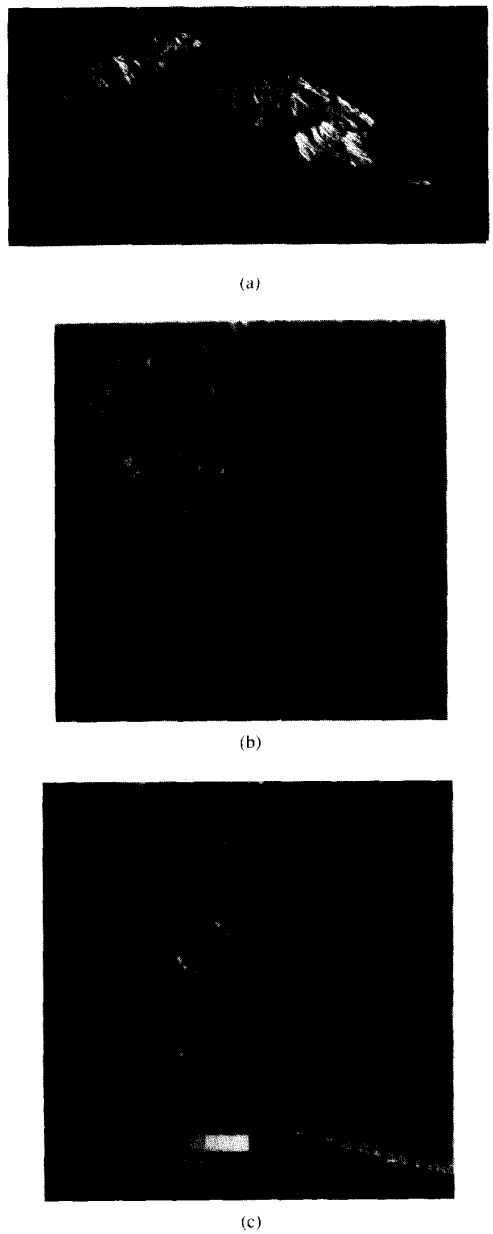
not designed specifically for interferometric applications, performance has not been optimized for topographic mapping. However, the application gracefully degrades with suboptimal geometry and under reasonable conditions of orbit alignment and surface temporal properties topographic maps may be derived.

We can estimate the system signal-to-noise ratio for each with the aid of a design control table (Table 2). Given the TOPSAR baseline parameters, assuming a nominal 200-m baseline length and horizontal alignment for ERS-1, and the signal-to-noise ratios we can then analyze performance of the interferometers. Equation (8) yields a phase noise of 2.8° and 9.1°, respectively, for the TOPSAR and ERS-1 systems; (4) then gives height uncertainties of 1.4 and 2.4 m, respectively.

Figure 3 illustrates a sample DEM acquired by TOPSAR. Here the elevation data are used to generate a perspective view of Walnut Gulch experimental watershed near Tombstone, AZ; the amplitude at each point in the image is determined by the radar backscatter coefficient. Since the amplitude and the phase information are carried together in the data processor, precise alignment of the radar brightness and terrain information is maintained. This makes it relatively easy to reference the location of points in the DEM to a known coordinate system as many features are identifiable in the radar backscatter image.

Figure 4(a) is a TOPSAR image acquired over Ft. Irwin, near Barstow, CA. This was the site of a verification experiment [13] where TOPSAR data were compared with a very accurate DEM produced by the U.S. Army Topographic Engineering Center (TEC). The stated 1-m or better accuracy of the TEC reference DEM (shown in shaded relief representation in Fig. 4(b)) was ensured by using many ground control points. In Fig. 4(c) we plot an error map of the height difference between the TEC DEM and that obtained by TOPSAR. Peak errors approach 10 m, but for rms error the agreement is to the 1 m level in the flat regions and 2–3 m in the mountainous regions, as expected by the theoretical models.

TOPSAR shows its usefulness by acquiring data over regions of scientific interest for which the existing data are poor or nonexistant. One such region is the Galapagos Islands, which are remote and difficult to measure by conventional means, however are of great interest to the volcanology community. It is expensive to deploy aircraft with stereo cameras for extended lengths of time, which would be required to obtain complete cloud-free coverage over the entire islands. TOPSAR covered several interesting



(a)

(b)

(c)

Fig. 4. (a) TOPSAR image of Ft. Irwin, CA. This was the site of a verification experiment that included a variety of terrains. A conventional stereo photogrammetric image was produced by the U.S. Army Topographic Engineering Center (TEC) as a reference. (b) Shaded relief representation of the reference TEC DEM. The TEC map does not cover all of the area in (a). (c) Error map of the difference between the TEC DEM and that from TOPSAR. Although peak errors approach 10 m, in an rms sense the agreement is to the meter level in the flat regions and 2–3 m in the mountainous areas.

islands in a matter of a few hours. In Fig. 5 we show a contour map derived from TOPSAR data over the Galapagos island of Fernandina. In this case there are no known elevation data of large parts of the island and for those data that exist the accuracy is not well characterized. This map is now the best available, and is in convenient digital format.

Fig. 5. Contour map of Isla Fernandina, Galapagos Islands, derived from TOPSAR DEM. The contour interval is 50 m, and the brightness represents radar backscatter coefficient. This image consists of four parallel strips mosaicked together.

It consists of four strips mosaicked into a single image and covers an area of about 900 km^2.

As mentioned above, the principal advantages of spaceborne platforms are large areal coverage and global data acquisition. We have selected an interferometric pair of scenes acquired over part of the Mojave Desert in California, where little temporal decorrelation occurs and we chose two orbits separated by about 100 m in space. From these data we produced the ERS-1 digital elevation model shown in Fig. 6. This image is 40 km on a side, and represents a subset of a single swath of radar data. While we have not completed a detailed performance analysis of this scene, analyses on similar scenes [12] indicates that 5-m vertical precision at 40-m spatial resolution may be achieved if five ground control points are identified within the image and used to constrain the solution.

V. TOPSAT Implementation Options

In this section we present two possible spacecraft implementations that could satisfy many of the global topographic mapping requirements as described above. Each approach has its own limitations based on available technology. Thus many of the limits to performance described here result from details of the implementation rather than by shortcomings of the technique. However, both of the proposed implementations are feasible and are an appropriate starting point for a realistic system design.

If we eliminate repeat pass implementations as unreliable for a global mapping mission requiring contiguous, uniform data over much of the Earth, two alternatives for interferometric radar instrument designs remain: i) a single spacecraft with two displaced antennas, and ii) two spacecraft, each with a synthetic aperture radar, flying in formation to form the interferometer baseline. For each system, the basic limitations to interferometer performance

Fig. 6. ERS-1 interferometric digital elevation model of part of the Mojave Desert, CA. The bright, flat region in the foreground is the Pisgah lava flow, and the dark dry lake in the background is Drinkwater Lake.

as described above remain. Signal-to-noise ratio must be maximized in the constrained spacecraft environment, the baseline must be of sufficient length to give the desired height sensitivity without causing too much baseline decorrelation, and the baseline attitude must be measured at the 1 arc second level.

The requirement to realize sufficient signal-to-noise ratio, coupled with technological limitations on transmitter efficiencies, practical device power ratings, and spacecraft prime power levels, translates into a need for a large antenna compared to the wavelength, producing a relatively narrow swath. The single-spacecraft design we present here (Tables 3 and 4) produces a 10.5-km swath after accounting for overlap in the mosaic process, and thus requires at least 241 days to map the world completely. A longer mapping cycle is more costly for two reasons: the design lifetime of the spacecraft must be longer and the operations phase of the mission lasts longer.

The single-spacecraft approach would achieve the required baseline lengths by mounting one or both of the antennas on a boom at a distance from the spacecraft. A boom would likely be limited to about 25-m length (see next section), therefore to form an adequate baseline the wavelength would have to be short, preferably 2 cm (Ku-band) or less. The effective boom length, however, can be doubled by "ping-ponging," or alternately transmitting from each antenna, at a cost of cutting pulse repetition rate and the average power per channel by a factor of two, which also increases azimuth ambiguities due to undersampling of the Doppler spectrum.

Baseline attitude determination would be achieved by measuring both the rigid body spacecraft attitude and the structural distortion between the two antennas. The spacecraft attitude can be measured by a star tracking system and the structural distortion by a laser metrology system. Of these two the measurement of spacecraft attitude to less than 1 arc-second is the more difficult. Current star trackers cannot measure absolute attitude to less than several arc-seconds, due to limitations in the trackers themselves, uncertainties in mechanical spacecraft flexure, and by limitations in absolute knowledge of star positions (private communication, S. J. Wang, JPL internal memorandum Jan. 10, 1994).

Several of the above limitations could be reduced by using two spacecraft to form the interferometer by flying in

Table 3 TOPSAT Ku- and L-Band Radar System Parameters

Parameter	Ku-band	L-band
Wavelength, m	0.02	0.24
Peak power, W	750	1600
Pulse rate, Hz	3800	2100
Pulse length, μs	60	50
Antenna length, m	5	9
Antenna width, m	0.65	3.5
Antenna gain, dB	49.9	38.4
Range bandwidth, MHz	20	20
Receiver noise temperature, K	700	600
Antenna baseline, m	25	800-2000 (variable)
Baseline angle (α), deg	30	0
Slant range resolution, m	7.5	7.5
Azimuth resolution, m	3.3	5.9
Orbit altitude, km	440	564
Look angle,	30	30
Orbit repeat interval, days	241	84

proper formation. Since the baseline can be much longer, a twin-spacecraft approach could utilize a much lower frequency, and hence technologically simpler, radar system (see Tables 3 and 4). We have chosen L-band with a wavelength of 24 cm as a nominal approach due to the relative technological maturity of the radar electronics—SEASAT, SIR-A, SIR-B, SIR-C, and the Japanese JERS-1 satellites all have L-band radar channels. Here we would navigate the two spacecraft in "parallel" orbits, identical except in node crossing, to form any desired baseline length. The baseline attitude knowledge requirement is still at the 1 arc-second level, but differential Global Positioning Satellite (GPS) techniques promise to determine the relative positions of the spacecraft to a precision of about 3 mm in all directions [21]. If the baseline length is nominally 1000 m, this translates to 0.62 second of arc, well within the requirements.

Successful measurement of the relative position of the interferometric antennas to an accuracy of 3 mm relies on extrapolation of current GPS system performance using knowledge of error sources gained from the TOPEX GPS precision orbit determination experiment. This experiment resulted in absolute position knowledge to the several centimeters level [22]. The total error was found to have four main sources: i) receiver thermal noise, ii) multipath effects, iii) satellite orbit knowledge limitations, and iv) ionospheric propagation effects. Since TOPSAT will require only the relative spacecraft positions at the very high precision, and since the spacecraft are identical in configuration and separated by only 1–2 km, the contributions of the last three error sources are reduced from the centimeter level to the millimeter level.

In the twin-spacecraft case, navigational complexity associated with two spacecraft orbiting within 1 km of each other, not to mention the additional cost of a second spacecraft, are the principal challenges. Spacecraft-to-spacecraft communications and synchronization are required and the data downlink problem may be more difficult. All of these problems can be solved with existing technologies, and the ultimate arbiter in the choice between one and two spacecraft will likely be the cost issue.

Table 4 TOPSAT Design Control Tables

Parameter	TOPSAR (dB/dBW)	ERS-1 (dB/dBW)
Peak power	28.8	32.0
Antenna directional gain	49.9	38.4
Antenna efficiency	−3	−3
$1/(4\pi)$	−11	−11
$1/R^2$	−114.1	−116.3
Illuminated area	75.8	84.2
$sigma^0$	−13	−20
$1/(4\pi)$	−11	−11
$1/R^2$	−114.1	−116.3
Antenna area	5.3	15
Antenna efficiency	−3	−3
System losses	−3	−3
Oversampling gain	2.1	2.1
Total	−110.3	−111.9
Thermal noise (kTB)	−125.2	−126.9
Signal-to-noise ratio	14.8	15.0

Table 5 TOPSAT Error Budgets (All values in meters.)

	Ku-Band		L-Band	
Parameter	Flat Terrain	Sloped Terrain	Flat Terrain	Sloped Terrain
Height errors, m				
Phase noise error	3.26	5.32	1.94	3.17
Baseline error	0.00	0.00	0.71	1.15
Attitude error	1.23	2.01	0.98	1.59
Orbit height error	0.10	0.10	0.10	0.10
Other	0.10	0.16	0.10	0.16
RSS total	3.48	5.68	2.29	3.73
Across-track position errors, m				
Phase noise error	5.65		3.36	
Baseline error	0.00		1.22	
Attitude error	2.13		1.69	
Navigation error	3.00		3.00	
Other	0.10		0.10	
RSS total	6.74		4.97	
Along-track position errors, m				
Orbit timing error	0.01		0.01	
Navigation error	3.00		3.00	

A complete error budget [19] for both implementations is shown in Table 5. Here we break down the total error into many components, only two of which (denoted height phase noise error and height attitude error) were described previously (4) and (5). Details of this procedure are beyond the scope of this paper, please consult the reference for more information.

Orbit Characteristics: Advantages accrue for both concepts if a sun-synchronous orbit operating near the terminator is utilized, as solar illumination remains constant for many months. The main advantages realized by this orbit choice are constant availability of power from the solar panels and few thermal shocks to the spacecraft. The twin satellites operate at an altitude of 565 km and inclination of

97.6° while the single satellite with dual antennas operates at an altitude of 440 km and inclination of 97.6°. The lower altitude is required by the single satellite with dual antennas to maintain adequate height resolution, since the height resolution is a function of antenna spacing and radar signal-to-noise ratio.

Navigation issues for the single-satellite implementation are straightforward and the major concern is to cover the entire Earth with a minimum of gaps. The twin-satellite mission design is more complex, involving two spacecraft; it has been described in detail [22] and will be summarized here. For the dual-satellite mission, the satellites are injected together into the 565-km orbit. After the correct orbit has been attained, the satellites are separated into two different orbit planes and at slightly different altitudes (a few meters). A good understanding of the satellite's flight properties will first be determined when the satellites are at a large lag distance. The differential drag experienced by the two satellites will be measured, after which the satellites will be set at the operational lag distance. After a radar calibration period, the mapping phase will begin. A map of the land masses between ±70° latitude is obtained after about 84 days. By launching at the proper time of the year, two global maps can be obtained before entering solar occultation, the time when Earth shadows disrupt the sun's illumination of the spacecraft solar panels. The baseline mission is defined as the completion of 60 days of in-orbit checkout followed by two complete 84-day surveys for a total mission duration of approximately 8 months.

Both implementations would be planned to conduct dual complete ground surveys. The second survey would fill in any gaps in the first. Data would be acquired only on the ascending pass in the first survey because of limitations of on-board storage and downlink data rate. In the second survey, data could be taken on the descending pass also so that the ground would be seen from the opposite look angle. This would help to locate errors in the data that are caused by high surface slope.

Twin-Satellite Orbit Configuration: A baseline distance (the distance between the two satellites measured perpendicular to the velocity vector) of 800 to 2000 m is required for proper single-pass interferometric results. Figure 7 shows that the two orbits are identical except for a 2020-m difference in the locations of the node crossings giving a baseline separation of 2000 m at the equator and 800 m at 65° latitude. Because the ground tracks are denser at the higher latitudes, good results can be obtained up to about 70° latitude despite the short relative separation. Coverage between ±70° includes almost all the land areas of topographic interest. By increasing the equatorial separation to 6 km, higher latitudes (about 80°) could be covered in an extended mission.

Data Storage and Downlink: The L-band radar design presented above produces data at a rate of 51.4 Mb/s per spacecraft, and the Ku-band radar at a rate of 64 Mb/s. These data must be stored on-board the spacecraft for transmission to the ground. Ideally, the data storage device would have a capacity of 100's of gigabits to

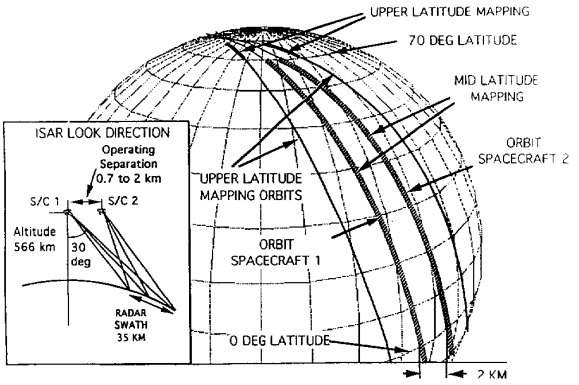


Fig. 7. Dual-satellite trajectory for mid latitude and upper latitude coverage. Note that the track separation is greatly exaggerated for clarity—the actual separation of 2 km at the equator would be too small to see at this scale.

provide ground station scheduling flexibility and backup for missed passes and on-board failures. It would also be possible to read out any desired random block of data in the same order as it was recorded. Presently, no recording device meets the TOPSAT required data rate and volume performance parameters. Although 100-Gb tape recorders are expected to be available in the next few years, they have the disadvantages of moving parts, reverse playback, and difficulty in randomly addressing recorded data. Solid-state recorders, which are being developed for the EOS-AM platform, would better meet the needs of TOPSAT. Because SAR's are such prodigious producers of data, there is almost no point where the amount of storage is considered enough. Any future developments in data storage technology will provide real value to TOPSAT and other future SAR missions.

There are two main options possible for data return. One option involves the use of 10-m, X-band ground stations located in Alaska and McMurdo, Antarctica, which have frequent opportunities to see the spacecraft as well as additional coverage by 11-m DSN stations. The downlink radar and altimeter data rate would be at 85 Mb/s. Downlink of GPS and spacecraft engineering data would be at a rate of 512 kb/s by S-band to either the DSN 26-m network or the McMurdo station. Uplink would be at 2 kb/s from the DSN 26-m stations. The second data return option would use the TDRSS satellite system in a K-band single-access mode and employ high-gain antenna on the spacecraft. Even in this case, a large on-board recorder would be needed as TDRSS is run as a facility shared by many users, and is not always available. Normal uplink would also use TDRSS.

Other Mission Issues: The single-spacecraft approach relies on precise position and attitude knowledge and control of two radar antennas separated by a long structure. Occultation-induced thermal changes could disrupt the pointing control as well as causing lower orbit-averaged power availability. The orbit would be designed to minimize occultations but when they occur data collection may be interrupted. When the occultation periods end

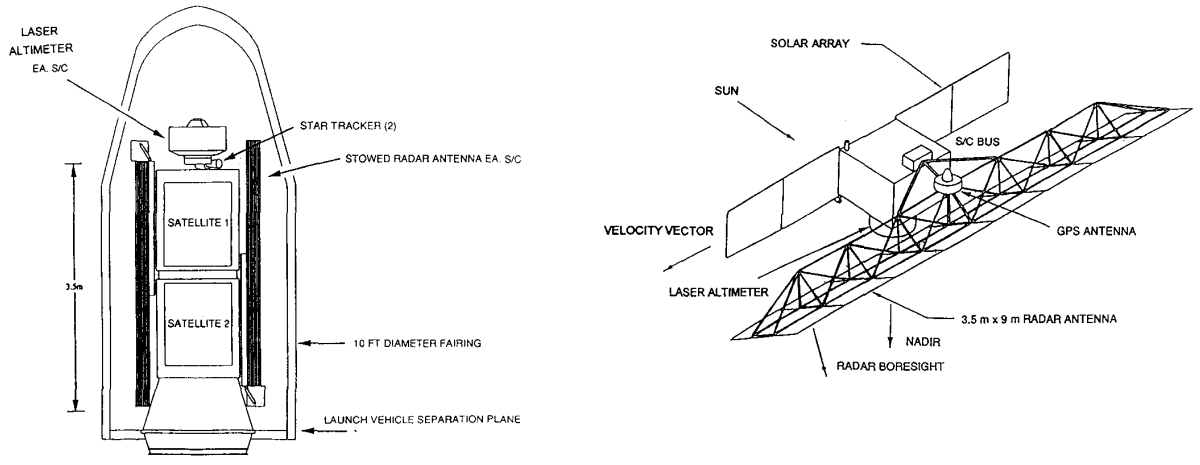


Fig. 8. Single-spacecraft configuration in launch vehicle shroud (left), and deployed (right).

the spacecraft is orbiting over a different set of ground tracks on the Earth, creating a gore in the map, and data acquisition could not be completed until the unmeasured regions become visible once again.

The overall mission could be enhanced, providing additional capabilities to a topographic mission, if a laser profiling instrument were included in the payload to enable polar ice volume studies. A practical design [24] has a swath width of 150 m and so cannot obtain a complete map except near the poles where the coverage is dense. Away from the poles the laser obtains a contiguous line of points for comparison with the radar data, both for validation and, if necessary, ground control point information. The lines of laser data from successive orbits in this approach are about 32 km apart at the equator.

In addition to providing surface height from measurement of the time of flight of the laser pulse, analysis of the return pulse waveform can provide information to help constrain measurements of surface slope, surface roughness, vegetation height, and surface reflectance at 1.06 μm.

VI. Flight System Characteristic Summaries

We present here brief descriptions of the complete flight systems as driven by our radar implementation options above. While not strictly descriptive of the radar system itself, it is useful to examine the demand on spacecraft resources from any proposed implementation approach. The major characteristics of both the Ku-band and the L-band system are summarized in Table 6. Finally, we show possible spacecraft configurations for each option in Figs. 8 and 9. The single-satellite launch and on-orbit configurations are depicted in Fig. 8. For the case of the twin satellites, both spacecraft can be launched on a single Delta II class vehicle. A possible configuration of the two spacecraft in the Delta shroud and on-orbit is shown in Fig. 9.

In the case of the single spacecraft, the solar array would deploy once and then be fixed in position. There would also be a one-time deployment of the outboard antenna boom. The star tracking system would be located on the spacecraft near the radar antenna and a laser metrology system would be used to measure the position of the second antenna realtive to the spacecraft body. The twin-spacecraft implementation also is designed to have a one-time deployment of the radar antenna and solar array. The sun synchronous orbit allows the solar array to be fixed in position, decreasing cost, and increasing reliability.

Table 6 Flight System Characteristics

	Single s/c	Twin s/c
Radar data rate	64 Mb/s	51 Mb/s (each s/c)
Radar power	922 W	694 W (each s/c)
Radar mass	250 kg	300 kg (each s/c)
Total flight system mass	1460 kg	2315 kg (both s/c)
Launch vehicle capability	3580	3420 kg
Launch vehicle margin	2120	1105 kg

VII. Conclusion

Global-scale topographic data are of fundamental importance to many Earth science studies, and obtaining these data are a priority for the Earth science community. Several groups have considered the requirements for such a data set, and a consensus assessment is that many critical studies would be enabled by the availability of a digital global topographic model with accuracies of 2 and 30 m in the vertical and horizontal directions, respectively. The necessity to acquire data globally in a time period of a year or less suggests that active microwave techniques, which are much less susceptible to atmospheric effects such as cloud cover than conventional optical techniques, are the preferred technology.

Radar interferometric techniques have been used to produce digital elevation models at these accuracies and are technologically feasible as the centerpiece of a spaceborne

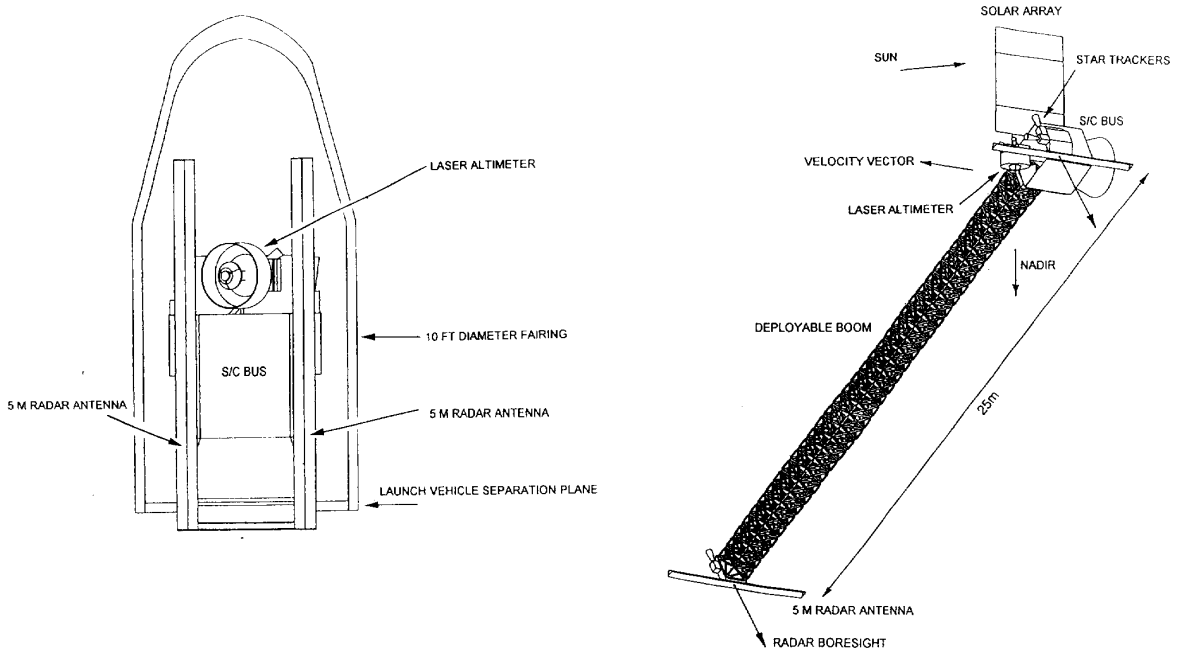


Fig. 9. Twin-spacecraft configuration in launch vehicle shroud (left), and deployed (right).

satellite mission designed to map the world's land masses. A radar interferometer is formed by combining the radar echoes received at a pair of antennas displaced across-track, and specialized data processing results in the elevation data. Two demonstration instruments, the TOPSAR airborne prototype and repeat track analysis of ERS-1 satellite data show that achieving the needed accuracies is feasible at modest cost. The accuracies of the prototype instruments depend mainly on signal-to-noise ratio of the radar echo and to knowledge of the precise imaging geometry

Two alternative implementations, one using a 2-cm-λ radar, and one using a 24-cm-λ radar, are technologically feasible. The former requires an interferometer baseline length of about 15 m to achieve the required accuracy, and thus could be built on a single spacecraft with a long extendable boom. The latter necessitates a kilometers-long baseline, and would thus be best implemented using two spacecraft flying in formation. Measurement errors are dominated by phase noise, due largely to signal-to-noise ratio considerations, and attitude errors in determining the baseline orientation. For the 2-m accuracy required by TOPSAT, the orientation must be known to 1 arc-second. For the single-spacecraft approach, where attitude would be determined by star tracking systems, this performance is just beyond the several arc-second range of existing instruments. For the dual-spacecraft systems, though, differential global positioning satellite measurements possess sufficient accuracy. Studies indicate that similar performance can be realized with either satellite system.

We have presented here preliminary mission concepts which have the potential to produce a global topographic data set in the near future at modest cost. The realization of a TOPSAT mission requires selection of an implementation approach and its detailed design and construction. There are, however, significant open issues with respect to end-to-end mission design. Therefore, ongoing work emphasizes refining the system error budgets and system requirements, defining the mission implementation approach, and technology requirements, as well as examining efficient methods of ground data processing. None of these are technologically infeasible and it is quite likely that a mature concept can be selected in the near future.

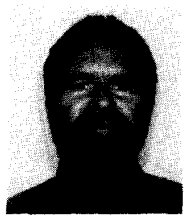
REFERENCES

[1] "Land Processes DAAC Science Advisory Panel," in *Topographic Data Requirements for EOS Global Change Research*, NASA Publication, in press, 1993.
[2] "Task Group on Earth Sciences," in *Space Science in the Twenty-First Century: Imperatives for the Decades 1995–2015*. Washington, DC: National Academy Press, 1988, 121 pp.
[3] "Topographic Science Working Group," Topographic Science Working Group Report to the Land Processes Branch, Earth Science and Applications Division, NASA Headquarters. Lunar and Planetary Institute, Houston, TX, 64 pp., 1988.
[4] "Committee on Earth Sciences," in *A Strategy for Earth Sciences from Space in the 1980's, Part 1: Solid Earth and Oceans*. Washington, DC: National Academy Press, 1982, p. 99.
[5] L. C. Graham., "Synthetic interferometer radar for topographic mapping," *Proc. IEEE*, vol. 62, pp. 763–768, 1974.
[6] H. Zebker and R. Goldstein, "Topographic mapping from interferometric SAR observations," *J. Geophys. Res.*, vol. 91, no. B5, pp. 4993–4999, 1986.
[7] R. M. Goldstein, H. A. Zebker, and C. Werner, "Satellite radar interferometry: two-dimensional phase unwrapping," *Radiosci.*, vol. 23, no. 4, pp. 713–720, July–Aug. 1988.
[8] F. Gatelli, A. Monti Guarnieri, F. Parizzi, P. Pasquali, C. Prati, and F, Rocca, "Use of the spectral shift in SAR interferometry: Applications to ERS-1," submitted to *IEEE Trans. Geosci. Remote Sensing*, 1993.

[9] H. A. Zebker, S. N. Madsen, J. Martin, K. B. Wheeler, T. Miller, Y. Lou, G. Alberti, S. Vetrella, and A. Cucci, "The TOPSAR interferometric radar topographic mapping instrument," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 5, pp. 933–940, Sept. 1992.
[10] D. L. Evans, T. G. Farr, H. A. Zebker, and P. J. Mouginis-Mark, "Radar interferometric studies of the Earth's topography," *EOS*, vol. 73, no. 52, pp. 553 and 557–558, Dec. 29, 1992.
[11] S. N. Madsen, H. A. Zebker, and J. Martin, "Topographic mapping using radar interferometry: processing techniques," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, no. 1, pp. 246–256, Jan. 1993.
[12] H. A. Zebker, C. L. Werner, P. Rosen, and S. Hensley, "Accuracy of topographic maps derived from ERS-1 radar interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 823–836, July 1994.
[13] S. N. Madsen, J. Martin, and H. A. Zebker, "Analysis and evaluation of the NASA/JPL TOPSAR across-track interferometric SAR system," submitted to *IEEE Trans. Geosci. Remote Sensing*, vol. 31, Oct. 1993.
[14] J.-P. Muller and P. Eales, "Global topography accuracy requirements for EOS," in *Proc. IEEE Geoscience and Remote Sensing Symp.* (Washington, DC. 1990).
[15] M. Shimada, M. Nakai, and T. Moriyama, "Derivation of topography from OPS stereoscopic data," in *Proc. CEOS Working Group on Calibration/Validation* (Pasadena, CA, 1992).
[16] A. K. Gabriel, R. M. Goldstein, and H. A. Zebker, "Mapping small elevation changes over large areas: differential radar interferometry," *J. Geophys. Res.* , vol. 94, no. B7, pp. 9183–9191, July 10, 1989.
[17] A. K. Gabriel and R. M. Goldstein, "Crossed orbit interferometry: theory and experimental results from SIR-B," *Int. J. Remote Sensing*, vol. 9, no. 8, pp. 857–872, 1988.
[18] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 5, pp 950–959, Sept. 1992.
[19] E. Rodriguez and J. Martin, "Theory and design of interferometric SARs," *Proc. Inst. Elec. Eng.*, vol. 139, no. 2, pp. 147–159, 1992.
[20] F. Li and R. M. Goldstein, "Studies of multi-baseline spaceborne interferometric synthetic aperture radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 1, pp. 88–97, Jan. 1990.
[21] T. P. Yunck, "TOPSAT Orbit Determination and Satellite-Satellite Vector Measurements with GPS," private communication and JPL internal memorandum, Jet Propulsion Laboratory, 1993.
[22] W. I. Bertiger *et al.*, "GPS precise tracking of TOPEX/Poseidon: Results and implications," submitted to *J. Geophys. Res.* (TOPEX/Poseidon Special Issue), Nov. 1993.
[23] M. A. Vincent, "Design of the TOPSAT mission," presented at AAS/AIAA Astrodynamics Conf., Victoria, B.C., Canada, Aug. 16–19, 1993, paper AAS 93–593.
[24] J. L. Bufton and D. J. Harding, "Multi-beam laser altimeter for the global topography mission," in *Instrument Description*. Greenbelt, MD: Goddard Space Flight Center, 1993.

Howard A. Zebker (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from the California Institute of Technology, University of California at Los Angeles, and Stanford University, respectively. His dissertation was on the solution of particle size distribution functions for Saturn's rings, as derived from Voyager radio occultation data.

At present, he is Assistant Manager of the Radio Science and Engineering Section at the NASA Jet Propulsion Laboratory, where he is involved in the development of techniques such as radar polarimetry and radar interferometry. His current research interests include planetary exploration, interferometric radar, and EM scattering theory and measurement.



Thomas G. Farr received the B.S. and M.S. degrees in 1974 and 1976, respectively, from the California Institute of Technology, and the Ph.D. degree from the University of Washington, in 1981, all in geology.

In 1975 he joined the Radar Science Group, now part of the Terrestrial Science Element, at the Jet Propulsion Laboratory. He was supervisor of the group from 1990 to 1993. At JPL, he has participated in the Arctic Ice Dynamics Joint Experiment (AIDJEX; 1976), the Seasat radar project (1978), the Shuttle Imaging Radar-A (SIR-A; 1981), and the Geological Remote Sensing Field Experiment (GRSFE; 1989–1990). He was Principal Investigator on the SIR-B project (1984) and is currently involved with SIR-C, in a study of the geomorphology of northern China. He is also a Principal Investigator on ERS-1. Recently, he has taken on the role of pre-project scientist for the Global Topography Mission (TOPSAT), which will use an interferometric SAR and a laser altimeter to map the topography of the Earth at high resolution.

Ronald P. Salazar received the B.S. degree from the University of Southern California and the Ph.D. degree from the University of California at Santa Barbara.

He joined the Jet Propulsion Laboratory in 1977 and has worked on a number of space projects as a thermal engineer and system engineer. Presently, he is the study leader for the TOPSAT pre-project study. He was awarded the NASA Exceptional Service Award for his work on the TOPEX/Poseidon ocean topography mission.

Timothy H. Dixon received the B.Sc. degree in geology from the University of Western Ontario, London, Ont., Canada, in 1974, and the Ph.D. degree in Earth sciences from Scripps Institution of Oceanography, University of California, San Diego, in 1979.

From 1979 to 1992 he worked in the Earth and Space Sciences Division of the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, and is currently Associate Professor in the Marine Geology and Geophysics Division, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL. Also, he currently chairs a NASA study group investigating scientific aspects of the proposed Global Topography Mission.

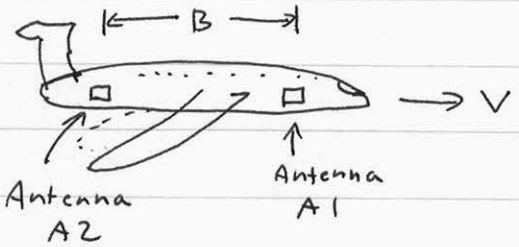
Dr. Dixon is a member of the American Geophysical Union.

Sensing motions

Today we will change our interferometer geometry to enable us to detect motions rather than topography. We will use many of the same processing techniques described in previous lectures but the imaging geometry will ensure that our baseline is described in time, rather than space.

Temporal baselines

To accommodate temporal measurements, our interferometer baseline must span time rather than space. Consider the following experimental setup:



The radar antennas A1 and A2 are mounted on the front and back sections of an aircraft fuselage, rather than across-track. We will form two radar images again, as we did in the earlier interferometer.

Suppose the aircraft flies at velocity v, and then the antennas are displaced by a distance B along track. Other than thermal noise effects, the two images formed in the two channels will be identical, except that the channel two image is delayed in time by the time it takes the aircraft to fly a distance B, $\frac{B}{v}$.

Expressing the image coordinates along-track and across track by x and y, as usual,

$$i_1(x,y) = i_2(x+B, y)$$

In azimuth time and slant range,

$$i_1(t,r) = i_2\left(t + \frac{B}{v}, r\right)$$

In both cases the images are identical except for a shift in position along-track. Now, suppose we are observing an object whose range from the flight track, not the sensor, is $r(t)$. For a stationary object $r(t)$ is constant.

Once again form an interferogram by multiplying i_1 by i_2^*, but shift the images to overlap so as to align scatterers between the images. We'll need to delay image i_2 by $\frac{B}{v}$, so

$$i(t,r) = \cancel{i_1(t,r) * i_2(}$$

$$i_1(t_1, r(t_1)) \cdot i_2^*(t_2, r(t_2))$$

$$= i_1(t_1, r(t_1)) \cdot i_2^*\left(t_1 + \frac{B}{v}, r\left(t_1 + \frac{B}{v}\right)\right)$$

Note that the phase of the object at t is

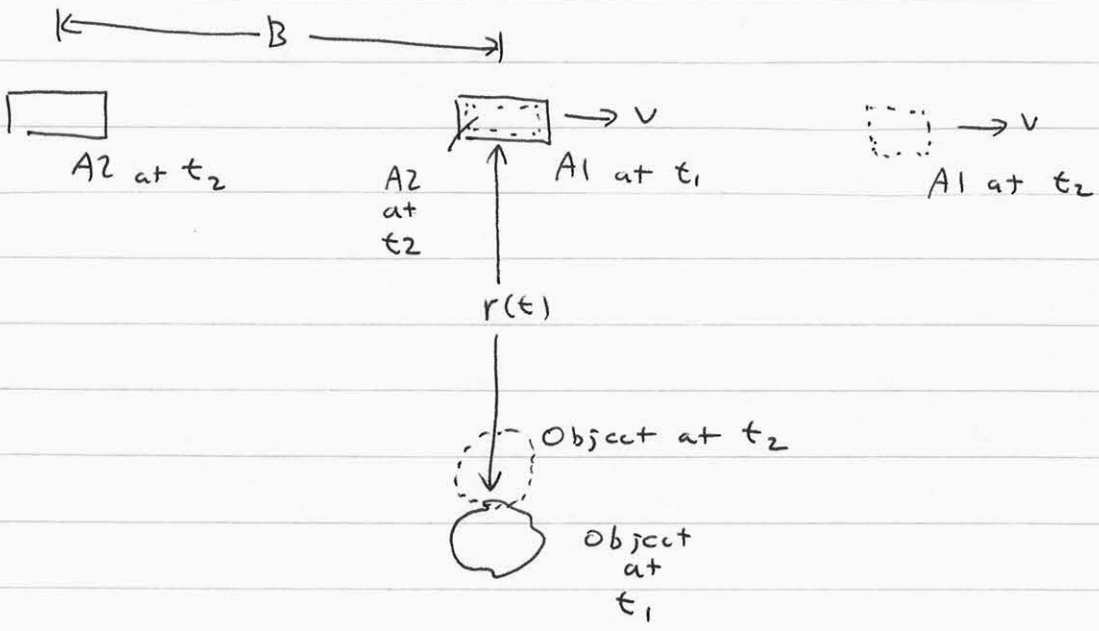
$$\phi = -\frac{4\pi}{\lambda} r(t)$$

hence the interferogram phase $\phi_1 - \phi_2$ is

$$\phi_1 - \phi_2 = -\frac{4\pi}{\lambda}\left(r(t_1) - r\left(t_1 + \frac{B}{v}\right)\right)$$

If the object is not moving, $\phi_1 - \phi_2 = 0$, or the image is identical at both times.

Now what happens if the object is moving?



Suppose the object is moving toward the radar at velocity v. Then

$$r(t+\Delta t) = r(t) - v\Delta t$$

and the interferogram phase ϕ_{int} is

$$\phi_{int} = -\frac{4\pi}{\lambda}\left[r(t) - r\left(t+\frac{B}{v}\right)\right]$$

$$= -\frac{4\pi}{\lambda}\left[r(t) - \left\{r(t) - v\frac{B}{v}\right\}\right]$$

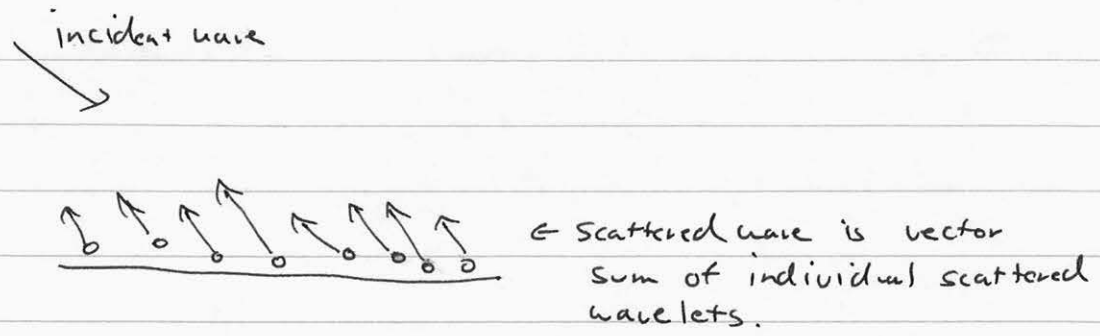
$$= -\frac{4\pi}{\lambda}\cdot\frac{v}{v}\cdot B$$

Thus the phase is directly proportional to the velocity of the object v.

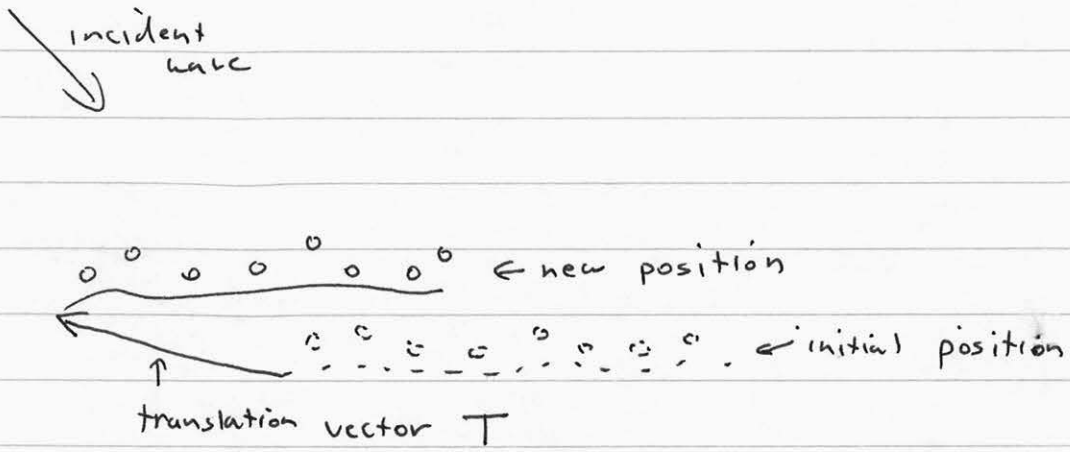
Ocean Currents

We have previously gone through a detailed argument showing that motion of the surface leads to decorrelation. Won't this target movement also lead to decorrelation?

Consider a set of scatterers as we had before for the discrete scattering model:



If the incident direction changes significantly (by the critical baseline) or if we allow random motion among the scatterers, we will observe decorrelation. But if we simply translate the full set of scatterers the distance to each scatterer changes by the same amount.



Thus the mean value of the phase of the echo changes, but the higher order moments are unchanged, and we get no decorrelation.

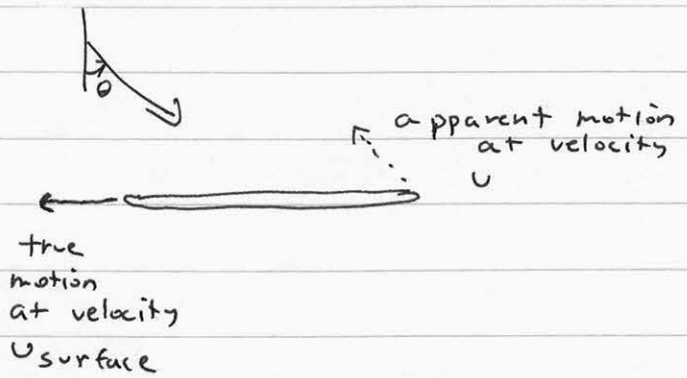
It should also be clear that the amount of phase shift is proportional to the component of the translation vector T that lies in the radar line of sight direction. Hence only one component of T is measured.

Therefore the value of v that is inferred from

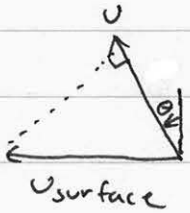
$$\phi = \frac{4\pi B}{\lambda} \frac{v}{V}$$

is the component in the look direction.

What does this mean for ocean current and wave imaging? For currents the entire surface moves together, so the interpretation is straight-forward. For the ocean a good approximation is that the surface is flat, so that motions are constrained to lie in a plane. This makes for easy interpretation of our measured v in terms of true surface velocity.



so the following construction works



$$\frac{v}{v_{surface}} = \sin\theta$$

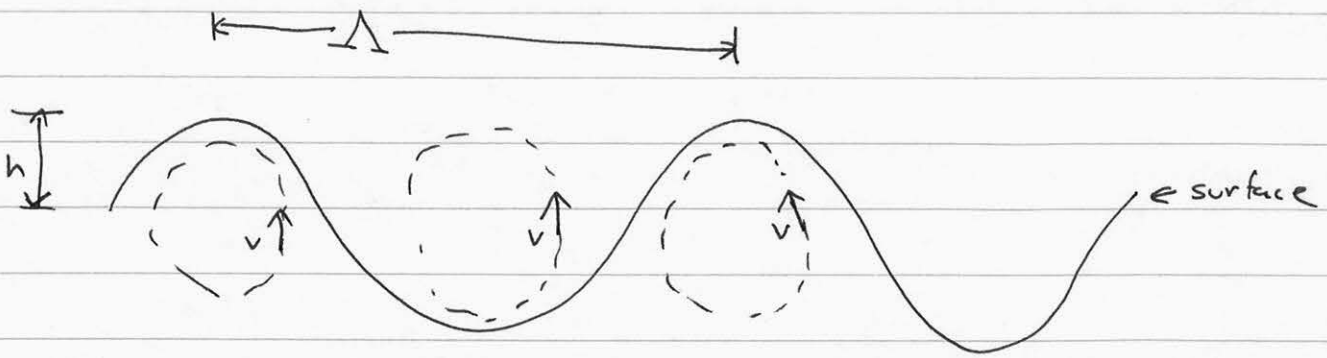
$$v_{surface} = \frac{v}{\sin\theta}$$

Thus our solution for surface velocities for ocean currents is

$$v_{surface} = \frac{\phi \lambda V}{4\pi B \sin\theta}$$

Swell - Large-scale ocean waves

A somewhat different situation applies if we image fields of large (compared to a resel) ocean waves, called swell. In that case individual parcels of water move in circular, or orbital, motions, as in:



For large-scale ocean waves, called gravity waves because gravity forms the restoring force, we can relate orbital motion to wavelength by considering basic wave equations:

$$c = \sqrt{\frac{\lambda g}{2\pi}} \quad \leftarrow \text{phase velocity of wave}$$

and the period

$$T = \frac{\lambda}{c}$$

Now, the orbital motion is such that a parcel of water traverses a circle of radius h in the same period, so the orbital velocity v_{orb} is

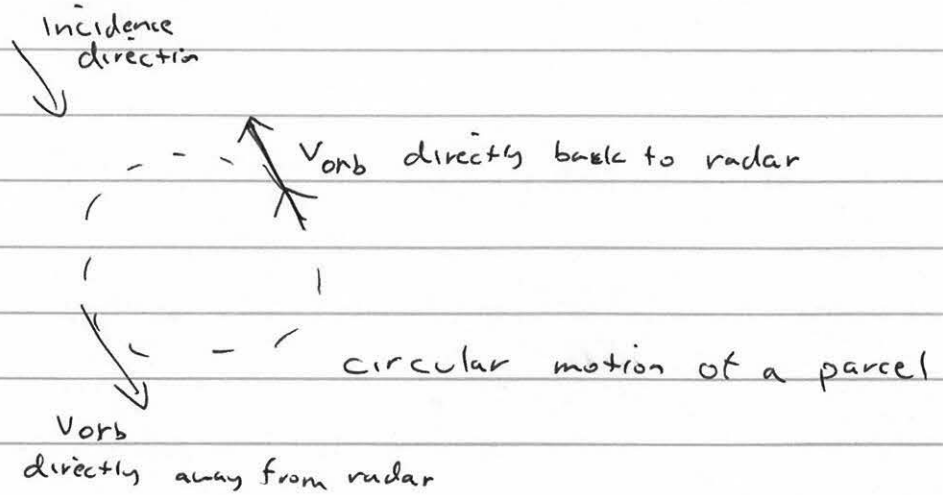
$$V_{orb} = \frac{2\pi h}{T}$$

Combining these, we get

$$V_{orb} = \frac{2\pi h c}{\Lambda} = \frac{2\pi h}{\Lambda}\sqrt{\frac{\Lambda g}{2\pi}}$$

$$= \sqrt{\frac{2\pi g h^2}{\Lambda}}$$

Note that we don't have the same obliquity problem with swell velocity that we had with surface currents, as the motions are circular and not constrained to lie in a plane.

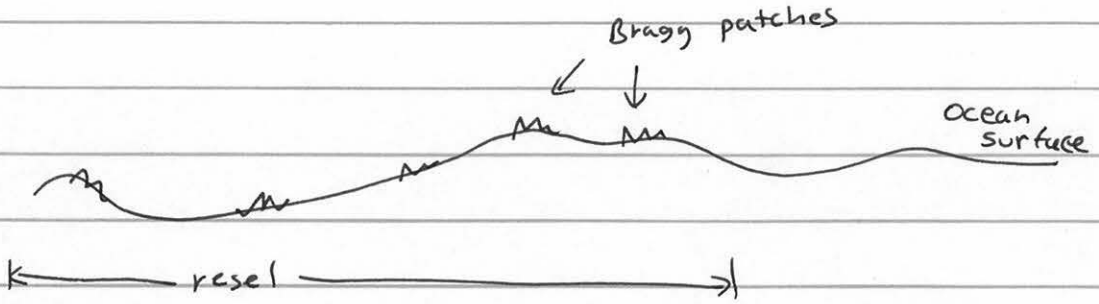


Now, since we can find V_{orb} directly from the interferogram phase, and Λ from the intensity image or phase, we can solve for h, the height of the swell. Hence using an interferometer for ocean remote sensing characterizes swell nicely.

Ocean scattering models

We have implicitly assumed that our discrete scatterer model was appropriate for the ocean surface in order to derive our interferometric relations. But it may be hard to picture the physical realization of these scattering centers on the surface of water.

A useful model for ocean surface scattering for interferometric purposes is to use the discrete scatterer approach, but form the scatterers by created small "patches" which scatter by the Bragg mechanism. These Bragg patches of matched radar wavelength are then scattered about the resolution element to complete the scattering picture.



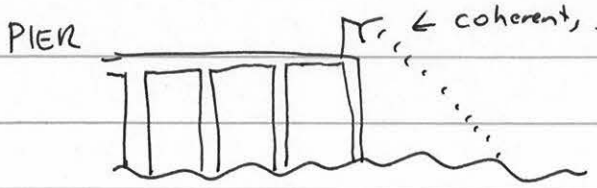
Each patch consists of a few cycles of wavelets with the Bragg condition

$$\Lambda_{Bragg} = \frac{\lambda}{2 \sin\theta}$$

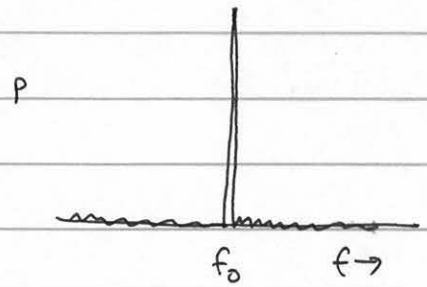
Two issues are raised by incorporating this model into interferometry. One, we have to worry about the lifetime of the scattering patches. Clearly a given patch on a real water surface will dissipate with time. Our temporal interferometer baseline cannot exceed this lifetime.

second, because these patches are waves they too will be moving and will bias our velocity results. We certainly must correct for this effect in interpreting our velocity field.

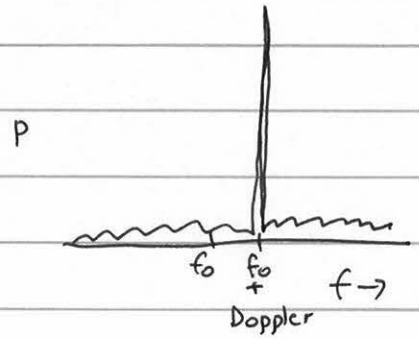
Lifetime - Several groups have measured the lifetime of Bragg patches by using single-frequency coherent measurement techniques. The experiment is as follows:



Consider the spectrum of the received echo. If we were looking at an unmoving land surface, the received spectrum would be very narrow, like this



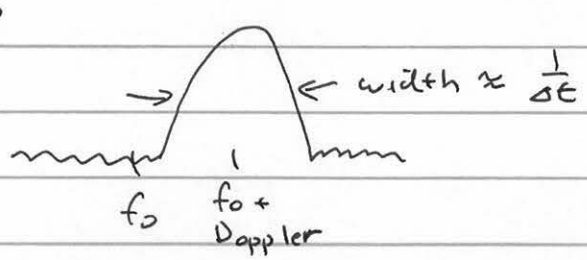
Now suppose we are looking at a series of swell or a fixed current. The entire spectrum will be Dopper shifted to reflect the motion toward the radar.



The width of the spectrum will be determined by the integration time in our receiver, and depicts the usual time-frequency uncertainty relation

$$\delta t = \frac{1}{\delta f}$$

What happens if the patch lifetime is less than our integration time? In this case the observed width of the spectrum will indicate the lifetime rather than the observation time. In other words, the integration time is limited by the coherence time, or lifetime, of the patch. Thus the spectrum will be



Such measurements seem to indicate lifetimes of $\sim \frac{1}{2}$ sec for L band, $\sim$ 0.1 sec for C-band for nominally smooth ocean states. It would decrease for rougher seas.

Is this long enough for us? A typical baseline might be 20 m for a jet aircraft, hence

$$\frac{B}{V} = \frac{20\ m}{250\ m/s} = 0.08\ s$$

which would be comfortably within the L-band limit but pushing things for C-band. A C-band instrument should have a smaller baseline, say two meters or so.

Motion effects Because the Bragg patches are composed of waves, they will be in motion on the water, either toward or away from the radar. Using

$$c = \sqrt{\frac{g\lambda}{2\pi}} = \sqrt{\frac{\Lambda g}{2\pi}}$$

and the Bragg wavelength $\Lambda = \frac{\lambda}{2\sin\theta}$

$$c_{Bragg} = \sqrt{\frac{\lambda g}{4\pi \sin\theta}}$$

so for L band ($\lambda = 24$ cm), and 30° incidence,

$$c_{Bragg} = 0.6 \text{ m/s}$$

So the patches themselves will have velocities on the surface of ± 0.6 m/s. The line-of-sight component will be either 0.3 m/s or -0.3 m/s. This motion will be seen as a bias on top of the desired current or swell velocity.

Hence, we measure the algebraic sum of all three of these effects, and each must be considered in any oceanographic application. But the ability to make these measurements at the resel scale for large swaths makes the instrument appropriate for a great many ocean studies.

Reports

Remote Sensing of Ocean Currents

R. M. GOLDSTEIN, T. P. BARNETT, H. A. ZEBKER

A method of remotely measuring near-surface ocean currents with a synthetic aperture radar (SAR) is described. The apparatus consists of a single SAR transmitter and two receiving antennas. The phase difference between SAR image scenes obtained from the antennas forms an interferogram that is directly proportional to the surface current. The first field test of this technique against conventional measurements gives estimates of mean currents accurate to order 20 percent, that is, root-mean-square errors of 5 to 10 centimeters per second in mean flows of 27 to 56 centimeters per second. If the full potential of the method could be realized with spacecraft, then it might be possible to routinely monitor the surface currents of the world's oceans.

MEASURING NEAR-SURFACE OCEAN currents is a difficult task under any circumstances. Conventional current meters give point values whose spatial representativeness is unknown. Shipboard acoustic systems give spatial information, but the data contain time lags and the very near-surface layer is not observed. Drifting buoys can be used to solve the latter problem but the observations obtained from them also can be hard to interpret. We have found that an SAR can be used to provide virtually instantaneous high-resolution snapshots of near-surface currents over large ocean regions (*1*). The basic idea is to use dual antennas to receive the SAR signal and then from the phase difference between the image scenes from the antennas to obtain an interferogram, which is directly proportional to the velocity field of the ocean surface. The method is distinctly different from earlier attempts to use SAR as a remote current meter (*2*) and is an extension of a two-antenna, stationary target–canceling scheme by Raney (*3*).

The main result of our work (Fig. 1) shows the radar-derived estimates of near-surface ocean currents off San Diego, California, compared with more conventional estimates derived from freely drifting "buoys" for approximately the same space-time locations. The two different estimates of surface current are linearly related with a slope of 1.12 ± 0.18 (95% confidence limits), have little bias (2.2 cm/s), and have a root-mean-square (rms) error of 12 cm/s. Well over one-half of the rms error is introduced by just 3 of the 20 data points. Given the experimental limitations discussed below, we believe the agreement between the currents estimated from the SAR interferometer and the buoy observations is quite encouraging.

In our experiment, the interferogram was formed with image scenes from two SAR receivers with their antennas located 19.3 m apart on the National Aeronautics and Space Administration (NASA) DC-8 radar aircraft. Ideally, one antenna should follow collinearly behind the other. If the reflecting elements of the scene were stationary, then the images from the two antennas would be identical except for a displacement along the track. Surface motion, however, produces a phase shift between corresponding picture elements (pixels) of the two images. The phase shift ($\Delta\phi$) is proportional to the change in distance of the resolution element that occurs in the time it takes for the rear antenna to move to a place of symmetry with respect to the front antenna, that is,

$$\Delta\phi = \frac{2\pi B}{\lambda}\frac{u}{v} \qquad (1)$$

where B is the distance between antennas (19.3 m), u is the line-of-sight component

R. M. Goldstein and H. A. Zebker, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.
T. P. Barnett, Climate Research Group, Scripps Institution of Oceanography, La Jolla, CA 92093.

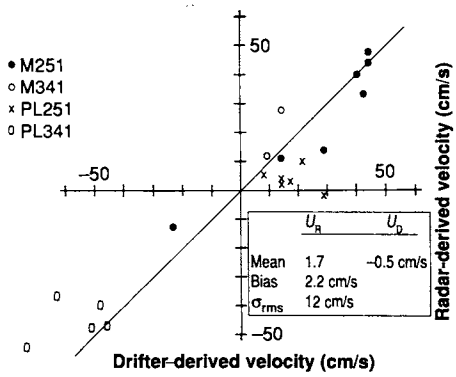


Fig. 1. Scatter diagram for observed versus radar-derived surface currents. The statistics of the data pairs are given in the lower right corner. 'M' and 'PL' refer to data taken off Mission Bay and Point Loma, respectively; the numbers that follow (251 and 341) refer to the current component direction (degrees from true north).

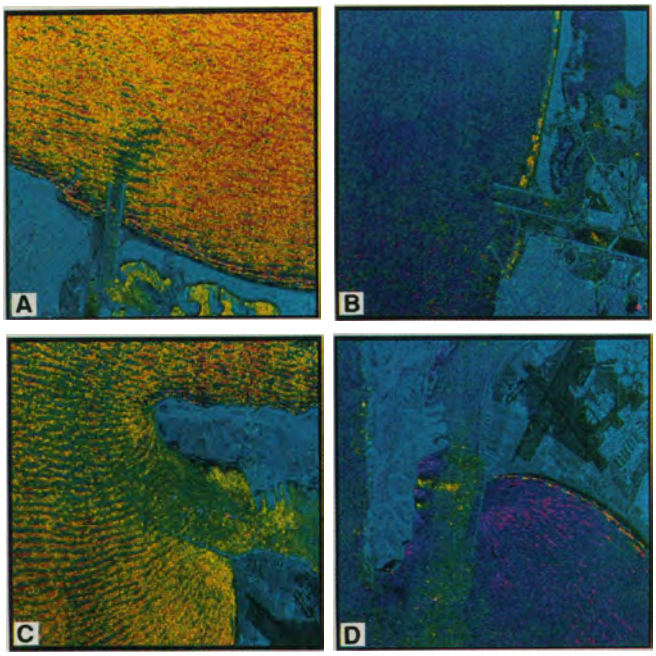


Fig. 2. (**A**) Radar interferogram of the Mission Bay area. Brightness represents echo strength, and hue represents surface motion. The aircraft flight path is from right to left along the top of the figure (toward 161°T). The light blue color of the land represents no motion. The predominantly yellow hue is associated with Bragg waves generated by the prevailing southwest winds that are traveling away from the aircraft. The deep blue streak extending out from the jetties is the outgoing tidal flow (see Fig. 3) and is directed toward the aircraft. The long shore striations are associated with wind wave–induced orbital motions. (**B**) Same as in (A) but for a flight path of 251°T. (**C**) Radar interferogram for Point Loma. The complex surface wave field diffraction around Point Loma is particularly clear in the interferogram but scarcely visible in the conventional SAR image (not shown). (**D**) Same as (B) but for Point Loma. The tidal flow out of San Diego Bay, the yellowest band, is clearly seen, as is the wind wave field and onshore flow (magenta) to the south (right) of Zuniga jetty (see Fig. 4).

of the ocean current velocity, v is the aircraft velocity (220 m/s), and λ is the radar wavelength (0.24 m).

Currents associated with larger gravity waves, internal waves, and other oceanic phenomena are visible to conventional microwave radar only in their ability to modulate the surface Bragg waves (*4*) imaged by the radar. In contrast, our interferometric technique measures the algebraic sum of small displacements of the Bragg waves: the phase velocity of the Bragg waves themselves, the orbital velocity associated with the swell upon which they ride, and any underlying current. Only the line-of-sight component of these velocities is observed.

Our experiment consisted of a flight of the NASA DC-8 radar aircraft from its home base at Moffett Field, California, to the San Diego area and return. The flight reached San Diego at the time of maximum expected outgoing tidal current, near 1230 local daylight time, on 16 April 1988. Radar data were gathered over the outlets of Mission Bay and San Diego Bay from two flight directions each. The directions were chosen to be nearly parallel, and nearly perpendicular, to the expected current directions.

The resulting interferograms are shown in Fig. 2. The first two images (Fig. 2, A and B) are of Mission Bay, as seen from flight directions of 161° (south-southeast) and 251° (west-southwest), respectively. The second two (Fig. 2, C and D) are of Point Loma along the same two flight paths. In these images, radar echo power is portrayed as brightness and phase (Eq. 1) as color. Thus the interferogram is presented coregistered with the conventional SAR image.

The color cyan represents a phase shift of 0° (zero line-of-sight velocity). Motion toward the radar is represented by shades from cyan through violet to magenta, motion away by shades from cyan through green to yellow. In all cases the aircraft flight line is across the top of the figure, toward the left. Most of the surface winds (less than 6 m/s) were from the southwest, producing the yellow bias observed in Fig. 2, A and C, and the violet bias in Fig. 2, B and D. This bias is the result of the phase velocity of the wind-driven Bragg waves.

An extensive wave system can be seen easily in these images, including considerable diffraction around Point Loma. This system is not so easily seen (if at all) in the conventional SAR images. The largest line-of-sight velocity associated with the wave system was 1.5 m/s. In addition to the Bragg wave phase velocity and the larger gravity wave orbital velocity, the outflowing tidal currents can be seen clearly, particularly in Fig. 2, A and D.

We made estimates of the in situ current simultaneously with the radar estimates by observing the motion of the freely drifting buoys (*5*). The buoys were deployed at 15-min intervals in and around the narrowest parts of the entrances to San Diego Bay and Mission Bay (Figs. 3 and 4); this positioning guaranteed that the buoys would likely "see" the largest current signals in the SAR

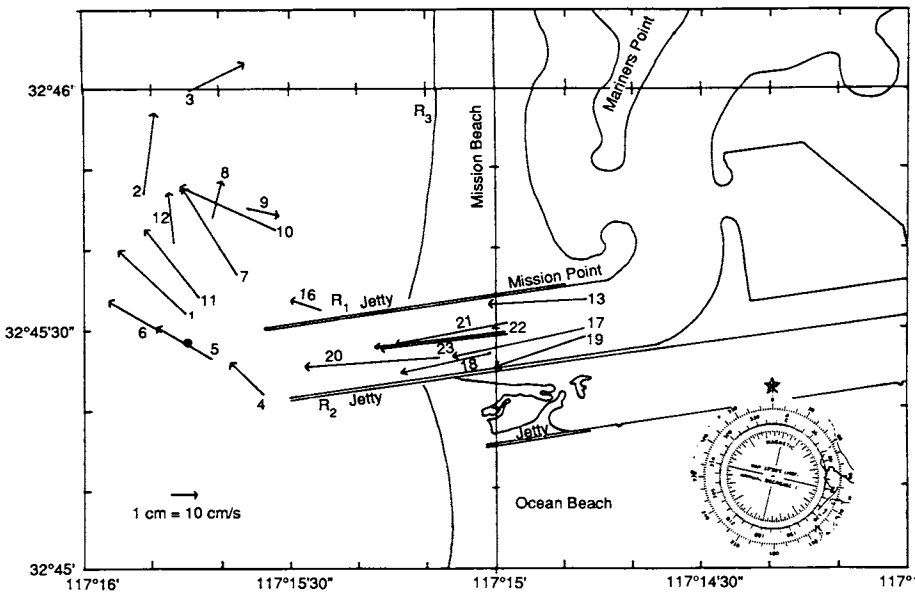


Fig. 3. Observed surface currents around the entrance to Mission Bay, California. The arrows point in the direction of the current, and the length of each arrow is proportional to the observed speed. The letters R_1, R_2, and R_3 designate reference areas for independent estimation of the Bragg wave speed. The solid dot denotes a navigation buoy.

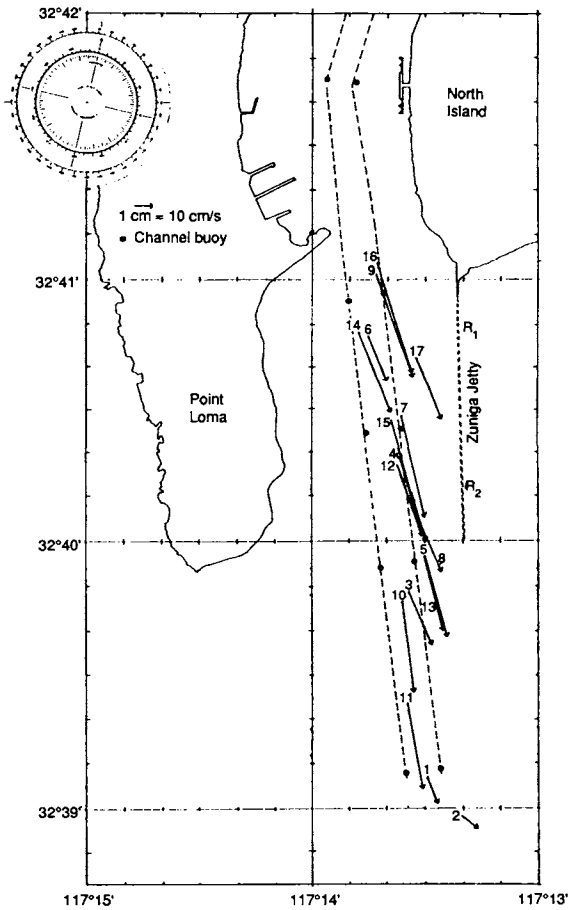


Fig. 4. Same as Fig. 3 but for Point Loma.

images, thus giving the largest signal-to-noise ratio for both current estimates.

The buoys were positioned by Loran-C (long-range navigation) approximately every 15 min. The lorans were on two swift boats and allowed full coverage of the (eventually) extensive buoy arrays at both Point Loma and Mission Bay. Tests of the lorans showed an estimated accuracy for repeat positions to within approximately ± 20 m of a fixed reference, say, a navigation buoy. Typical buoy displacements between positionings were of order 500 m, so current estimates should be accurate to about 4 to 5% or better.

The currents observed by the drifters and radar are tabulated in Table 1. The result of the experiment (the last two columns of Table 1) are shown in Fig. 1. This illustration, Fig. 2, and the associated numerics suggest a number of conclusions:

1) The method is most reliable when one is looking along the direction of strongest flow (Mission Bay, 251°, and Point Loma, 161°).

2) The method seems to be more accurate when one is looking upwind or downwind. In these cases, the Bragg waves responsible for the radar signal are less ambiguous.

3) The estimates of the reference velocity are in relatively good accord with those expected theoretically.

4) The interferograms contain a wealth of information on the ocean wave field and numerous other features in the flow field. We have largely neglected, or processed out, this information in the present study, but the potential of the technique for determination of the surface wave field appears promising.

Using the SAR in the interferometer mode did introduce a problem in our experiment. The collinear condition for the dual antennas was not quite met because of aircraft motion. For example, yaw will cause an effective sideways motion of the antennas, thereby inducing a phase shift and hence making the image scene appear to move even when it is stationary. During the flight over San Diego, the aircraft experienced a good deal of turbulence and so the motion errors were prominent in our data. We corrected for such influences by making use of the aircraft inertial navigation system (INS), which provided frequent readings of pitch, yaw, pitch and yaw rates, horizontal and vertical velocity, and other flight parameters. With the INS data we were able to calculate the expected phase shifts from a stationary field of view and to compensate the measurements accordingly.

We mitigated the effects of the swell first by averaging (filtering) the phase images over a square of 9 by 9 pixels (110 by 110 m on the sea surface). We then averaged the phase over the observed paths of the drifters to make the radar-estimated currents as spatially similar in nature as possible to the directly observed currents. The bias introduced by the Bragg wave phase velocity was estimated for a nearby calm location for which there was little or no tidal current. Theoretically, the value should be 52 cm/s for Bragg waves traveling directly along the line of sight, whereas our calibrated values varied between 50 and 64 cm/s for "look" directions more or less aligned with the wind direction. The modest difference (2 to 12 cm/s) between these values and the theoretical value are measures of the errors inherent in using the INS to correct for plane motion.

The method we have described offers an opportunity for estimating the large-scale ocean surface current field due to the combined effects of tides, wind, and waves on a resolution (12 m) never before obtained. Velocity fields associated with surface waves, internal waves, and Langmuir cells all seem amenable to study with the interferometric technique. On these smaller scales,

Table 1. Radar-derived and observed surface currents (in centimeters per second). Minimum and maximum velocities refer to the 81-pixel averages along the observed path of the drifters. The first column lists the drifter numbers (see Figs. 3 and 4). Average velocities are the grand average over the drifter path. The reference number indicates which area was used to calibrate the Bragg wave phase velocity effect (see Figs. 3 and 4). The resultant radar velocity is the horizontal current velocity near the sea surface in the direction given in the heading. The buoy measurements of the same component are the in situ velocity measurements.

Location number	Velocity			Reference number	Radar velocity	In situ velocity
	Min.	Max.	Average			
			Mission Bay 251° component			
13, 17, 19	−27	−9	−14	*1, 2*	40	40
21, 22, 23	−32	2	−9	*1, 2*	44	44
18	−32	2	−12	*1, 2*	42	33
20	−21	23	−10	*1, 2*	44	48
5,6	−64	11	−28	*1*	29 *2*	14
1, 7, 11	−58	−33	−40	*1, 2*	14	11
2	−88	−53	−73	*2*	−23	13
(R_1)			−57			
(R_2)			−50			
Average					27	25
Bias					2	
rms					8	
			Mission Bay 341° component			
20	−5	23	6	*1*	9	12
2	6	34	21	*3*	14	28
(R_1)			−3			
(R_3)			7			
Average					12	20
Bias					−8	
rms						
			Point Loma 251° component			
9, 16, 17	−55	−40	−50	*1*	14	2
6, 14	−46	−25	−35	*1*	29	−2
4, 7, 12, 15	−47	−40	−44	*2*	17	3
3, 5, 10, 13	−49	−46	−47	*2*	14	4
11	−42	−32	−37	*2*	21	10
2	−56	−46	−53	*2*	8	−5
(R_1)			−64			
(R_2)			−61			
Average					17	2
Bias					15	
rms					7	
			Point Loma 341° component			
9, 16, 17	−38	−14	−29	1	−74	−57
6, 14	−33	−10	−18	1	−63	−45
4, 7, 12, 15	−23	−12	−17	2	−51	−48
3, 5, 10, 13	−23	−9	−14	2	−48	−40
11	−13	−10	−12	2	−46	−48
(R_1)			45			
(R_2)			34			
Average					−56	−48
Bias					−8	
rms					8	
Grand average					1.7	−0.5
Grand bias					2.2	
Grand rms					12	

the problems of aircraft motion will likely be less important than in the present study. This SAR device, if rigged on an orbiting platform, might be subject to fewer problems than experienced on the aircraft. If this were done, the interferometer technique might make feasible the remote sensing of the surface current field for the world oceans.

REFERENCES AND NOTES

1. R. M. Goldstein and H. A. Zebker, *Nature* **328**, 707 (1987).
2. F. I. Gonzales, C. L. Rufenbach, R. A. Shuchman, in *Oceanography from Space*, J. F. R. Gower, Ed. (Plenum, New York, 1981), pp. 511–523.
3. R. K. Raney, *IEEE Trans. Aerosp. Electron. Syst.* **AES-7**, 499 (1971).
4. SAR theory is described in R. Stewart, *Methods of Satellite Oceanography* (Univ. of California Press, Berkeley, 1985), p. 352.
5. Bragg waves observed by the SAR most effectively feel currents at a depth of approximately 2 cm below the mean surface [R. Stewart and J. Joy, *Deep Sea Res.* 21, 1039 (1974)]. We measured these (very) near-surface currents with drifting pieces of plywood (1.2 m long by 0.6 m wide by 2 cm thick) with an identification number on the top. Their overall horizontal dimensions made them rather insensitive to vertical motions of the short waves, whereas their thickness ensured response to the very near-surface currents felt by the Bragg waves. The windage of these simple floats was virtually nil, a fact we checked by droguing several buoys with large, 20-liter, water-filled plastic bags.

6. The research described in this paper was carried out, in part, by the Jet Propulsion Laboratory (JPL) of the California Institute of Technology and was sponsored by JPL and the Defense Advanced Research Projects Agency through an agreement with the National Aeronautics and Space Administration. The University of California Space Institute and the Scripps Institution of Oceanography also sponsored this work. We thank B. and W. Barnett, D. Fuhrman, and A. Saraspe for assisting with the ocean portion of the field program.

17 July 1989; accepted 28 September 1989

Magnitude of Late Quaternary Left-Lateral Displacements Along the North Edge of Tibet

GILLES PELTZER,* PAUL TAPPONNIER, ROLANDO ARMIJO

Images taken by the earth observation satellite SPOT of the Quaternary morphology at 18 sites on the 2000-kilometer-long Altyn Tagh fault at the north edge of Tibet demonstrate that it is outstandingly active. Long-term, left-lateral strike-slip offsets of stream channels, alluvial terrace edges, and glacial moraines along the fault cluster between 100 and 400 meters. The high elevation of the sites, mostly above 4000 meters in the periglacial zone, suggests that most offsets resulted from slip on the fault since the beginning of the Holocene. These data imply that slip rates are 2 to 3 centimeters per year along much of the fault length and support the hypothesis that the continuing penetration of India into Asia forces Tibet rapidly toward the east.

FOR NEARLY 2000 KM, THE ALTYN Tagh fault follows the northwestern edge of the Tibet-Qinghai highlands (Fig. 1). From studies of Landsat images, the fault has been inferred to be active, left-lateral (*1, 2*), and to absorb an important fraction of the present-day convergence between India and Asia by allowing the Tibet-Qinghai plateau to move northeastward relative to the Tarim basin (*1*). Little else is known of this fault, however. The historical seismic record in its vicinity is scanty (*3*). Strike-slip surface breaks have been found in the field only at two localities, near 90° and 96°E (*4*). Estimates of the rate of horizontal slip on the fault vary widely (*5*). Thus, although the length and morphology of the Altyn Tagh fault (*6*) imply that it may be the largest Quaternary strike-slip fault of Asia, further constraints on the amount and history of recent movement are required for a quantitative understanding of recent tectonics north of the Himalayas.

With a pixel size of 10 m in the panchromatic mode, images taken by the satellite SPOT (*7*) allow identification of cumulative Holocene offsets (that is, of features ~10,000 years old) on all faults moving at rates faster than 1 mm per year. In order to measure long-term offsets of late Quaternary morphological features and deposits along the Altyn Tagh fault, we studied seven scenes selected on the basis of earlier work with Landsat (*1*). In these scenes, late Quaternary deformation is spectacular at 18 sites (Fig. 1), 11 of which allow unambiguous estimates of recent left-lateral offsets (Table 1). We describe some of these sites below.

At site 1, the fault trace is marked in abundant, hummocky glacial till. It cuts the valley of a glacier and offsets its lateral moraines (Fig. 2, A and B). The left-lateral offset of the eastern, well-preserved, moraine ridge is 100 ± 20 m. The tip of the present glacier tongue in the valley lies only 1750 m upstream from the fault trace, thus only a few hundred meters above it if the valley slope is 10 to 20% (Fig. 2, A and B).

Between 78° and 79°E, the fault trace lies in the Karakax He valley, for the most part above 4000 m (Fig. 1). Here this valley is a gently west-sloping trough, several kilometers wide, flanked on either side by high mountains, which have been dissected by numerous, mostly extinct valley glaciers (*8*). Streams now flowing down such valleys north of the Karakax have built large alluvial fans. On the SPOT images, the fan deposits appear to form three major terraces, which are cut by the fault trace (Fig. 2, C to F). At site 3 (Fig. 2, C and D), for instance, the edges of the upper and middle terraces on the east bank of the stream are offset left-laterally by 185 ± 20 m. At sites 5 and 6 (Fig. 1 and 2, E and F), two ancient fans are offset 210 ± 20 m and 240 ± 20 m, respectively, in a left-lateral sense. The upper surfaces of these fans are incised by small channels and are no longer depositional surfaces (Fig. 2, E and F). The height of the fan surfaces above the Karakax flood plain and their degree of erosion suggest that they correspond to the middle terrace level at site 3 (Fig. 2, B and D). Close examination of the SPOT image reveals that the fault scarp in the fan surface at site 5 (Fig. 2E) faces toward the north across the eastern half of the fan and toward the south across the western half of this fan (*9*). This geometry, which is also clear across the fan at site 6 and across the terraces at site 3, attests to

Table 1. Values of 15 late Quaternary offsets measured along the Altyn Tagh fault. Errors are taken to be ±2 pixels [pixel sizes are 10 m and 20 m for panchromatic and multispectral (XS) images, respectively] for offsets of sharp, fossil morphological features (for example terrace or moraine edges), and standard deviations for average offsets of stream channels. Approximate elevations of sites, from (*6*), and inferred ages of offsets are also indicated.

Site	Offset (m)	Elevation (m)	Age (10^3 years)
1	100 ± 20	5000	<8
2	175 ± 20	4000	10 ± 2
3	185 ± 20	4000	10 ± 2
	185 ± 20		
5	210 ± 20	4000	10 ± 2
6	240 ± 20	4000	10 ± 2
7	~120	4000	<12
	~250		
	~120		
10	250 ± 80	4600	?
11	195 ± 95	3600	10 ± 2 ?
13	~400	4400	10 ± 2 ?
15	125 ± 40	3500	<8 ?
	125 ± 40		
18	60 ± 40	3300	10 ± 2 ?

Laboratoire de Tectonique, Mécanique de la Lithosphère, Institut de Physique du Globe de Paris, 4, place Jussieu, 75252 Paris, Cedex 05, France.

*Now at Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

Longer time scale motions

We have just shown that we can operate an interferometer that is sensitive to motions on the surface of the ocean by acquiring two images separated in time by a tenth of a second or so. If we can measure our phase to $\frac{1}{100}$ cycle or so, this means that we can measure currents at a precision of

$$\sigma_v = \frac{\lambda V}{4\pi B}\sigma_\phi \quad , \sigma_\phi \text{ in radians}$$

$$= \frac{\lambda V}{4\pi B} \cdot \frac{2\pi}{100}$$

$$= \frac{0.24 \cdot 250}{2 \cdot 20 \cdot 100} = 1.5 \text{ cm/s} \quad \text{for } \lambda = 24\text{cm}, B = 20\text{m}, V = 250\text{ m/s}.$$

What if instead of acquiring our two images on the same platform and pass, we use a satellite in an orbit that repeats exactly every month or so? Our velocity sensitivity could be

$$\sigma_v = \frac{\lambda}{4\pi}\frac{\sigma_\phi}{\Delta t}$$

where we have replaced $\frac{B}{V}$ by the time interval Δt. For a 24 cm system with 1/100 cycle phase accuracy, and $\Delta t = 30$ days, we have

$$\sigma_v = \frac{0.24}{4\pi} \cdot \frac{1}{30\text{days}} \cdot \frac{2\pi}{100}$$

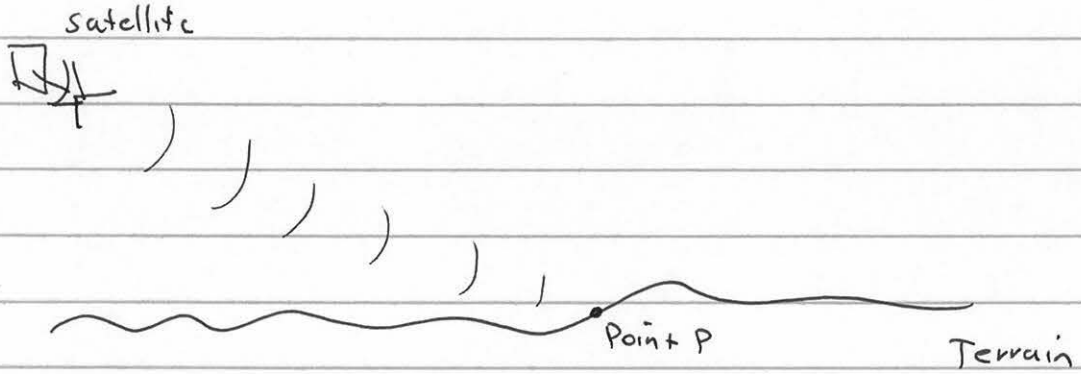
$$= 4 \times 10^{-5} \text{ m/day} \quad , \text{or} \quad 1.5 \text{ cm/year}!$$

Now, tectonic motions of the major plates on the surface of the Earth are about a few cm/yr, so we potentially can map even these very slow processes.

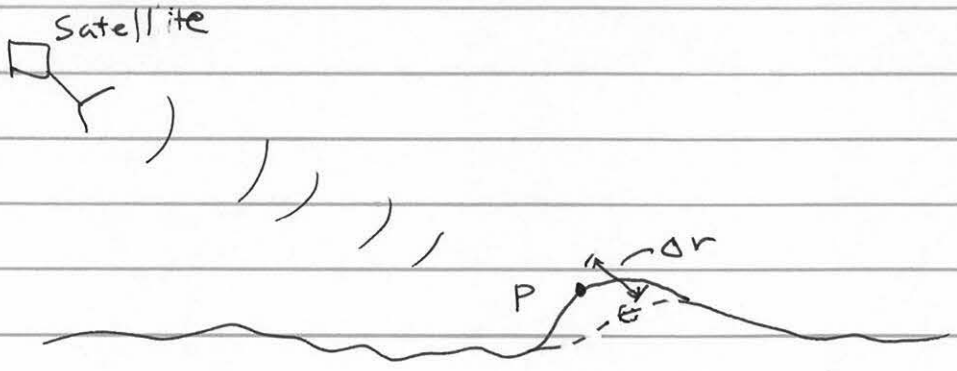
Interferometry for surface deformation

There is a more intuitive way to picture surface deformation measurements, looking at the displacements directly rather than inferring velocity. This is also important because many motions are episodic and don't occur as smooth, continuous motions.

Instead, let's consider the following experiment. Illuminate a scene at some time t_0, and record the radar image.



Now, some time later, repeat the same experiment, but displace the point P toward the radar by a distance Δr:



Something has deformed the surface in the vicinity of P and caused it to be closer to the radar at new time $t_1 = t_0 + \Delta t$. The phase of the pixel in the two scenes is

$$\phi(t_1) = \frac{-4\pi}{\lambda} r(t_1) + \phi_{scatt}$$

$$\phi(t_2) = -\frac{4\pi}{\lambda} r(t_2) + \phi_{scatt}$$

where ϕ_{scatt} is the phase associated with the random sum of the waves from each scattering center in the resolution cell. If the surface is unchanged between observations, ϕ_{scatt} will be the same in the two images. If in addition the radar is at exactly the same place in the sky at time t_2, the interferogram phase will be

$$\phi_{int} = \phi_{t_1} - \phi_{t_2}$$

$$= -\frac{4\pi}{\lambda}(r(t_1) - r(t_2)$$

$$= -\frac{4\pi}{\lambda}\Delta r$$

Thus the phase of this interferogram at every point is the line of sight component Δr of any surface motion that occurred between times t_1 and t_2.

Complicating factors

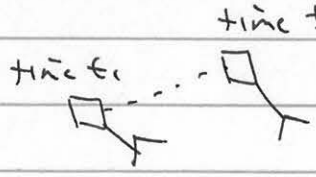
Two things can confuse the above simple picture: surface changes can cause ϕ_{scatt} to be unequal in the two images, or the satellite can return to a point not quite exactly where it was on the previous pass.

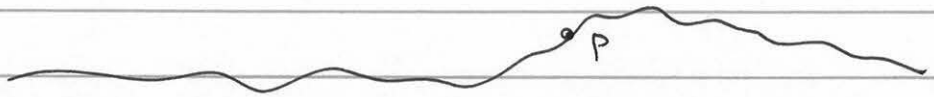
Surface changes altering ϕ_{scatt} are the temporal decorrelation phenomenon we have already discussed. We can minimize the temporal effect by using shorter temporal bandwidths

or longer imaging wavelengths. But both of these affect our motion sensitivity.

→ Why is that?

The second problem is associated with imprecision in the satellite orbits. Note that if the satellite returns to a spot slightly different in space to the position of the first acquisition, we reproduce our topography interferometer.





We know that the interferogram in this situation contains terms dependent on local topography. Hence even if there was motion around P, the phase signature of the deformation is mixed with that of the topography. Thus we need a method to distinguish the two types of phase modulation.

Equations for deformation interferometry

Begin with our equations for topography as before:

$$\phi_{int} = \frac{-4\pi}{\lambda}\left(r_1(t_1) - r_2(t_2)\right)$$

where we have explicitly identified the r's as derived from separate

orbits. Thus

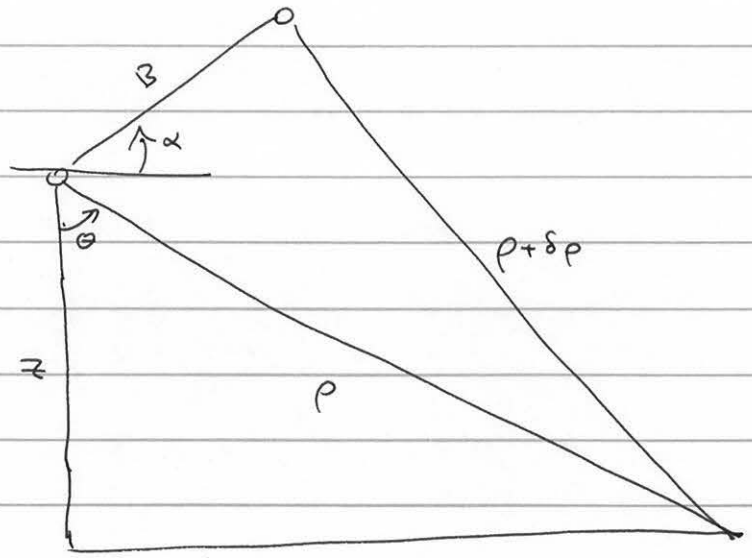
$$\phi_{int} = \frac{-4\pi}{\lambda} \delta\rho$$

where $\delta\rho$ is the change in range due to a difference in orbit geometry, not motion.

We used a law of cosines expansion to obtain

$$(\rho + \delta\rho)^2 = \rho^2 + B^2 - 2\rho B \sin(\theta - \alpha)$$

from the construction



Neglecting the term of order $(\delta\rho)^2$, we found

$$\delta\rho = B \sin(\theta - \alpha) + \frac{B^2}{2\rho}$$

Furthermore, the parallel ray approximation leads to a further simplification

$$\delta\rho = B \sin(\theta - \alpha)$$

or

$$\delta\rho = B_{\perp}$$

or that the topographic phase is proportional to the parallel component of the baseline at each point:

$$\phi_{topo} = \frac{-4\pi}{\lambda} B_{\parallel}$$

Now, suppose that in addition we have a displacement of a point P by Δr toward the radar. The interferogram phase will be altered by the corresponding amount

$$\phi_{int} = \frac{-4\pi}{\lambda} B_{\parallel} - \frac{4\pi}{\lambda} \Delta r$$

We need to be able to separate the two effects. In the initial topography case, we assumed $\Delta r = 0$, so only the first term contributed. In the velocity case, we assumed $B_{\parallel} = 0$, so only the second term mattered. How can ~~that~~ we accomodate both?

Elimination of topographic phase

Consider the more complete form of the phase equation where dependence on topography is shown explicitly.

$$\phi_{int} = -\frac{4\pi}{\lambda} B \sin(\theta - \alpha) - \frac{4\pi}{\lambda} \Delta r$$

and $z = r \cos\theta$

$$\phi_{int} = -\frac{4\pi}{\lambda} B \sin\left(\cos^{-1}\frac{z}{r} - \alpha\right) - \frac{4\pi}{\lambda} \Delta r$$

Suppose first of all that we have independent knowledge of

the topography of an area. After all, many areas of the world have been mapped. We could calculate the topography term from knowledge of z and the imaging geometry, r, B, and α. Then we could subtract this phase from the measured interferogram and we would be left with

$$\phi_{def} = \frac{-4\pi}{\lambda} \Delta r$$

This is the approach pioneered by the French at CNES and it often works well (we'll address limitations below). So if we have a map of an area in digital form with sufficient accuracy this approach is viable.

Three-pass method

Sometimes we do not have maps of a region available to remove the topographic term directly. In this case we can eliminate the topo dependency by using another radar pass over the scene. In essence, we use the radar to map topography and then subtract off that term.

In fact we will not have to solve for the topography explicitly to use this approach.

We'll start by altering our geometric construction to show two interferograms formed by three passes, where one pass serves as a common reference to the other two.

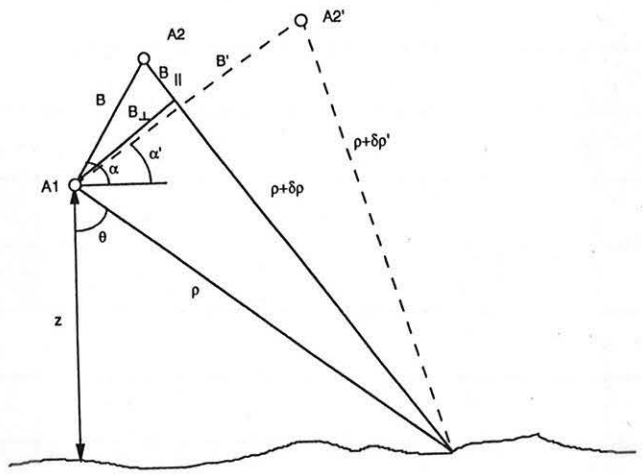


Figure 1. Radar imaging geometry. The solid lines show that radar signal paths for the first interferogram pair formed by antennas at A1 and A2. Dashed lines show signal path for second interferogram acquired over the same site but with antennas located at A1 and A2′.

The phase from the solid line system, assuming for the moment absence of deformation, was given above

$$\phi = \frac{-4\pi}{\lambda} B_{||}$$

so from the primed system, with the dashed lines,

$$\phi' = \frac{-4\pi}{\lambda} B'_{||}$$

In general $\phi \neq \phi'$ as the baselines are different. But, the underlying topography is constant in both pairs. So we can express the ratio

$$\frac{\phi}{\phi'} = \frac{B_{||}}{B'_{||}}$$

which depends only very weakly on topography.

Suppose for the moment that $\frac{\phi}{\phi'}$ or $\frac{B_{\parallel}}{B'_{\parallel}}$ is truly a constant value with respect to topography. Now consider how ϕ' would change if we had a surface deformation occurring ~~at~~ during the interval we collect the primed interferogram over but no displacement over the unprimed interval. Then

$$\phi' = \frac{-4\pi}{\lambda}\left(B'_{\parallel} + \Delta r\right)$$

and since $\phi = \frac{-4\pi}{\lambda} B_{\parallel}$, if we form the quantity

$$\phi' - \frac{B'_{\parallel}}{B_{\parallel}}\phi = \frac{-4\pi}{\lambda}\Delta r$$

Thus if we scale the unprimed measurement by the ratio of the parallel baselines, and subtract it from the primed interferogram, we obtain the deformation interferogram directly.

Thus if we use three radar passes, we get the same result as if we had used the two pass plus map method.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. B10, PAGES 19,617–19,634, OCTOBER 10, 1994

On the derivation of coseismic displacement fields using differential radar interferometry: The Landers earthquake

Howard A. Zebker, Paul A. Rosen, Richard M. Goldstein,
Andrew Gabriel, and Charles L. Werner
Jet Propulsion Laboratory, California Institute of Technology, Pasadena

Abstract. We present a map of the coseismic displacement field resulting from the Landers, California, June 28, 1992, earthquake derived using data acquired from an orbiting high-resolution radar system. We achieve results more accurate than previous space studies and similar in accuracy to those obtained by conventional field survey techniques. Data from the ERS 1 synthetic aperture radar instrument acquired in April, July, and August 1992 are used to generate a high-resolution, wide area map of the displacements. The data represent the motion in the direction of the radar line of sight to centimeter level precision of each 30-m resolution element in a 113 km by 90 km image. Our coseismic displacement contour map gives a lobed pattern consistent with theoretical models of the displacement field from the earthquake. Fine structure observed as displacement tiling in regions several kilometers from the fault appears to be the result of local surface fracturing. Comparison of these data with Global Positioning System and electronic distance measurement survey data yield a correlation of 0.96; thus the radar measurements are a means to extend the point measurements acquired by traditional techniques to an area map format. The technique we use is (1) more automatic, (2) more precise, and (3) better validated than previous similar applications of differential radar interferometry. Since we require only remotely sensed satellite data with no additional requirements for ancillary information, the technique is well suited for global seismic monitoring and analysis.

Introduction

Interferometric radar techniques for the generation of highly accurate digital elevation models (DEMs) by now have been well documented in the literature [*Zebker and Goldstein*, 1986; *Goldstein et al.*, 1988; *Prati et al.*, 1990; *Zebker et al.*, 1992; *Evans et al.*, 1992; *Madsen et al.*, 1993, also personal communication, 1993; H. A. Zebker et al., personal communication, 1993]. A related application of such techniques allows the measurement of the motion of all resolved points in a remotely sensed image [*Goldstein and Zebker*, 1987; *Goldstein et al.*, 1989]. These similar techniques both follow from analysis and interpretation of interferograms, which consist of the phase differences between two radar images of the same scene acquired at separate locations or times: a sensor location change gives sensitivity to topography and a sensor temporal change gives motion sensitivity. A combination of the two approaches, denoted differential radar interferometry since the phase measurements of interest result from the difference of two interferograms, has previously been used by *Gabriel et al.* [1989] to map the changes in surface elevation of agricultural fields over a large area to centimeter-level sensitivities.

More recently, there has been activity by at least two groups applying the capabilities of radar interferometry to the study of seismic phenomena. *Massonet et al.* [1993] of Centre National d'Etudes Spatiales (CNES) in Toulouse, France, used an interferometric digital elevation model derived from the European Space Agency (ESA) ERS 1 satellite data for analysis of the magnitude 7.3 earthquake centered near Landers, California, on June 28, 1992. In this study a single interferogram which contained phase signals from the local topography and from the earthquake displacements was subtracted from a manipulated U.S. Geological Survey (USGS) 15 arc min DEM of the area. The residual phases were interpreted as ground displacements from the event. The interferogram, when corrected for topographic effects, shows a displaced dual-lobed pattern of fringes emanating from the fault zone, where each fringe represents about 2.8 cm of motion in the radar line of sight direction. They also derive a theoretical fringe pattern from a model of the earthquake motion which matches the observations fairly closely.

Despite its success, there are several important limitations in the technique used for the above study. Although a USGS 90-m spacing DEM was available for this site, for many sites in the world, no DEM exists. In addition, an existing DEM may not be sufficiently accurate to yield the desired precision. DEMs typically contain errors and distortions of the order of the phenomena being investigated. The CNES team estimates a precision in their measurements of about 2.8 cm in the radar line of sight motion, limited mainly by imprecision in the USGS DEM plus radar system noise. Also, the DEM must be precisely coregistered to the radar image, which itself may be a difficult task. (S. N. Madsen et al. (personal communication, 1993) provide more on errors induced by DEM misregistration.) Finally, since the interferogram phases are all measured modulo 2π, the absolute, or even relative, phase relationship between arbitrary points

Paper number 94JB01179.
0148-0227/94/94JB-01179$05.00

in the scene is difficult to determine. Thus it is virtually impossible to fit continuous two-dimensional models of the displacement field to the observations.

These limitations aside, it is important to realize that the phase displacements due to motion in an interferometric DEM can be hundreds of times more sensitive than simply differencing the actual height measurements before and after an event (see below). More complete use of phase information allows the interferometric approach to map centimeter scale distortions over a region many tens of kilometers in size at a resolution of a few meters.

In this paper, we approach the Landers analysis differently from Massonet et al. by utilizing only data acquired by the ERS 1 satellite. Our approach overcomes the aforementioned limitations and hence is more readily quantifiable given the radar system parameters, and the quality of the result can be measured "up front." Specifically, imprecision introduced by the USGS DEM in the CNES study is not present, coregistration occurs automatically in forming the interferograms, and the entire usable phase field is "unwrapped," meaning that the displacement at each point is known digitally in an absolute sense. Unwrapping renders the displacement field more amenable to computer modeling and analysis and permits the precision of the technique to be increased from the 2.8-cm radar line of sight reported by Massonet et al. to about 0.2 cm obtained here. Further, we verify the accuracy of the measurements by comparing to a displacement field derived from conventional surveying techniques. These survey data were derived from a combination of electronic distance measurement (EDM) lines and Global Positioning System (GPS) satellite receivers. The methods and results presented here can serve as a baseline for the design of a seismic monitoring program.

The structure of the paper is as follows. We begin by summarizing the theory of radar interferometry and differential interferometry, error sources, and expectations of performance for seismic studies. Next, we present the set of differential radar interferometric observations of the Landers earthquake and discuss their accuracy. Finally, we compare our results with those of the earlier study and with the in situ measurements made by GPS techniques.

Summary of Theory

In this section we derive the equations needed for calculating ground displacement fields from interferometric synthetic aperture radar measurements. Here we assume that the reader has a general knowledge of radar remote sensing systems. The interested reader may consult a general text on radar remote sensing such as that by *Elachi* [1988] or by *Curlander and McDonough* [1991] for questions on radar system operation and processing. As for information on the technique of radar interferometry, much of the work is still too new for general textbooks, and thus the technical literature is the only source available. We cite the major relevant papers in this text, and the reader may consult these when appropriate.

A side-looking spaceborne synthetic aperture radar system may map a continuous swath many tens of kilometers in width as the satellite progresses along its orbit track, yielding measurements of the amplitude and phase of radar echoes associated with independent patches on the ground perhaps 10 m in size: this size is the resolution of the radar. We first examine the case where no ground movement between radar observations occurs. Consider two radar systems observing the same ground swath from two positions A1 and A2, respectively, as illustrated in Figure 1. The measured phase at each point in each of the two radar images may be taken as equal to the sum of a propagation part proportional to the round-trip distance traveled and a scattering part due to the interaction of the wave with the ground. If each resolution element on the ground behaves the same for each observation (see more on this important condition below), then calculating the difference in the phases removes dependence on the scattering mechanism and gives a quantity dependent only on geometry. If the two path lengths are taken to be ρ and $\rho + \delta\rho$, the measured phase difference ϕ will be

$$\phi = \frac{4\pi}{\lambda}\,\delta\rho \tag{1}$$

or 2π times the round-trip distance difference in wavelengths. The law of cosines permits solution for $\delta\rho$ in terms of the imaging geometry as follows. Then

$$(\rho + \delta\rho)^2 = \rho^2 + B^2 - 2\rho B \sin(\theta - \alpha) \tag{2}$$

where the baseline length is B, the range to a point on the ground is ρ, the look angle is θ, and the angle of the baseline with respect to horizontal at the sensor is α. Neglecting the term of order $(\delta\rho)^2$ yields

$$\delta\rho \approx B \sin(\theta - \alpha) + \frac{B^2}{2\rho}. \tag{3}$$

For simplicity in describing the approach we used, we can make a second approximation, although it is not necessary for the analyses presented below. In the case of spaceborne geometries we can ignore the second term on the right-hand side of (3) and obtain

$$\delta\rho \approx B \sin(\theta - \alpha) \tag{4}$$

or

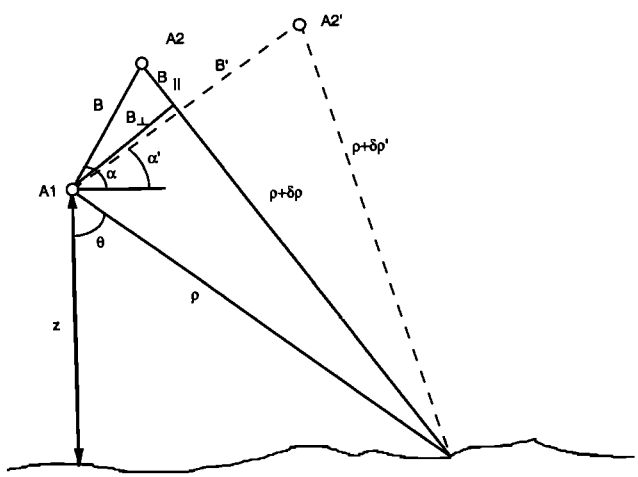


Figure 1. Radar imaging geometry. The solid lines show that radar signal paths for the first interferogram pair formed by antennas at A1 and A2. Dashed lines show signal path for second interferogram acquired over the same site but with antennas located at A1 and A2′.

$$\delta\rho \approx B_{\parallel}. \tag{5}$$

$B_{\parallel} = B \sin(\theta - \alpha)$ is simply the component of the baseline parallel to the look direction. This is the parallel-ray approximation used by *Zebker and Goldstein* [1986] in their initial paper on topographic mapping.

Equations (1) and (4) show that the measured phase of an interferometer is the component of the interferometer baseline parallel to the look direction to a given point on the surface measured in wavelengths, multiplied by two for round-trip travel. We note that the height sensitivity of the instrument enters through the dependence of the exact look angle θ on the altitude $z = h - \rho\cos\theta$, where h is the height of the sensor above the reference surface.

If a second (denoted by a prime) interferogram is acquired over the same area, sharing one orbit with the previous pair so that ρ and θ are unchanged (dashed lines in Figure 1), we can compare the interferogram phases with each other. This second interferogram is acquired with a different baseline B' and baseline orientation α', thus a different $B'_{\parallel}$. Combining (1) and (4) above, we obtain

$$\phi' = \frac{4\pi}{\lambda} B'_{\parallel}. \tag{6}$$

Examination of the ratio of the two phases yields

$$\phi/\phi' = B_{\parallel}/B'_{\parallel}. \tag{7}$$

In other words, the ratio of the phases is equal to the ratio of the parallel components of the baseline, independent of the topography.

Now consider the situation of two interferograms acquired over the same region as before but in this case an earthquake has displaced each resolution element between observations for the primed interferogram. The displacements are assumed small with respect to a resolution cell so that the radar echoes remain correlated. Here in addition to the phase dependence on topography there is a phase change due to the radar line of sight component of the displacement $\Delta\rho$. In this interferogram the phase ϕ' will be given by

$$\phi' = \frac{4\pi}{\lambda}(B'_{\parallel} + \Delta\rho). \tag{8}$$

The displacement term $\Delta\rho$ adds to the topographic phase term, creating confusion in the interpretation of the result. However, if the data from the initial unprimed interferogram are scaled by the ratio of the parallel components of the baseline and subtracted from the primed interferogram, we can obtain a solution dependent only on the displacement of the surface, as follows

$$\phi' - \frac{B'_{\parallel}}{B_{\parallel}}\phi = \frac{4\pi}{\lambda}\Delta\rho. \tag{9}$$

Since the quantity on the left is determined entirely by the phases of the interferograms and the orbit geometries, the line of sight component of the displacement $\Delta\rho$, is measurable for each point in the scene.

The ratio

$$\frac{B'_{\parallel}}{B_{\parallel}} = \frac{B' \sin(\theta - \alpha')}{B \sin(\theta - \alpha)} \tag{10}$$

is a function of the angle θ, which depends both on the illumination geometry and also the topography at each point in the radar image. To evaluate (9) via (10) directly, we must solve for the topographic map of the area of interest from the interferometric data or obtain the elevation data from another source. In the interests of simplicity and accuracy, we have devised an indirect approach for which it is not necessary to implement the step of either topographic solution or registration of dissimilar data sets.

We remove from the interferogram phase a term that would exist even in the absence of topography on a spherical Earth. The phase corrected for the "curved Earth" effect, denoted ϕ_{flat}, is given by

$$\phi_{\text{flat}} = \frac{4\pi}{\lambda}[B \sin(\theta - \alpha) - B \sin(\theta_0 - \alpha)], \tag{11}$$

where θ_0 is the look angle to each point in the image assuming zero local height. The interferogram phase after this correction represents the distortion of the interference grating pattern due to topographic variation relative to a spherical surface and displacements due to motion in the scene. The interferograms shown throughout this paper have been "flattened" according to (11).

Noting that the deviation of the exact θ from θ_0 is small, we can expand the first term on the right-hand side of (11), leading to

$$\phi_{\text{flat}} = \frac{4\pi}{\lambda}\delta\theta B \cos(\theta_0 - \alpha), \tag{12}$$

where $\delta\theta = \theta - \theta_0$. Numerically, ϕ_{flat} is equal to the product of the perpendicular component of the baseline $B_{\perp}$, assuming no topography is present on the surface, and the topographic angular distortion $\delta\theta$. Thus the ratio $\phi_{\text{flat}}/\phi'_{\text{flat}}$ is now in terms of θ_0 rather than θ and depends only on the viewing geometry and the baseline. If we now restate the differential phase equation (9) above in terms of the flattened phase ϕ_{flat}, we obtain

$$\phi'_{\text{flat}} - \frac{B'_{\perp}}{B_{\perp}}\phi_{\text{flat}} = \frac{4\pi}{\lambda}\Delta\rho. \tag{13}$$

With this function, we can now solve directly for the displacement $\Delta\rho$ without requiring the exact values of θ, and hence the topographic information, at an intermediate step. We have used this procedure (equation (13)) for the reduction of the data presented in this paper.

We note that if the baseline used in the flattening operation (equation (11)) is not exactly the true baseline value, (12) will contain error terms and the subsequent displacement maps will be distorted. This condition is described in detail in the appendix; for the rest of this paper we will assume that the correct baseline values are used.

We have shown that the phase in radar interferograms depends both on the local topography and on any motion that may occur between viewing instances. We may compare the sensitivity of the phase measurement to the phenomena of topography and displacement, which may be derived by differentiating (8) with respect to height through $B_{\parallel}$ and displacement. In the first case, using $dz = \rho \sin\theta\, d\theta$, obtained from the dependence of height on angle described above, we find

$$d\phi' = \frac{4\pi}{\lambda} B \cos(\theta - \alpha)\, d\theta \tag{14}$$

and

$$\frac{d\phi'}{dz} = \frac{4\pi}{\lambda} \frac{B \cos(\theta - \alpha)}{\rho \sin\theta}. \tag{15}$$

For the displacement case we have

$$\frac{d\phi'}{d\Delta\rho} = \frac{4\pi}{\lambda}. \tag{16}$$

Since the distance ρ typically is very much greater than the baseline distance B, it is evident from (15) and (16) that a much more sensitive dependence of phase results from displacements than from topographic variation. In other words, the system is more sensitive in an absolute sense to surficial change than to the topography itself. Comparing the two results numerically, for the April–August ERS 1 case described here (see next section), 1 m of topography gives a phase signature of 4.3°, while for the same pass pair a 1-m surface displacement yields a phase signature of 12800°, or nearly 3000 times greater sensitivity. Thus, while radar interferometry can be used to measure topography to an accuracy of meters, displacements may be determined to the centimeter or millimeter level.

This ratio of sensitivities illustrates the power of the interferometric technique to detect small changes. If, for example, we chose to map seismic displacements by differencing DEMs, whether acquired interferometrically or by conventional stereo photogrammetry, changes would only be visible if they were significant in size compared to the uncertainty of the DEM measurement, which is typically meters. For the interferometric case in the previous paragraph, for example, system noise limits the useful signatures to those causing a phase shift greater than about 20°, or 4.6 m. While thus permitting topographic mapping with a vertical precision of 4.6 m (H. A. Zebker et al. (personal communication, 1993) give a discussion of ERS 1 DEMs with this precision), a worthwhile result for many applications, it is not particularly useful for the study of earthquakes. In contrast, if data are acquired with an interferometric pair that spans the seismic event, even 1 cm of line of sight displacement results in a signature of 64°, easily detectable in ERS 1 data.

There are, however, two very important limitations to the interferometric technique. First, radar echoes acquired on the three passes must correlate with each other; that is, the signals must be substantially similar over a significant period of time. Physically, this translates to a requirement that the ground scattering surface be relatively undisturbed at the radar wavelength scale between measurements. Several studies have addressed this phenomenon, both theoretically [e.g., *Li and Goldstein*, 1990], and experimentally [e.g., *Gray et al.*, 1993]. *Zebker and Villasenor* [1992] were able to model and quantify the temporal decorrelation process and found that different surfaces decorrelate at different rates. This limits the applicability of the approach to areas that do not change much with time. Some regions, such as desert areas, may exhibit very little decorrelation over long periods. In the data presented here, correlation was usably high even after 105 days, the longest time period examined. From this we conclude that the radar properties of the desert surface change little over months if weather and other environmental factors are not altering the condition of the ground significantly.

The second limitation, more important for this study, is that the phases must be "unwrapped" before data from one interferogram may be used to correct the second interferogram (equation (9)) to estimate the displacement phases. The measurements of each phase are known only modulo 2π, and various techniques exist [*Goldstein et al.*, 1988; *Ghiglia and Romero*, 1993; A. Hiramatsu, personal communication, 1992] to determine the absolute phase relationship between all arbitrary points in a data set (that is, unwrapping). While not fully characterized in any of the existing literature, it is apparent that the ability to unwrap arbitrary phase fields depends on two factors: the noise level in the system and the interferometric fringe spacing. For the July–August pair described here in particular, the interferometric baseline is quite large, being 40% of the critical baseline at which no correlation between signals is possible. (For a more complete discussion on baseline decorrelation, see *Zebker and Villasenor* [1992]). Since the fringe rate depends on local surface slope, typically it is more difficult to estimate phases reliably in rough terrain than in flat terrain if the fringe rate is high to begin with. The result of this is that we were unable to obtain reliable phase estimates in the rougher regions, as will be seen in the data presented below.

ERS 1 Interferograms of the Landers Earthquake

The ERS 1 radar system, operating at a wavelength of 5.67 cm, images Earth from an altitude of about 790 km and produces radar backscatter maps of 100-km-wide swaths at a resolution of about 25 m across track and 6 m along track. We obtained raw ERS 1 radar signal samples acquired over the Landers region on April 24, July 3, and August 7, 1992. We combined these to form two interferograms, one from the April–August pair and one from the July–August pair. The April–August pair spans the June 28 earthquake and was chosen over the April–July pair which exhibited an exceptionally large baseline. No data were acquired on May 29 when the satellite again passed over the site. Orbit reconstructions provided by the European Space Agency (ESA) enabled us to determine the geometrical parameters for the pairs chosen as given in Table 1. The parallel baseline components given in Table 1 are for a look angle of 21°. Since the radar swath is quite wide, the actual look angle varies from about 17° to 23° and the parallel components vary somewhat.

The Landers area is shown in Figures 2a and 2b, where the faults shown, illustrated by heavy lines, are those affected by the Landers earthquake and imaged by the ERS 1 radar.

Table 1. ERS 1 Landers Interferometer Baseline Parameters

Pair	Baseline B, m	Orientation α, deg	Parallel Component $B_\parallel$, m	Perpendicular Component $B_\perp$, m
April–August	146.1	152	110.3	95.8
July–August	503.1	175	220.5	452.2

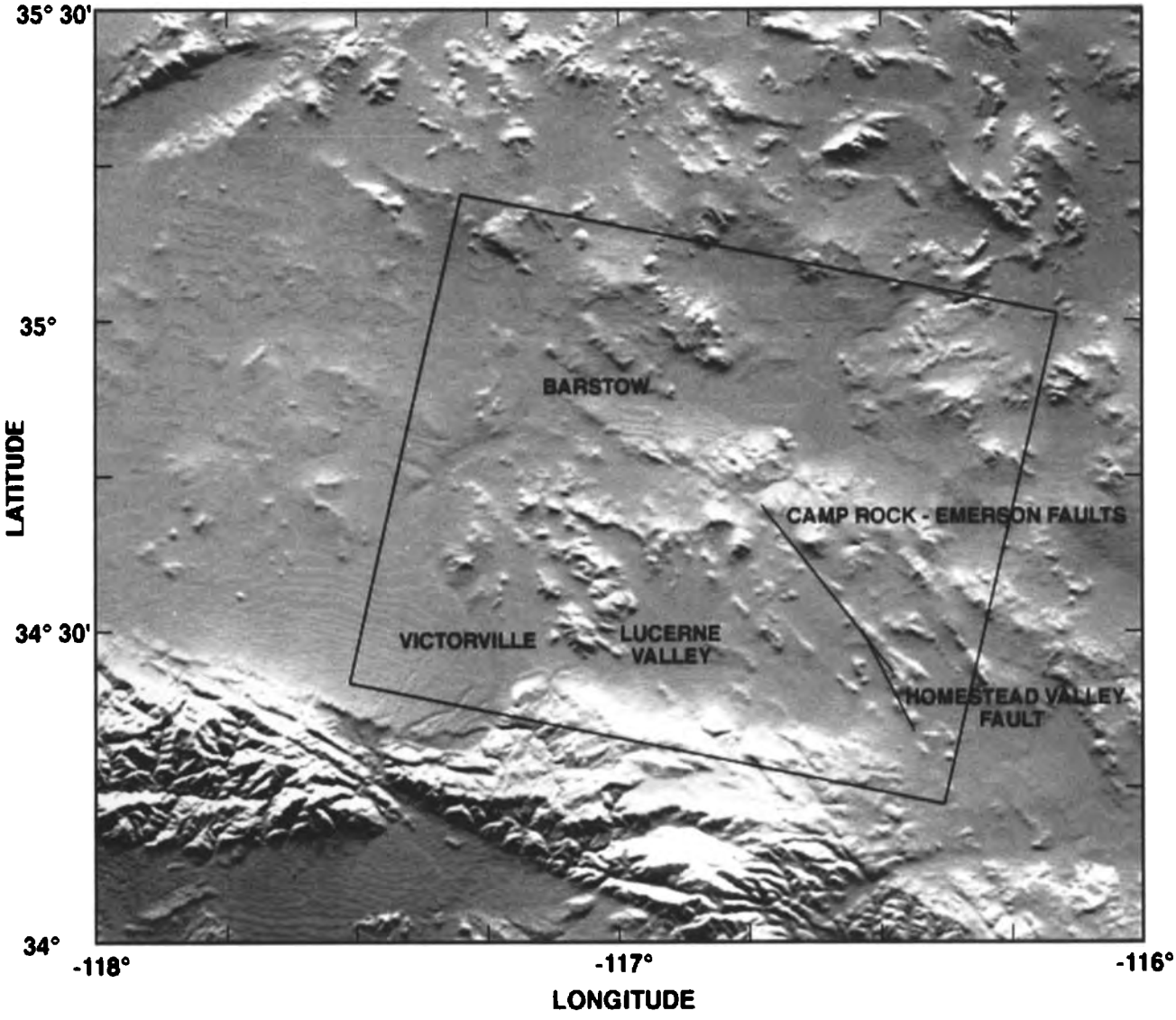


Figure 2a. Shaded relief map derived from USGS DEM with geographic features shown for reference. The inset rectangle is the region of the ERS 1 radar swath analyzed here. Also noted are the approximate position of the Camp Rock–Emerson and Homestead Valley faults. The cities of Barstow, Victorville, and also Lucerne Valley area are shown for reference.

Figure 2a is a shaded relief representation of the region: limits of the ERS 1 data we analyzed are indicated by the rectangle. Figure 2b is the radar backscatter image with two further areas denoted in addition to the faults. It represents an area roughly 113 km by 90 km. These data, as well as the radar images below, are in a radar slant range and along-track direction coordinate system. Radar slant range, denoted by ρ in Figure 1, here means that the across-track distances given are in terms of line of sight distance of each point to the radar rather than that distance projected on the ground. That is, the data have not been geocoded, or placed in map coordinates. We have preserved the "natural" spacing of the data points in order to maintain the highest possible signal fidelity throughout the processing procedure. However, we do apply a geocoding transformation before comparison with the field survey results described in the following section.

We processed the radar signal samples at the Jet Propulsion Laboratory (JPL) using a software processor constructed specifically by us for ERS 1 interferometric applications. The data were processed using a range-Doppler algorithm, but the range-compressed signals were filtered for the July–August pair using the method suggested by F. Gatelli et al. (personal communication, 1993) to reduce baseline decorrelation. We found that this approach yielded about 5–10% greater correlation in some regions at the expense of a slight reduction in range resolution. More information on radar processing approaches can be found in the general radar textbooks described previously.

The interferograms obtained in this process are shown in Plate 1, with the corresponding correlation coefficient maps shown in Plate 2. The top image in each case represents the April–August interferogram, while the bottom image shows data from the July–August pair. The June 28 earthquake effects are found in the April–August pair. In these plots the fringe signature of a curved Earth surface, as described in

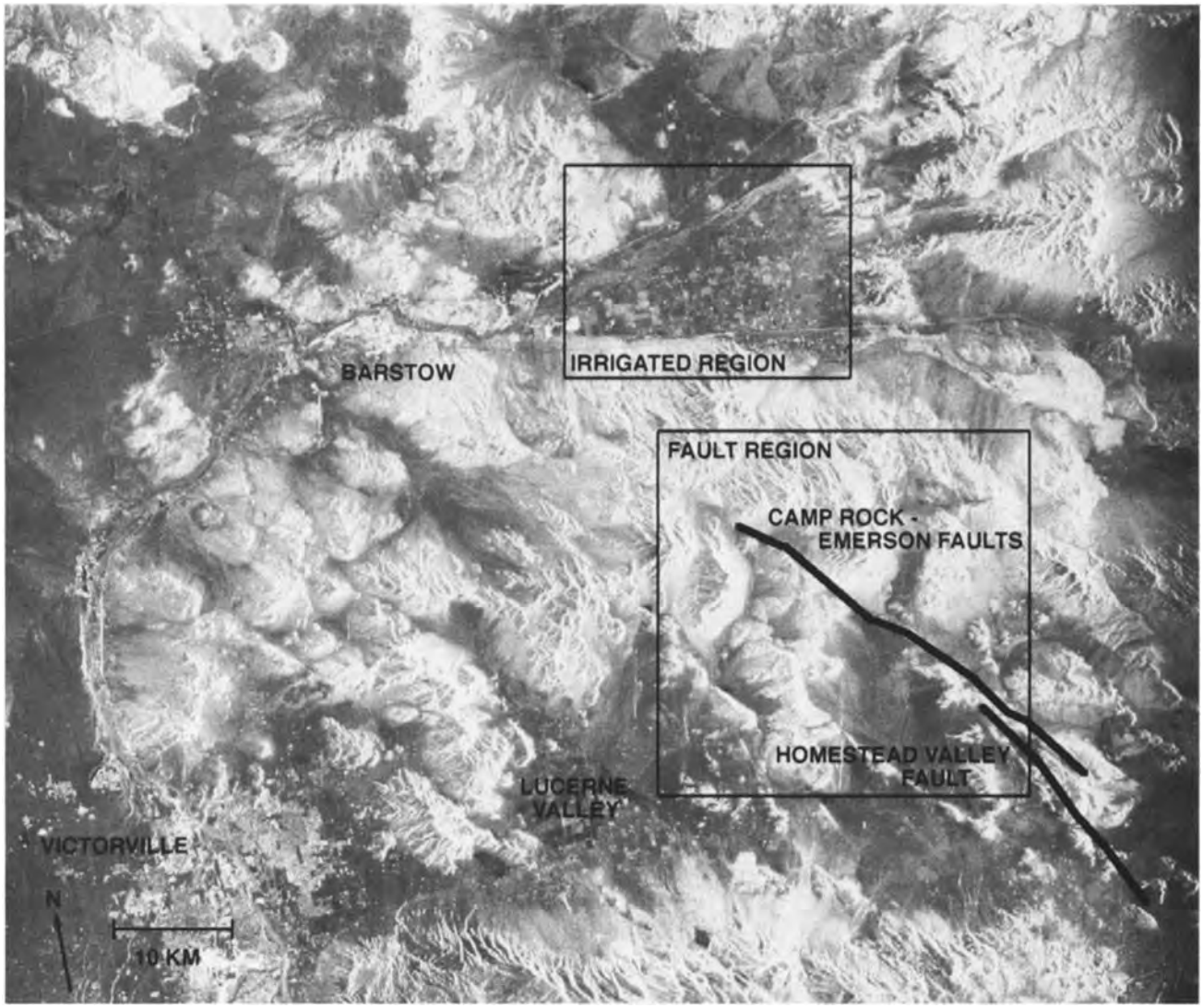


Figure 2b. Radar image of the Landers area, where the fault locations are illustrated by the heavy lines. The radar image covers an area on the ground approximately 113 km by 90 km. Insets of an irrigated region as well as a fault zone are shown for later reference. These data, as well as the radar images in the remaining figures, are in a radar slant range and along track direction coordinate system.

the section on theory above, has been removed from the interferograms for clearer display. Note that the very high fringe rates, and corresponding loss of correlation, in the mountainous regions for the July–August pair, lead to our inability to unwrap the phase in these regions. Also note in the April–August pair a similar loss of correlation in the fault zone, presumably due to (1) very high fringe rates of greater than one cycle per resolution element, (2) large ground shifts resulting in lack of precision alignment of the pixels from pass to pass, and (3) rearrangement of the surface at the wavelength scale from the earthquake itself.

These interferograms were filtered using a spatially variable bandpass filter that selected the optimal fringe rate passband in each 32 by 32 pixel subregion in the interferogram. In this process we also identified areas of low fringe visibility to serve as a mask in the final product, eliminating regions where we felt we could not trust the phase estimates. The data were then unwrapped using the method of A. Hiramatsu (personal communication, 1992), which is an extension of the method first presented by *Goldstein et al.* [1988].

Finally, the differential interferogram was calculated by scaling the July–August measurement by the ratio of the parallel baseline components for each look angle and subtracting that value from the corresponding value in the April–August pair. The result is a map of the displacements of the ground in the radar line of sight direction (equation (9)), shown in Plate 3, where the shift is coded by color and the brightness at each point is the radar image brightness. In addition, contour lines representing line of sight displacements spaced every 5 cm are shown.

It must be noted that the earthquake is not the only process affecting the phase measurements in this region of the Mojave. Plate 4 is an enlargement of the April–August interferogram plus the correlation coefficients for the region east of Barstow indicated in Figure 2 where center pivot irrigation has been employed. The irrigated circles, and some other agricultural fields, show a clear loss of correla-

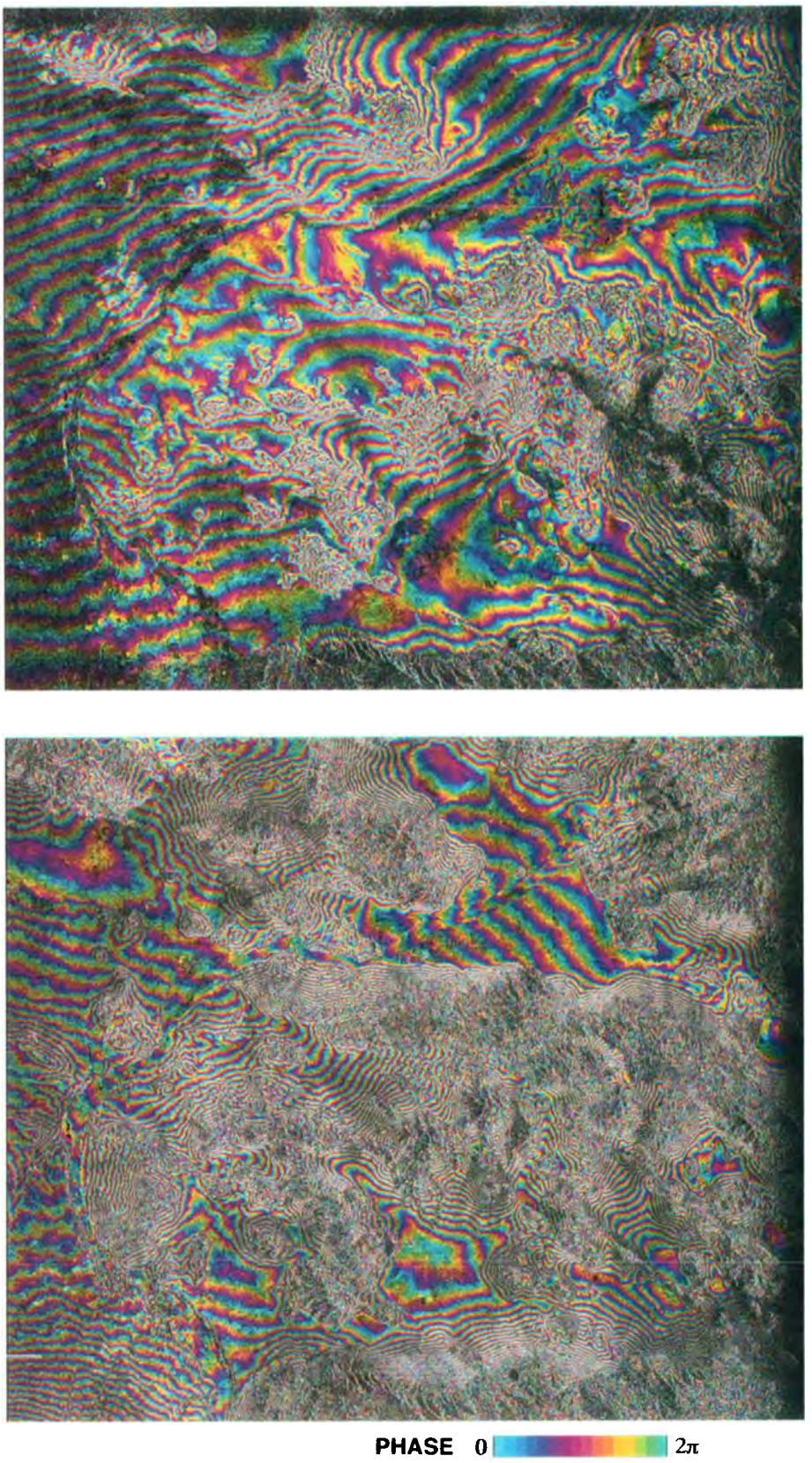


Plate 1. Interferograms of the Landers area. (top) April–August interferogram; (bottom) July–August pair. The June 28 earthquake effects are found in the April–August pair. The fringe signature of a curved Earth surface has been removed from the interferograms for clearer display. Note the very high fringe rates in the mountainous regions for the July–August pair, leading to our inability to unwrap the phase in these regions.

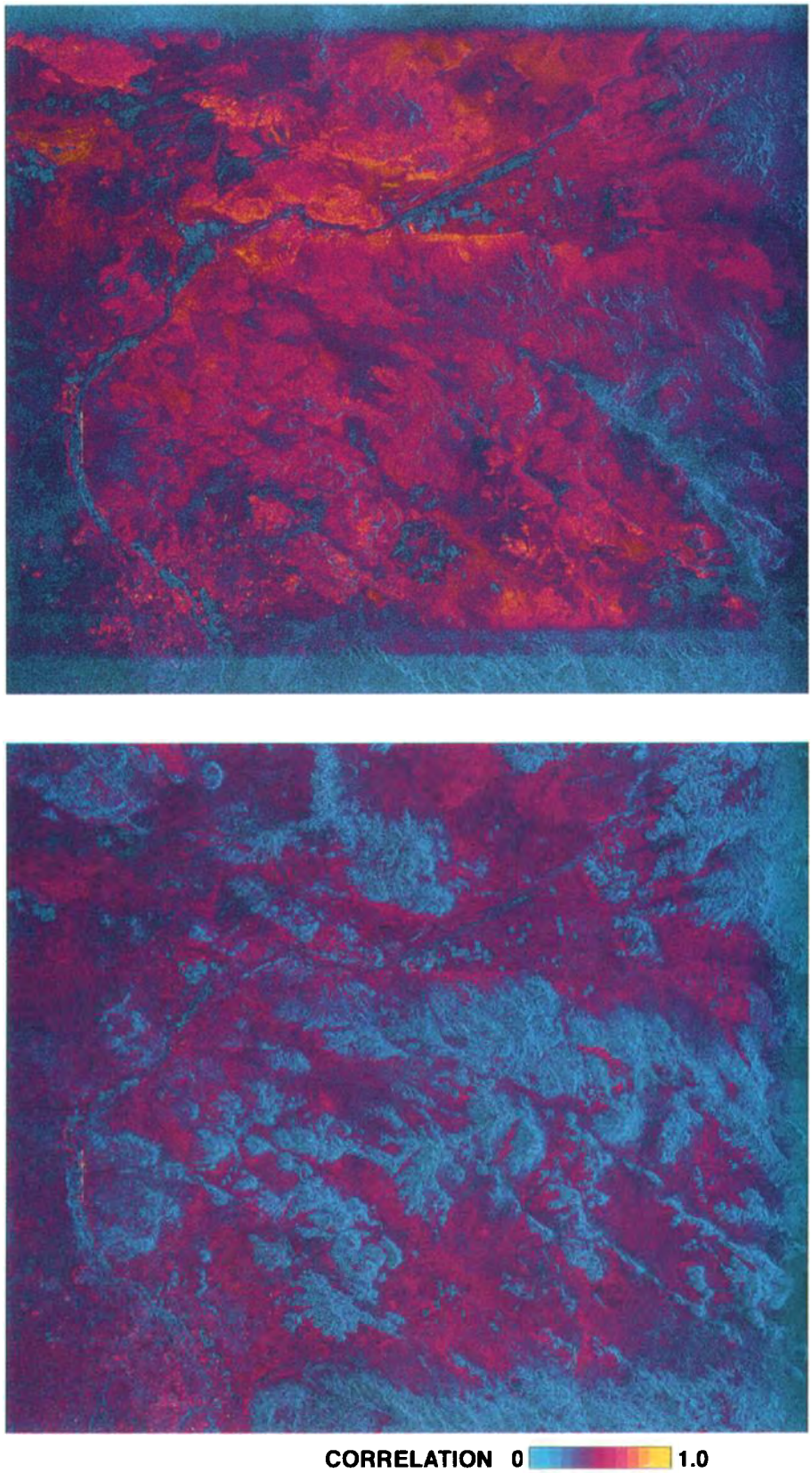


Plate 2. Corresponding correlation coefficient maps to interferograms of Plate 1. (top) April–August; (bottom) July–August. Note in the April–August pair a loss of correlation in the fault zone, presumably due to (1) very high fringe rates of greater than one cycle per resolution element, (2) large ground shifts resulting in lack of precision alignment of the pixels from pass to pass, and (3) stirring up of the surface at the wavelength scale from the earthquake itself.

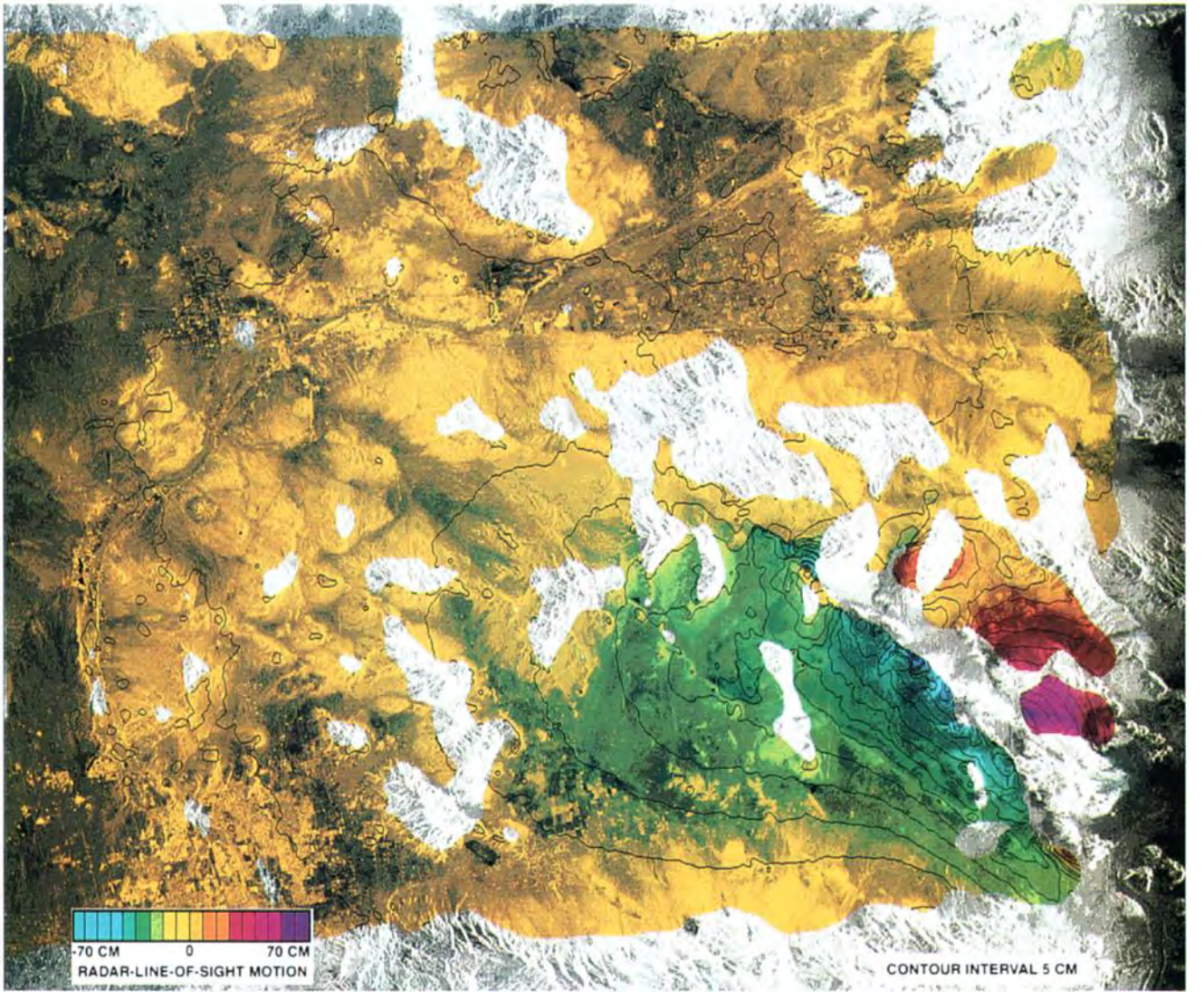
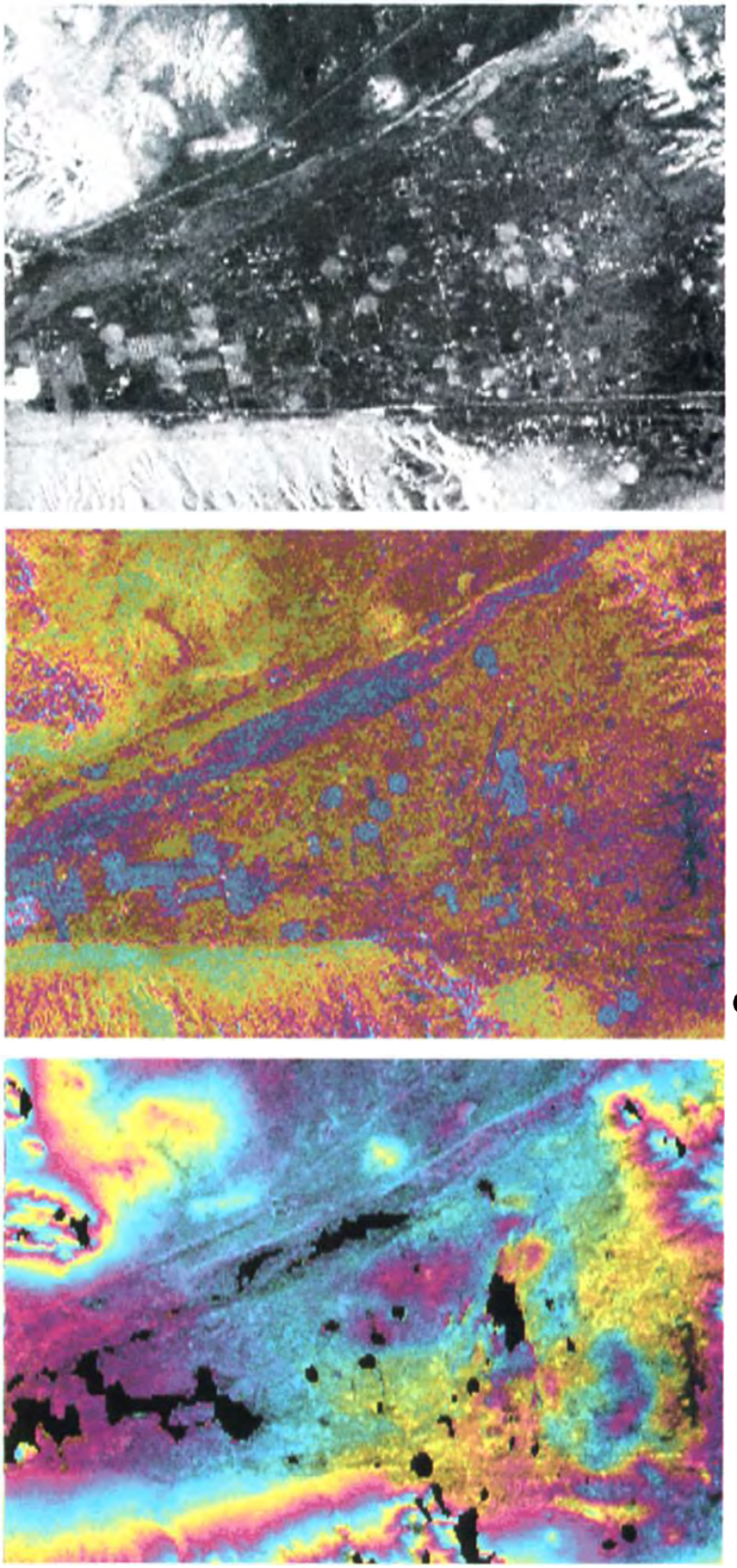


Plate 3. Differential interferogram of the Landers earthquake region. Radar line of sight displacements are coded in color, ranging from −70 to 70 cm, while the radar reflectivity of the surface is shown as brightness. Contours indicating each 5 cm of displacement are drawn in black.

tion presumably due to crop growth and phase shifts which are due to motion, not topography (it is a flat area). *Gabriel et al.* [1989] found similar surface displacements of several centimeters in fields that had been irrigated over a 9-day period. The motions observed in this image as well could be caused by changes in the surface elevations from pumping underground water or other hydrologic effects.

Since one of the strengths of this technique is its intrinsic high spatial resolution, we also show in Plate 5 an enlargement in the April–August interferogram of the region around the fault zone shown in Figure 2b. The phases in an interferogram are not unwrapped and so should not exhibit discontinuities except in regions of severe layover (where the surface slopes are greater than or equal to the radar incidence angle, resulting in severe image distortion) unless spatially discontinuous motions (breaks) occurred during the period spanned by the interferogram pair. Nevertheless, Plate 5 shows clear discontinuities in relatively flat areas. For example, the region denoted A in Plate 5 shows a clear break in the phase measurements. A similar break does not occur in the July–August pair, and therefore the April–August discontinuity must be due to a displacement of the surface where one piece moved more than the other. This cracking effect is more pronounced in the region denoted B, shown enlarged again in Plate 6, where the cracking is so extensive that it seems the ground has been broken into many tiles each several hundred meters across. These data are shown in unwrapped form. The phase unwrapping algorithm we use must identify phase discontinuities before calculating the absolute phase values; the locations of cuts determined automatically by our algorithm are shown in black. Presumably, these phase discontinuities are representations of centimeter-scale displacement discontinuities on the surface resulting from the earthquake. It would be an interesting field exercise to compare the computer generated cuts with any visible surface scars.

We also present in Plate 7 a perspective view of the entire area shown in Figure 2b where the vertical scale is proportional to the displacement in the radar line of sight of the surface. As usual, the brightness at each point is related to


1
0
CORRELATION
4π
2π
0
PHASE

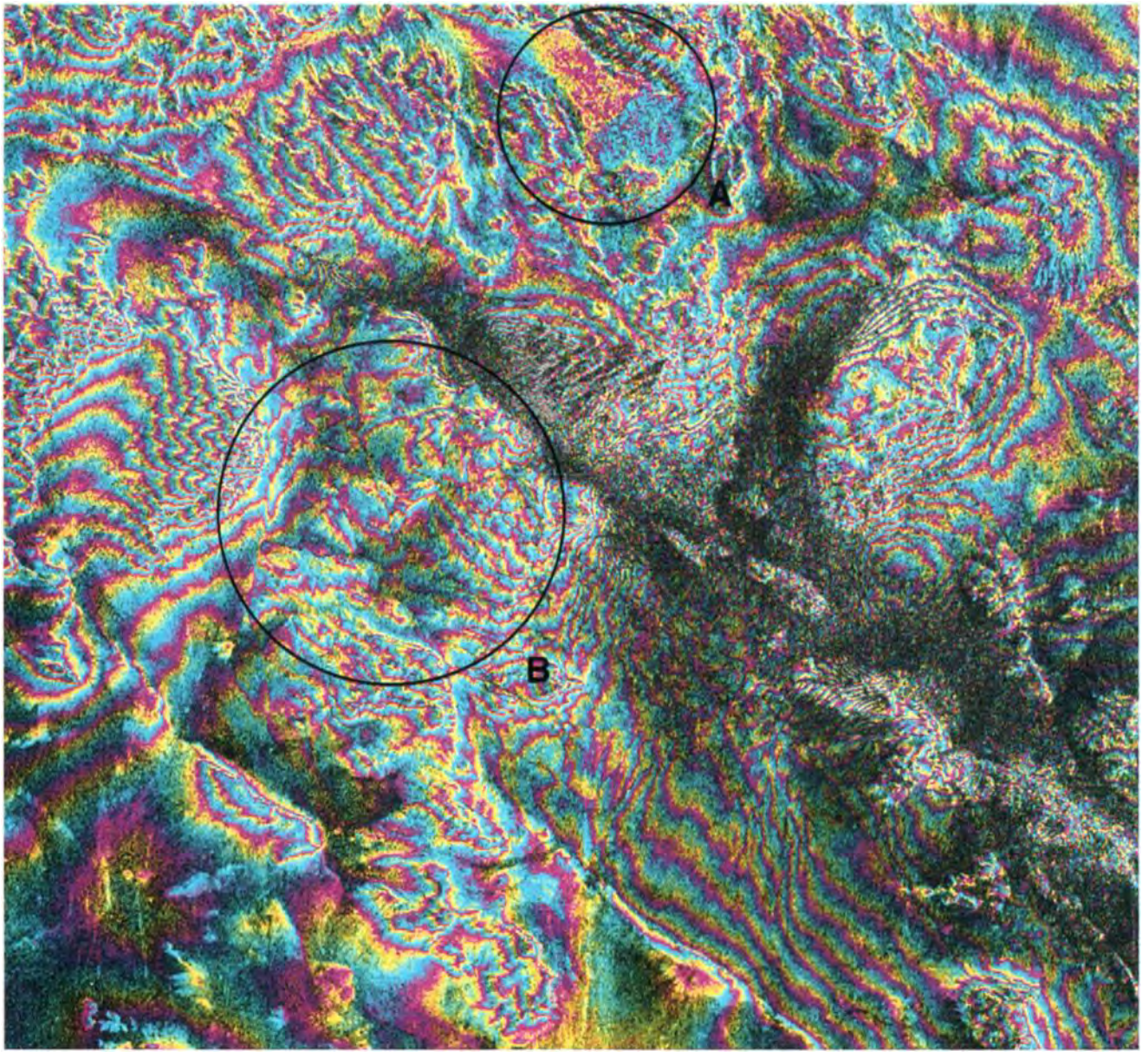


Plate 5. Enlargement in the April–August interferogram of a region around the fault zone. The region denoted A shows a clear break in the phase measurements, which must be due to a displacement of the surface where one piece was displaced more than the other. More extensive cracking is found in region B (see Plate 6).

Plate 4. (opposite) Enlargement of the April–August interferogram plus the correlation coefficients for a region east of Barstow where center pivot irrigation has been employed. (top) Radar reflectivity, (middle) correlation coefficient, and (bottom) unwrapped interferogram. Black spots in the lower image are where correlation was insufficient for reliable phase estimates. The irrigated circles show a clear loss of correlation, presumably due to crop growth, and phase shifts which are due to motion, not topography. Examination of the mountains at the top left of the image shows that a topographic change of over 150 m is necessary to cause a one cycle change in phase, and the area in question shows less than 30 m topographic variation. These phase changes could be caused by changes in the surface elevations from pumping underground water or other hydrologic effects.

radar reflectivity, while the color is the displacement mapped into a repeating color table to accentuate the visibility of the changes to produce a contour-like map. From this view one can see that the displacement increases as the fault is approached at which point there is an abrupt break in the surface; from this point hence the surface displacement is of opposite sign.

We assess the internal consistency and accuracy of the measurements presented here by three separate calculations. First, we calculate the expected errors due to statistical variation of the phase estimates. Assuming a radar signal to noise ratio of 6 dB for the flat desert surfaces, our 20 equivalent look processing (20 resolution elements are spatially averaged to reduce statistical noise) yields a stan-

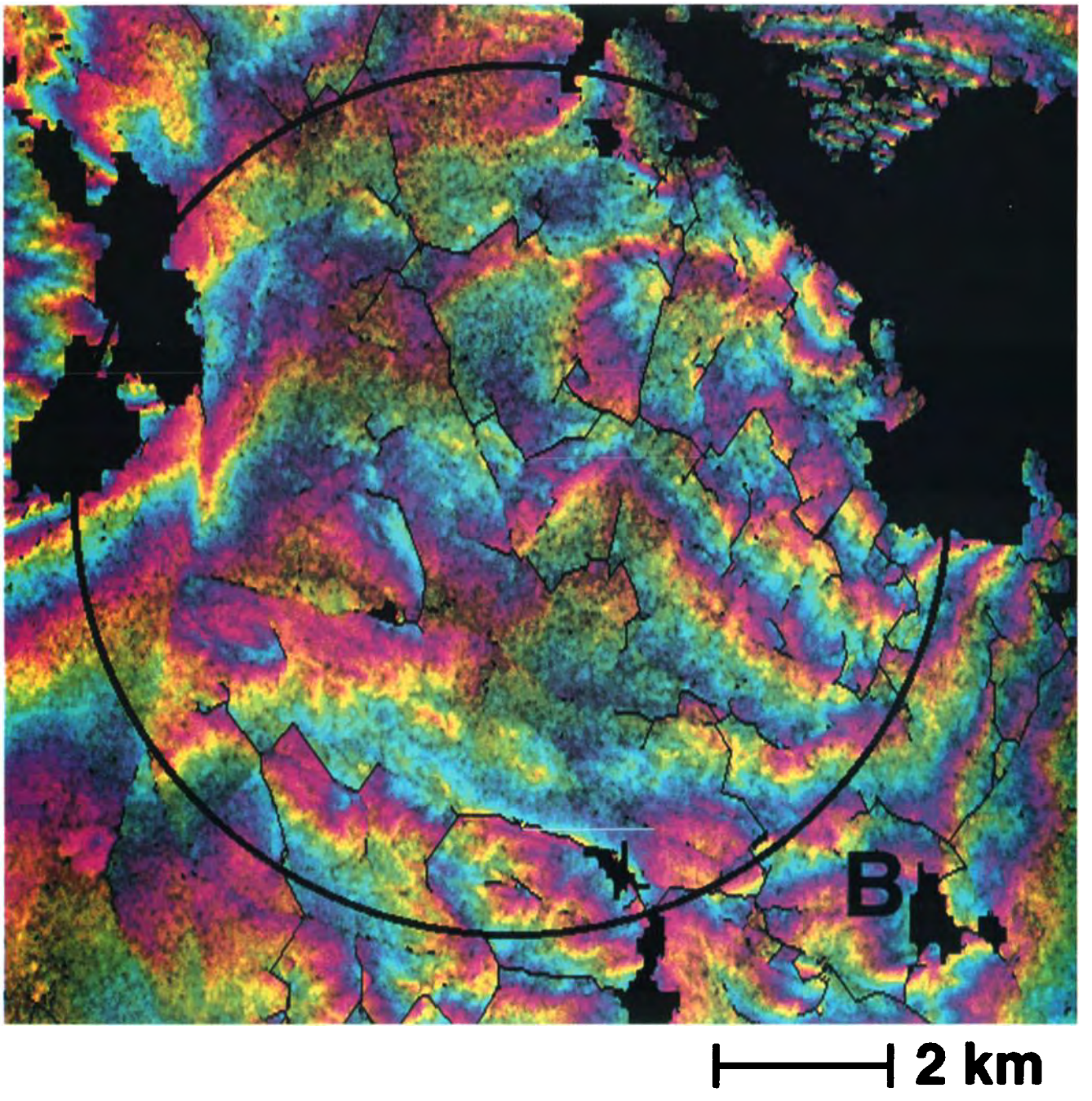


Plate 6. Region B from Plate 5, showing phase data in unwrapped form. The phase unwrapping algorithm we use identifies phase discontinuities before calculating the absolute phase values, and the locations of cuts determined automatically by our algorithm are shown in black. These phase jumps likely correspond to ground discontinuities at the cm level that appeared between April and July 1992, probably coincident in time with the earthquake. Cracking is so extensive that it seems the ground has been broken into many tiles each several hundred meters across.

dard deviation of 9.5° in the phase for the geometry of the April–August interferogram and 14.5° for the July–August interferogram; these values follow from using a target radar cross section of −17 dB and accounting for losses accruing from illuminating the ground off the boresight of the antenna. Combining these yields an expected phase error of 10° rms for the differential interferogram, equivalent to a horizontal displacement noise due to finite signal to noise ratio and baseline decorrelation of 0.2 cm. We would expect this value to be an underestimate as it does not take into account any temporal decorrelation due to surface disturbances or additional processing artifacts such as misregistration or other sampling and interpolation errors.

Second, we empirically determined statistical variations by measuring the observed phase standard deviations and converting the result to horizontal displacement errors. Choosing boxes corresponding to about 400 m by 400 m of the surface in areas of little seismic variation yielded an

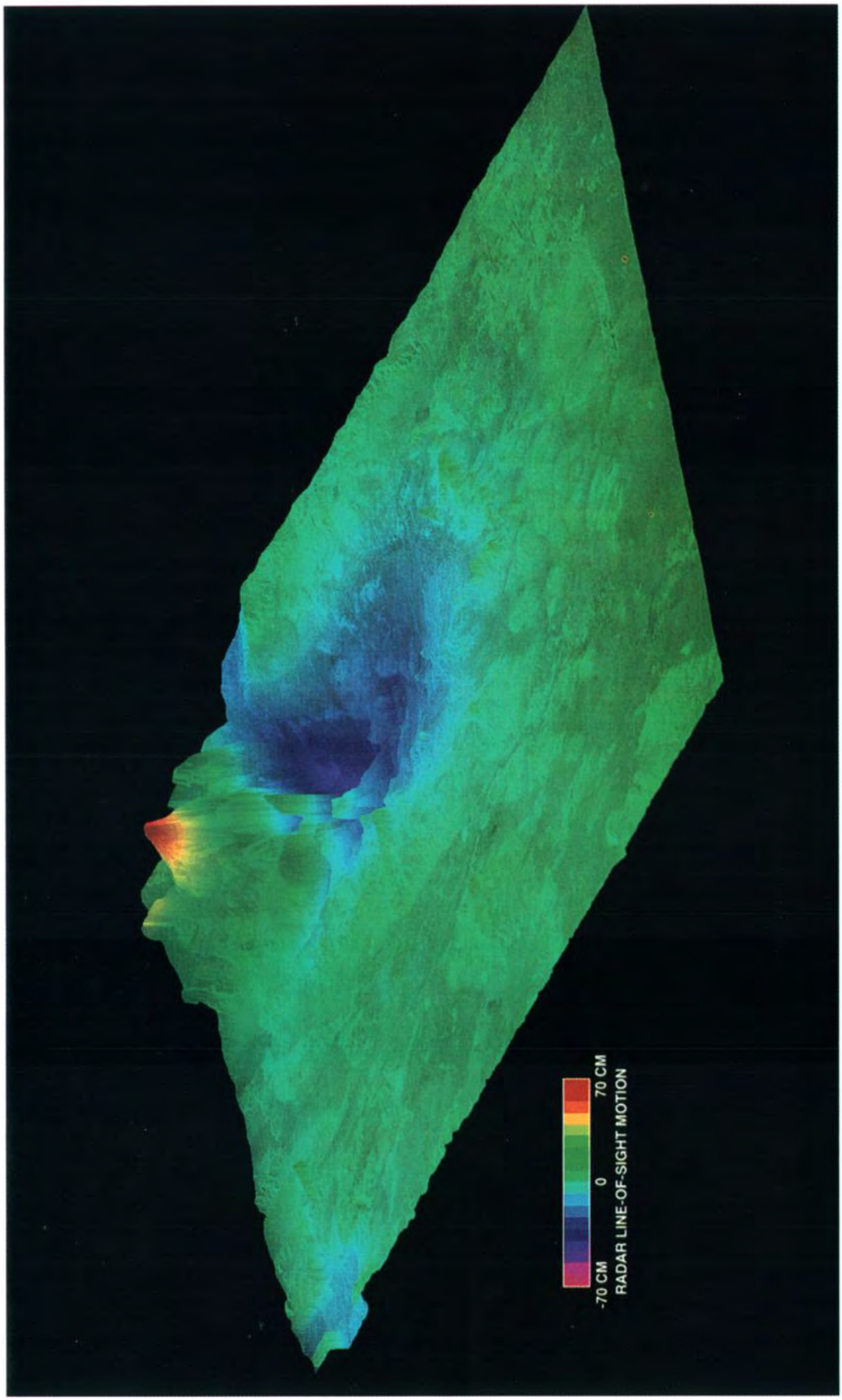


Plate 7. Perspective view of the region where the vertical scale is proportional to the surface displacement in the radar line of sight. The brightness at each point is related to radar reflectivity, while the color is the displacement. The color repeat interval is 140 cm. The displacement increases as the fault is approached at which point there is an abrupt break in the surface.

Table 2. Comparison of Radar and GPS Motion Estimates

Site	Latitude, deg	Longitude, deg	Horizontal Displacement for Observed Radar Motion, cm	GPS Vector in Radar Direction, cm	Difference, cm
6052	34.52	−116.84	47.8	33.2	14.6
6056	34.37	−116.65	18.1	21.9	−3.8
7000	34.68	−116.72	36.8	91.1	−54.3
7001	34.56	−116.47	−37.1	−70.2	33.1
HECT	34.79	−116.42	9.7	−5.2	14.9
LAZY	34.34	−116.51	62.9	49.4	13.5
LUCS	34.44	−116.88	26.4	20.7	5.7
POIN	34.45	−117.07	13.4	9.5	3.9
SOAP	34.90	−116.98	12.3	1.7	10.6
STIM	34.54	−117.24	7.8	7.4	0.4
FLASH	34.82	−117.02	14.1	12.1	2.0
HARVARD	34.94	−116.67	7.5	−0.4	7.9
BOULDER	34.51	−116.56	176.1	210.8	−34.7
FRY	34.50	−116.72	66.0	74.6	−8.6
MEANS	34.41	−116.55	82.2	69.8	12.4
OLD WOMN	34.39	−116.75	25.0	17.1	7.9
ORD	34.68	−116.81	44.4	48.3	−3.9
ROCK	34.54	−116.77	63.9	69.1	−5.2

average horizontal displacement of 0.4 cm rms for the high-frequency component of variations.

Finally, we attempted to address larger-scale variations by measuring the displacement at 10 widely separated locations far from the fault, and we determined their standard deviation. In this case the boxes were separated by 10 km or so, so that sensitivity to larger-scale variations would dominate. This calculation gave a horizontal displacement error component of 0.6 cm rms for these medium frequency variations, where medium frequency here refers to irregularities occurring with a spatial frequency of several cycles across the radar image.

Comparison With Field Measurements

In this section we discuss the accuracy of our measurements and compare the results to those obtained in the field using Global Positioning Satellite (GPS) and electronic distance measurement (EDM) survey data. As a basis of comparison we will use the coseismic displacement field solution as derived by J. Freymueller et al. (personal communication, 1993), data which were compiled by K. W. Hudnut et al. (personal communication, 1993). Hudnut et al. also analyzed these data and obtained a slightly different, but consistent solution. These calculated displacements were derived from a combination of GPS data from several sources and EDM line lengths obtained by the USGS (please see the above references for a more detailed description of the data sources and techniques).

The area of overlap between the field survey and our image contains 18 points at which both field data and radar estimates of the motion are available. Three additional site measurements of field data exist in the overlap region, but we were not able to obtain reliable radar phase estimates for them (they occur in the gray regions of Plate 3). As can be seen from Plate 3, however, the radar data are generally valid over a wide area and should future surveys or analyses produce additional field points, they may be easily compared with the present analysis.

As stated previously, the radar technique is sensitive to the line of sight component of motion. We therefore calculated the component of the GPS motion vectors in the direction of the projection on the ground of the radar sensor boresight, the vector from the sensor to a point on the Earth's surface. As for the radar measurements, since the line of sight direction is not in the plane defined by the local Earth surface, we derived the equivalent horizontal surface motion to yield the observed slant range displacement using

$$\Delta y = \frac{\Delta \rho}{\sin \theta_{\text{inc}}} \tag{17}$$

which relates the horizontal displacement Δy to slant range displacement $\Delta \rho$ and the incidence angle θ_{inc}. This angle is equal to the look angle (denoted θ in Figure 1) for a flat Earth approximation, and is approximately equal for a curved Earth model. We use a curved Earth model for its improved accuracy. The results of both of these calculations are shown in Table 2 and Figures 3 and 4.

Because the orbit of the ERS 1 satellite is known only approximately, as discussed above there are residual tilts in the derived radar displacement field. Therefore we have removed this distortion by solving, in a least squares sense, for the planar tilt that minimizes disagreement between the radar and GPS/EDM measurements. This nicely illustrates one aspect of the complementary nature of the two techniques for analyzing ground motions: the radar measures a widespread displacement field while the GPS/EDM data provide accurate point measurements which are used to refine the radar estimates.

The mean value of the differences in Table 2 is 0.9 cm, and the rms difference is 18.9 cm. The formal correlation of the data is 0.96, which we illustrate in Figure 3, a scatter plot comparing the radar and GPS/EDM measurements. Note that the best fit through the data evidences a slight bias.

Figure 4 shows the same data of Table 2 presented graphically. For each survey site, denoted by a triangle, we illustrate vectors corresponding to motion as determined by

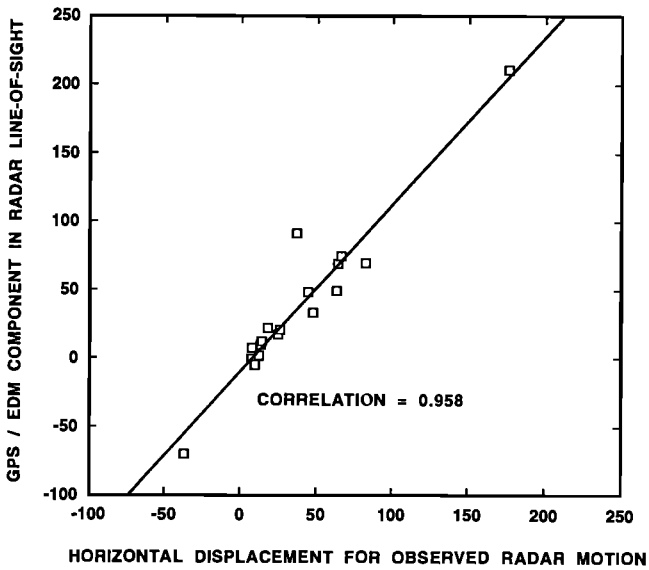


Figure 3. Scatter plot of displacement measurements with GPS/EDM data on vertical axis and radar measurements on horizontal axis. The correlation of the data sets is 0.96, however a slight bias is observed as the slope of the line is not 1.

survey techniques (diamond-headed arrows) and as determined by the radar (cross-headed arrows). Note that the radar vectors are all parallel to the edge of the radar image, as only the component of motion in the line of sight is measured.

From each of these presentations it is apparent that at most sites, with significant exception of sites 7000 and 7001, the measurements are in rough agreement. The absolute disagreement is also large at BOULDER, but the motion here is quite large, and on a relative scale the agreement is comparable to the values for the remaining sites. It is interesting to note that in the deviant cases a large motion is observed by the GPS technique, while a smaller displacement is visible by the radar technique. In each case where a small motion is detected by the field survey, a small motion is measured by the radar interferometer. Figure 4 also suggests that there is a degree of spatial correlation in the regions of agreement, that is, the amount of agreement is spatially dependent.

There are several possible causes for the disagreements in the measurements. First, the radar technique is highly sensitive to vertical motions which are not expressed in the GPS displacement field. While this is likely to affect the differences on the centimeter scale, it is probably not a significant

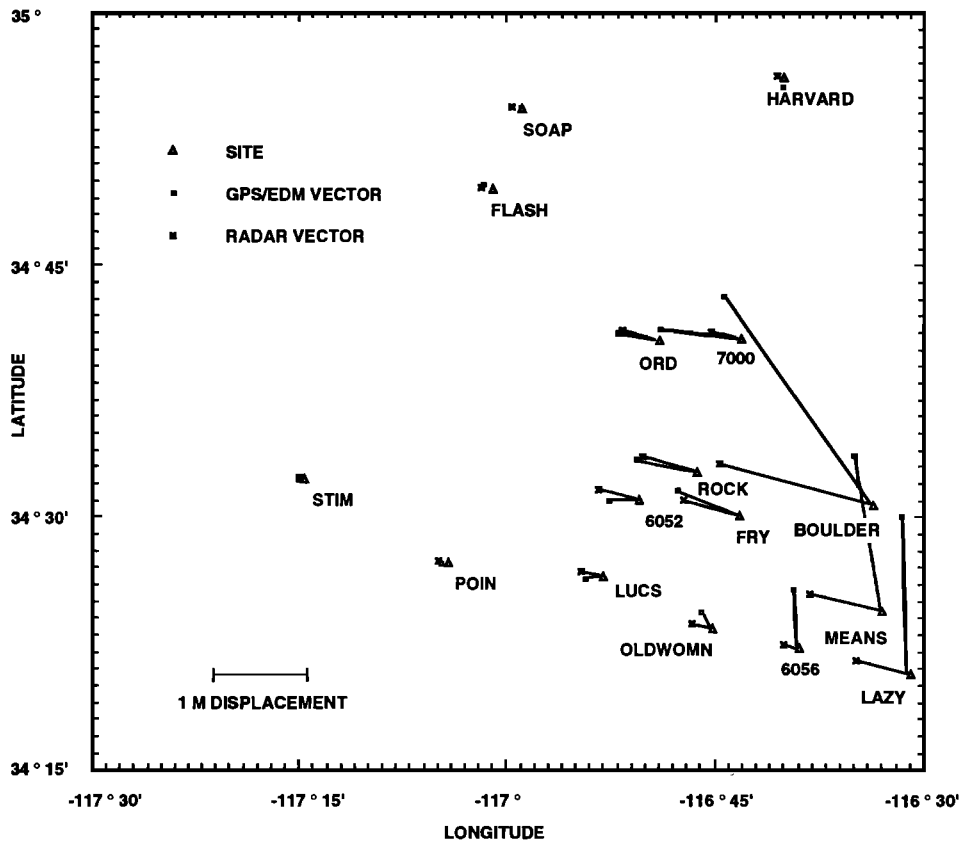


Figure 4. Displacement vectors as measured by GPS/EDM data and by radar interferometry. Each GPS or EDM site is denoted by a triangle, and a vector ending with a square (GPS/EDM measurement) and a vector ending with a cross (radar measurement) are shown in the direction of motion. Note that for the radar case only the component in the radar line of sight direction is determined and thus all measurements are parallel. Vectors are correlated at 0.96 level and show that radar and field surveys are measuring similar phenomena.

factor in the radar underestimation of the motions. This follows from the unlikelihood that vertical motions would just happen to be in the direction with respect to the radar to cancel out any horizontal shifts.

The GPS sites, particularly the dual-frequency sites, in fact yield vertical components to the displacement. The rms vertical displacement for the GPS sites is 17.2 cm, but most of this is associated with site 7000; when this site is removed, the remaining rms displacement is only 5.1 cm. Ignoring vertical displacements, as was done in our analysis, results in a misinterpretation of lateral shift of magnitude equal to the true vertical movement divided by the tangent of the incidence angle. The errors in the above cases then become 40.5 cm and 12.0 cm, respectively. However, we must note that for nine of the 10 sites the $1 - \sigma$ error in the vertical displacement is larger than that of the measurement itself, so these data must not be overly interpreted.

The second cause for disagreement is error in the measurements. As discussed above, the radar data exhibit statistical errors less than 1 cm rms on both small and medium scales and thus would be insignificant for this comparison. However, large-scale warping of the radar image remains a possibility. We were able to remove most of these effects by minimizing the errors with the least squares removal of planar tilts as described previously. That this correction was approximately correct may be verified by examining the residual motion in the upper and left hand portions of the radar image, those portions farthest from the fault. The observed motion here is very small, as we would expect. If the ERS 1 coverage had been such that the fault was positioned in the center of the radar swath, we could have verified the lack of displacement more accurately all the way around the image. The possibility of a long-scale error thus still exists and may to some degree explain the observed spatial correlation of the errors.

The errors in the GPS/EDM data themselves account for part of the disagreement. The 18 sites listed in Table 2 exhibit an rms error of 9.1 cm in the Stanford analysis, while the 10 GPS-only sites have a 7.7-cm rms error of J. W. Hudnut et al. (personal communication, 1993).

A third possibility is the existence of phase unwrapping errors in the radar data. As each unwrap error results in a one cycle phase error in one interferogram, these errors would appear as $\lambda/2$ errors in $\Delta\rho$, or 8 cm in horizontal shift if it occurred in the April–August pair or 2 cm in the July–August pair. However, we have examined the data for signs of unwrapping errors and believe that the regions near the GPS sites are unwrapped correctly. In addition, it is unlikely again that phase unwrapping mistakes would nearly correct for GPS-observed displacements.

Finally, the locations of the GPS sites are known only to a few tens of meters in the radar image as the radar data are not accurately geocoded, thus leading to estimates at the wrong places. However, we have analyzed the regions around the sites in the radar data and have determined that the displacement does not change rapidly in those areas. Thus even a slight positional shift would not result in a significant error.

Discussion

We have shown that it is possible to map a coseismic displacement field resulting from a major earthquake using only data acquired from an orbiting high-resolution radar system and to achieve results comparable in magnitude to those obtained by conventional field survey techniques. Data from the ERS 1 synthetic aperture radar instrument acquired at three separate instances of time are sufficient to generate a high-resolution, wide area map of the displacements. Comparison of these data with GPS and EDM survey data indicates a high degree of confidence in the radar measurements. We are confident that the differences between the radar and GPS measurements are reconcilable and do not point to a fundamental limitation in the radar technique. Further work is needed along these lines however.

The power of the differential interferometry technique for seismological applications lies in its centimeter-scale measurement sensitivity of line of site displacements over a wide area. The derived displacement fields can be used as a tight constraint in the modeling of earthquake motion. The fine accuracy, fine spatial resolution, and large areal coverage will likely allow increasingly detailed models to be explored, on both large and small spatial scales. The promise of a system to map small-scale fractures in the Earth's surface over a wide region automatically with a remote sensing system will greatly facilitate field activities by permitting concentration in the most important areas.

What of earthquake prediction? Current understanding of the behavior of earthquakes suggests that differential interferometry may not have the accuracy required to detect precursory seismic motions necessary for prediction. Calculations based on theoretical seismic deformation models show small but steady deformation rates in fault zones with a change in the rates occurring within a period of months to years prior to a seismic event [*Lorenzetti and Tullis*, 1989; *Stuart et al.*, 1985]. While the steady deformations have been observed and are well studied, no precursory rate changes have been measured. Furthermore, the steady rates themselves are probably at or below the limits of detectability by differential interferometry, perhaps 10 mm/yr, while the precursory signal is expected to be smaller. Wide area mapping of the surface distribution of these small deformation rates afforded by differential interferometry may provide new insights into local accumulation of strain close to and along a fault, but the possibility that radar interferometry can be used as a predictive tool now appears to be remote. This is not to say that likely future technological advances in spaceborne radar such as higher-resolution, increased signal to noise ratio, and multiple frequency operation will not close this gap and permit the sensitivities required for millimeter-level surface characterization. Nonetheless, in addition to after the fact seismic event modeling, currently radar interferometry can aid in monitoring, forecasting, and in some cases predicting a range of hazardous events. For example, volcanoes are known to bulge prior to eruption at a scale suitable for radar interferometry.

In the short run, existing and planned radar missions such NASA's Shuttle Imaging Radar, the European Space Agency's ERS 2, Canada's Radarsat, and the Japanese JERS 1 system could be operated to emphasize repeat-pass observations at the largest acceptable incidence angles, providing a very large suite of instruments collecting data that may be processed for change detection analysis. For the future one can envision a global seismic satellite mission designed to detect and forecast earthquakes and other natural hazards: a single satellite in a short repeat period orbit similar in design

to that proposed by H. A. Zebker et al. (personal communication, 1993) for global topographic mapping. The repeat cycle of the orbit should be short, of the order of 1 day, to minimize the effects of temporal decorrelation. Precise satellite ephemeris from GPS measurements can ensure automatic construction of interferograms and displacement fields. Only three repeat periods of data need to be stored at any time; the processing can proceed in real time, and results can be perused automatically for evidence of anomalous displacements. Detailed design of the radar system and orbital scenario plus the establishment of detection and false alarm thresholds, must await interest by the global community. Given the enormous cost in lives and resources inflicted by earthquakes, interest is sure to follow any evidence that radar interferometry can be used predictively in assessing natural hazards such as earthquakes.

Appendix: Baseline-Induced Displacement Errors

Equation (13) is the displacement determined from the flattened interferometric phase assuming perfect knowledge of the baselines. Reiterating,

$$\Delta\rho = \frac{\lambda}{4\pi}\left[\phi'_{\text{flat}} - \gamma\phi_{\text{flat}}\right], \tag{A1}$$

where

$$\frac{\lambda}{4\pi}\phi_{\text{flat}} = B\sin(\theta - \alpha) - B\sin(\theta_0 - \alpha), \tag{A2}$$

$$\gamma = \frac{B'\cos(\theta_0 - \alpha')}{B\cos(\theta_0 - \alpha)}. \tag{A3}$$

With imperfect knowledge of the baselines $\hat{B}$ and $\hat{\alpha}$, errors are introduced in both the phase, denoted $\hat{\phi}_{\text{flat}}$, and the scale factor, denoted $\hat{\gamma}$. Defining

$$\hat{B} = B + \delta B \tag{A4}$$

$$\hat{\alpha} = \alpha + \delta\alpha, \tag{A5}$$

we have to first order

$$\frac{\lambda}{4\pi}\hat{\phi}_{\text{flat}} = B\sin(\theta - \alpha) - \hat{B}\sin(\theta_0 - \hat{\alpha}) = \frac{\lambda}{4\pi}\phi_{\text{flat}} - \delta B\sin(\theta - \alpha) - \delta\alpha B\cos(\theta_0 - \alpha) \tag{A6}$$

and

$$\hat{\gamma} = \frac{\hat{B}'\cos(\theta_0 - \hat{\alpha}')}{\hat{B}\cos(\theta_0 - \hat{\alpha})} = \frac{B'\cos(\theta_0 - \alpha') + \delta B'\cos(\theta_0 - \alpha') + \delta\alpha' B'\sin(\theta_0 - \alpha')}{B\cos(\theta_0 - \alpha) + \delta B\cos(\theta_0 - \alpha) + \delta\alpha B\sin(\theta_0 - \alpha)}. \tag{A7}$$

Note that for $\hat{\gamma}$, baseline length error and angle error are complementary: baseline length error is weighted highly when the baseline is orthogonal to the look direction ($\theta_0 - \alpha_i = 0$), whereas angle error is weighted highly with the baseline aligned with the look direction. For baselines that are not nearly aligned with the look direction, the ratio in (A7) may be expanded to give

$$\hat{\gamma} = \gamma\left[1 + \frac{\delta B'}{B'} - \frac{\delta B}{B} + \delta\alpha'\tan(\theta_0 - \alpha') - \delta\alpha\tan(\theta_0 - \alpha)\right]. \tag{A8}$$

Using (A6) and (A8) to evaluate the displacement gives

$$\begin{aligned}\Delta\hat{\rho} &= \frac{\lambda}{4\pi}\left[\hat{\phi}'_{\text{flat}} - \hat{\gamma}\hat{\phi}_{\text{flat}}\right] \\ &= \Delta\rho - \delta B'\sin(\theta_0 - \alpha') + \delta\alpha' B'\cos(\theta_0 - \alpha') \\ &\quad - \gamma\left[-\delta B'\sin(\theta_0 - \alpha') + \delta\alpha' B'\cos(\theta_0 - \alpha')\right] \\ &\quad - \gamma\frac{\lambda}{4\pi}\phi_{\text{flat}}\left[\frac{\delta B'}{B'} - \frac{\delta B}{B} + \delta\alpha'\tan(\theta_0 - \alpha') - \delta\alpha\tan(\theta_0 - \alpha)\right]\end{aligned} \tag{A9}$$

Equation (A9) shows that in addition to the desired term $\Delta\rho$, there are slowly varying (fraction of a cycle) sinusoidal artifacts across the displacement field and topographic residuals dependent on baseline length and angle errors. Even if the slowly varying artifacts are removed empirically, accurate estimates of the displacements, to fractions of a cycle, require fairly accurate baseline knowledge. We can estimate the scale of the topographic term as follows. Assume $B \approx B'$ and likewise for the uncertainties δB and $\delta\alpha$. Then the final term in (A9) becomes

$$\Delta\hat{\rho}_{\text{topo}} \sim \frac{\lambda}{4\pi}\phi_{\text{flat}}\frac{\delta B_{\text{net}}}{B_{\text{net}}} \tag{A10}$$

where we have assumed a worse case $\gamma = 1$ and an equivalent net baseline error δB_{net} including δB and $\delta\alpha B$. Expanding $\phi_{\text{flat}}(\theta)$ about θ_0,

$$\Delta\hat{\rho}_{\text{topo}} \sim \delta B_{\text{net}}\delta\theta,$$

where $\delta\theta = z/\rho$ is the angular deviation of the look direction due to topography. Thus, to limit displacement errors due to residual topography, $\Delta\hat{\rho}_{\text{topo}}$, to say 1/4 wavelength, the error in δB_{net} must satisfy the inequality

$$\delta B_{\text{net}} < 0.25\lambda\rho/z_{\max}$$

where $z_{\max}$ is the maximum topographic extent over the scene. For ρ = 800,000 m, λ = 0.0566 m, $z_{\max}$ = 5000 m, $\delta B_{\text{net}} < 2.3$ m.

Acknowledgments. We would like to acknowledge Paul Segall for supplying the GPS/EDM measurements and for several useful discussions regarding the intercomparison of the data sets. We would also like to acknowledge discussions with Ken Hudnut for discussions prompting a reexamination of the comparison of our results with the GPS/EDM data. The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

Curlander, J. C., and R. N. McDonough, *Synthetic Aperture Radar*, Wiley-Interscience, New York, 1991.

Elachi, C., *Spaceborne Radar Remote Sensing: Applications and Techniques*, IEEE Press, New York, 1987.

Evans, D. L., T. G. Farr, H. A. Zebker, J. J. van Zyl, and P. J. Mouginis-Mark, Radar interferometric studies of the Earth's topography, *Eos Trans. AGU*, *73*, 553, 557–558, 1992.

Gabriel, A. G., R. M. Goldstein, and H. A. Zebker, Mapping small elevation changes over large areas: Differential radar interferometry, *J. Geophys. Res.*, *94*, 9183–9191, 1989.

Ghiglia, D. C., and L. Romero, Robust two-dimensional weighted and unweighted phase unwrapping using fast transforms and iterative methods, *J. Opt. Soc. Am.*, *11*, 107–117, 1993.

Goldstein, R. M., and H. A. Zebker, Interferometric radar measurement of ocean surface currents, *Nature*, *328*, 707–709, 1987.

Goldstein, R. M., H. A. Zebker, and C. L. Werner, Satellite radar interferometry: two dimensional phase unwrapping, *Radio Sci.*, *23*, 713–720, 1988.

Goldstein, R. M., T. P. Barnett, and H. A. Zebker, Remote sensing of ocean currents, *Science*, *246*, 1282–1285, 1989.

Gray, A. Laurence, and P. J. Farris-Manning, Two-pass interferometry with airborne synthetic aperture radar, *IEEE Trans. Geosci. Remote Sens.*, *31*, 180–191, 1993.

Li, F., and R. M. Goldstein, Studies of multi-baseline spaceborne interferometric synthetic aperture radars, *IEEE Trans. Geosci. Remote Sens.*, *28*, 88–97, 1990.

Lorenzetti, E., and T. E. Tullis, Geodetic predictions of a strike-slip fault model: Implications for intermediate and short-term earthquake prediction, *J. Geophys. Res.*, *94*, 12,343–12,361, 1989.

Madsen, S. N., H. A. Zebker, and J. Martin, Topographic mapping using radar interferometry: Processing techniques, *IEEE Trans. Geosci. Remote Sens.*, *31*, 246–256, 1993.

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

Prati, C., F. Rocca, A. Monti Guarnieri, and E. Damonti, Seismic migration for SAR focusing: Interferometrical applications, *IEEE Trans. Geosci. Rem. Sens.*, *28*, 627–640, 1990.

Stuart, W. D., R. J. Archuleta, and A. G. Lindh, Forecast model for moderate earthquakes near Parkfield, California, *J. Geophys. Res.*, *90*, 592–604, 1985.

Zebker, H. A., and R. M. Goldstein, Topographic mapping derived from synthetic aperture radar measurements, *J. Geophys. Res.*, *91*, 4993–4999, 1986.

Zebker, H. A., and J. Villasenor, Decorrelation in interferometric radar echoes, *IEEE Trans. Geosci. Remote Sens.*, *30*, 950–959, 1992.

Zebker, H. A., S. N. Madsen, J. Martin, K. B. Wheeler, T. Miller, Y. Lou, G. Alberti, S. Vetrella, and A. Cucci, The TOPSAR interferometric radar topographic mapping instrument, *IEEE Trans. Geosci. Rem. Sens.*, *30*, 933–940, 1992.

A. Gabriel, R. M. Goldstein, P. Rosen, C. L. Werner, and H. A. Zebker, MS 300-227, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109.

(Received December 6, 1993; revised April 21, 1994; accepted April 29, 1994.)

We've shown that no additional info on topography is needed to infer deformation if three radar passes are available. We simply calculation the ratio of the $B_{\perp}$'s and found the combined product

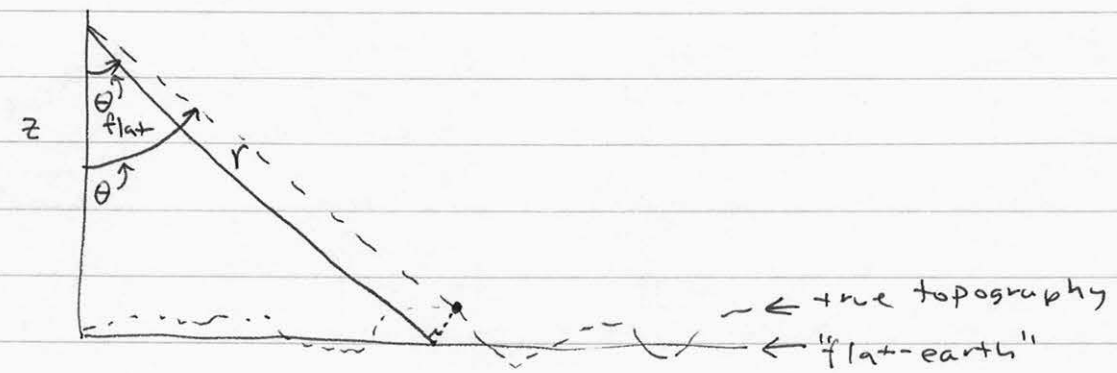
$$\phi' - \frac{B'_{\perp}}{B_{\perp}}\phi = -\frac{4\pi}{\lambda}\Delta\rho$$

Consider the baseline ratio in more detail:

$$\frac{B'_{\perp}}{B_{\perp}} = \frac{B' \sin(\theta - \alpha')}{B \sin(\theta - \alpha)}$$

Now, this ratio depends on θ, which in turn depends on topography. Once again we are limited by knowledge of the topography of a region. However, by making certain approximations it is possible to proceed quite nicely.

Let's correct the measured fringe patterns ϕ and ϕ' for the "flat-earth" correction, that is remove the phase pattern ~~xxx~~ which would result from an earth surface devoid of topography. In that case we can calculate the look angle exactly, denoting it by θ_{flat}:



Define as ϕ_{flat} the phase after "flattening", or correction by the expected flat-earth fringe pattern. Using the same baseline length and orientation we have

$$\phi_{flat} = -\frac{4\pi}{\lambda}\left[B\sin(\theta-\alpha) - B\sin(\theta_{flat}-\alpha)\right]$$

$B_{\parallel}$ assuming topography

$B_{\parallel}$ in the case of no topography

Since $\theta_{flat} \approx \theta$, we can simplify the above expression using

$$\theta = \theta_{flat} + \delta\theta$$

$$\sin(\theta-\alpha) = \sin(\underset{flat}{\theta-\alpha} + \delta\theta)$$

$$\approx \sin(\theta_{flat}-\alpha) + \cos(\theta_{flat}-\alpha)\,\delta\theta$$

so that

$$\phi_{flat} = \frac{-4\pi}{\lambda} B\cos(\theta_{flat}-\alpha)\,\delta\theta$$

Under these approximations the flattened phase is proportional now to the perpendicular component of the baseline and to the angular topographic distortion $\delta\theta$ directly. So now when we form the ratio

$$\frac{\phi'_{flat}}{\phi_{flat}} = \frac{\frac{-4\pi}{\lambda} B'\cos(\theta_{flat}-\alpha')\,\delta\theta}{\frac{-4\pi}{\lambda} B\cos(\theta_{flat}-\alpha)\,\delta\theta} = \frac{B'_{\perp}}{B_{\perp}}$$

which doesn't depend on topography at all.

Hence we can restate the differential phase equation in terms of the flattened phase

$$\phi'_{flat} - \frac{B'_\perp}{B_\perp}\phi_{flat} = -\frac{4\pi}{\lambda}\Delta\rho$$

and now there is no topographic dependence at all. Hence we scale flattened interferograms by the ratio of the perpendicular baselines rather than scaling unflattened interferograms by parallel components of the baselines.

Sensitivity of deformation maps

Now consider the sensitivity of our deformation measurements as compared to topographic measurements. Begin by determining the sensitivity of the phase measurement ϕ to height and deformation in our total phase equation

$$\phi = -\frac{4\pi}{\lambda}\left[B\sin(\theta-\alpha) + \Delta\rho\right]$$

We've done $\frac{d\phi}{dz}$ for the sensitivity many times now:

$$\frac{d\phi}{dz} = -\frac{4\pi B}{\lambda r}\frac{\cos(\theta-\alpha)}{\sin\theta}$$

Add to this

$$\frac{d\phi}{d\Delta\rho} = \frac{-4\pi}{\lambda}$$

Because usually $\frac{B}{r} << 1$, our interferogram will be very much

more sensitive to small changes in deformation rather than topography. We can illustrate this by looking at values for typical ERS geometries:

$r = 800$ km

$\theta = 23°$

$\alpha = 0°$

$B = 200$ m

$\lambda = 6$ cm

$$\frac{d\phi}{dz} = \frac{-4\pi \cdot 200}{0.06 \cdot 800 \times 10^3} \cdot \frac{\cos 23°}{\sin 23°}$$

$$= -0.123 \text{ rad/m} = -7°/\text{m}$$

Hence a 1 m topography signature results in 7° of interferometric phase, near the limit of detectability. Thus meter-scale topographic accuracy is possible.

How about for deformation?

$$\frac{d\phi}{d\Delta\rho} = \frac{-4\pi}{0.06}$$

$$= -209 \text{ rad/m} = 11,938°/\text{m}$$

A 1m deformation signature gives nearly 12000° of phase, about 30 cycles. Hence a very small amount of deformation is detectable with interferometric techniques.

This argument should help illustrate that deformation enters into the interferogram phase in a fundamentally different way than does topography. In the topographic case a position is scaled by a ratio of about $\frac{B}{r}$ before adding to the phase, while for deformation it enters directly.

Therefore deformation measurements are $\frac{r}{B}$ times more sensitive than simple differencing of topography at two different times would suggest.

A case example - Landers earthquake

We will illustrate the deformation measurement by looking in more detail at the magnitude 7.3 earthquake centered near Landers, CA, in June 1992. This was the first earthquake studied in detail by radar interferometry.

The data for this study were acquired by ERS-1 on April 24, July 3, and August 7, 1992. The earthquake occured on June 28.

We can form several interferograms from these data, consider one pair from the April and August images, and one from the July and August images.

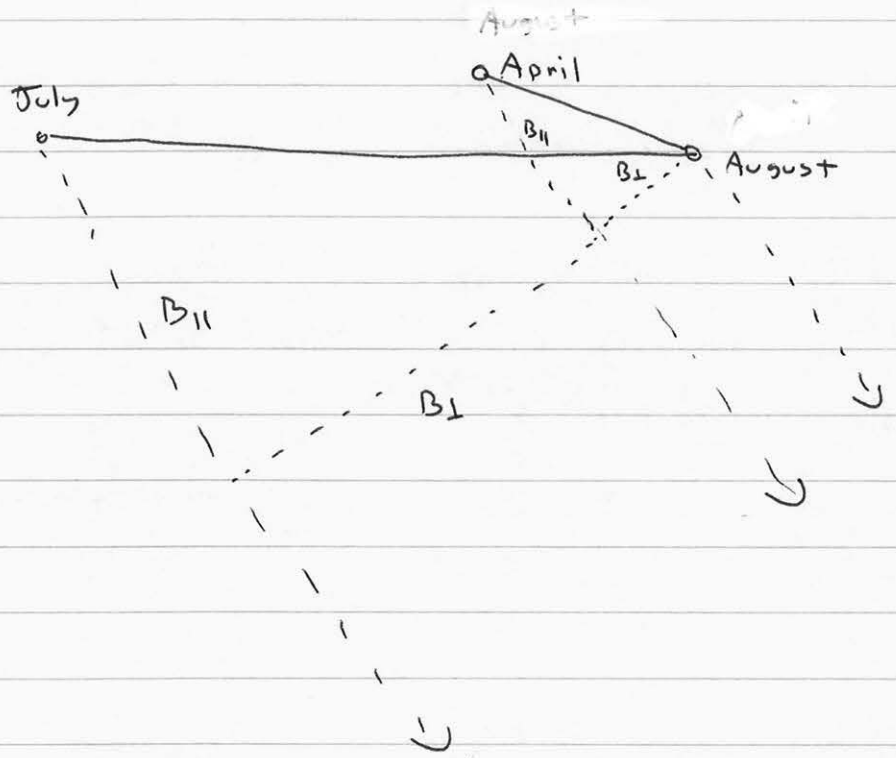
The July-August pair will form an interferogram dominated by topography since no other large motions occured during this period. But the April-August pair span the large earthquake, and we would expect a significant ϕ_{eq} term in our phase expression.

We could also use the April-July pair to form an earthquake-containing interferogram. However, this pair has a

very large baseline associated with it, and the high fringe rate makes it difficult to read the phase values accurately. Our baseline data are as follows:

Pair	B	α	$B_{\parallel}$	$B_{\perp}$
April - August	146	152	110	96
July - August	503	175	220	452

We can sketch the locations in the sky as follows:

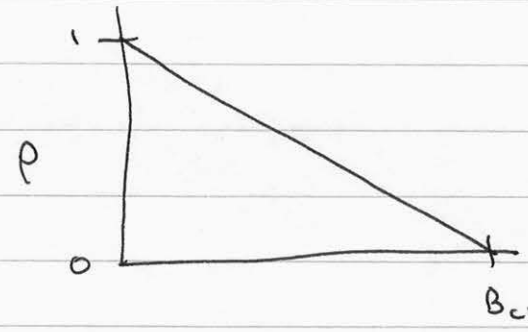


We are stuck with the long-baseline July-August pair since it is the only one with no deformation. We choose April-August over April-July because it has a shorter baseline.

When we examine these two interferograms, we expect a

much higher fringe rate for July-August than for April-August, and this is indeed what we see.

We also would expect a higher correlation in the shorter baseline pair. Recall how we modeled correlation as a function of baseline length:



It fell off linearly with baseline to a maximum baseline of $B_{critical}$. Evaluating $B_{critical}$ for ERS:

$$B_{critical} = \frac{\lambda r}{2\cos\theta\, dy}$$

$$= \frac{6\text{ cm} \cdot 800\text{ km}}{2 \cdot \cos 23^\circ \cdot 25\text{ m}}$$

$$\approx 1050\text{ m}$$

As $B_\perp$ for July-August is 452 m, the correlation ought to be decreased by about half, as we observe.

→ Viewgraphs in class depict these and the resulting 3-pass deformation interferogram ←

In the case at Landers we were fortunate to also have GPS satellite measurements of the deformation available. We might compare the radar interferometer measurements with the GPS data.

Now, note that while the GPS measurements are vector quantities the radar only is sensitive to motion in the line of sight of the radar antenna. Thus only one component is measured in the interferogram. To compare them we calculate the dot product of the GPS measurement with a unit vector in the radar line of sight. The results look like this:

Table 2. Comparison of Radar and GPS Motion Estimates

Site	Latitude, deg	Longitude, deg	Horizontal Displacement for Observed Radar Motion, cm	GPS Vector in Radar Direction, cm	Difference, cm
6052	34.52	−116.84	47.8	33.2	14.6
6056	34.37	−116.65	18.1	21.9	−3.8
7000	34.68	−116.72	36.8	91.1	−54.3
7001	34.56	−116.47	−37.1	−70.2	33.1
HECT	34.79	−116.42	9.7	−5.2	14.9
LAZY	34.34	−116.51	62.9	49.4	13.5
LUCS	34.44	−116.88	26.4	20.7	5.7
POIN	34.45	−117.07	13.4	9.5	3.9
SOAP	34.90	−116.98	12.3	1.7	10.6
STIM	34.54	−117.24	7.8	7.4	0.4
FLASH	34.82	−117.02	14.1	12.1	2.0
HARVARD	34.94	−116.67	7.5	−0.4	7.9
BOULDER	34.51	−116.56	176.1	210.8	−34.7
FRY	34.50	−116.72	66.0	74.6	−8.6
MEANS	34.41	−116.55	82.2	69.8	12.4
OLD WOMN	34.39	−116.75	25.0	17.1	7.9
ORD	34.68	−116.81	44.4	48.3	−3.9
ROCK	34.54	−116.77	63.9	69.1	−5.2

where we have related a radar horizontal displacent estimate to slant range measurements by the usual

$$\Delta y = \frac{\Delta \rho}{\sin \theta}$$

If we plot both radar and GPS displacements we see clearly the difference between vector and scalar measurements:

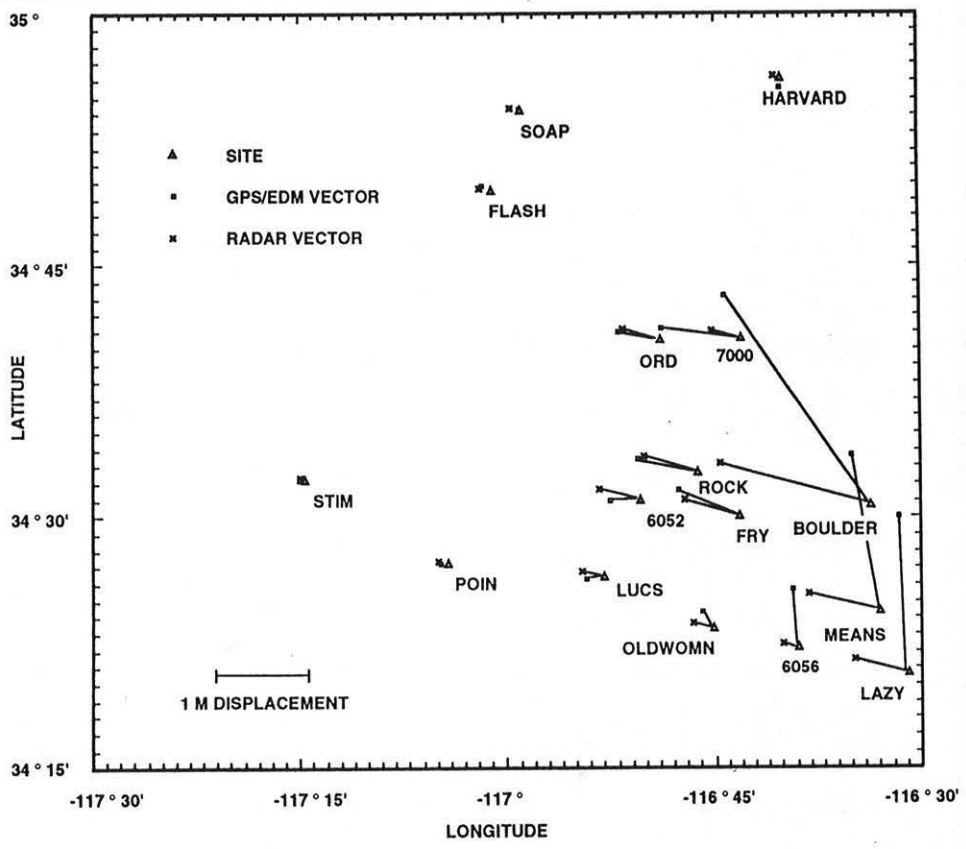


Figure 4. Displacement vectors as measured by GPS/EDM data and by radar interferometry. Each GPS or EDM site is denoted by a triangle, and a vector ending with a square (GPS/EDM measurement) and a vector ending with a cross (radar measurement) are shown in the direction of motion. Note that for the radar case only the component in the radar line of sight direction is determined and thus all measurements are parallel. Vectors are correlated at 0.96 level and show that radar and field surveys are measuring similar phenomena.

We can also look at the statistical correlation (not interferogram correlation) and see that the two sets of measurements are 96% correlated:

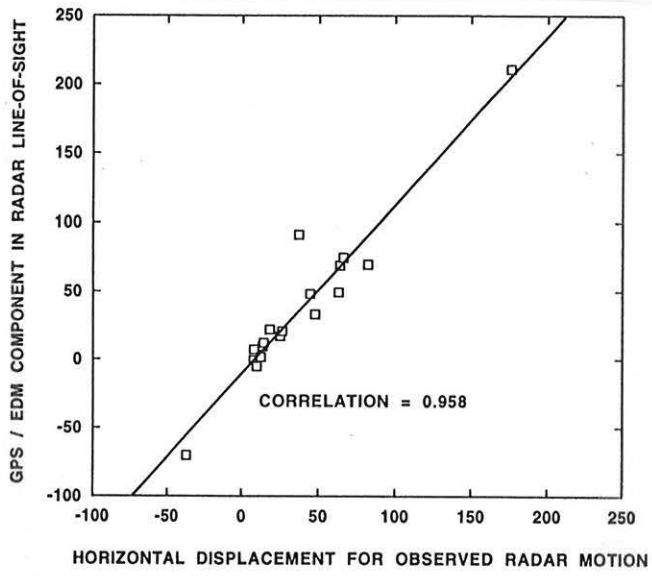
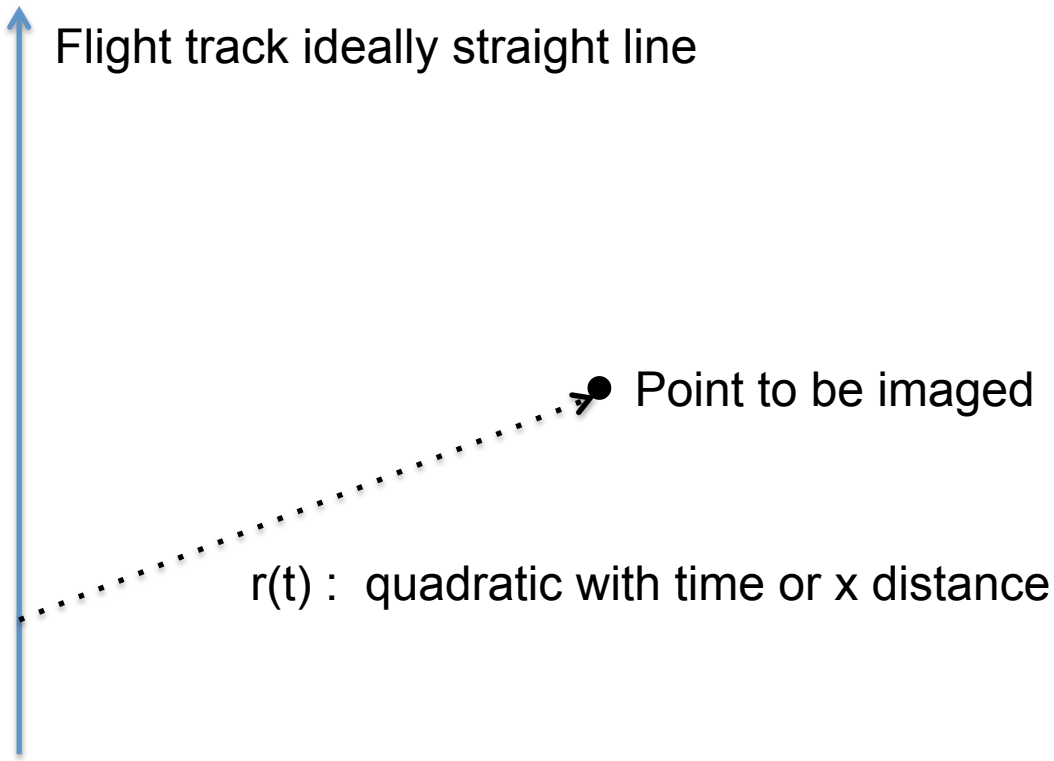


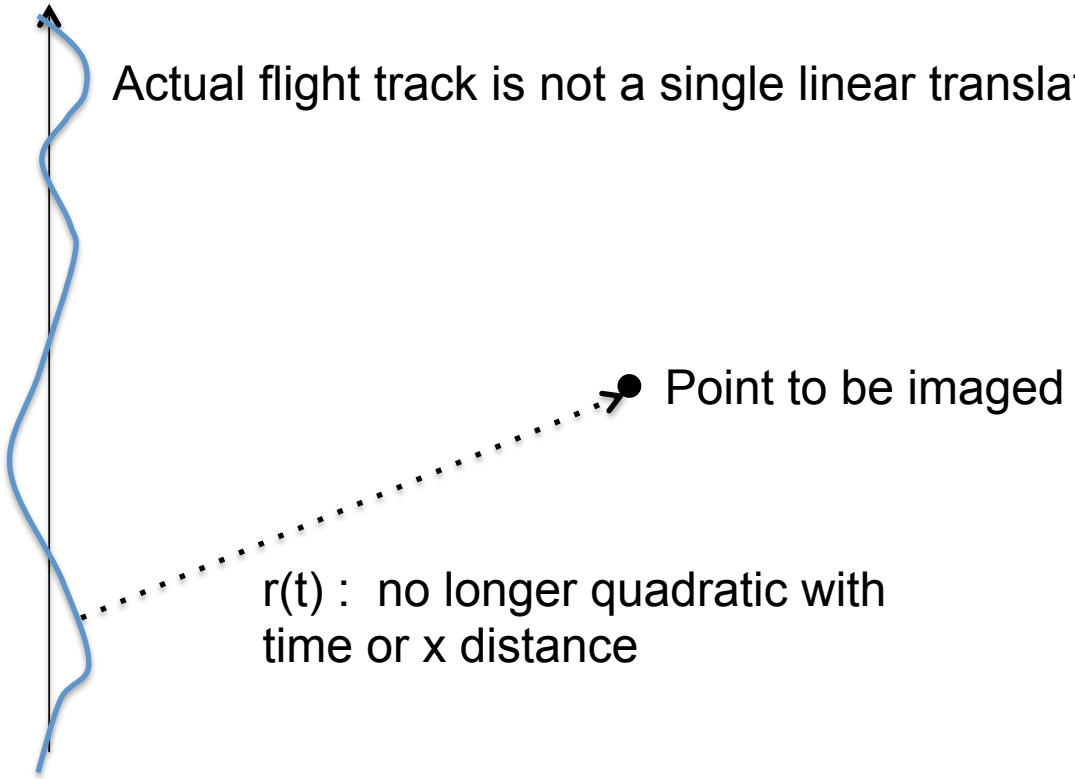
Figure 3. Scatter plot of displacement measurements with GPS/EDM data on vertical axis and radar measurements on horizontal axis. The correlation of the data sets is 0.96, however a slight bias is observed as the slope of the line is not 1.

Motion compensation InSAR processing

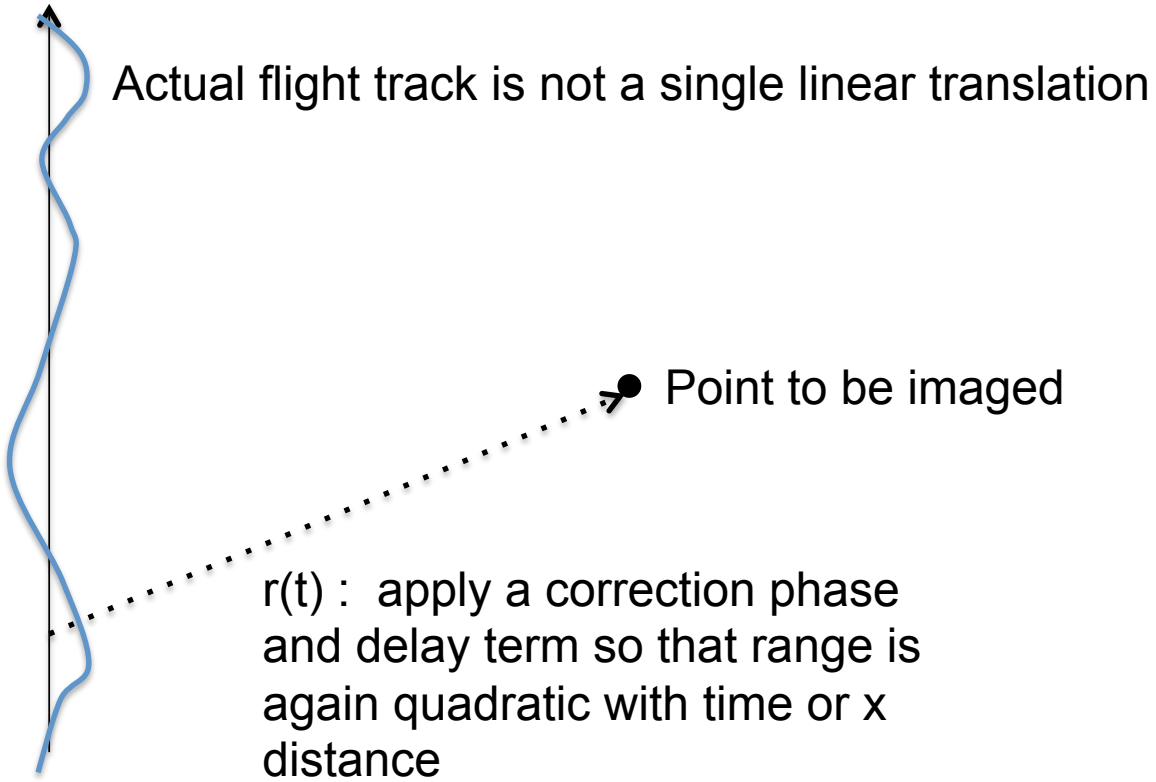
Historical motivation for motion compensation



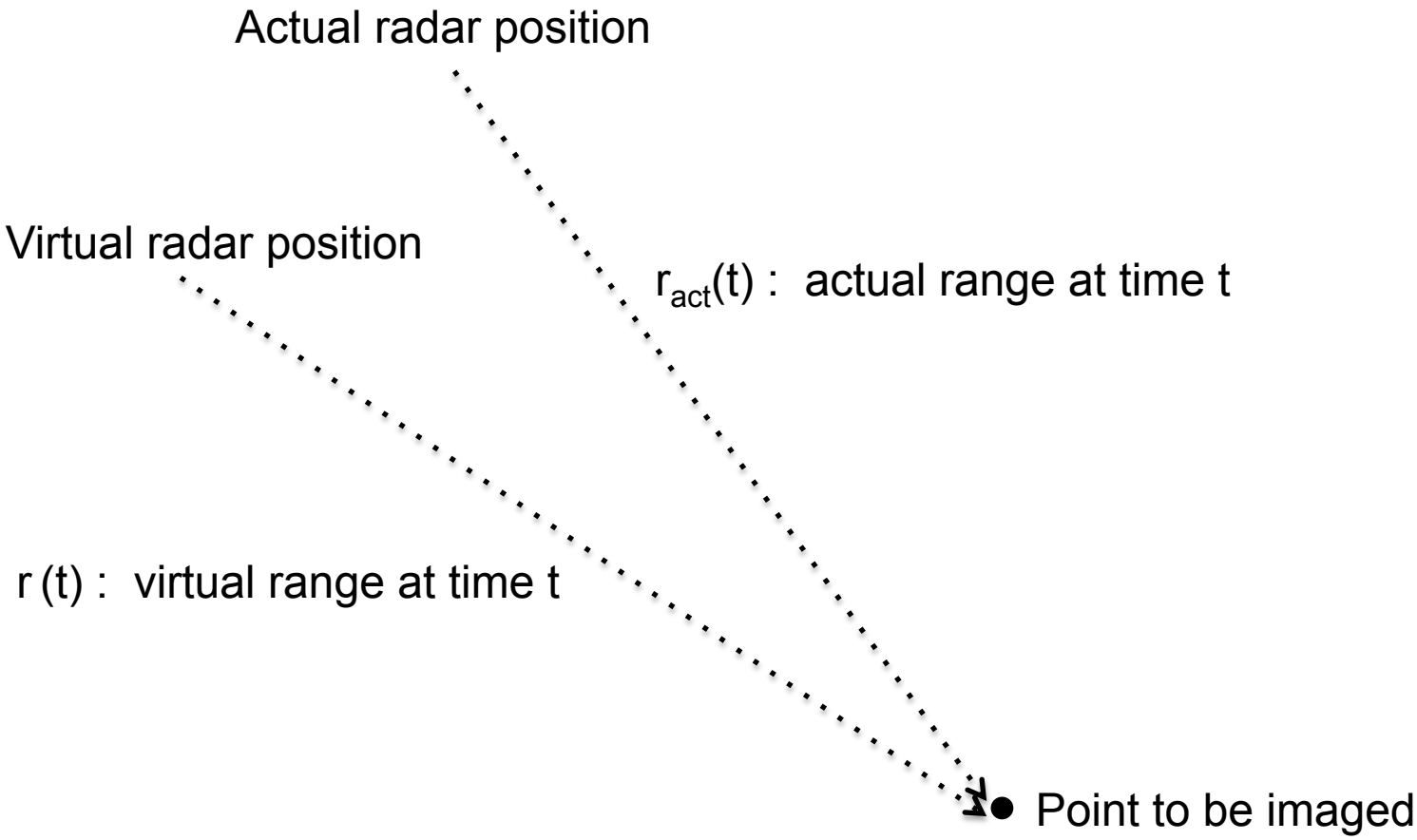
Need for motion compensation



The motion compensation correction



Computing motion compensation correction



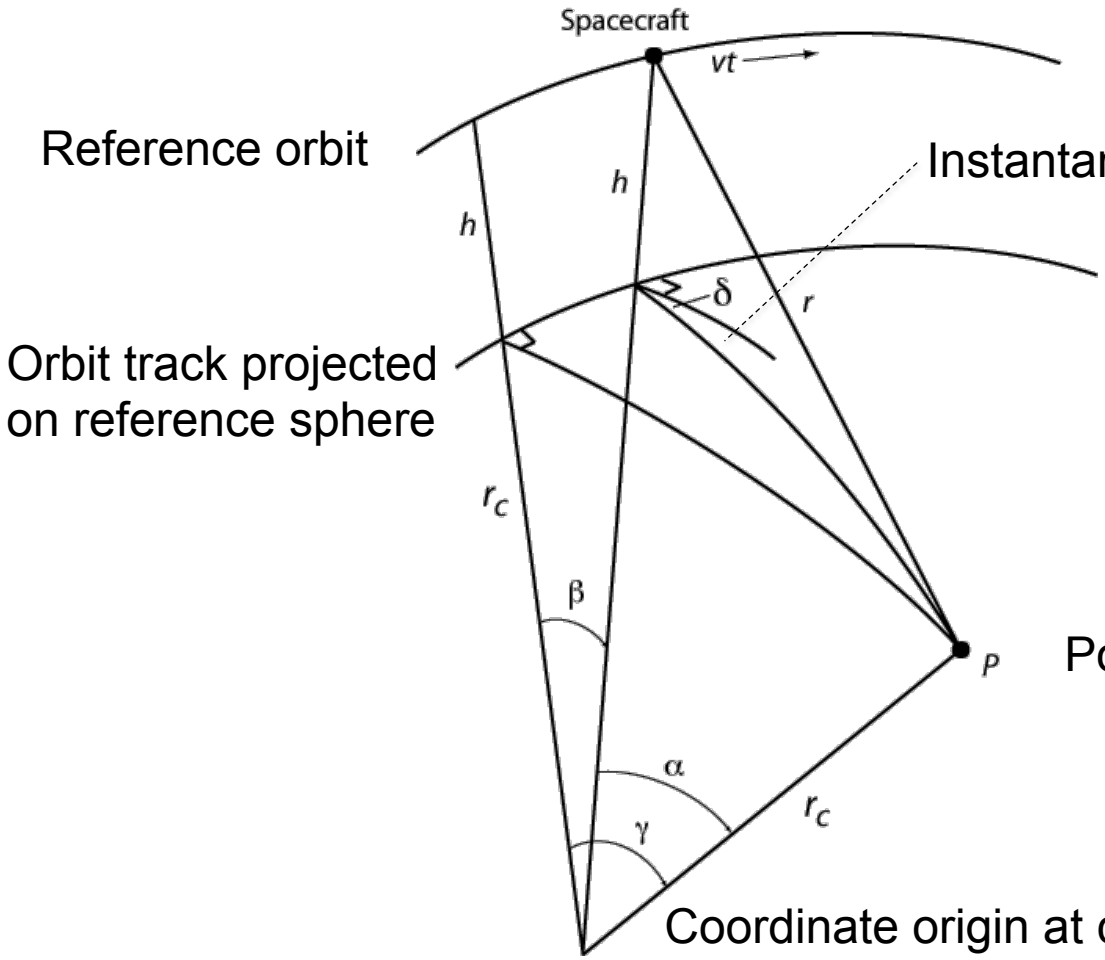
Motion compensation phase and delay

- Motion compensation baseline and time delay
  - $b=r_{act}-r$
  - $\tau=2b/c$
- Motion compensation phase shift
  - $\phi_{baseline}=4\pi/\lambda\ (r_{act}-r)$

Application for InSAR

- Can use motion compensation approach to coregister slcs by processing to single reference track
- Define perfectly spherical reference orbit to processing geometry equations simple
- Avoid coregistration problems

Definitions for orbital geometry



Remember range/Doppler basics

Phase and range relations

$$\phi(t) = -\frac{4\pi}{\lambda} r(t)$$

$$2\pi f(t) = -\frac{4\pi}{\lambda} \dot{r}(t)$$

Doppler relations

$$f_D = -\frac{2}{\lambda} \dot{r}$$

$$f_{rate} = -\frac{2}{\lambda} \ddot{r}$$

Focus and position equations in our geometry

$$f_D = -\frac{2}{r\lambda} r_c (h + r_c) \cos\gamma \sin\beta\dot{\beta}$$

$$f_{rate} = \frac{2}{\lambda}\left[\frac{\dot{r}^2}{r} - \dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta}\right]$$

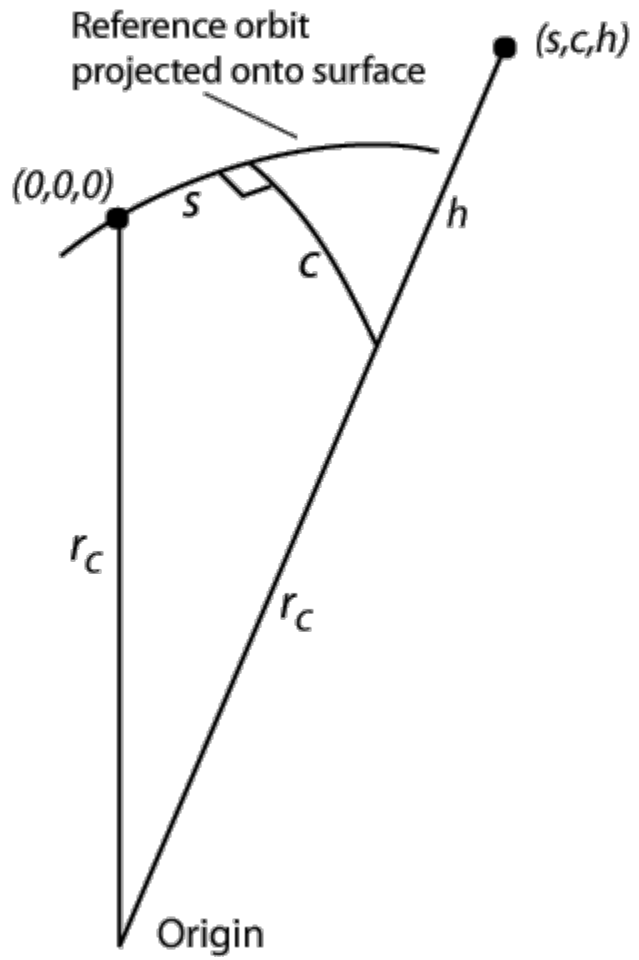
SCH coordinate system

r_c – local radius of curvature, not Earth radius

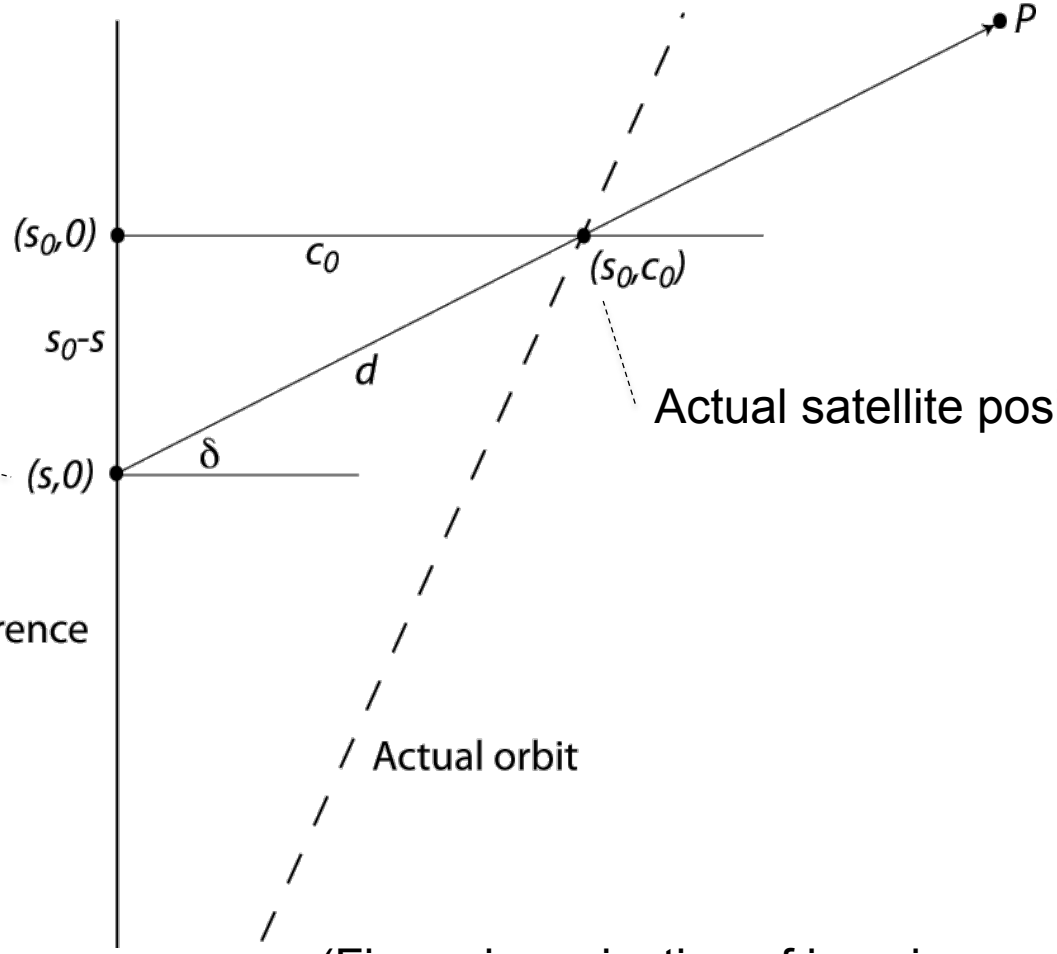
s – along track distance on local sphere from reference point

c – across-track distance on local sphere

h – height above local sphere



Geometry for motion compensation distance and phase



Satellite position on reference orbit at same squint direction

Actual satellite position

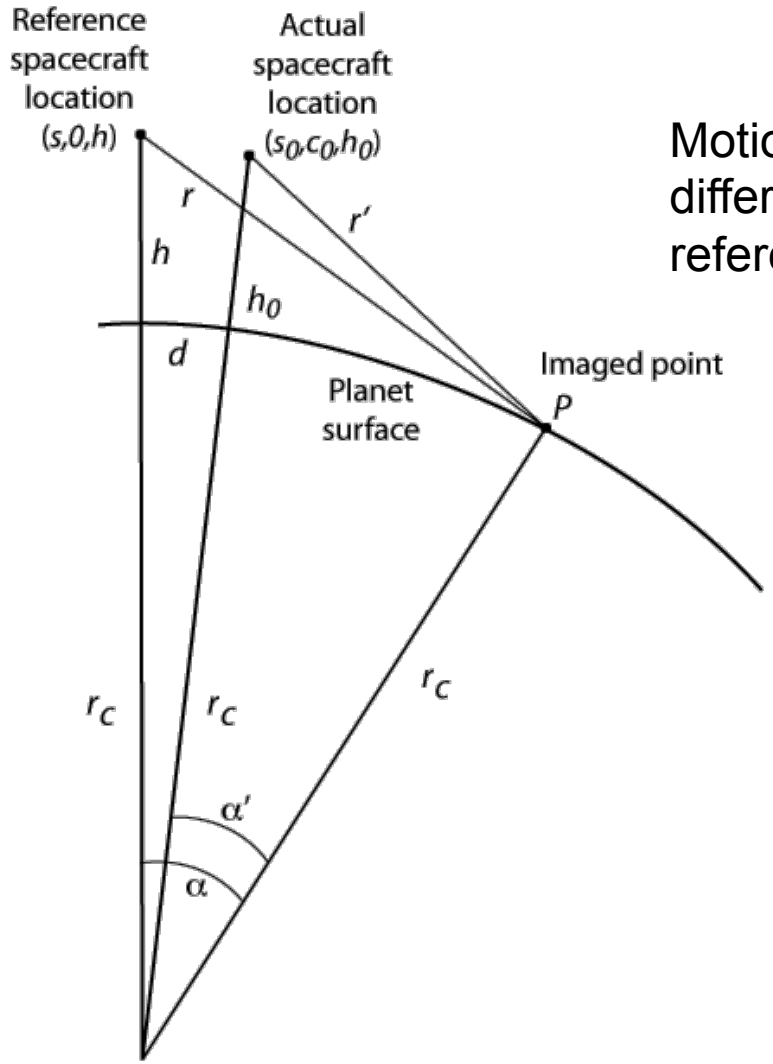
(Figure is projection of imaging geometry onto the reference sphere)

Finding the position on the reference orbit for an actual spacecraft location

$$\sin\delta = \frac{\left(\cos\gamma - \cos\beta\cos\alpha\right)}{\sin\beta\sin\alpha}$$

$$s = s_0 - r_c \sin^{-1}\left(\tan\delta\tan\frac{c_0}{r_c}\right)$$

Motion compensation distance calculation



Motion compensation baseline is difference between actual range r’ and reference orbit range r

Motion compensation algorithm

Derivation of reference distance r:

$$\cos\frac{d}{r_c} = \cos\frac{s_0 - s}{r_c}\cos\frac{c_0}{r_c}$$

$$\cos\alpha = \cos\alpha'\cos\frac{d}{r_c} - \sin\alpha'\sin\frac{d}{r_c}$$

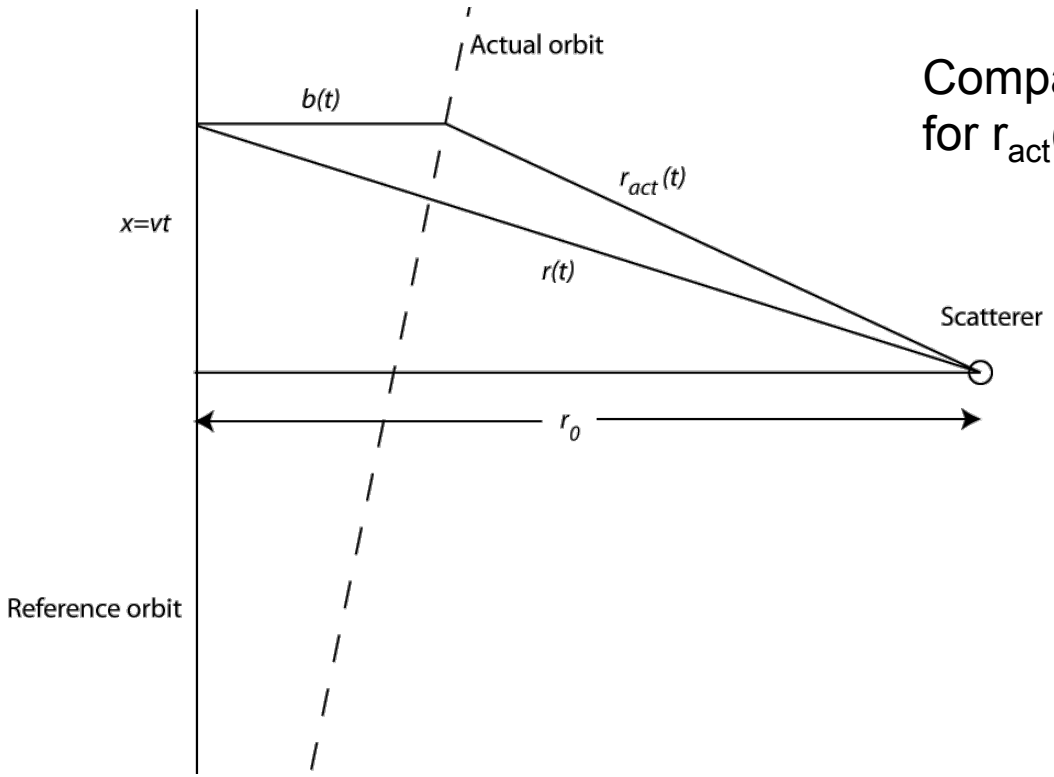
$$r = \sqrt{\left(r_c + h\right)^2 + r_c^{\,2} - 2\left(r_c + h\right)r_c\cos\alpha}$$

Mocomp distance and phase corrections:

$$b = r'(r) - r$$

$$\phi_{baseline} = \frac{4\pi}{\lambda}\left(r'(r) - r\right)$$

Phase history for mocomped scatterer



Compare phase histories for $r_{act}(t)$ and r(t)

Focus corrections

Quadratic phase correction from processing at wrong distance:

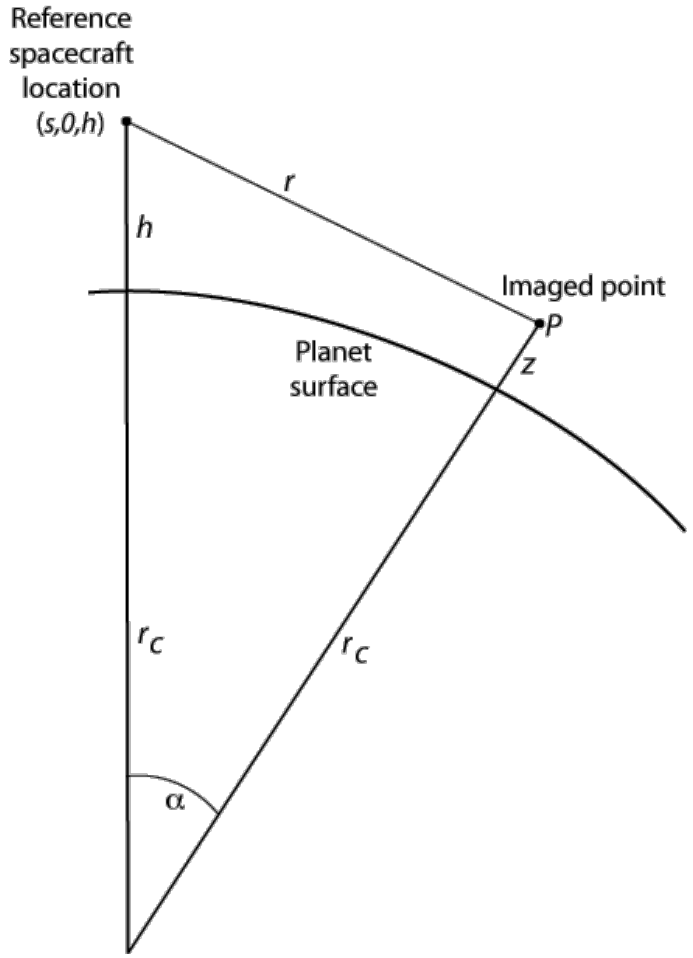
$$r_{mocomp}(t) = r_0 + \frac{1}{2}\frac{v^2 t^2}{r_0}(1 + \frac{b(t)}{r_0})$$

Frequency domain phase term from range-varying motion compensation phase:

$$\phi_{correction} = r_{migration} \cdot \frac{\partial}{\partial r}\left(\frac{4\pi}{\lambda}(r'(r) - r)\right)\Big|_{r=r_0}$$

$$= \frac{\pi}{f_{rate}} \cdot f^2 \cdot \frac{\partial}{\partial r}\left(\frac{4\pi}{\lambda}(r'(r) - r)\right)\Big|_{r=r_0}$$

Topographic correction



- Processor computes SLCs assuming perfectly spherical Earth

- No easy closed form solution for position so use iterative method to find pixel location in 3-space

- Apply phase correction based on pixel elevation

Iterative topography correction

$$\cos\alpha = \frac{(h + r_c)^2 + (r_c + z)^2 - \rho^2}{2(h + r_c)(r_c + z)}$$

$$s = s_{satellite} + r_c \tan^{-1}\left(\frac{f_d(r_c + h)\lambda r}{v(r_c{}^2 + (h + r_c)^2 - r^2)}\right)$$

$$c = -r_c \cos^{-1}\left(\frac{\cos\alpha}{\cos\beta}\right)$$

$$h = z$$

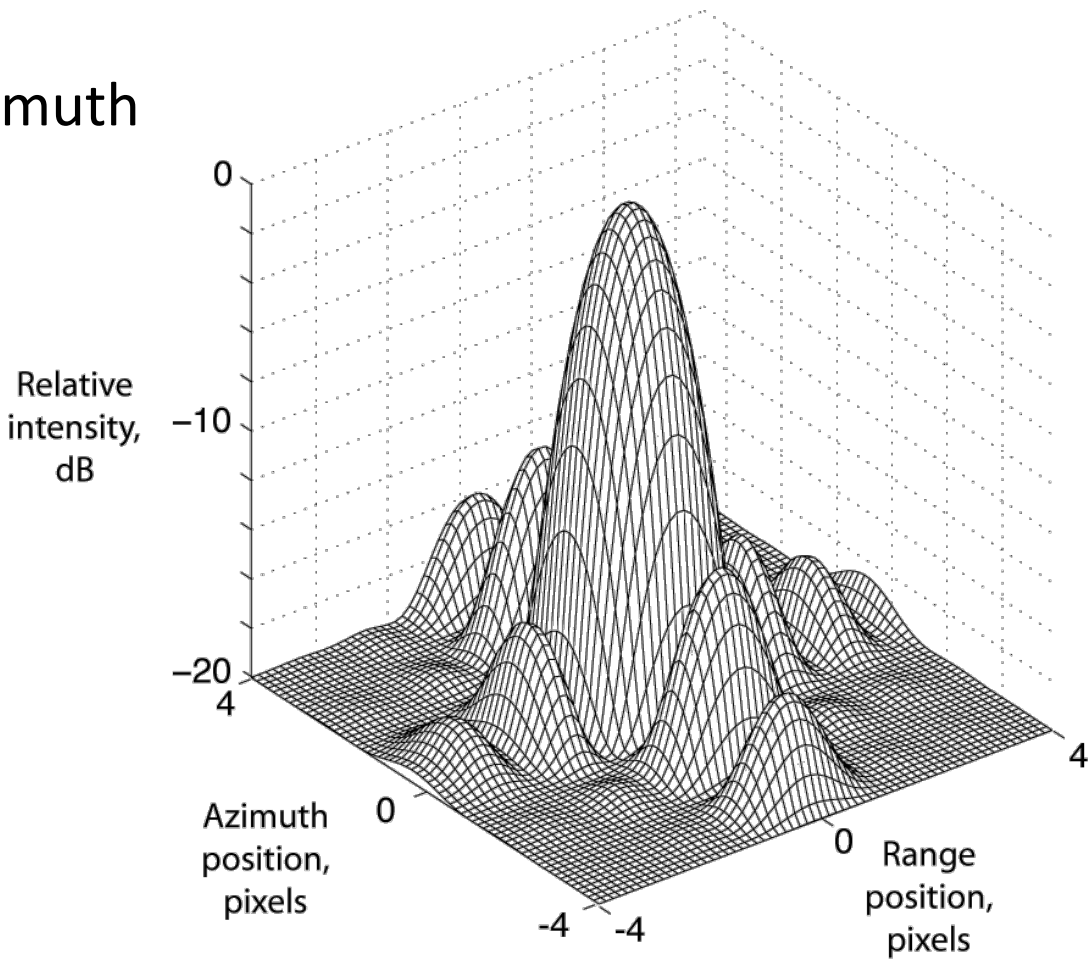
Topographic phase correction:

$$\phi_{elevation} = \frac{4\pi}{\lambda}\left(u^{elevation}_{line-of-sight} - u^{zero\ height}_{line-of-sight}\right) \bullet \mathbf{b}(t)$$

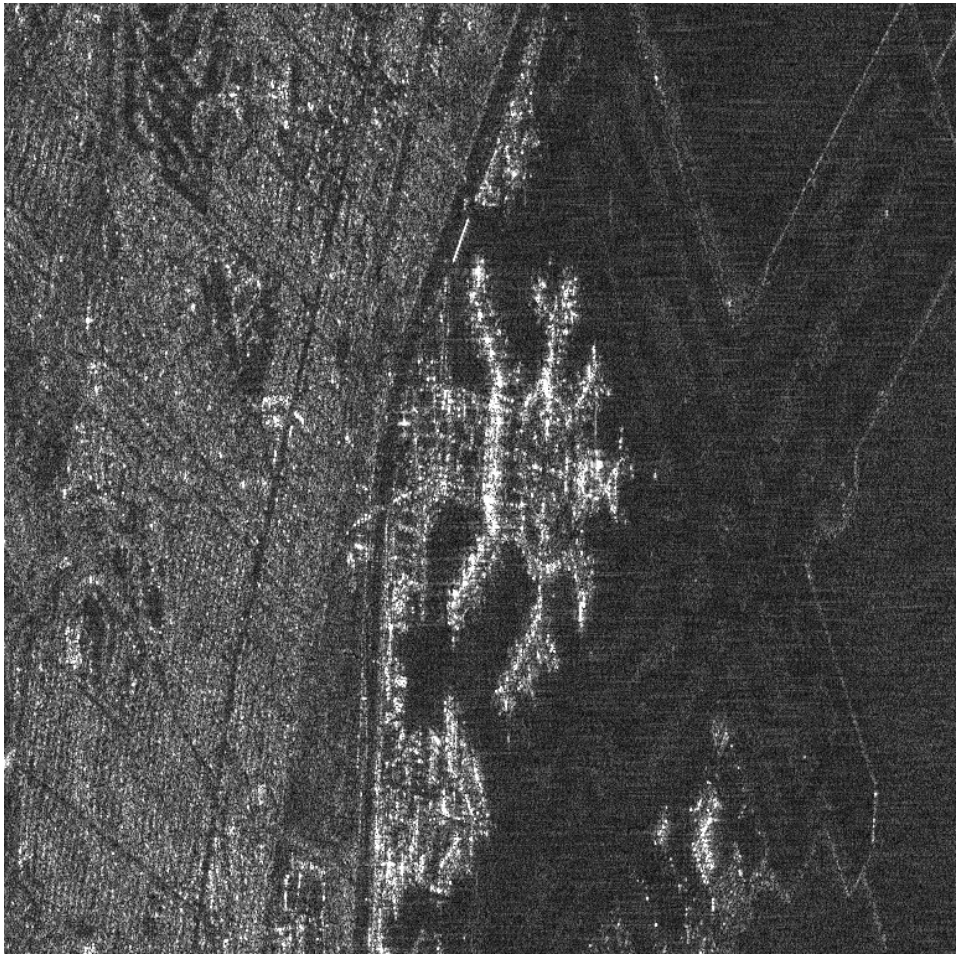
Impulse response

Impulse resolution:
5.3 m range, 4.0 m azimuth

Figure for mocomp
baseline of 1500 m
(InSAR baseline 3km)



Single look complex image of SFO



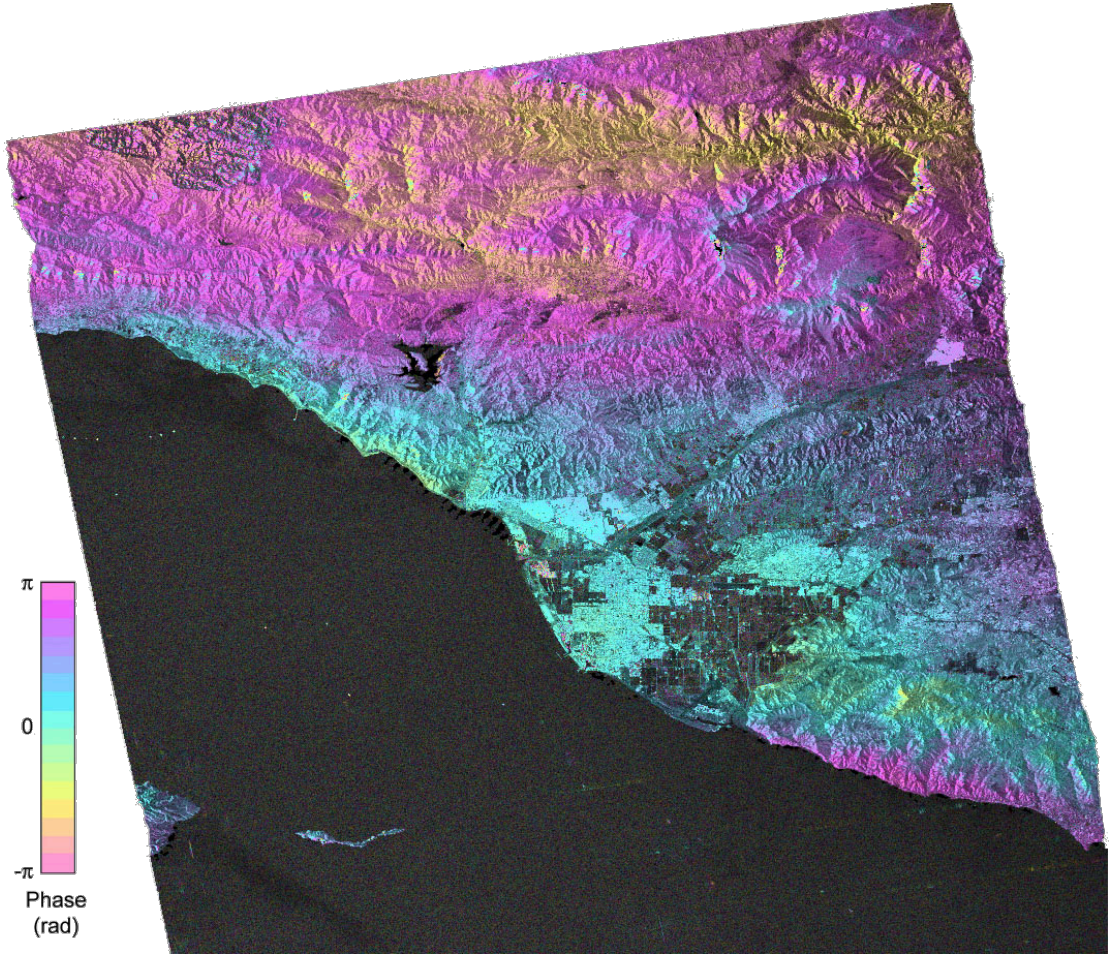
Geodetic accuracy – Pinon Flat Corner Reflector Locations

Measurement	Latitude (deg)	Longitude (deg)	Latitude error (m)	Longitude error (m)
Reflector aligned with ascending orbit				
InSAR location, unregistered image	33.61233	-116.4570	9	-18
InSAR location, registered image	33.61215	-116.4567	-11	9
Ground GPS survey	33.61225	-116.4568	--	--
Reflectors aligned with descending orbit				
InSAR location, unregistered image	33.61215	-116.4579	-11	0
InSAR location, registered image	33.61213	-116.4577	-13	18
Ground GPS Survey	33.61225	-116.4579	--	--
InSAR location, unregistered image	33.60729	-116.4517	-9	9
InSAR location, registered image	33.60727	-116.4516	-11	18
Ground GPS survey	33.60737	-116.4518	--	--

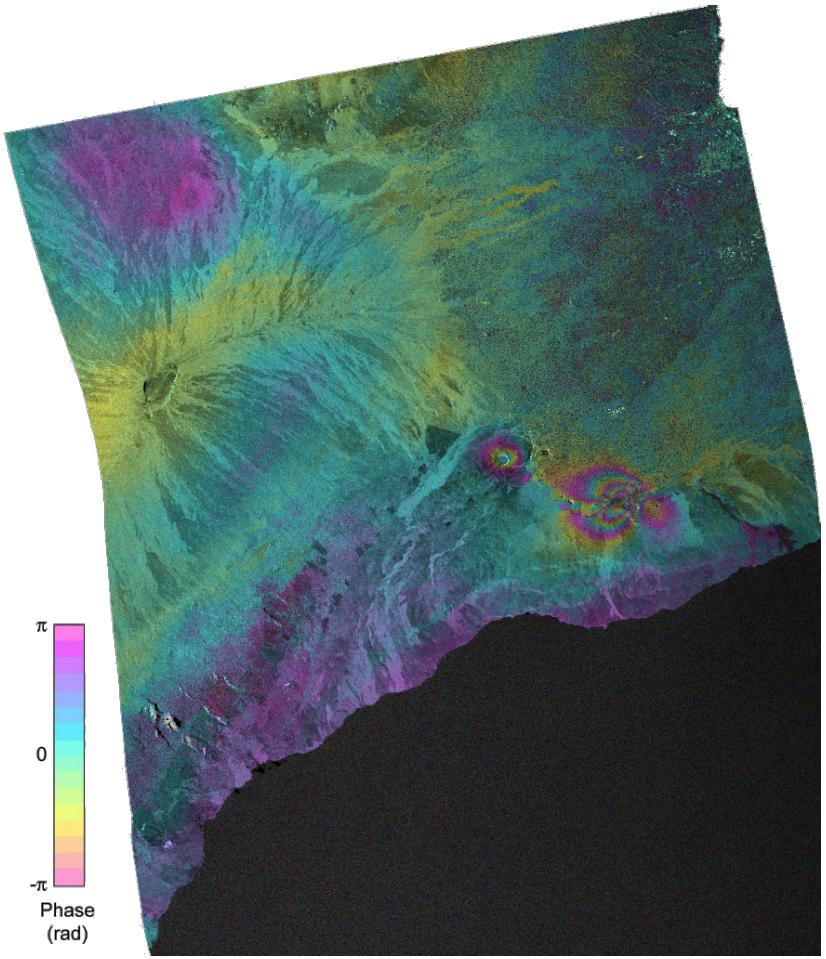
Geodetic accuracy – Image offsets from SRTM DEM

Scene	Range offset at center (m)	Azimuth offset at center (m)	Additional stretch Range (m)	Additional stretch Azimuth (m)
Ventura	-15.8	18.2	9.4	15.2
Hawaii	-21.5	24.0	14.1	25.4
Iceland	2.0	2.9	44.0	29.4

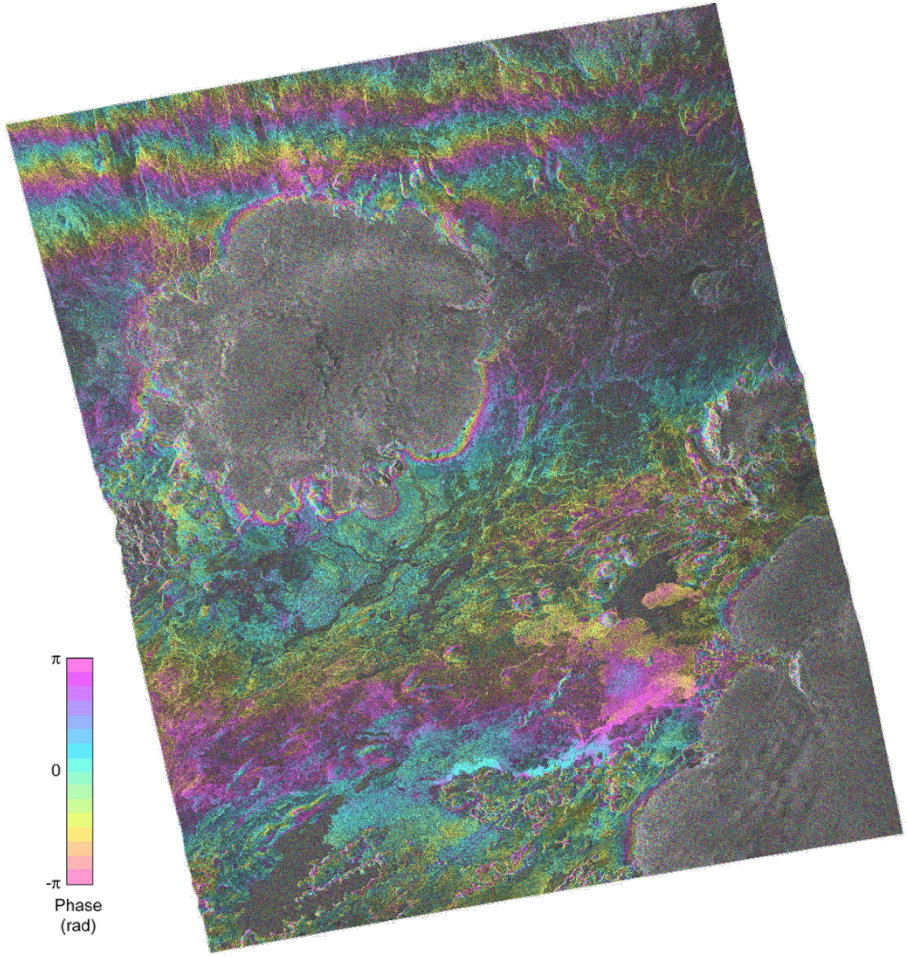
Ventura, CA – Atmospheric phases



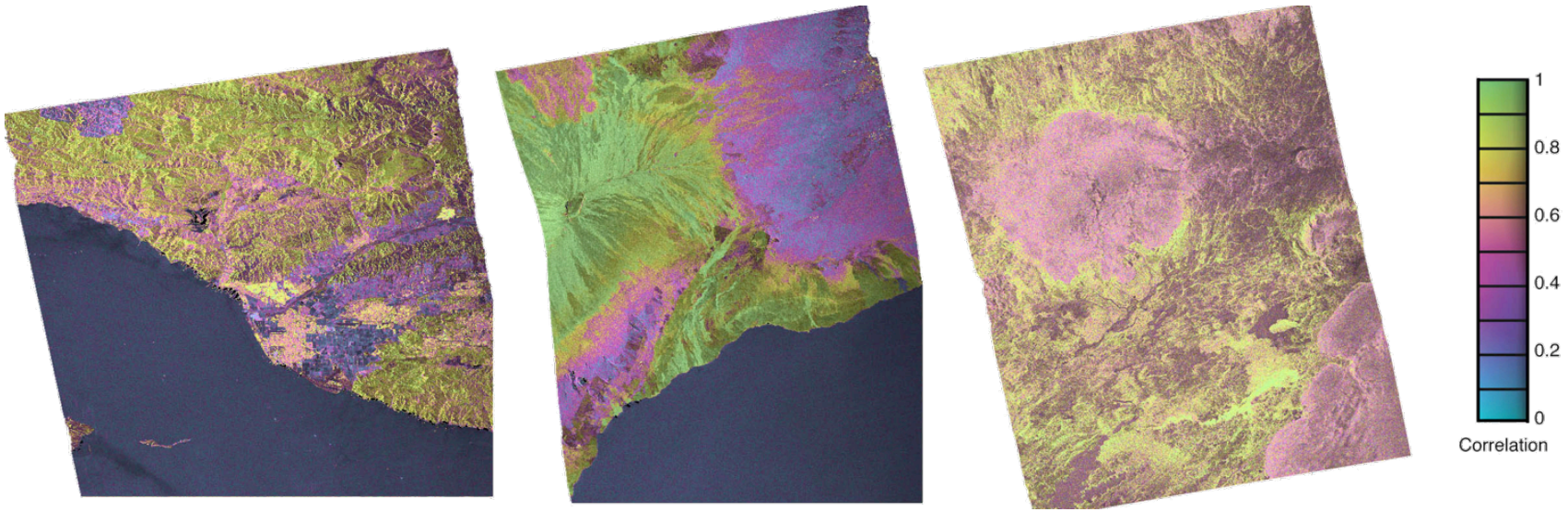
Hawaii – deformation plus atmosphere



Iceland – significant ionospheric artifact



Correlation images



Ventura

Hawaii

Iceland

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 48, NO. 12, DECEMBER 2010 4309

Geodetically Accurate InSAR Data Processor

Howard A. Zebker, *Fellow, IEEE*, Scott Hensley, Piyush Shanker, and Cody Wortham

***Abstract*—We present a new interferometric synthetic aperture radar (InSAR) processing approach that capitalizes on the precise orbit tracking that is available with modern radar satellites. Our method uses an accurate orbit information along with motion-compensation techniques to propagate the radar echoes to positions along a noninertial virtual orbit frame in which the location and focusing equations are particularly simple, so that images are focused without requiring autofocus techniques and are computed efficiently. Motion compensation requires two additional focus correction phase terms that are implemented in the frequency domain. If the images from an interferometric pair or stack are all computed along the same reference orbit, flat-Earth topographic correction is not needed, and image coregistration is simplified, obviating many difficulties that are often encountered in InSAR processing. We process several data sets collected by the ALOS PALSAR instrument and find that the geodetic accuracy of the radar images is 10–20 m, with up to 20 m of additional image distortion needed to align 100 km × 100 km scenes with reference digital elevation models. We validated the accuracy by using both known radar corner reflector locations and by the registration of the interferograms with digital maps. The topography-corrected interferograms are free from all geometric phase terms, and they clearly show the geophysical observables of crustal deformation, atmospheric phase, and ionospheric phase.**

***Index Terms*—Interferometric synthetic aperture radar (InSAR), motion compensation, radar interferometry, SAR processing, synthetic aperture radar (SAR).**

I. Introduction

INTERFEROMETRIC synthetic aperture radar (InSAR) has evolved into a common tool for analysis of crustal deformation [1]–[8], ice motion and structure [9]–[14], hydrologic modeling [15]–[17], vegetation canopy characterization [18]–[20], and generation of topographic data [21]–[24]. The InSAR technique is defined by computing the phase difference of complex radar echoes at each resolution element in a radar image, resulting in millimeter-scale displacement images at meter-level postings over wide areas (typically 100-km scales). Recent developments in satellite tracking and radar signal processing now permit the generation of InSAR images or interferograms that are, in addition, geodetically quite accurate. The geodetic accuracy not only provides data products in better known coordinate systems but also facilitates routine processing by avoiding many of the image registration and resampling steps incorporated into existing processing systems. The algorithms presented here are computationally efficient and more robust than many traditional processing approaches, enabling advanced approaches to data interpretation such as time series analysis of surface change.

The continuing advances in the accuracy of orbit determination now routinely produce spacecraft position estimates with submeter uncertainties. Satellites such as the European ERS-1 and ERS-2, Envisat, and Japanese ALOS produce operational products with standard errors of tens of centimeters or less [25]–[28]. We have designed a new software radar processing system based on these accurate orbit measurements so that the radar pixels may be located on the surface with accuracies of tens of meters or less. In our approach, we use radar motion-compensation techniques to propagate radar echoes from their actual acquisition locations to ideal orbits in which the focusing and positioning equations are particularly simple. Thus, when processing multiple images, for a case that is as simple as two scenes to be formed into a single interferogram or as complex as hundreds of scenes to form a persistent scattering estimate of temporal evolution of crustal deformation, all of the data passes may be processed to a common coordinate system. This facilitates the resampling of the individual single-look complex scenes to common locations, which is a step that is often problematic in geodetically inaccurate processing methods. The accuracy of the orbits is such that autofocus or other image refinement steps are not necessary, significantly increasing both the efficiency of the processor and the accuracy to which pixels may be located on the surface. Another advantage to motion-compensated processing using a common reference orbit is that the "curved-Earth" range phase term is not present in the interferograms as the effective InSAR baseline, as regards Earth curvature, is zero.

Here, we describe our processing method, starting with the definition of our reference orbit and the equations needed for radar image focusing and pixel location. We then summarize our motion-compensation approach and show that two focus correction phase histories must be added to the radar echo to properly focus the image. We then describe an iterative algorithm for mapping the interferograms, as expressed in radar coordinates, to evenly spaced and known geodetic coordinates so that the images may be readily combined with other data types. We assess the geodetic accuracy of the system by analyzing the data acquired over a set of GPS-surveyed radar corner reflectors. Finally, we present several interferograms from L-band ALOS PALSAR data in order to demonstrate applicability to a variety of applications.

We note that our method is not necessarily more geodetically accurate than other InSAR software that has been similarly motivated. We have attempted to design a processing system

Manuscript received October 23, 2009; revised March 18, 2010. Date of publication July 15, 2010; date of current version November 24, 2010.

H. A. Zebker, P. Shanker, and C. Wortham are with the Departments of Electrical Engineering and Geophysics, Stanford University, Stanford, CA 94305-2155 USA (e-mail: zebker@stanford.edu).

S. Hensley is with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2010.2051333

with geodetic accuracy as a fundamental design consideration. Hence, a major emphasis in this paper is on useful coordinate systems to facilitate geodetic accuracy in both the SAR processing and the derivation of InSAR products. Geodetic accuracy is not only important for many applications but also readily feasible in today's precise-orbit world, as we demonstrate in the succeeding discussion.

Previous presentations of geodetically accurate radar processing [29], [30] also show accuracies that are roughly the size of a 10-m radar resolution cell. The work at Scripps Institution of Oceanography [30] has already shown that orbit accuracy for the ALOS satellite is fine enough to obviate the need for autofocus modules in the software. We find similarly that this code is unneeded. In addition, several groups [31]–[34] have experimented with aligning time sequences of images precisely during processing to a single master image. In our approach, we do the same, but the coordinate system used is not a physically realizable system for a satellite in an inertial orbit. In essence, we use a virtual coordinate system to simplify postprocessing codes that implement the InSAR product generation.

It is worth noting that the motion-compensation approach that we present here is a critical aspect for processing SAR data from airborne platforms, where interactions with the atmosphere lead to turbulent flight trajectories that defocus SAR images and lead to InSAR phase errors. These short-period orbit errors are usually less significant for spaceborne platforms that generally orbit well above the atmosphere, although they can be present for certain system configurations or imaging geometries. In this paper, we meet precise geometrical standards using an approach that can work even with very irregular orbits, such as those that might arise from low-altitude satellites or from platforms with extremely long synthetic apertures, as required for long-wavelength radar systems.

II. Phase History for Spacecraft in Perfect Circular Orbit

We begin by developing the equations needed to properly locate and focus a SAR image from an orbiting radar sensor. Assume that we have a satellite in a perfect circular orbit above a nonrotating planet. Since all known planets rotate, such an orbital is noninertial, it cannot exist in a physical sense without continuous accelerations applied to the spacecraft, and, thus, it is not feasible for satellites in use today. Nonetheless, we can define such an orbit and translate the actual radar echoes to the ideal reference trajectory using motion-compensation methods. We define the geometry of the spacecraft radar observing a point P as shown in Fig. 1.

Here, the spacecraft travels along an orbit path at a constant velocity v, at a constant height h, and above a spherical planet with a radius of curvature r_c. The range to the imaged point r varies as a function of time t. The usual relations for radar phase history ϕ and instantaneous frequency f hold

$$\phi(t) = -\frac{4\pi}{\lambda} r(t) \tag{1}$$

$$2\pi f(t) = -\frac{4\pi}{\lambda} \dot{r}(t). \tag{2}$$

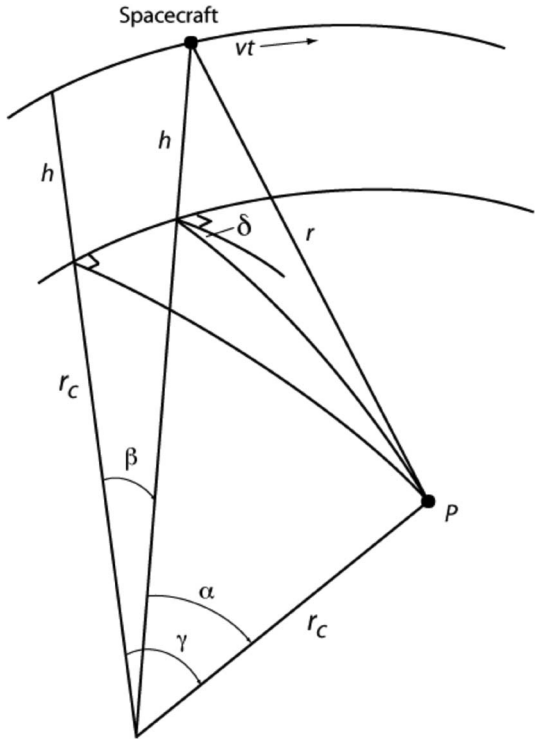


Fig. 1. Definitions of angles and distances for the reference orbit. The planet is assumed to be locally spherical with a radius of curvature r_c, and note that r_c is not necessarily equal to the local distance from the surface to the center of mass of the Earth. The spacecraft flies at an altitude h above the surface and on a perfect circular path about a point r_c below the surface. The point P that is to be imaged lies at an origin-centered angle γ from the satellite orbit at the closest approach. The distance along the orbit is equal to the spacecraft velocity v multiplied by time t, so that at any given time, the satellite-origin-closest approach point angle is β. The origin-centered angle α is formed by the point to be imaged, the origin, and the satellite location. Angle δ is the squint angle from the satellite to the imaged point.

The Doppler frequency f_D and the Doppler rate $f_{\rm rate}$, respectively, can be written as

$$f_D = -\frac{2}{\lambda} \dot{r} \tag{3}$$

$$f_{\rm rate} = -\frac{2}{\lambda} \ddot{r}. \tag{4}$$

We now relate these general expressions to the geometry of Fig. 1. From the law of cosines

$$r^2 = (h + r_c)^2 + r_c^2 - 2r_c(h + r_c)\cos\alpha. \tag{5}$$

The spherical law of cosines considering the right angle shown in the figure is equal to

$$\cos\alpha = \cos\beta\cos\gamma. \tag{6}$$

Let us rewrite this as

$$r^2 = (h + r_c)^2 + r_c^2 - 2r_c(h + r_c)\cos\beta\cos\gamma. \tag{7}$$

By noting that $vt/(h + r_c) = \beta$, and $\dot{\beta} = v/(h + r_c)$ is a constant, we can differentiate (7) with respect to time

$$2r\dot{r} = -2r_c(h + r_c)\cos\gamma(-\sin\beta)\dot{\beta}. \tag{8}$$

Thus

$$\dot{r} = \frac{r_c(h + r_c)\cos\gamma\sin\beta\dot{\beta}}{r} \tag{9}$$

or it is expressed as f_D as a function of the along-track angle β

$$f_D = -\frac{2}{r\lambda} r_c(h + r_c)\cos\gamma\sin\beta\dot{\beta}. \tag{10}$$

To determine the SAR focusing parameter $f_{\rm rate}$, we start with (8) and again differentiate with respect to time, obtaining

$$\begin{aligned} r\ddot{r} + \dot{r}^2 &= r_c(h + r_c)\cos\gamma\dot{\beta}\cos\beta\dot{\beta} \\ &= r_c(h + r_c)\cos\gamma\sin\beta\dot{\beta}\cdot\frac{\cos\beta}{\sin\beta}\dot{\beta} \\ &= r\dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta}. \end{aligned} \tag{11}$$

Then

$$r\ddot{r} = r\dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta} - \dot{r}^2 \tag{12}$$

$$\ddot{r} = \dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta} - \frac{\dot{r}^2}{r} \tag{13}$$

finally yielding

$$f_{\rm rate} = \frac{2}{\lambda}\left[\frac{\dot{r}^2}{r} - \dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta}\right]. \tag{14}$$

The expressions for f_D and $f_{\rm rate}$, which are (10) and (14), respectively, supply the information that is necessary for locating the along-track position and the optimal chirp rate to image point P, as in many range Doppler processing implementations (see, for example, [35]).

III. Motion-Compensation Approach and Geometry

Thus, if the spacecraft was flying in the above ideal orbit, we can readily construct the matched filter for an along-track location, as given by the Doppler centroid f_D and azimuth chirp rate $f_{\rm rate}$. However, it would be very wasteful of fuel to force a satellite into this noninertial orbit, and existing sensors do not follow such an orbital trajectory. We therefore apply a motion-compensation algorithm to the received echoes so that the data are similar to what the sensor would have recorded if it had flown along the reference track. In addition to allowing ready focusing and identifying the location of the image using the equations of the previous section, the motion-compensation step allows us to process multiple acquisitions to the same coordinate system.

We introduce here a coordinate system that is defined with respect to the projected ground track of the ideal satellite orbit. This coordinate system, referred to as sch, was developed and used for the NASA Shuttle Radar Topography Mission at the Jet Propulsion Laboratory [36]. In this coordinate system, s is the along-track coordinate along the surface projection of the satellite path, c is the cross-track coordinate along the surface to the projection of a point, and h is the height of the point above the surface (Fig. 2).

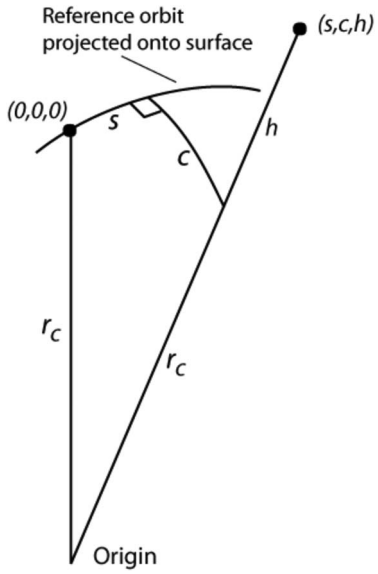


Fig. 2. Definition of the sch coordinate system. A reference point along the reference orbit has coordinates (0,0,0), and the local Earth radius of curvature is r_c. An arbitrary point in space (s, c, h) is located at a height h above the spherical surface and at a surface distance c from the projection of the reference orbit and is displaced along track by s. The reference orbit is assumed to be perfectly circular and centered about the origin shown.

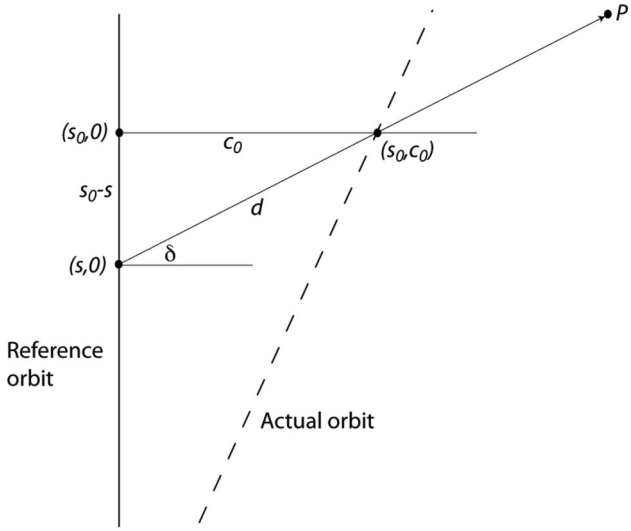


Fig. 3. Motion-compensation geometry and position definitions. The projection of the actual satellite orbit on the planet surface is the dashed line, and the desired reference orbit projects as the solid line. The spacecraft observes a point P when it is located above (s_0, c_0). The squint angle δ is defined by the Doppler centroid. The virtual position of the spacecraft after motion compensation is at point $(s, 0)$, and distance d is the surface projection of the amount where the echo must be propagated to represent what the sensor would have measured if it had indeed been located at $(s, 0)$.

Consider a top-down view of the motion-compensation reference orbit and the actual ground track of the satellite projected on a spherical planet, as shown in Fig. 3.

Here, the actual location of the satellite is (s_0, c_0, h_0); P is the point to be imaged; δ is the squint angle, given the Doppler centroid of the point; the desired position of the motion-compensated satellite is $(s, 0, h)$; and c_0, $s - s_0$, and

d are all distances along great circles on the planet surface. Once again, starting with the spherical law of cosines, we see

$$\cos\frac{d}{r_c} = \cos\frac{s_0 - s}{r_c}\cos\frac{c_0}{r_c} \tag{15}$$

while the law of sines yields

$$\frac{\sin(\pi/2 - \delta)}{\sin\frac{c_0}{r_c}} = \frac{\sin\pi/2}{\sin\frac{d}{r_c}}. \tag{16}$$

In addition, we can also write using the spherical law of cosines

$$\cos(\pi/2-\delta)\sin\frac{s_0 - s}{r_c}\sin\frac{d}{r_c} + \cos\frac{s_0 - s}{r_c}\cos\frac{d}{r_c} = \cos\frac{c_0}{r_c}. \tag{17}$$

By using (15) earlier

$$\cos(\pi/2-\delta)\sin\frac{s_0-s}{r_c}\sin\frac{d}{r_c} + \cos^2\frac{s_0-s}{r_c}\cos\frac{c_0}{r_c} = \cos\frac{c_0}{r_c} \tag{18}$$

or

$$\begin{aligned}\cos(\pi/2-\delta)\sin\frac{s_0-s}{r_c}\sin\frac{d}{r_c} &= \cos\frac{c_0}{r_c}\left(1-\cos^2\frac{s_0-s}{r_c}\right)\\ &= \cos\frac{c_0}{r_c}\sin^2\frac{s_0-s}{r_c}\end{aligned} \tag{19}$$

and thus

$$\cos(\pi/2 - \delta)\sin\frac{d}{r_c} = \cos\frac{c_0}{r_c}\sin\frac{s_0 - s}{r_c}. \tag{20}$$

Moreover, from (16)

$$\cos(\pi/2 - \delta)\frac{\sin\frac{c_0}{r_c}}{\sin(\pi/2 - \delta)} = \cos\frac{c_0}{r_c}\sin\frac{s_0 - s}{r_c} \tag{21}$$

from which

$$\tan\delta\sin\frac{c_0}{r_c} = \cos\frac{c_0}{r_c}\sin\frac{s_0 - s}{r_c}. \tag{22}$$

Finally

$$\tan\delta\tan\frac{c_0}{r_c} = \sin\frac{s_0 - s}{r_c}. \tag{23}$$

Now, we can solve for the desired spacecraft position s

$$s = s_0 - r_c\sin^{-1}\left(\tan\delta\tan\frac{c_0}{r_c}\right). \tag{24}$$

The desired cross-track position and height are 0 and h, respectively, so now, we know both the actual and motion-compensated spacecraft locations. The solution for the composite squint angle δ can be found from the aforementioned α, β, and γ since (again from the spherical law of cosines)

$$\sin\delta = \frac{(\cos\gamma - \cos\beta\cos\alpha)}{\sin\beta\sin\alpha}. \tag{25}$$

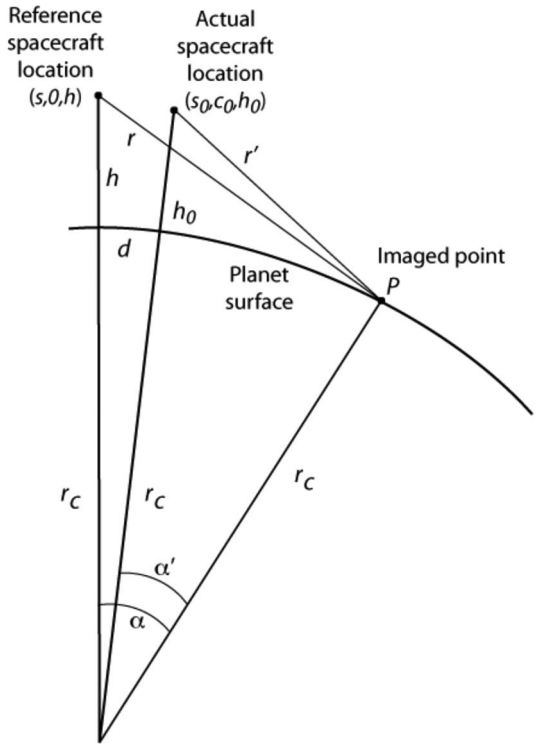


Fig. 4. Motion-compensation distances. The echo must be propagated from an actual distance r' to a desired distance r. The shift corresponds to a change in angle $\alpha - \alpha'$ that is relevant to the surface shift d in Fig. 3.

IV. Motion-Compensation Algorithm

The equations in the previous section provide the relationship between the actual spacecraft locations and the desired imaging locations along the reference orbit. The satellite is not flying in this ideal orbit, of course, so we propagate the radar echoes from the actual satellite position to the ideal reference track to make it appear as if the radar had flown the perfect circular path described earlier, which is a procedure known as motion compensation (a good description is found in [37]).

We use an approximate form for motion compensation by assuming that the appropriate displacement for a scatterer is the one that is associated with its location at the point where the scatterer passes through the antenna boresight. This approximation is most valid for systems with a narrow antenna beam. In other words, we assume that most of the backscattered energy comes from a direction that corresponds to the Doppler centroid of the echo. This approximation is quite good for many existing spaceborne radar systems, although it leads to a focusing error because echoes from the scatterer that are away from the boresight are slightly shifted in phase. This phase error is corrected in the azimuth focusing step, which we explain in the following focusing section.

We apply the motion-compensation resampling to the measured data by adding to the echo the appropriate phase and by shifting its position in time according to the motion-compensation distance. The distance where the echo must be shifted is readily seen in the following (Fig. 4).

For the motion-compensation algorithm, we use the actual distance from the spacecraft (s_0, c_0, h_0) to the point P (denoted as r') as measured by the radar and the calculated

distance (denoted as r in Figs. 1 and 4) from the motion-compensated reference track $(s, 0, h)$. We determine the origin-centered angles α and α' for the reference and actual positions starting with

$$\cos\alpha' = \frac{(r_c + h_0)^2 + r_c^2 - r'^2}{2r_c(h_0 + r_c)} \tag{26}$$

and by using the difference $\alpha - \alpha'$ (which is equal to d/r_c), we find that the cosine of the origin-centered angle α for the reference location is

$$\cos\alpha = \cos\alpha' \cos\frac{d}{r_c} - \sin\alpha' \sin\frac{d}{r_c} \tag{27}$$

so that

$$r = \sqrt{(r_c + h)^2 + r_c^2 - 2(r_c + h)r_c\cos\alpha}. \tag{28}$$

We then shift the position of the return echo by $r' - r$ and its phase by $(4\pi/\lambda)(r' - r)$. Since r' is a function of r, we can express the motion-compensation baseline distance and phase, respectively, as

$$b = r'(r) - r \tag{29}$$

$$\phi_{\text{baseline}} = \frac{4\pi}{\lambda}\left(r'(r) - r\right). \tag{30}$$

V. Focus Corrections

Two phase correction terms are needed to properly focus the motion-compensated SAR images. The first correction is a change in the Doppler frequency rate f_{rate} resulting from the motion-compensation shift in the scatterer distance from the radar. The second correction is a phase term added to the phase history to account for range dependence of the motion-compensation phase shift.

The echo signal after motion compensation is moved to a different range if the motion-compensation baseline (the difference between the actual spacecraft position and the position projected onto the reference orbit as described earlier) is nonzero. The phase history of a scatterer, which depends on actual imaging geometry, is thus located at a different range than its original position, and it differs from the history that is expected at its motion-compensated range. While motion compensation adequately corrects the constant and linear terms for the phase history, the second order term requires an additional correction. The first correction factor changes the frequency rate of the matched filter by the ratio of the motion-compensation baseline (the distance that the echo moved in the motion-compensation step) to the scatterer range. Consider the following (Fig. 5).

This figure shows a simplified motion-compensation geometry that is constrained so that the reference orbit, the actual orbit, and the scattering point are all coplanar, and we assume that we are processing the echo that is centered at the zero Doppler point. The distance $b(t)$ is the motion-compensation baseline. The range history for the scatterer with respect to the reference orbit satisfies

$$r^2(t) = r_0^2 + v^2t^2 \tag{31}$$

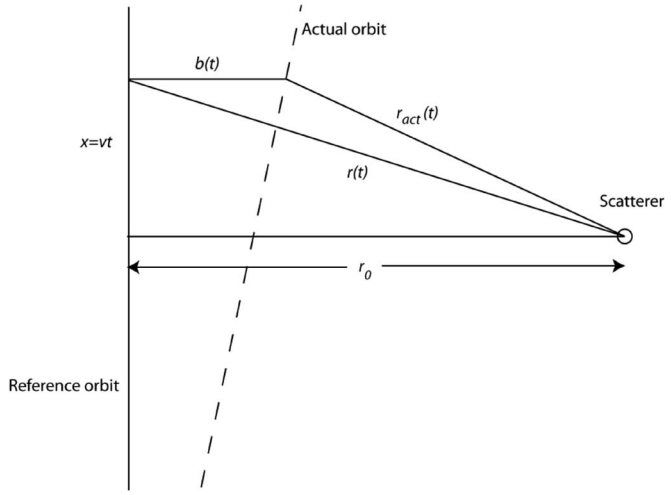


Fig. 5. Geometric construction to understand the Doppler rate correction required after motion compensation. In this simplified geometry, the reference orbit, the actual orbit, and the scatterer all lie in the plane of the page. The actual distance from the spacecraft to the scatterer is $r_{\text{act}}(t)$, and the distance of the motion-compensated spacecraft along the reference orbit to the scatterer is $r(t)$. The distance from the reference spacecraft to the scatterer at the closest approach is r_0, and $b(t)$ is the motion-compensation baseline as a function of time.

which, under the usual SAR approximation, can be written as

$$r(t) \approx r_0 + \frac{1}{2}\frac{v^2t^2}{r_0} \tag{32}$$

leading to a Doppler rate $f_{\text{rate}} = -(2v^2/\lambda r_0)$.

The range history of the actual return from the scatterer, again under the SAR approximation, is

$$\begin{aligned} r_{\text{act}}(t) &\approx r_0 - b(t) + \frac{1}{2}\frac{v^2t^2}{(r_0 - b(t))} \\ &\approx r_0 - b(t) + \frac{1}{2}\frac{v^2t^2}{r_0}\left(1 + \frac{b(t)}{r_0}\right). \end{aligned} \tag{33}$$

Next, note that, in the motion-compensation step, we add the value of the baseline to this to form the motion-compensated range history, resulting in

$$r_{\text{mocomp}}(t) = r_0 + \frac{1}{2}\frac{v^2t^2}{r_0}\left(1 + \frac{b(t)}{r_0}\right). \tag{34}$$

Comparing with the reference range history earlier shows that the Doppler rate that is needed to focus the echo is the same as the reference rate, scaled by a factor that depends on the ratio of the motion-compensation baseline to the range. In other words, since the scatterer is moved to a different range than its original location, the focus must be corrected to account for this distortion. The correction factor does depend on the varying baseline with time, but in practice, we find that using a constant value for each processed patch of data is sufficiently precise for many radar satellites.

To calculate the change in f_{rate} for the full geometry rather than the simplified case of Fig. 5, note that we

can write

$$\begin{aligned} f_{\text{rate}} &= \frac{2}{\lambda}\left[\frac{\dot{r}^2}{r} - \dot{r}\frac{\cos\beta}{\sin\beta}\dot{\beta}\right] \\ &= \frac{2}{\lambda}\left[\frac{\lambda^2 f_D^2}{4r} + \frac{\lambda}{2}\frac{f_D}{\tan\beta}\dot{\beta}\right] \\ &= \frac{\lambda f_D^2}{2r} - \frac{2v^2}{\lambda r}\frac{r_c}{(h+r_c)}\cos\alpha. \end{aligned} \tag{35}$$

Thus, the difference in f_{rate} for the scatterer at its original position (primed coordinates) and its motion-compensated position is

$$\begin{aligned} f_{\text{rate}} - f'_{\text{rate}} &= \frac{\lambda f_D^2}{2}\left(\frac{1}{r} - \frac{1}{r'}\right) - \frac{2v^2}{\lambda}\frac{r_c}{(h+r_c)} \\ &\quad \times \left(\frac{\cos\alpha}{r} - \frac{\cos\alpha'}{r'}\right) \\ &= \frac{\lambda f_D^2}{2}\left(\frac{r'-r}{rr'}\right) - \frac{2v^2}{\lambda}\frac{r_c}{(h+r_c)} \\ &\quad \times \left(\frac{r'\cos\alpha - r\cos\alpha'}{rr'}\right) \\ &\approx \frac{\lambda f_D^2}{2r}\frac{b(t)}{r} - \frac{2v^2}{\lambda r}\frac{r_c}{(h+r_c)}\cos\alpha\frac{b(t)}{r} \\ &= f_{\text{rate}}\frac{b(t)}{r} \end{aligned} \tag{36}$$

which is the same relation that we had in the simplified case for $r = r_0$ and which holds under the same approximation of slowly changing $b(t)$ for each patch and for $h \approx h'$.

A second focus correction factor is required as well to compensate for the phase added to each radar echo during motion compensation. Recall that each echo has been altered by a range-varying phase of the form given in (30). Due to range migration, this phase varies as a function of the range migration distance for each scatterer, so that every scatterer has a range-dependent phase added to its phase history. When the phase history is reconstructed during processing to form the matched filter, this range-dependent term is still present. Thus, we must remove this phase term in order to focus the image properly.

This additional phase is introduced in the motion-compensation step because each echo is repositioned in range by a distance defined by the actual and reference orbit locations. In the motion-compensation step, the phase is advanced by an amount that corresponds to the distance of the scatterer at the location defined by the antenna boresight. Of course, for most of the range history, the scatterer is at a different distance from the antenna. Thus, the phase that is applied in motion compensation is only approximately correct over the phase history. Since we can calculate how much extra phase is added to the radar echoes at each position in the synthetic aperture, we remove that extra phase in this step by applying the second focus correction factor.

The magnitude of the phase correction depends on the amount of range migration for each scatterer at each point in time, and in the time-domain signal, echoes from many scatterers at differing azimuth locations are present at each azimuth position. Thus, we cannot apply a single correction term to the time-domain signal. However, if we consider, instead, range migration as a function of frequency after applying a Fourier transform in the azimuth direction, we can apply a single correction to all scatterers at the same reference range simultaneously, which is analogous to the range migration resampling needed for range-Doppler processing. Since in the frequency domain we can represent the range migration as

$$r_{\text{migration}} = \frac{\lambda}{4\pi}\cdot\pi\cdot\frac{1}{f_{\text{rate}}}f^2 \tag{37}$$

we can apply the correction based on the idea that the range history is a function of the Doppler frequency. The required correction phase is the product of the migration distance and the gradient in the range of the motion-compensation phase

$$\begin{aligned} \phi_{\text{correction}} &= r_{\text{migration}}\cdot\frac{\partial}{\partial r}\left(\frac{4\pi}{\lambda}(r'(r)-r)\right)\Big|_{r=r_0} \\ &= \frac{\lambda}{4\cdot f_{\text{rate}}}\cdot f^2\cdot\frac{\partial}{\partial r}\left(\frac{4\pi}{\lambda}(r'(r)-r)\right)\Big|_{r=r_0}. \end{aligned} \tag{38}$$

These two phase corrections suffice to focus the image properly.

VI. Summary of the Processing Steps

In summary, our radar image generation steps include the following: Select a circular reference orbit, range compress each echo, apply motion compensation to move each echo to the reference track, Fourier transform the data in the azimuth direction, form the azimuth matched filter whose quadratic term reflects the motion-compensation baseline, remove the residual azimuth phase that results from the motion-compensation step, and inverse Fourier transform the data in azimuth. This produces the single-look complex data set that is needed for subsequent analysis.

VII. Limitations

Our approach for focusing the radar images will be less accurate under several different conditions, which must be assessed for each radar system configuration. These are the following.

1) If the motion-compensation baseline varies significantly over a single patch of raw data, the motion-compensation focus correction will not be correct everywhere. In our implementation, we assume that a single baseline is representative for the entire patch for the focus correction. While each echo is shifted in position for the instantaneous value of the baseline, focus correction is applied to the entire patch at once, which is a consequence of its frequency-domain implementation. If the variation in baseline is such that the chirp rate varies by more than about one part in the azimuth time–bandwidth product over the patch, some defocusing will occur.

2) If the baseline is not measured accurately enough, i.e., the trajectory is not known well enough, then the resulting interferogram will exhibit phase artifacts that are related to the error in position. For spaceborne sensors, which tend to be quite stable, this is not a significant problem unless the errors are very large. However, for airborne systems where there is a great deal of platform motion on short time scales, the phase artifacts will be quite visible and may mask the underlying desired phase signature.
3) Finally, if the orbit trajectory is such that there are very large velocities in the c or h direction, our estimates of the Doppler frequency will differ significantly for the actual and reference orbit cases. Processing the data at the incorrect Doppler centroid leads to defocusing and loss of signal. In our current implementation, we simply use the Doppler centroid as estimated from the raw data. Future implementations could include a refined Doppler estimation using the orbit data to avoid this problem.

Despite these known limitations, our method works quite well for existing spaceborne radar systems over the range of wavelengths and resolutions used in today's environmental radars.

VIII. Interferogram Formation

The aforementioned steps lead to well-focused single-look complex SAR images with known coordinates for each point. The next step in most InSAR processors is the formation of the interferogram from a pair of these images. Interferogram formation is particularly simple if the coordinates of the two single-look images coincide, eliminating the difficult and time-consuming resampling step. Since we are free to choose any reference orbit for each image, selecting the same reference for both images of the InSAR pair leads directly to coincident images. Typically, we choose, as a reference, an orbit at the average height of the two scenes, with the average heading of the two scenes, and an along-track spacing set by the average velocity of the two scenes.

The processing equations as presented earlier locate the pixels, assuming that the planet surface is a perfect sphere with no topography. Thus, the images do not quite align perfectly, and offsets of up to a pixel are common. In addition, propagation delays through the ionosphere and troposphere are not yet accounted, leading to additional errors in pixel location. Thus, we apply a resampling based on image cross correlation to align the images optimally. However, because the misposition error is small, typically a pixel or two, this step is efficient, and the interferogram formation may be implemented without detailed topographic or propagation medium delay knowledge.

For more advanced processing methods, such as time series analysis, persistent scatterers [32], [33], [38]–[42], or small baseline analysis [34], [43], many images are required, rather than a single interferogram pair. In these cases, we still choose a single reference orbit based on the collection of scenes to be combined. Selection of an orbit that approximates an average of all of the orbits used is somewhat arbitrary but straightforward. For many applications, the exact reference orbit used is unimportant as long as the same orbit is used for all scenes.

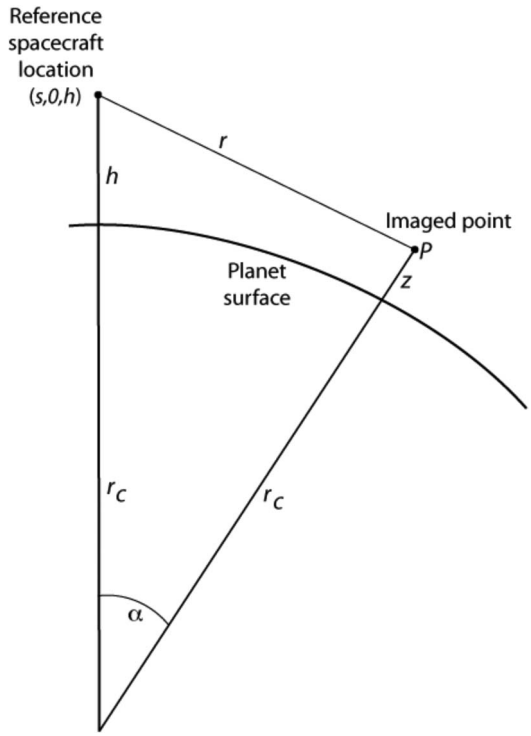


Fig. 6. Geometric construction for topographic correction. Imaged point P actually lies at an elevation z above the reference surface of a sphere of radius r_c. The Earth-centered angle α and the spacecraft height h are the same as defined in Fig. 1.

IX. Topographic Correction

The interferogram formed, as described in the previous section, does not contain the background phase pattern due to the general curvature of the Earth surface, since the motion-compensation method generates an effective InSAR baseline of zero for scatterers located on the surface. However, since the imaged area generally has topographic relief, the topographic phase contribution is still present. Therefore, for deformation applications, we must compensate for this phase term so that the "flattened" interferogram has only signals that are related to surface or propagation medium change.

Because the digital topographic data are often available for our study areas, we use the two-pass [1], [2] method for topographic phase compensation. In this method, we compute the latitude and longitude for each radar pixel, retrieve the elevation for that location from a digital elevation model (DEM), and compute and subtract the phase associated with a pixel at that elevation.

There is no closed-form solution to yield latitude, longitude, and elevation from range and azimuth radar coordinates, so we have developed an iterative approach that converges quickly to compute the elevation and location for each pixel. Consider the geometric construction of Fig. 6 in the following.

We initially let the scatterer height z be equal to zero, although another initial estimate will suffice as well. Given the reference orbit height, we can solve for the Earth-centered angle α as

$$\cos\alpha = \frac{(h+r_c)^2 + (r_c+z)^2 - \rho^2}{2(h+r_c)(r_c+z)}. \tag{39}$$

We next compute the sch coordinates of the pixel. The along-track coordinate s is related to the s coordinate of the satellite $s_{\text{satellite}}$ in the reference orbit coordinates by

$$s = s_{\text{satellite}} + r_c \tan^{-1}\left(\frac{f_d(r_c+h)\lambda r}{v\,(r_c^2+(h+r_c)^2-r^2)}\right) \quad (40)$$

where the second term on the right is the along-track distance for a pixel of the Doppler shift f_d. In our implementation, we compute the single-look complex images in a "skewed" geometry, so a range line of data corresponds to a constant squint direction so that only a single InSAR baseline vector is needed at each range line. This simplifies the bookkeeping requirements for topographic correction, but one could use deskewed images in which case the second term in (40) is not needed. The c and h coordinates follow from

$$c = -r_c \cos^{-1}\left(\frac{\cos\alpha}{\cos\beta}\right) \quad (41)$$

$$h = z \quad (42)$$

where β is the same as that defined in Fig. 1. Given an estimate of the pixel coordinates in the sch system, we convert the location to latitude/longitude/height coordinates. With this estimate of latitude and longitude, we then retrieve the elevation of the location from the DEM. This becomes our new estimate of z, and we repeat the process (39)–(42) to refine the estimate. We can iterate until the sequence converges, which typically takes two or three iterations. At the convergence of the iterative loop, we have the latitude, longitude, and elevation for each point in the image.

Finally, given the elevation of the pixel, we evaluate the phase expected from moving a scatterer from the reference sphere to its true elevation as

$$\phi_{\text{elevation}} = \frac{4\pi}{\lambda}\left(u_{\text{line-of-sight}}^{\text{elevation}} - u_{\text{line-of-sight}}^{\text{zero height}}\right) \bullet b(t) \quad (43)$$

where the u's are the unit vectors to the pixel at elevation and on the reference sphere, respectively, and $b(t)$ is the InSAR baseline vector. Subtracting this phase from the interferogram at each point removes the topographic signature, leaving only the deformation and propagation variation phases.

Errors in orbit determination, plus unmodeled delays in the propagation medium, can lead to slight distortions in the topographically corrected interferograms. Thus, we further correct the images by registering them to the DEM transformed into radar coordinates. This registration corrects for additional pixel-scale shifts in the interferograms to yield a more geodetically precise result. Typical final registration shifts observed here are on the order of one pixel in the range direction and one to three pixels in the azimuth direction. The results for the sample ALOS data sets are tabulated in the following geodetic accuracy section (Table II).

X. Geocoding

The final step that we apply in data processing is a resampling of the terrain-corrected interferogram onto an orthorectified grid. The interferogram before this stage is still sampled in a uniform radar sensor coordinate system. It is essentially a range-Doppler coordinate system image where the along-track distance is expressed in meters rather than the Doppler frequency. The coordinates for each pixel are known in absolute position, as the transformation between the sch and absolute coordinates in an Earth-fixed rotating frame (which we refer to as the xyz coordinates) is well characterized once the topographic correction is applied. It is nonetheless more useful to resample the images to uniformly sampled latitude–longitude or Universal Transverse Mercator coordinates so that the data are more easily related to other data types. Our algorithms are not unusual; however, we are currently using a nearest neighbor interpolation during this step to avoid amplitude artifacts that follow from nonband-limited multilook interferogram data. This is an implementation choice, and if desired, the full single-look complex images may be resampled using the steps outlined earlier. We have chosen to work with multilook data to minimize disk and memory requirements.

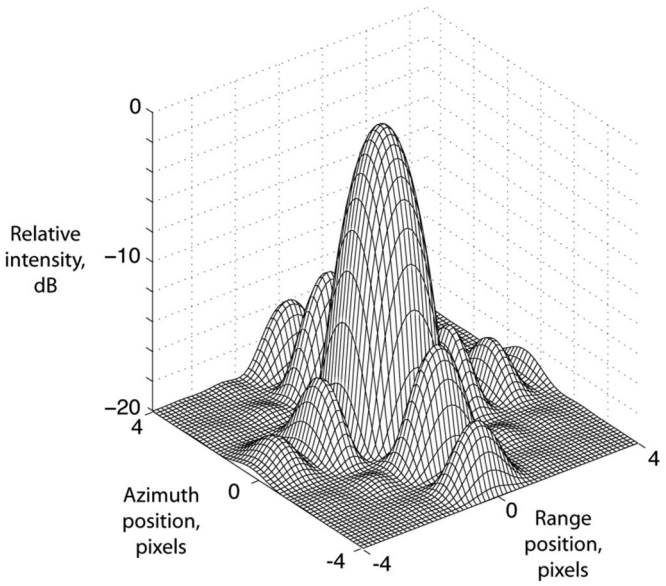


Fig. 7. Impulse response of the SAR processing module for the simulated ALOS data in a nominal geometry with an incidence angle of 34°. We computed the theoretical echoes for a point target corresponding to a 1500-m motion-compensation baseline and an InSAR baseline of about 3000 m. The half-power widths are 5.3 and 4.0 m in the slant range and in azimuth, respectively.

XI. Geodetic Accuracy and Examples

In this section, we present several images from our new processing system to illustrate its performance, especially as regards geodetic accuracy. Fig. 7 shows the impulse response of the SAR compression module for a simulated ALOS echo, where we have assumed an InSAR baseline of 3000 m, corresponding a motion-compensation baseline for each image of about 1500 m. The measured widths at half-power of the impulse response are roughly 5.3 m in the range dimension and 4.0 m in the azimuth dimension; theoretically, we would expect 5.35 m in the slant range and around 5 m (half the antenna length) in azimuth. The azimuth resolution is finer than we

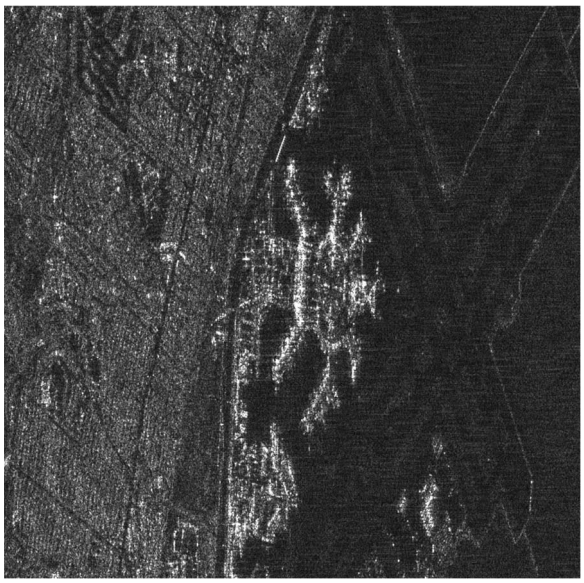


Fig. 8. Single-look complex image of the San Francisco airport as seen in the ALOS satellite data. The L-band wavelength leads to very dark runways at the right-hand side of the image, while the terminal structures stand out well. The slant range and azimuth pixel spacings of the image are 4.7 and 3.2 m, respectively. The image is well focused at this scale. Range (across) artifacts include sidelobes from the bright reflectors in the terminal area and interfering L-band signals that are visible over the darker parts of the image.

might expect as our simulator computes a phase history over a longer interval than the antenna illuminates.

In Fig. 8, we display a portion of a single-look complex image of the San Francisco airport from the ALOS satellite data acquired on February 28, 2008. The orbit and frame designators for this scene are 11162 and 740. The slant range (across) and azimuth (vertical) pixel spacings are 4.7 and 3.2 m, respectively. The resolution of the processed image is roughly one pixel, i.e., the image shows no appreciable blurring at this scale. Range sidelobes are visible from the strong reflectors on the airport terminal buildings. The horizontal artifacts that are visible mainly over the darker regions of the image are due to the L-band interference signals, which are many in the region. We have not filtered these interfering signals in this processing.

We assess the geodetic accuracy of our system in two ways: by processing an image containing known survey markers and by evaluating the shift in position between our images and reference data from existing DEMs. Three corner reflectors were installed in the Piñon Flat area in southern California by investigators at the Scripps Institution of Oceanography, University of California, San Diego, La Jolla [44]. One of these was aligned to return echoes in the direction of the ALOS satellite on its orbit track 213 in frame 660, which is an ascending orbit. We processed an interferogram on this track from orbits 7588 and 8259, acquired on June 28 and August 13, 2007. Two other reflectors were aligned with the descending orbit track, and we processed data from orbits 9360 and 10031, October 27 and December 12, 2007, track 534, and frame 2940. For all of these reflectors, we measured the inferred location from the interferograms and compared the results to a GPS ground survey done by scientists at Scripps. In Table I, we summarize our corner reflector location measurements from the

TABLE I
PIÑON FLAT CORNER REFLECTOR LOCATIONS

Measurement	Latitude (deg)	Longitude (deg)	Latitude error (m)	Longitude error (m)
Reflector aligned with ascending orbit				
InSAR location, unregistered image	33.61233	-116.4570	9	-18
InSAR location, registered image	33.61215	-116.4567	-11	9
Ground GPS survey	33.61225	-116.4568	--	--
Reflectors aligned with descending orbit				
InSAR location, unregistered image	33.61215	-116.4579	-11	0
InSAR location, registered image	33.61213	-116.4577	-13	18
Ground GPS survey	33.61225	-116.4579	--	--
InSAR location, unregistered image	33.60729	-116.4517	-9	9
InSAR location, registered image	33.60727	-116.4516	-11	18
Ground GPS survey	33.60737	-116.4518	--	--

ALOS data and from the Scripps ground survey. The top line in each section of the table gives the observed location from our processor before alignment with a DEM, which is the "dead-reckoning" result. The middle line gives the location after a cross-correlation registration with a DEM, while the third line gives the location as determined from the ground geodetic survey. The disagreements here are on the order of 10–15 m, and the corner reflector is imaged with similar accuracy with and without registration to the DEM. Note that these results are quantized to the pixel spacing, though, because we use the nearest neighbor interpolation algorithm in this step of the implementation. We believe that this result is better than our typical accuracy across the entire image, however, as we discuss further in the following.

We present several processed interferogram data sets from ALOS measurements in Figs. 9–12. Fig. 9 shows an interferogram formed from the data acquired over southern California, centered over the town of Ventura. The image center latitude

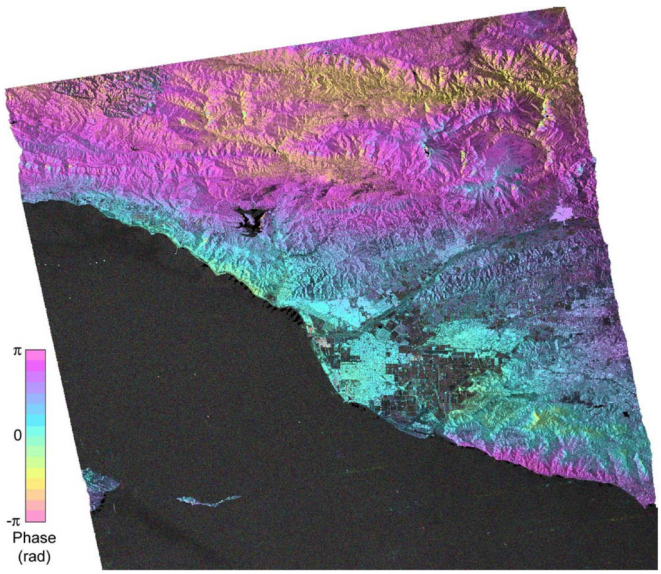


Fig. 9. ALOS interferogram of the Ventura area from the data acquired on June 22 and September 21, 2007. The phase signature is likely the variability of the atmosphere on the two days. The InSAR baseline is 100 m, and the illumination is from the left. The spacecraft motion is from south to north.

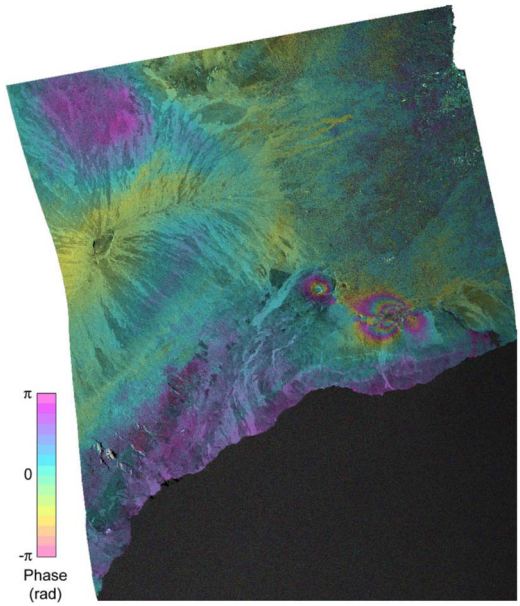


Fig. 10. ALOS interferogram from May 5 to June 20, 2007, over the Kilauea region of the island of Hawaii. An intrusion occurred in June 17–19 along the East Rift of the Kilauea Volcano 19° 25′ N, 155° 18′ W) and produced the clear crustal deformation phase signal. Additional phase variation, which is visible at the top portion and elsewhere on Mauna Loa, plus some signals along the coast, most likely results from atmospheric change. Illumination is from the left. The InSAR baseline is about 200 m.

and longitude are roughly 34° N, 119° 15′ W, and the data were acquired on June 22, 2007 (orbit 7486 and frame 670) and September 21, 2007 (orbit 8828). In this interferogram, we see that there is a phase signature that is locally correlated with topography. However, globally, the phase does not depend on elevation, so we speculate that this is predominantly an atmospheric signature. The magnitude of the phase is comparable to previous reports of tropospheric phase in interferograms

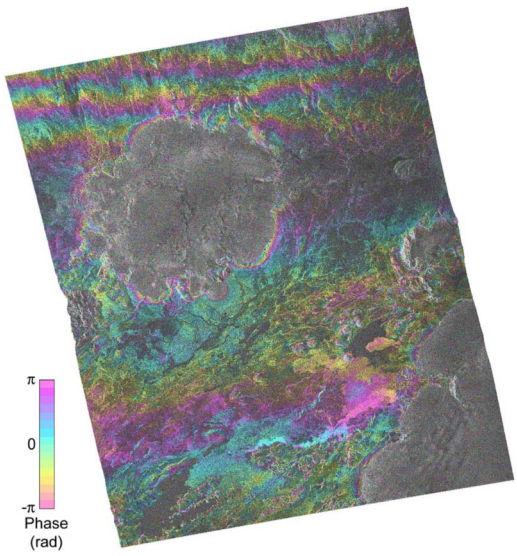


Fig. 11. ALOS interferogram of an area in Iceland, showing two glaciers and strong ionospheric artifacts. The change in the total electron content in the ionosphere is directly proportional to the phase delay and is visible as phase "bars." The gradient in electron content is higher at the north than in the south. Data were acquired on September 2 and October 18, 2007, on orbits 8561 and 9232 (frame 1290). The InSAR baseline is about 300 m, and again, illumination is from the left.

[45]–[49]. The spacecraft moves from south to north, and the illumination is from the left.

In Fig. 10, we present an ALOS interferogram from orbits 6802 and 7473 (May 5 and June 20, 2007) over the Kilauea region of the island of Hawaii. Illumination again is from the left. In this image, there is a clear crustal deformation signal from an intrusion in June 17–19 along the East Rift of the Kilauea Volcano. The intrusion along the rift is accompanied by a deflation at the Kilauea Caldera. Additional phase signals, such as at the topmost portion of Mauna Loa and elsewhere on this volcano, plus some signals along the coast, most likely result from atmospheric change.

A different sort of artifact is visible in Fig. 11, which is an ALOS interferogram of glaciated terrain in Iceland (image center is approximately 64° 40′ N, 18° 30 W). The illumination is again from the left. This image shows the phase "bars" that are aligned roughly with the range direction, mainly in the top part of the image, but are also visible in the southern third of the image. The northern artifacts are narrower than those in the south. We speculate that these phase patterns are due to the variations in the ionosphere rather than the troposphere, because they are also associated with the azimuth pixel shifts that would result from the gradients in the ionospheric electron content. The pixel shifts are most easily seen in the correlation image (see Fig. 12), where they cause similar bars of decorrelation as the two single look complex images do not align well. Tropospheric phase patterns would not be significant here because the surface temperature is low, so that the partial pressure of water vapor, which is responsible for most of the variable atmospheric signal [48], is very low. The interferogram decorrelates significantly over the glaciers near the image center and the southeast corner of the image, likely due to surface melt or motion.

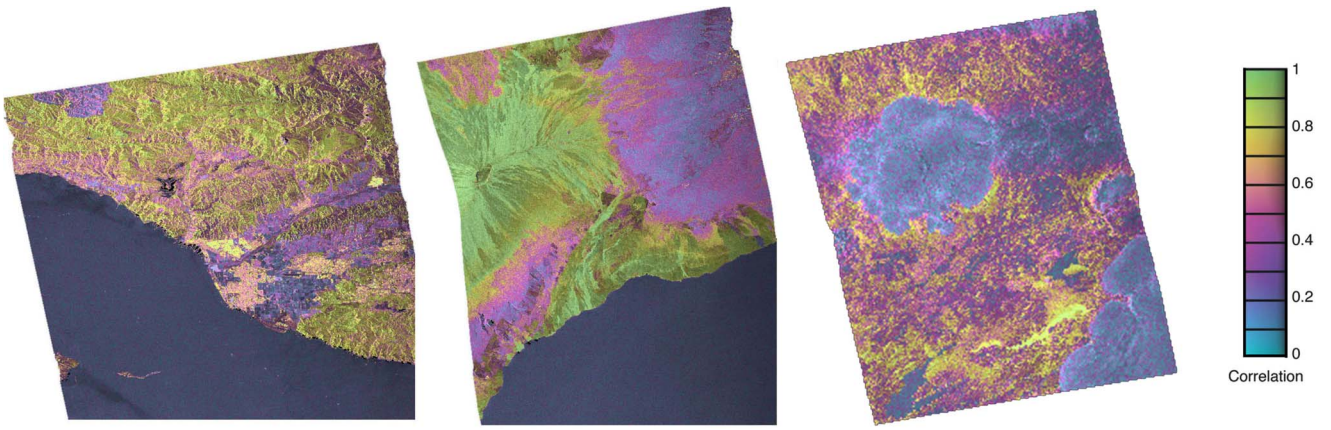


Fig. 12. Correlation images for the three scenes shown in Figs. 9–11. The Ventura image is at the left, Hawaii is at the center, and Iceland is at the right. Correlations range from near unity to essentially uncorrelated; hence, they are a representative of surface effects. The decorrelation band running across the Iceland image over the upper glacier is probably from pixel misposition due to electron density gradients in the ionosphere.

TABLE II
IMAGE REGISTRATION OFFSETS

Scene	Range offset at center (m)	Azimuth offset at center (m)	Additional stretch Range (m)	Azimuth (m)
Ventura	-15.8	18.2	9.4	15.2
Hawaii	-21.5	24.0	14.1	25.4
Iceland	2.0	2.9	44.0	29.4

The correlation images for all three scenes of Figs. 9–11 are shown in Fig. 12. The correlations are generally high, except over water and over the glaciers in the Iceland image. The correlation drops over the vegetated area in the Hawaii image due to surface change as the vegetation changes over time. The Iceland image, in addition, shows a significant band of decorrelation near the northern glacier where there is the greatest shift in azimuth position due to electron content gradients in the ionosphere.

We can further characterize the geodetic accuracy of the processing system by examining the image registration shift needed to align the processed interferograms with the DEM that is used to correct the interferogram for elevation. In Table II, we list the offsets in range and azimuth (in meters), which are required to align the image with the DEM. We also list in the right-hand columns the additional amount of stretch that is required to align the remainder of the points with the DEM. This extra stretch tends to be great at the image corners. We can see that the image-center offset ranges from −22 to 24 m, and the additional image distortion that is needed to align with a reference DEM can be up to 44 m at one corner of the Iceland image.

XII. CONCLUSION

Modern satellite orbit determination produces trajectories that are extremely accurate. A precise knowledge of the satellite position enables processing of the InSAR images that are accurate in position at the 10-m level and are focused to the pixel level without requiring autofocus procedures. We have developed a processing approach that capitalizes on accurate orbit information to implement an efficient and robust InSAR processing package of software.

We have used the motion-compensation algorithms to propagate the raw radar echoes from their initial locations to a reference orbit that is chosen to simplify the pixel location and to focus equations, so that the implementation is both accurate and efficient. By choosing a single reference orbit for a collection of radar passes, pairs of scenes for interferogram formation or stacks of scenes as needed for persistent scattering or small baseline subset analysis are all produced in the same coordinate system so that the coregistration of the scenes is very easy, and it does not require the detailed image matching that haunts many InSAR processing runs. Motion compensation introduces two phase terms in the scatterer phase histories, which require correction in the processor, but these are easily applied using frequency-domain methods. The resulting single-look complex radar images are very well focused using only the orbit information for the radar satellite.

We have processed several radar images from the data acquired by the ALOS PALSAR instrument and L-band radar satellite. We have first assessed the geodetic accuracy by comparing the observed locations of a set of radar corner reflectors located in the Piñon Flat area in California. Corner reflector positions were accurate at the 10–20-m level in our images. We have also processed interferograms from California, Hawaii, and Iceland and calculated the image location error by coregistering the images with DEMs. These images also showed 10-m errors in position. We have also found out that the images had to be stretched up to 40 m so that all points in the radar scene matched the locations in the elevation model data. We speculate that these offsets are mainly due to unmodeled ionospheric and tropospheric effects or other unknown instrumental errors. Nonetheless, the data products are sufficiently accurate for many geophysical surface studies.

The overall set of processing equations may be implemented efficiently on modern multicore desktop computers so that, combined with the robustness of the approach, a reliable desktop generation of interferograms on cheap hardware is realized.

REFERENCES

[1] D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, and T. Rabaute, "The displacement field of the Landers earthquake mapped by radar interferometry," *Nature*, vol. 364, no. 6433, pp. 138–142, Jul. 1993.

[2] H. A. Zebker, P. A. Rosen, R. M. Goldstein, A. Gabriel, and C. Werner, "On the derivation of coseismic displacement fields using differential radar interferometry: The Landers earthquake," *J. Geophys. Res.—Solid Earth*, vol. 99, no. B10, pp. 19 617–19 634, Oct. 10, 1994.

[3] G. Peltzer, P. Rosen, F. Rogez, and K. Hudnut, "Postseismic rebound in fault step-overs caused by pore fluid flow," *Science*, vol. 273, no. 5279, pp. 1202–1204, Aug. 30, 1996.

[4] W. Thatcher and D. Massonnet, "Crustal deformation at Long Valley Caldera, eastern California, 1992–1996 inferred from satellite radar interferometry," *Geophys. Res. Lett.*, vol. 24, no. 20, pp. 2519–2522, 1997.

[5] C. Wicks, W. Thatcher, and D. Dzurisin, "Migration of fluids beneath Yellowstone Caldera inferred from satellite radar interferometry," *Science*, vol. 282, no. 5388, pp. 458–462, Oct. 1998.

[6] F. Amelung, S. Jonsson, H. A. Zebker, and P. Segall, "Widespread uplift and trapdoor faulting on Galapagos volcanoes observed with radar interferometry," *Nature*, vol. 407, no. 6807, pp. 993–996, Oct. 26, 2000.

[7] S. Jonsson, H. Zebker, P. Segall, and F. Amelung, "Fault slip distribution of the 1999 Mw 7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements," *Bull. Seismol. Soc. Amer.*, vol. 92, no. 4, pp. 1377–1389, May 2002.

[8] M. Pritchard and M. Simons, "A satellite geodetic survey of large-scale deformation of volcanic centers in the central Andes," *Nature*, vol. 418, no. 6894, pp. 167–171, Jul. 2002.

[9] R. M. Goldstein, H. Engelhardt, B. Kamb, and R. M. Frolich, "Satellite radar interferometry for monitoring ice sheet motion: Application to an Arctic ice stream," *Science*, vol. 262, no. 5139, pp. 1525–1530, Dec. 3, 1993.

[10] E. Rignot, K. C. Jezek, and H. G. Sohn, "Ice flow dynamics of the Greenland ice sheet from SAR interferometry," *Geophys. Res. Lett.*, vol. 22, no. 5, pp. 575–578, 1995.

[11] I. R. Joughin, R. Kwok, and M. A. Fahnestock, "Interferometric estimation of three-dimensional ice-flow using ascending and descending passes," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 25–37, Jan. 1998.

[12] E. Rignot, "Fast recession of an Antarctic Glacier," *Science*, vol. 281, no. 5376, pp. 549–551, Jul. 1998.

[13] R. R. Forster, E. Rignot, B. L. Isacks, and K. Jezek, "Interferometric radar observations of Glaciares Europa and Penguin, Hielo Patagonico Sur, Chile," *J. Glaciol.*, vol. 45, no. 150, pp. 325–337, 1999.

[14] E. Hoen and H. A. Zebker, "Penetration depths inferred from interferometric volume decorrelation observed over the Greenland ice sheet," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 6, pp. 2571–2583, Nov. 2000.

[15] D. L. Galloway, K. W. Hudnut, S. E. Ingebritsen, S. R. Phillips, G. Peltzer, F. Rogez, and P. A. Rosen, "Detection of aquifer system compaction and land subsidence using interferometric synthetic aperture radar, Antelope Valley, Mojave Desert, California," *Water Resour. Res.*, vol. 34, no. 10, pp. 2573–2585, 1998.

[16] D. E. Alsdorf, J. M. Melack, T. Dunne, A. K. Mertes, L. L. Hess, and L. C. Smith, "Interferometric measurements of water level changes on the Amazon flood plain," *Nature*, vol. 404, no. 6774, pp. 174–177, Mar. 2000.

[17] J. Hoffmann, H. A. Zebker, D. L. Galloway, and F. Amelung, "Seasonal subsidence and rebound in Las Vegas Valley, Nevada, observed by synthetic aperture radar interferometry," *Water Resour. Res.*, vol. 37, no. 6, pp. 1551–1566, 2001.

[18] R. N. Treuhaft and P. R. Siqueira, "Vertical structure of vegetated land surfaces from interferometric and polarimetric radar," *Radio Sci.*, vol. 35, no. 1, pp. 141–177, 2000.

[19] K. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2352–2363, Nov. 2001.

[20] J. Askne, M. Santoro, G. Smith, and J. E. S. Fransson, "Multitemporal repeat-pass SAR interferometry of boreal forests," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1540–1550, Jul. 2003.

[21] H. A. Zebker and R. M. Goldstein, "Topographic mapping from interferometric synthetic aperture radar observations," *J. Geophys. Res.*, vol. 91, no. B5, pp. 4993–4999, 1986.

[22] J. O. Hagberg and L. M. H. Ulander, "On the optimization of interferometric SAR for topographic mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 1, pp. 303–306, Jan. 1993.

[23] J. Moreira, M. Schwabisch, G. Fornaro, R. Lanari, R. Bamler, D. Just, U. Steinbrecher, H. Breit, M. Eineder, G. Granceschetti, D. Geudtner, and H. Rinkel, "X-SAR interferometry: First results," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 950–956, Jul. 1995.

[24] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf, "The Shuttle Radar Topography Mission," *Rev. Geophys.*, vol. 45, no. 2, p. RG2 004, 2007.

[25] B. D. Tapley, J. C. Ries, G. W. Davis, R. J. Eanes, B. E. Schultz, C. K. Shum, M. M. Watkins, J. Marshall, A. Nerem, and R. S. Putney, "Precision orbit determination for TOPEX/POSEIDON," *J. Geophys. Res.*, vol. 99, no. C12, pp. 24 383–24 404, Dec. 1994.

[26] R. Scharroo and P. Visser, "Precise orbit determination and gravity field improvement for the ERS satellites," *J. Geophys. Res.*, vol. 103, no. C4, pp. 8113–8128, 1998.

[27] P. H. Andersen, K. Aksnes, and H. Skonnord, "Precise ERS-2 orbit determination using SLR, PRARE, and RA observations," *J. Geodesy*, vol. 72, no. 7/8, pp. 421–429, 1998.

[28] S. Katagiri and Y. Yamamoto, "Technology of precise orbit determination," *Fujitsu Sci. Tech. J.*, vol. 44, no. 4, pp. 401–409, Oct. 2008.

[29] J. J. Mohr and S. N. Madsen, "Geometric calibration of ERS satellite SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 4, pp. 842–850, Apr. 2001.

[30] D. T. Sandwell, D. Myer, R. Mellors, M. Shimada, B. Brooks, and J. Foster, "Accuracy and resolution of ALOS interferometry: Vector deformation maps of the Father's Day Intrusion at Kilauea," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, pt. 1, no. 11, pp. 3524–3534, Nov. 2008.

[31] D. T. Sandwell and L. Sichoix, "Topographic phase recovery from stacked ERS interferometry and a low resolution digital elevation model," *J. Geophys. Res.*, vol. 105, no. B12, pp. 28 211–28 222, Dec. 2000.

[32] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[33] A. Ferretti, C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, pt. 1, no. 5, pp. 2202–2212, Sep. 2000.

[34] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002.

[35] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar*. New York: Wiley-Interscience, 1991.

[36] S. M. Buckley, "Radar interferometry measurement of land subsidence," Ph.D. Dissertation, Univ. Texas, Austin, TX, 2000.

[37] C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Synthetic Aperture Radar: A Processing Approach*. Boston, MA: Kluwer, 1996.

[38] A. Hooper, P. Segall, and H. Zebker, "Persistent scatterer interferometric synthetic aperture radar for crustal deformation analysis, with application to Volcan Alcedo, Galapagos," *J. Geophys. Res.*, vol. 112, no. B7, p. B07 407, 2007.

[39] S. Lyons and D. Sandwell, "Fault creep along the southern San Andreas from interferometric synthetic aperture radar, permanent scatterers, and stacking," *J. Geophys. Res.*, vol. 108, no. B1, p. 2047, 2003.

[40] C. Colesanti, A. Ferretti, C. Prati, and F. Rocca, "Monitoring landslides and tectonic motions with the permanent scatterers technique," *Eng. Geol.*, vol. 68, no. 1/2, pp. 3–14, Feb. 2003.

[41] G. E. Hilley, R. Bürgmann, A. Ferretti, F. Novali, and F. Rocca, "Dynamics of slow-moving landslides from permanent scatterer analysis," *Science*, vol. 304, no. 5679, pp. 1952–1955, Jun. 2004.

[42] P. Shanker and H. Zebker, "Persistent scatterer selection using maximum likelihood estimation," *Geophys. Res. Lett.*, vol. 34, no. 22, p. L22 301, 2007.

[43] R. Lanari, F. Casu, M. Manzo, G. Zeni, P. Berardino, M. Manuta, and A. Pepe, "An overview of the small baseline subset algorithm: A DInSAR technique for surface deformation analysis," *Pure Appl. Geophys.*, vol. 164, no. 4, pp. 637–661, Apr. 2007.

[44] D. Sandwell, personal communication.

[45] H. A. Zebker, P. A. Rosen, and S. Hensley, "Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps," *J. Geophys. Res.*, vol. 102, no. B4, pp. 7547–7563, 1997.

[46] R. F. Hanssen, T. M. Weckwerth, H. A. Zebker, and R. Klees, "High-resolution water vapor mapping from interferometric radar measurements," *Science*, vol. 283, no. 5406, pp. 1297–1299, Feb. 1999.

[47] R. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*. Dordrecht, Netherlands: Kluwer, 2001.

[48] F. Onn and H. A. Zebker, "Correction for interferometric synthetic aperture radar atmospheric phase artifacts using time series of zenith wet delay observations from a GPS network," *J. Geophys. Res.*, vol. 111, no. B9, p. B09 102, Sep. 23, 2006.

[49] J. Foster, B. Brooks, T. Cherubini, C. Shacat, S. Businger, and C. L. Werner, "Mitigating atmospheric noise for InSAR using a high resolution weather model," *Geophys. Res. Lett.*, vol. 33, no. 16, p. L16 304, 2006.

Howard A. Zebker (M'87–SM'89–F'99) received the B.S. degree from the California Institute of Technology, Pasadena, in 1976, the M.S. degree from the University of California, Los Angeles, in 1979, and the Ph.D. degree from Stanford University, Stanford, CA, in 1984.

He is currently a Professor of geophysics and electrical engineering with Stanford University, where his research group specializes in radar remote sensing, especially InSAR, of the Earth and planets. He has more than 30 years of experience in spaceborne radar, beginning with the SEASAT program in the 1970s and currently with the radar science team on the Cassini mission of the National Aeronautics and Space Administration (NASA). Other satellite programs where he has contributed include NASA's Spaceborne Imaging Radar-A, B, and C; the Magellan Mission to Venus; and the Voyager Mission to the outer planets. His major research focus is on the development of radar as a technique and the application of this technique to basic Earth science. He has published over 300 scholarly articles on remote sensing, and he teaches several courses on remote sensing at Stanford University to undergraduate and graduate students alike.

Dr. Zebker is a Fellow of the Electromagnetics Academy and a member of the American Geophysical Union.

Scott Hensley received the B.S. degrees in mathematics and physics from the University of California, Irvine, and the Ph.D. degree in mathematics from the State University of New York, Stony Brook, where he specialized in the study of differential geometry.

Subsequent to graduating, he was with Hughes Aircraft Company on a variety of radar systems, including the Magellan radar. Since 1992, he has been with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, where he studies advanced radar techniques for geophysical applications. His research has involved using both the stereo and interferometric data acquired by the Magellan spacecraft at Venus and using the European Remote Sensing Satellite-1, Japanese Earth Resources Satellite-1, and Spaceborne Imaging Radar-C data for differential interferometry studies of earthquakes and volcanoes. His current research also includes studying the amount of penetration into the vegetation canopy using simultaneous L- and C-band topographic synthetic aperture radar measurements and repeat-pass airborne interferometry data collected at lower frequencies (P-band). He was the Technical Lead of the Shuttle Radar Topography Mission Interferometric Processor Development Team for a shuttle-based interferometric radar that is used to map the Earth's topography in the $\pm 60^\circ$ latitude. He is currently the Principal Investigator of the Uninhabited Aerial Vehicle Synthetic Aperture Radar program that is developing a repeat-pass radar interferometry capability for use on conventional or unmanned aerial vehicles. Recently, he began working with the Goldstone Solar System Radar to generate topographic maps of the lunar surface.

Piyush Shanker, photograph and biography not available at the time of publication.

Cody Wortham received the B.S. and M.S. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 2005 and 2007, where he focused on space-time signal processing for radar interference cancellation. Since 2007, he was been working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA.

His research interests include InSAR processing and vector time-series from multiple InSAR geometries.

User-friendly InSAR Data Products: Fast and Simple Timeseries (FAST) Processing

Howard A. Zebker, *Fellow, IEEE*

***Abstract*— Interferometric Synthetic Aperture Radar (InSAR) methods provide high resolution maps of surface deformation applicable to many scientific, engineering and management studies. Despite its utility, the specialized skills and computer resources required for InSAR analysis remain as barriers for truly widespread use of the technique. Reduction of radar scenes to maps of temporal deformation evolution requires not only detailed metadata describing the exact radar and surface acquisition geometries, but also a software package that can combine these for the specific scenes of interest. Furthermore, the range-Doppler reference frame and radar coordinate system itself are confusing, so that many users find it hard to incorporate even useful products in their customary analyses. Finally, the sheer data volume needed for interferogram time series makes InSAR analysis challenging for many analysis systems. We show here that it is possible to deliver radar data products to users that address all of these difficulties, so that the data acquired by large, modern satellite systems are ready to use in more natural coordinates, without requiring further processing, and in as small volume as possible.**

***Index Terms*—InSAR, Synthetic Aperture Radar, Data Products, Geodesy**

I. Need for user-friendly products

Using InSAR we can measure mm-cm level surface deformation over large areas at fine resolution, and it has been extensively applied in studies such as earthquake and volcano modeling [1-4], glacier mechanics [5,6], hydrology [7,8], and topographic mapping [9,10]. Modern spaceborne satellites, for example Sentinel-1A/B, provide long sequences of observations that produce very many interferograms, which in turn provide the deformation histories of many points on the surface. The InSAR technique combines interferometry and conventional synthetic aperture radar (SAR) to compute the phase differences between two single look complex (SLC) SAR images. Since the resulting interferometric phase is proportional to the change in range between two observations at different sensor locations and times and a given point on the surface, a single interferogram contains phase signals from i) the local topography due to the spatial separation of the two satellite locations and ii) any radar line of sight displacements of the point occurring between the two SAR acquisition times.

Yet preparation of InSAR data for analysis remains difficult for the non-specialist, requiring InSAR data users to be considerably familiar with the detailed SAR imaging geometry for each acquisition, and also experienced in InSAR processing techniques. In addition, differential interferometric SAR techniques for investigating temporal evolution of surface deformation, such as the small baseline subset (SBAS) [11] and persistent scatterer [12,13] approaches use a large number of SAR acquisitions and an even larger number of pairwise interferograms. The sheer data volume needed makes InSAR analysis challenging: a sequence of 100 radar acquisitions yields 4950 interferograms, and it is much easier to download the 100 SLCs rather than transfer the huge volume of all of those interferograms. Recognizing these restrictions, applying topographic corrections to all interferograms in a given analysis can require not only sophisticated software but also significant computational resources. Finally, for many would-be users of deformation time series, for example professional hydrogeologists mapping the extent and storage of underground reservoirs or engineers assessing degradation of soil underlying critical infrastructure, the range-Doppler radar coordinate system is perplexing, so that many users find it hard to ingest even useful products in their own customary analyses. Nonetheless InSAR data are valuable in the assessment of these systems. Our goal here is to make access to InSAR methods and data easy for most users, relieving them of the burden of understanding the processing details and the need for large computational resources. Supplying the products in regularly gridded coordinates removes one more barrier to the use of the radar observations.

Here we advocate delivering semi-reduced SLC data directly to users, so that those who can benefit greatly from the analytical methods can readily use data in well-defined coordinate and reference systems such as latitude/longitude or UTM. In some cases, e.g. when the surface moves greatly over time as in a glacier or near a plate boundary, a local coordinate system may be warranted. Single look complex images, if precorrected for imaging geometry, make formation of hundreds of interferograms from an observation sequence both reliable and efficient. In our approach, enabled by the 10's of cm-level accuracy of orbit solutions of today's platforms, we begin with SLC images in a zero-Doppler geometry. If the data are delivered as raw measurements, we first process them to SLC form. Whether products are delivered as unfocused data or in SLC format, such as Sentinel-1 TOPS images, the orbit accuracy supports resampling of products directly to a well-defined grid with automatic viewpoint and topographic compensation. In both cases we fully correct for the topographic phase terms so that simple cross multiplication yields the needed interferograms directly in map coordinates.

Note that there is no technical case to be made for analyzing data on a regular grid rather than range-Doppler coordinates. And there will always be those who need to process data in

other formats for particular studies. Our proposed method is no more accurate than existing implementations, and realizes similar errors from uncertainty in the orbit control and its solution, poorly characterized satellite sensors, and signal to noise limitations. Rather we are resampling the data to a different set of coordinates that facilitates ease of use and allows geocoding plus elevation and orbit phase compensations in the SLC product generation, so that the InSAR experts at the processing facilities for each sensor, and not the end users, apply the corrections.

While the main motivation for this work is to expand the use of InSAR to those less familiar with radar processing and analysis, even sophisticated users can benefit from the geocoded and compensated data products. In fact, those requiring very long time series and hence 1000's of interferograms can save the most computational time and tedious correction burden that requires many ancillary data sets. The availability of reduced SLCs enables these scientists, as well as more casual InSAR users, to readily generate temporal change images through a much lessened data volume and simplified processing flow.

II. Technical Details

A. SAR processing algorithms

Focusing a synthetic aperture radar (SAR) image requires an azimuth-direction matched filter. There are many ways to implement this, but each addresses the phase history of an imaged scatterer as the radar sensor flies along its track. The phase history quantifies the time dependence of range to the point of interest, forming the signal basis of the matched filter. A general form of the phase history is

$$\phi(t) = -{}^{4\pi}/_{\lambda}\, r(t) \qquad (1)$$

where $r(t)$ is the distance from sensor along its orbit to the scatterer as a function of time, $\phi(t)$ is the propagation part of the received signal phase, and λ is the radar wavelength. While the principles of SAR were initially developed from an antenna theory perspective [14], early mathematical models of the characteristics of the radar echo [15] and eventually digital implementations [16] thereof recognized that the matched filter could be parameterized by the Doppler frequency f_d and Doppler rate f_r:

$$f_d = -{}^{2}/_{\lambda}\,\dot{r} \qquad (2)$$

$$f_r = -{}^{2}/_{\lambda}\,\ddot{r} \qquad (3)$$

Expanding the phase history (eq. 1) in a Taylor series about a reference point t_0, generally either the point of closest approach or the peak of the antenna response, yields

$$\phi(t) = -{}^{4\pi}/_{\lambda}\left[r(t_0) + \dot{r}\,(t-t_0) + {}^{1}/_{2}\,\ddot{r}(t-t_0)^2 + \cdots\right] \qquad (4)$$

showing that the matched filter can be computed from the Doppler and Doppler rate values. Differentiating the phase history to obtain f_d and f_r loses the constant phase term, so under this approximation the matched filter does not compensate for the propagation phase constant $-\frac{4\pi}{\lambda}r(t_0)$, and each output pixel is tagged with this reference phase. Furthermore, the phase history is usually computed as if the imaged point lies on a reference elevation surface, a necessary approximation for a convolutional processing algorithm. Hence the SLC image contains, in addition to any displacement phase, terms due to signal propagation from orbit location to the scatterer at its topographic elevation. Consequently, the interferograms formed from these images must contain corrections for both of these distortions. Here we compute each SLC with these corrections applied, so that the end user does not need either the orbit location or topographic elevation knowledge in order to study the deformation history of a scene.

B. Commonly implemented InSAR algorithms

Once we have focused the images to form the SLCs, we pairwise combine them to form interferograms. In fig. 1 we depict the geometric definitions for the various corrections needed to produce an interferogram, whose phase in the lack of any deformation is simply the projection of the InSAR baseline B in the direction of the line of sight of the point imaged u_{elev}, scaled by wavelength:

$$\phi_{obs} = -{}^{4\pi}/_{\lambda}\; u_{elev} \bullet B \qquad (5)$$

Most zero Doppler SLC images are focused to a reference elevation surface (T_{ref} in fig. 1), so that the interferogram requires corrections for the 'flat-Earth' interference pattern and for the difference phase due to elevation changes. Assuming the flat-Earth correction has been applied to each interferogram either explicitly or implicitly through motion compensation processing [17], we next subtract the expected phase from an elevated point on the surface. This is the topographic correction term that 'flattens' the interferogram:

$$\phi_{topo} = -{}^{4\pi}/_{\lambda}\;(u_{elev} - u_{ref}) \bullet B \qquad (6)$$

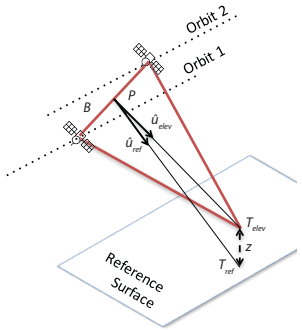


Figure 1. InSAR imaging geometry highlighting key parameters needed to create a deformation interferogram. The inner product of InSAR baseline vector B for an observation point P, the midpoint of the two satellite locations at the time of observation for a point T_{elev} at elevation z, compared to the same for the point on the reference elevation, yields the interferogram topographic correction.

After interferogram flattening, the observation reflects only the deformation signal plus phases due to atmospheric propagation and any changes in the electrical properties of the surface.

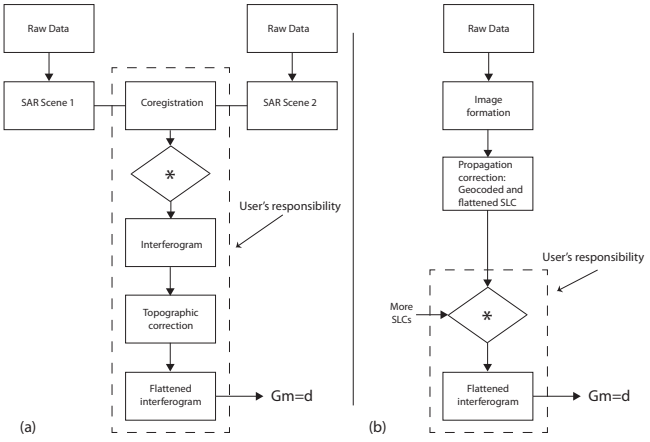


Figure 2. Left (a), an algorithm similar to those used in many common InSAR processors. Right (b), proposed data flow showing increased complexity on supplier data system but simplified steps by end user. In both figures the steps contained within the dotted box represent the operations typically carried out by the end user. In the proposed algorithm, the only InSAR expertise required is cross multiplication.

We illustrate data flow in a typical InSAR processor in fig. 2a. Data from two satellite passes, often in SLC format, are delivered to the end-user. The user's responsibility is to coregister them accurately, cross multiply the results to form an initial interferogram, correct the interferogram for topographic elevation and specific imaging geometry for each scene, and only then proceed to analyze the residual phases geophysically. These steps necessitate knowledge of the precise interferogram baseline and a digital elevation model (DEM) of the surface. This is both considerable work and requires processing expertise to produce accurate phase images.

In contrast, our proposed approach (fig. 2b) moves most of the work from the dotted box of fig. 2a upstream to the data producer, where the real expertise in InSAR product resides. Thus the user is relieved of this duty and can proceed directly with straight-forward cross-multiplication and the analysis for the investigation. Others have discussed correcting SLC phases individually and not as interferograms, for tomographic applications [18] and for general InSAR analysis [19]. Here we use that basic idea and show that it is particularly simple to apply if it is combined with the geocoding step in producing understandable radar products (see below).

The critical items for phase compensation are usually the orbit solutions and the elevation of each point on the surface as recorded in a DEM. Other factors affecting the observed phases, such as drift in the sensor electronics or flexure of the radar antenna, are generally small and can be removed through accurate calibration of the radar system. Thus traditional cal/val activities, along with attention to these system parameters, are still necessary to enable our approach to efficient processing. Other limitations of radar imaging, such as layover from the mapping of a 3D surface into 2 dimensions, are not eliminated in this approach. Rather, the data still possess features that require an understanding of the basic imaging process.

Summarizing, it is incumbent on the user to obtain the needed ancillary data sets, and understand the software package used well enough to apply the corrections properly.

C. Proposed algorithm and data flow

When the orbit is very well known, we can directly resample the natural-coordinate SLCs to a regular grid, simultaneously applying a phase correction that compensates for both the specific viewing geometry and the topographic elevation of the surface. Since we need to align images to a small fraction of a pixel in order to maintain high InSAR correlation, this method is most effective when the orbit is accurate to a few 10's of cm. Many if not most modern systems achieve this regularly.

The high-level data flow has been illustrated in fig. 2b above. The detail is straight-forward and easily implemented in most processing environments, hence we denote it FAST (fast and simple time series) processing. In our case the steps are

1. Create a zero-Doppler SLC for each scene using its particular orbit solution
2. Extract a DEM covering the region at desired postings
3. For each point in the DEM, find the corresponding zero-Doppler satellite observation location along the orbit
4. Resample the SLC to that DEM point, applying a phase correction for the propagation distance from the sensor zero-Doppler point to the elevated surface point

Steps 1 and 2 are self-explanatory. Step 3 in our system uses orbit state vectors and Newton-Raphson iteration to find the zero-Doppler orbit location for each output image point, following Agram [20]. Our code returns the satellite position vector $s_{zero}(x,y,z)$, the radar range r_0 and azimuth time t_{az} for the point located at Earth fixed coordinates p(x',y',z'). For step 4 we apply the propagation phase for distance r_0 according to

$$\phi_{prop} = {}^{-4\pi}/_{\lambda} \cdot |s_{zero}(x,y,z) - p(x',y',z')| \quad (7)$$

to adjust the phase of the interpolated SLC image. This removes all propagation phase, with both the imaging geometry and elevation distortions compensated. Since all SLCs are resampled to the same grid, they are coregistered in position. Thus formation of interferograms is accomplished using point by point cross multiplication of the geocoded SLCs.

Sentinel 1 data pose a secondary challenge resulting from the TOPS scanning that minimizes amplitude scalloping in the images. Each Sentinel-1 SLC contains not only the propagation phase, but also the TOPS scanning phase. Thus we first subtract this phase from each burst in each image so that the data can be resampled with a low-pass interpolator, apply steps 1-4 above, and finally reinsert the scanning phase after resampling to the desired output grid. At this point cross multiplication results in interferograms as with other systems.

Slight inaccuracies in burst positioning cause additional phase artifacts as the Doppler centroid of pass to pass matched pixels can vary. The very fine accuracy needed (about 0.001 pixel) can require a secondary phase compensation generally referred to as an enhanced spectral diversity correction [21], which in some cases is already known to be unnecessary [22]. Whether or not this is applied, regularly gridded SLC images allow interferogram formation by simple cross multiplication. Our experience is that often the magnitude of the enhanced spectral diversity correction is small, so that considering atmospheric phase variations of ~2 cm rms from water vapor irregularities it is often not needed.

The resampling may be chosen to account for Sentinel's anisotropic resolution to best preserve the data quality, and care must be taken to ensure that the burst used for points near the seams is consistent between scenes. In addition, if there is a need for split-spectral analysis such as ionospheric correction then each subband may be processed independently as per our procedure, the correction applied, and then recombined in the regridded image domain.

D. Effects of DEM errors

The phase corrections above rely on knowledge of the topography of the Earth's surface. Several good DEMs with fairly wide coverage exist, including the NASA SRTM DEM, and the Tandem-X DEM that is somewhat more accurate. The effect of a DEM error translates directly into a mispositioning of SLC pixels, and an unwanted elevation-dependent phase error. We calculate the phase error expressed as deformation error as a function of DEM uncertainty and InSAR baseline:

$$\lambda/_{4\pi}\,\Delta\phi_{topo} = 1/_{r\sin\theta}\,B_{perp}\Delta z \qquad (8)$$

and find that even the modest accuracy of SRTM suffices for InSAR reduction if baselines can be maintained within a few hundred meters. Extreme precision requirements can be met either through use of the higher quality Tandem-X DEM, by better orbit control, or both.

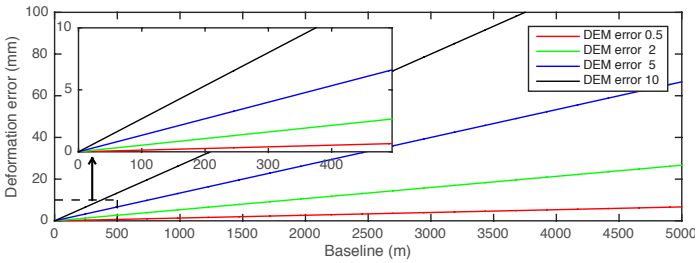


Figure 3. Line of sight deformation error as a function of baseline for several levels of DEM accuracy, for nominal radar parameters of 750 km range and 30° incidence. The inset is an expanded view of the curves for small baselines. Sentinel 1 satellites, e.g., rarely produce interferograms with baselines greater than 150 m, so that even a 10 m error in a DEM leads to a deformation error of 4.0 mm. If the DEM error averages 2 m, the same precise orbit control yields 0.8 mm error.

In fig. 3 we plot the left hand side of eq. 8, the error in deformation, where r is range, θ is the incidence angle, B_{perp} is the perpendicular baseline, and Δz is the DEM elevation error; in this case for a nominal radar geometry of 750 km range and 30° incidence angle. The artifacts resulting from DEM errors can be further corrected using methods developed for persistent scatterer analysis [12], where the residual phases from a time series are scaled by the interferogram baselines and the DEM error is compensated.

III. Examples

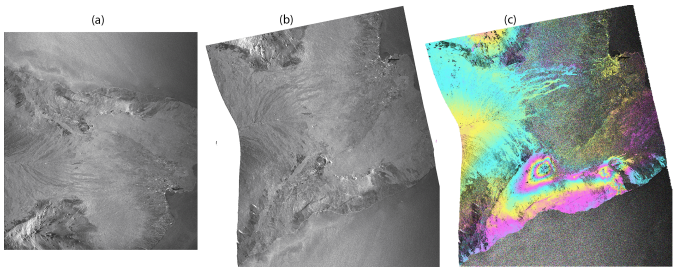


Figure 4. (a) ALOS range-Doppler image of Kilauea region of Hawaii, frame 380, path 287, acquired Jan. 14, 2008, (b) same in geocoded, phase corrected SLC format, (c) interferogram formed by cross multiplication with data from same region acquired on March 6, 2010. Deformation is seen at Kilauea caldera and along chain of craters area, while significant decorrelation occurs in vegetated areas.

In this section we show example interferograms produced from compensated, geocoded SLC images of the Kilauea region of Hawaii. In each case we compute a series of SLCs, and form interferograms by cross-multiplication, with no added coregistration or phase removal operations. Fig. 4 shows ALOS-1 L-band data (path 287 and frame 380) processed to range-Doppler images (fig. 4a), and with phase compensation and geocoding as described above (fig. 4b). We formed an interferogram by straight cross multiplication of scenes acquired on 1/14/2008 and 3/06/2010 (fig. 4c). The data lie on a regular latitude/longitude grid, ready for ingestion and easy manipulation in a GIS system. It is worth noting that some ALOS passes exhibit timing errors, which must be corrected before the pixels can be properly geolocated.

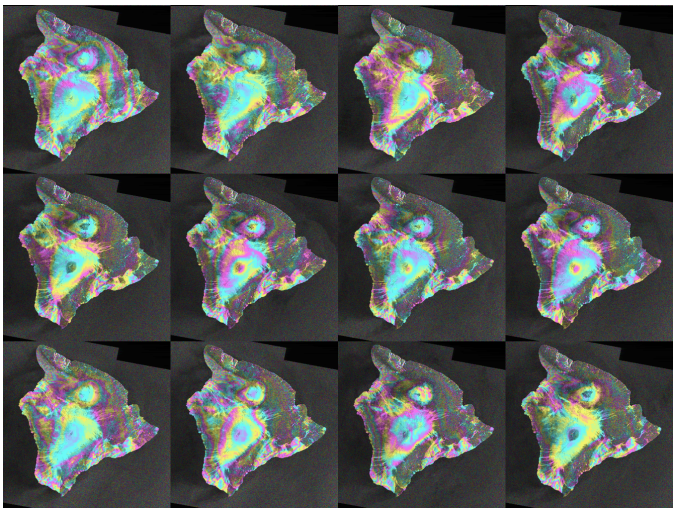


Figure 5. Subset of the 194 interferograms we formed from 30 Sentinel-1 acquisitions over Hawaii from May 2016 – June 2017, with elevation-dependent atmospheric phase removed through regression. Some residual fringes persist due to variations in atmospheric water vapor.

Our next example, using 30 Sentinel-1A/B acquisitions from May 2016 through June 2017, creates 194 interferograms with temporal baselines less than 100 days and spatial baselines less than 200 m. Several interferograms are shown in fig. 5. All exhibit high correlation over the less vegetated areas, and some atmospheric water vapor related residual fringes persist. These interferograms are now ready for time series analysis.

IV. Discussion

The significant barrier posed by i) the need for InSAR expertise by the end user and ii) unfamiliar coordinates, to widespread adoption of InSAR methods for many different studies can be addressed by delivery of user-friendly data products. Here we have shown that the needed phase compensations for both InSAR baseline and surface topography can be applied in a single step during resampling to a regular output grid, thanks to the availability of precise orbit solutions and moderate quality digital elevation models. The approach also enables the delivery of data in SLC rather than interferogram format, greatly reducing the requirements on network transfer of large data sets. As such the method promises to make life easier even for sophisticated users.

The processing flow we propose not only addresses these concerns, but makes it possible to process each scene essentially independently of others, so that it is not necessary to recompute the set of interferograms each time a new scene is added. There is no master scene or orbit needed, as the output coordinate grid defines the common master coordinate system. This approach also facilitates rapid, quick look analysis of a data set, by permitting an initial set of output products on a coarse grid defined by subsampling the DEM and its SLCs. This means that data sets can be 'pyramided,' a common technique in optical remote sensing product delivery so that users may select the scenes of greatest interest for further study.

We have demonstrated the approach using data supplied by the sensor in both raw (ALOS-1) and focused (Sentinel 1) format. Adoption of the geocoded, phase compensated data format as a standard product implies that InSAR expertise need reside almost entirely in the data provider, so that users from many familiarities and experiences can concentrate on their specific study needs, democratizing the widespread use on this powerful remote sensing technique.

References

[1] 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(6433), 138-142.

[2] Zebker, H. A., Rosen, P. A., Goldstein, R. M., Gabriel, A., & C. L. Werner (1994). On the derivation of coseismic displacement fields using differential radar interferometry: The Landers earthquake. J. Geophys. Res.: Solid Earth (1978–2012), 99(B10), 19617-19634.

[3] Wicks, C., Thatcher, W., & Dzurisin, D. (1998). Migration of fluids beneath Yellowstone caldera inferred from satellite radar interferometry. Science, 282(5388), 458-462.

[4] Amelung, F., Jónsson, S., Zebker, H., & Segall, P. (2000). Widespread uplift and 'trapdoor' faulting on Galapagos volcanoes observed with radar interferometry. Nature, 407(6807), 993-996.

[5] Goldstein, R. M., Engelhardt, H., Kamb, B., & Frolich, R. M. (1993). Satellite radar interferometry for monitoring ice sheet motion: application to an Antarctic ice stream. Science, 262(5139), 1525-1530.

[6] Fatland, D. R., & Lingle, C. S. (2002). InSAR observations of the 1993-95 Bering Glacier (Alaska, USA) surge and a surge hypothesis. Journal of Glaciology, 48(162), 439-451.

[7] Hoffmann, J., Zebker, H. A., Galloway, D. L., & Amelung, F. (2001). Seasonal subsidence and rebound in Las Vegas Valley, Nevada, observed by synthetic aperture radar interferometry. Water Resources Research, 37(6), 1551-1566.

[8] Galloway, D. L., & Hoffmann, J. (2007). The application of satellite differential SAR interferometry-derived ground displacements in hydrogeology. Hydrogeology Journal, 15(1), 133-154.

[9] Zebker, H. A., & Goldstein, R. M. (1986). Topographic mapping from interferometric synthetic aperture radar observations. Journal of Geophysical Research: Solid Earth, 91(B5), 4993-4999.

[10] Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., & Alsdorf, D. (2007). The shuttle radar topography mission. Reviews of geophysics, 45(2).

[11] Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. Geoscience and Remote Sensing, IEEE Transactions on,40(11), 2375- 2383.

[12] Ferretti, A., Prati, C., & Rocca, F. (2001). Permanent scatterers in SAR interferometry. Geoscience and Remote Sensing, IEEE Transactions on, 39(1), 8-20.

[13] Hooper, A., Zebker, H., Segall, P., & Kampes, B. (2004). A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers. Geophysical Research Letters, 31(23).

[14] Cutrona, L.J., W.E. Vivian, E.N. Lieth, and G.O. Hall (1961), A high resolution radar combat surveillance system, IRE Trans. on Military Electronics, Vol. MIL-5, Issue 2, pp. 127-131.

[15] Brown, W.M. (1967), Synthetic Aperture Radar, IEEE Trans. Aerosp. Electron. Syst., Vol. AES-3, no. 2, pp. 217-229.

[16] Wu, C., K. Liu, and M. Jin (1982), Modeling and a correlation algorithm for spaceborne SAR signals, IEEE Trans. Aerosp. Electron. Syst., Vol. AES-18, Issue 5, pp. 563-575.

[17] Zebker, H., Hensley, S., Shanker, P., & Wortham, C. (2010), Geodetically accurate InSAR data processor. Geoscience and Remote Sensing, IEEE Transactions on, 48(12), 4309-4321.

[18] Schmitt, M., and U. Stilla (2014), Maximum-likelihood-based approach for single-pass synthetic aperture radar tomography over urban areas, ET Radar Sonar Navig., Vol. 8, Iss. 9, pp. 1145–1153.

[19] Zheng, Y.J., & H.A. Zebker (2017), Phase Correction of Single-Look Complex Radar Images for User-Friendly Efficient Interferogram Formation, J-STARS, Volume: PP, Issue: 99, pp. 1-8.

[20] Agram, P., Jet Propulsion Laboratory, personal communication.

[21] R. Scheiber & A. Moreira (2000), Coregistration of interferometric SAR images using spectral diversity, IEEE Trans. Geosci. Remote Sens., vol. 38, no. 5, pp. 2179–2191.

[22] Xu, X., D. Sandwell, E. Tymofyeyeva, A. González-Ortega, & X. Tong, (2016). Tectonic and Anthropogenic Deformation at the Cerro Prieto Geothermal Step-Over Revealed by Sentinel-1A InSAR. IEEE Trans. Geosci. Rem. Sens. 10.1109/TGRS.2017.2704593.

InSAR Mission-level Products On Demand – Do We Need Range-Doppler?

Howard A. Zebker
Stanford University
zebker@stanford.edu
USA

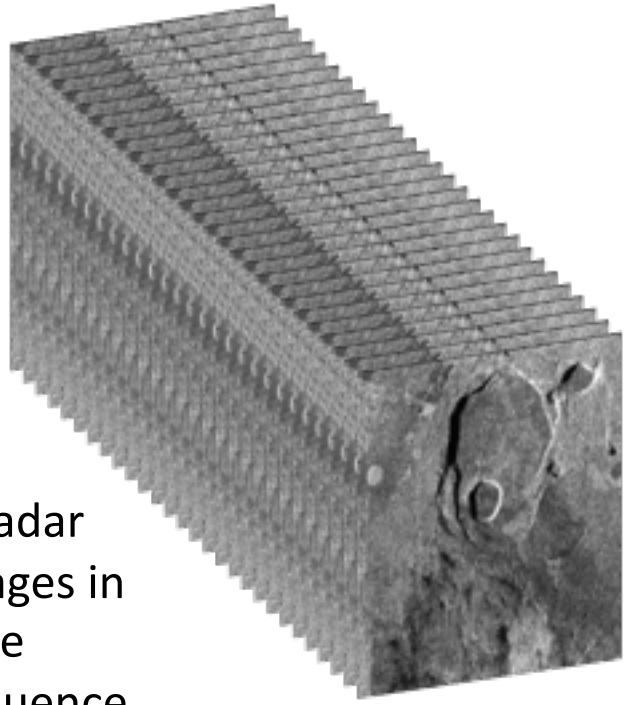
Background: User-friendly products

- Standard data products should be user-friendly
  - Geocoded to common coordinates
  - InSAR phase corrections applied
  - InSAR version of “Analysis ready data”
- SLCs consistent with these allow analysis by non-specialists
- Note: above needs never mention data in radar coordinate format, i.e. no range-Doppler

➢How can we provide products efficiently to many different users who have varying needs?

Modern system considerations

- Today's constellations produce huge data volumes
- All data need to be available but few are studied



N radar images in time sequence

- Computing is cheap
- Storage is expensive

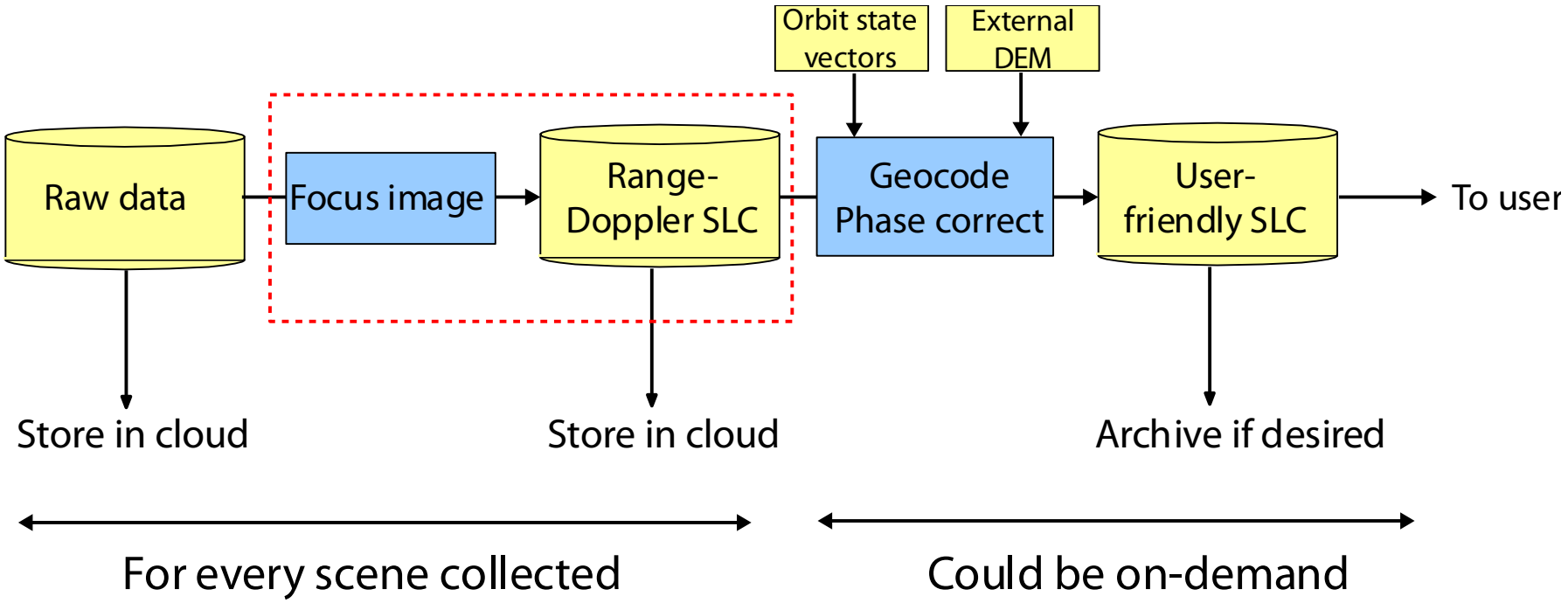
✓On-demand system has many advantages

What is today's usual approach?

- Generally some sort of convolutional processing applies the matched filter
  - Computationally efficient due to FFT
  - Robust for poorly known orbits
  - But applies same filter everywhere so phase corrections needed
- Data are produced in range-Doppler space and need phase compensation
- InSAR analysis requires precise coregistration and viewing/topographic corrections

Processor flow satisfying user needs

- One possible implementation with on-demand generation
- Cost driver likely the storage of range-Doppler SLCs
  - Needed for every imaged scene
  - Larger than raw data files



Backprojection SLC formation

- One of the earliest algorithms proposed for SAR imaging
- Was impractical due to computational inefficiency and lack of accurate platform knowledge
- Modern computers and orbit tracking enable approach

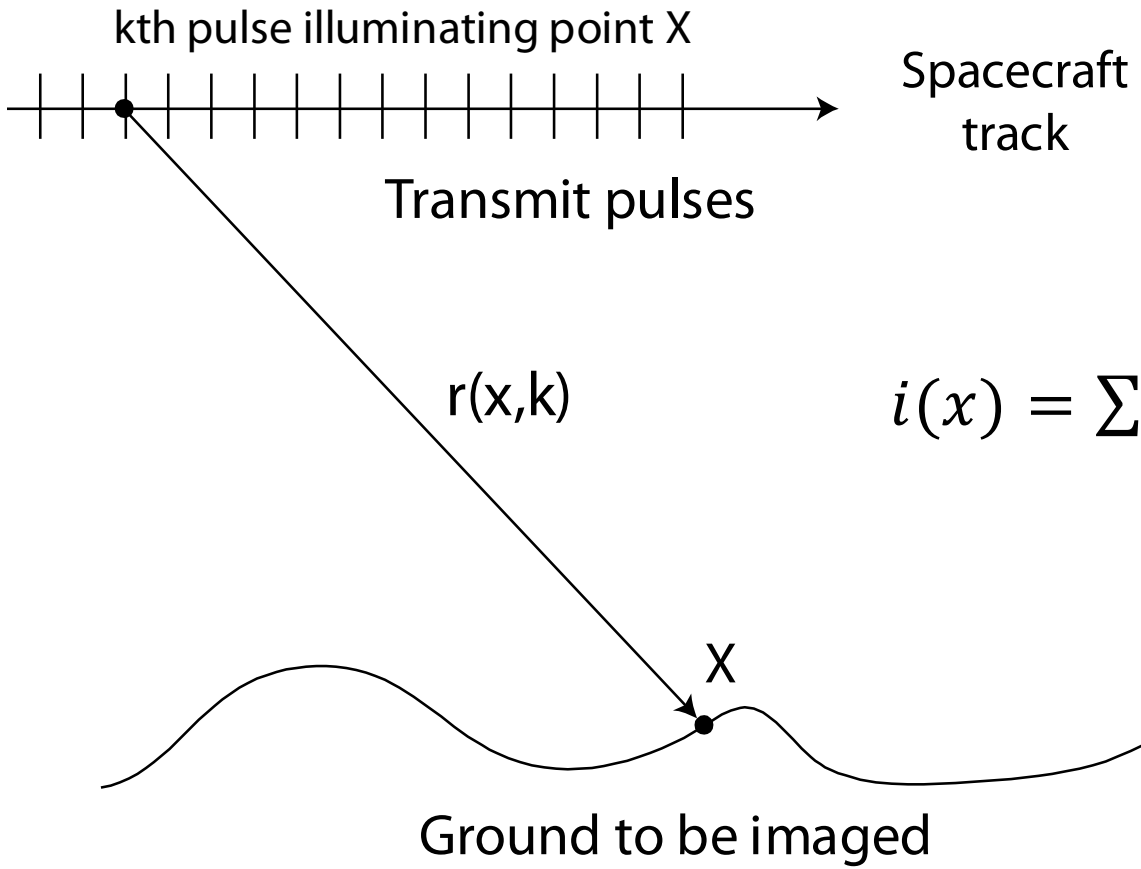
✓Forms ideal matched filter

✓Automatically applies phase corrections if DEM used

✓SLCs produced in lat/lon or other desired geometry

Calculating the SLC

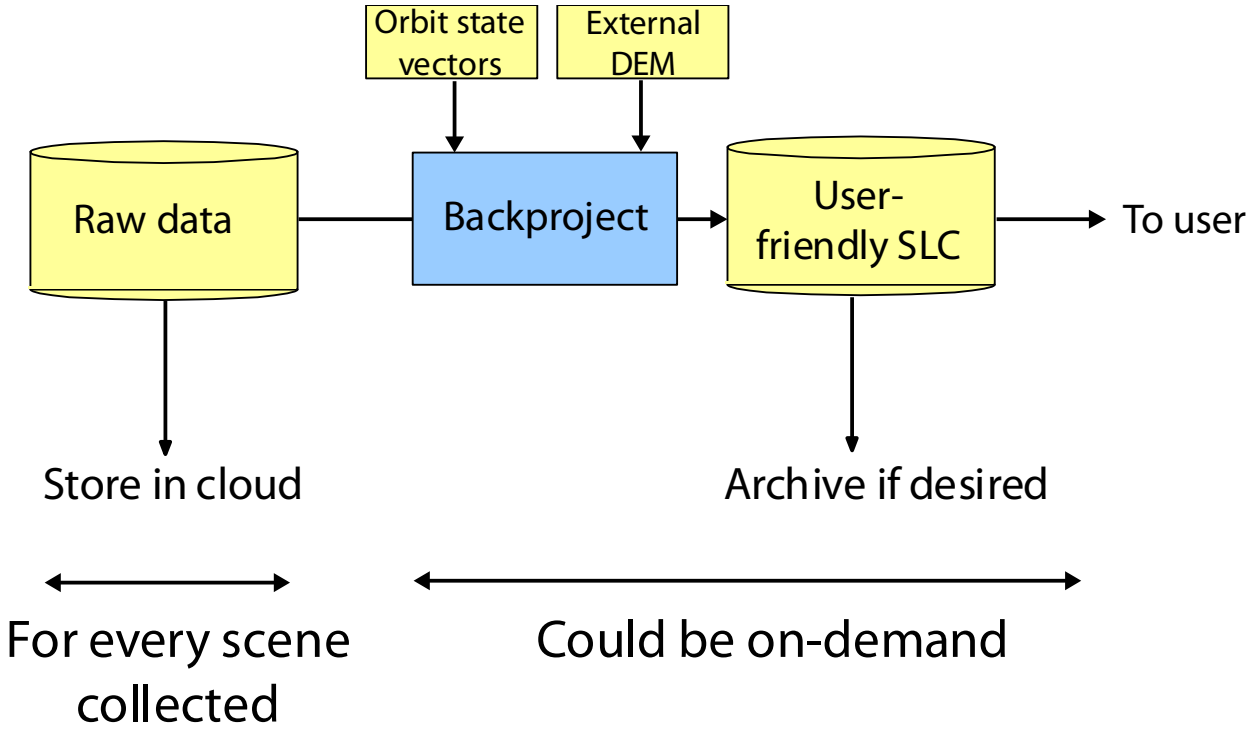
- Algorithm: add all echoes illuminating a point in phase
- Automatically compensates phase if DEM included



$$i(x) = \sum_k s\big(r(x,k)\big)\, \mathrm{e}^{-j\, r(x,k) \bullet \frac{4\pi}{\lambda}}$$

Processor flow - backprojection

- Backprojection avoids creating/storing range-Doppler intermediate products
- Greatly reduces archive costs



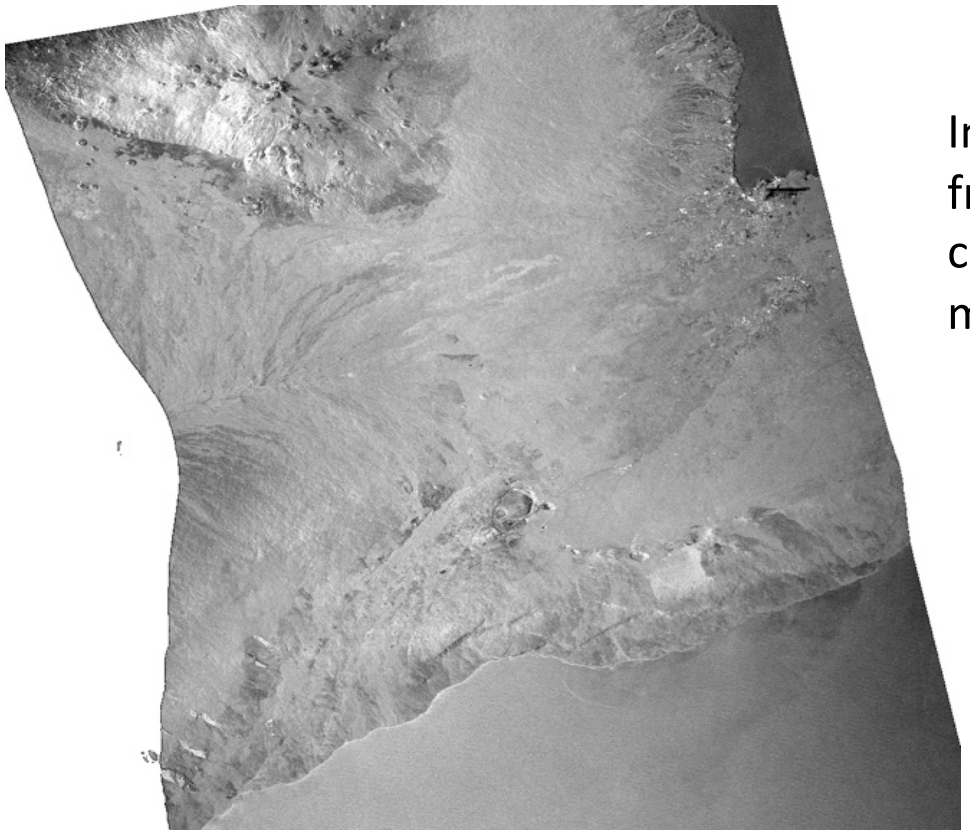
Let’s look at practical cases

- Examples using ALOS strip map and Sentinel burst processing for wide swaths
- In both cases we avoid storing intermediate products
- Added image formation processing burden eased through GPU architecture (cheap if not competing with currency miners)
- Data products more amenable to InSAR analysis

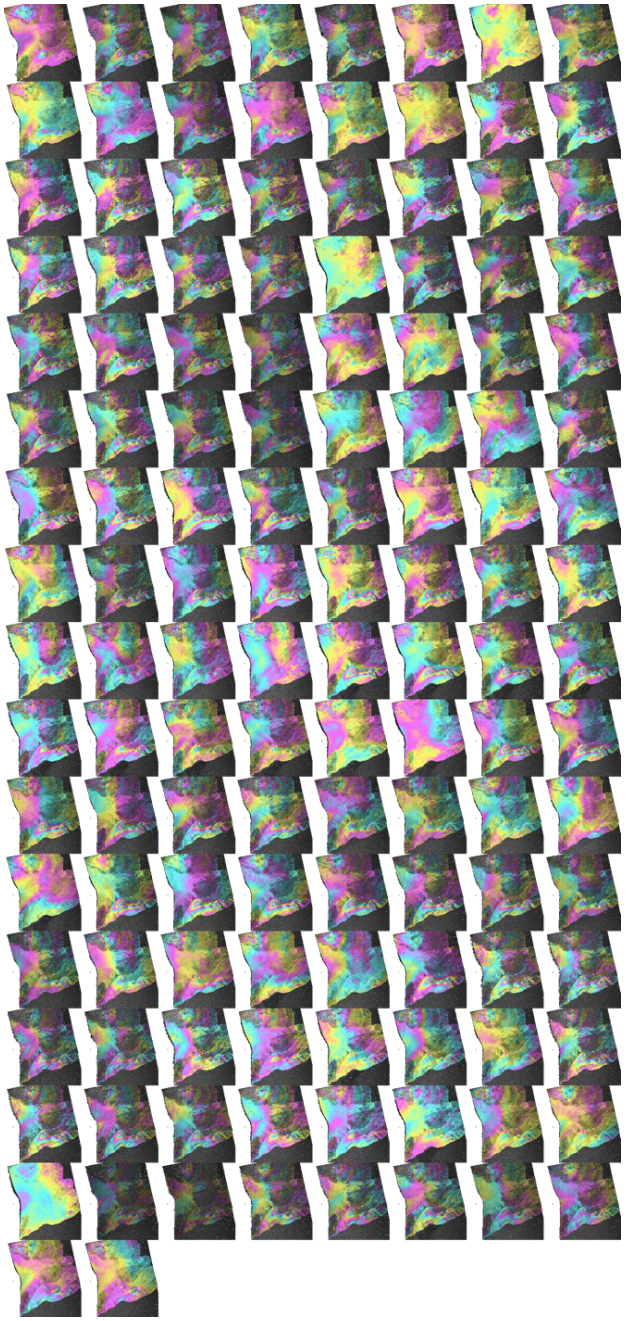
ALOS: A simple example

- L-band, strip mode, 20 km aperture
- GPU pipelined implementation

Geocoded amplitude image, Kilauea



Interferograms from simple cross multiplication

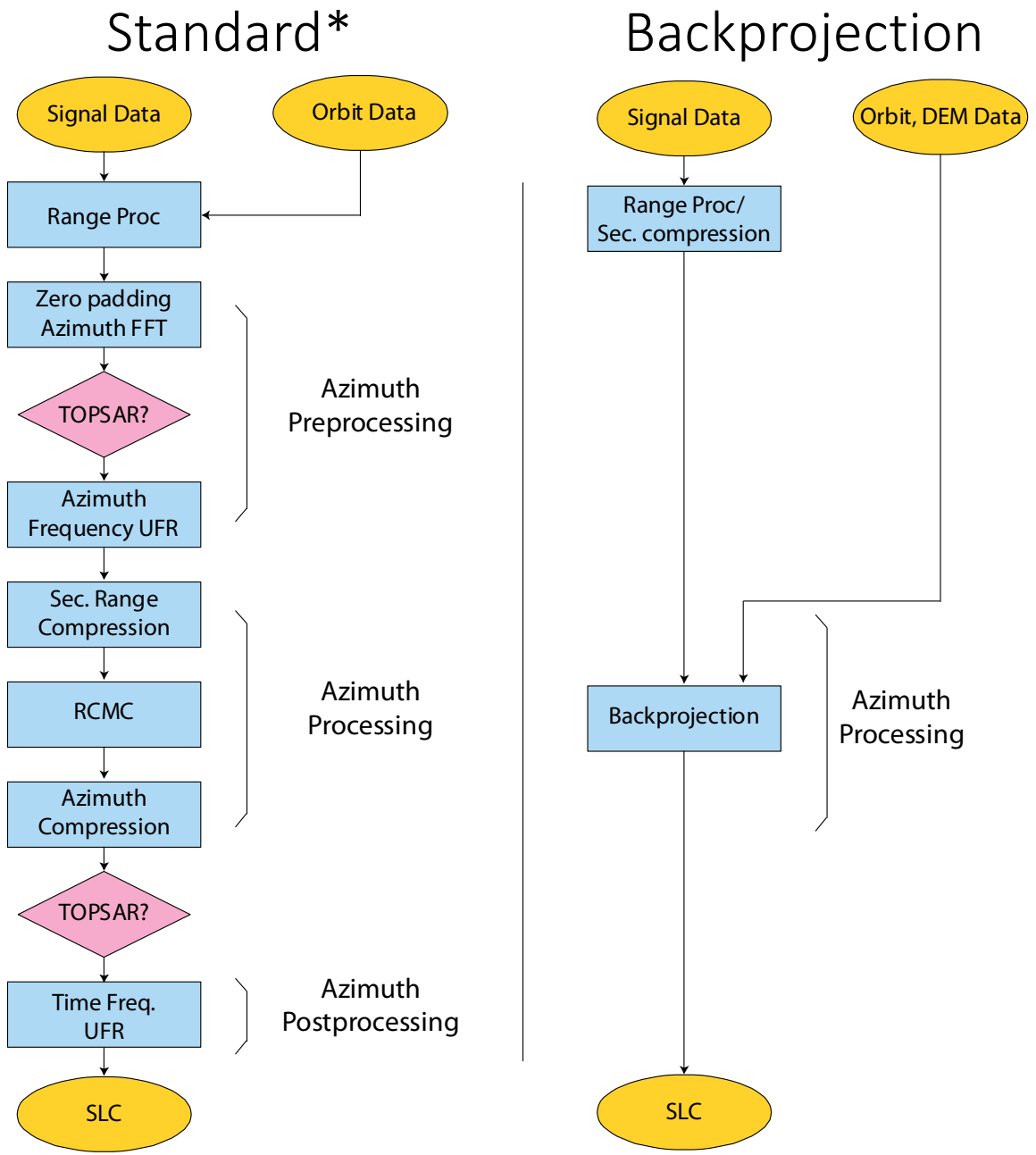


Sentinel 1A/B - TOPS imaging

- Sentinel 1 TOPS mode permits large and frequent coverage but products complex
- Standard product hard to use due to carrier phase
- Very precise coregistration needed
- Products can be resampled to common coordinates but phase compensation requires expertise

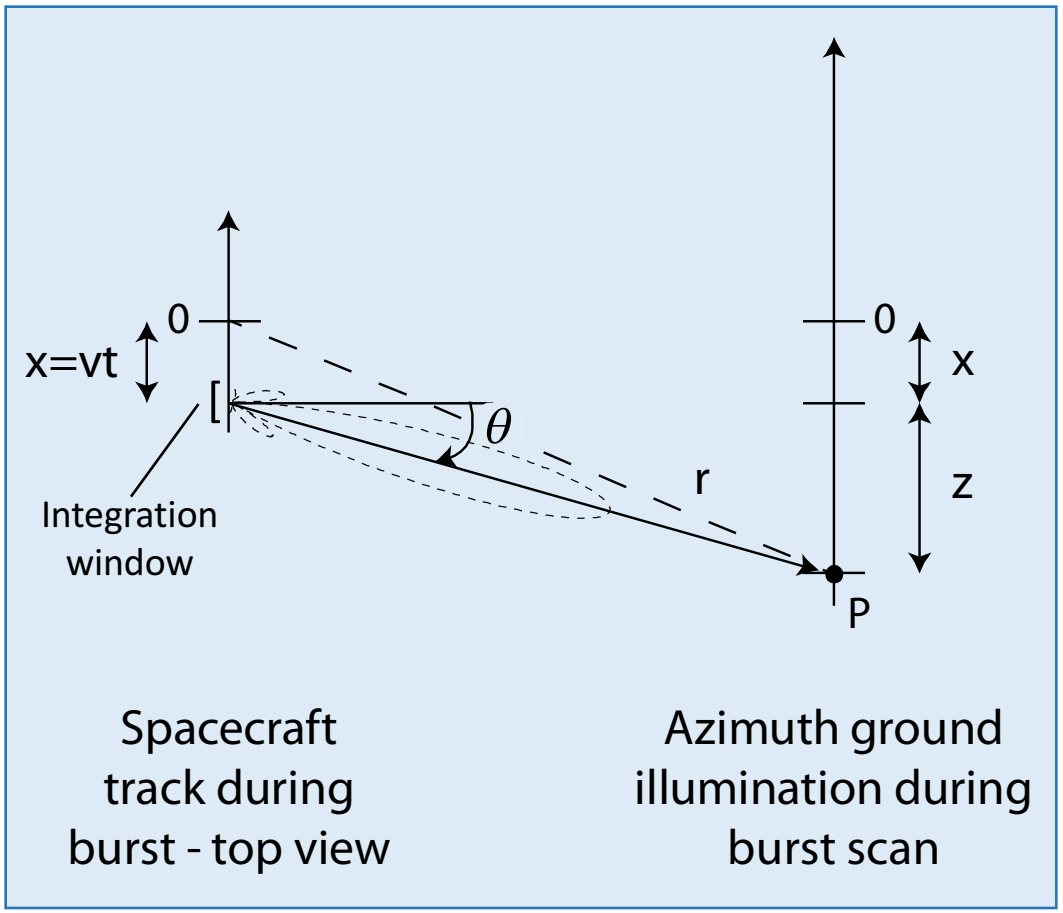
L0 SLC processing

- Dealiasing computationally expensive and very intricate though elegant
- Backprojection same for all modes
- Short integration time makes backprojection efficient



*Sentinel 1 SAR Technical Guide

TOPS integration window



For target P at $f_{Dop,0}$ wrt burst midpoint and TOPS steering angle θ:

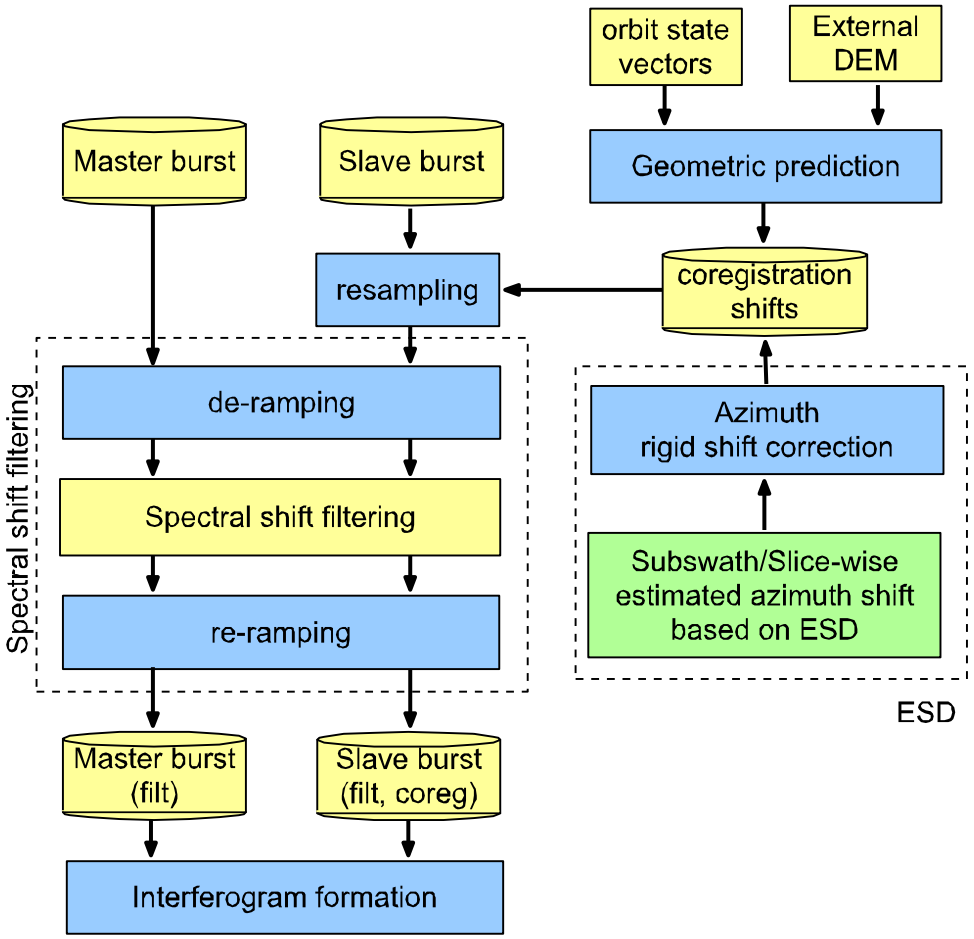
$$f_{Dop,0} = \frac{2v}{\lambda}\frac{x}{r} + \frac{2v}{\lambda}\theta$$

$$where\ \theta = c_1 t + c_0$$

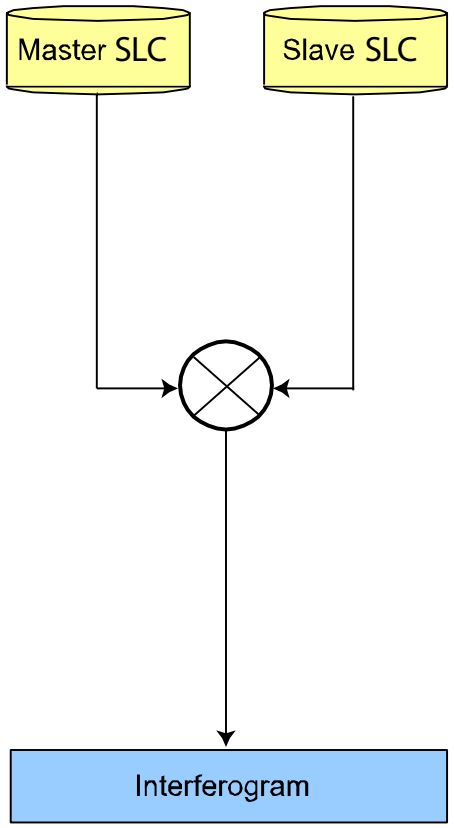
$$\rightarrow t = \frac{\frac{r\lambda}{2v^2} f_{Dop,0} - \frac{rc_0}{v}}{1 + \frac{rc_1}{v}}$$

Interferogram formation – by user

Standard (Yague-Martinez et al.,2016)



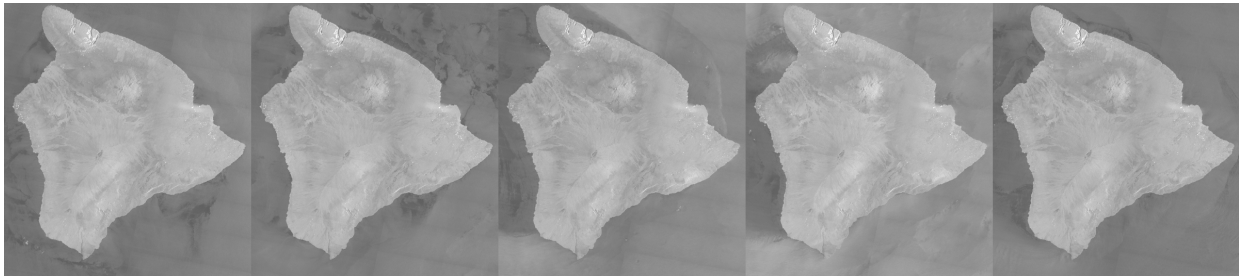
Backprojection



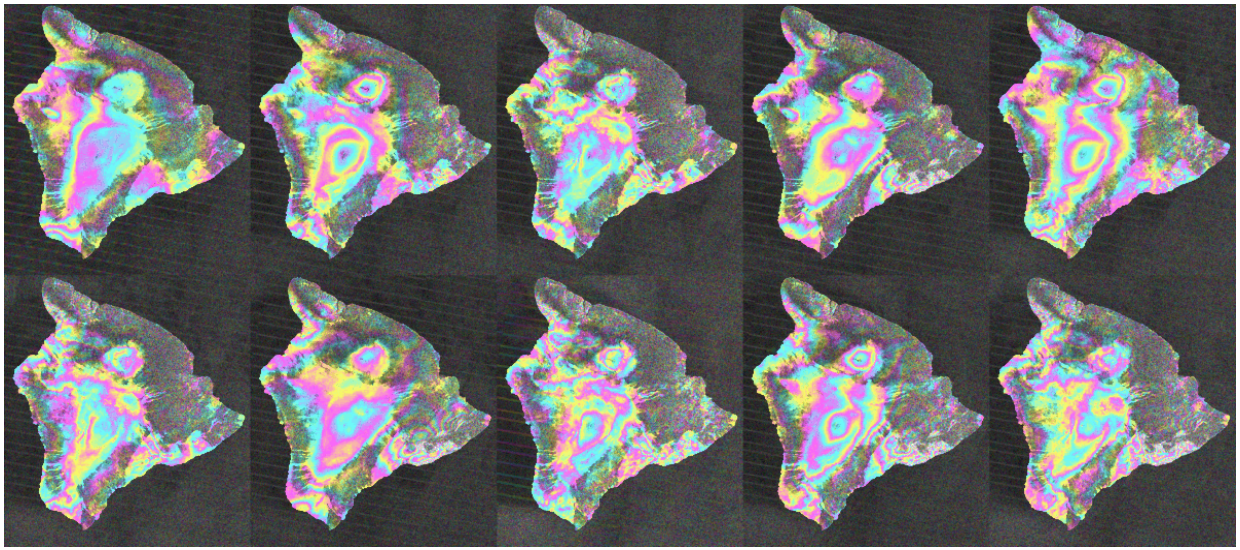
Sentinel 1B backprojection products

- Standardized, geocoded, phase corrected InSAR enables huge potential community

Island of Hawaii



Geocoded SLCs

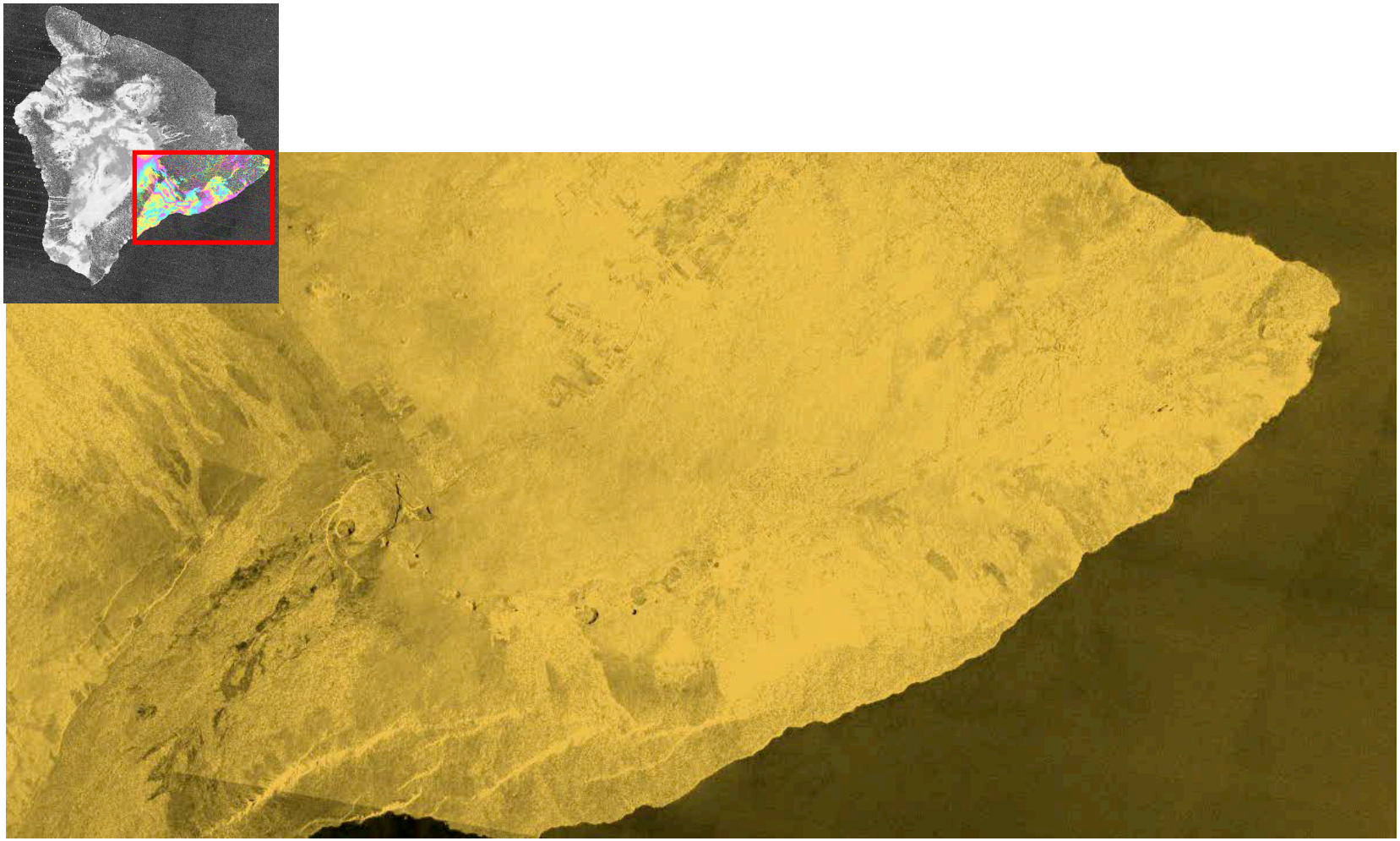


Backprojected phase corrected Sentinel-1B data facilitate easy interferogram analysis

User-friendly interferograms

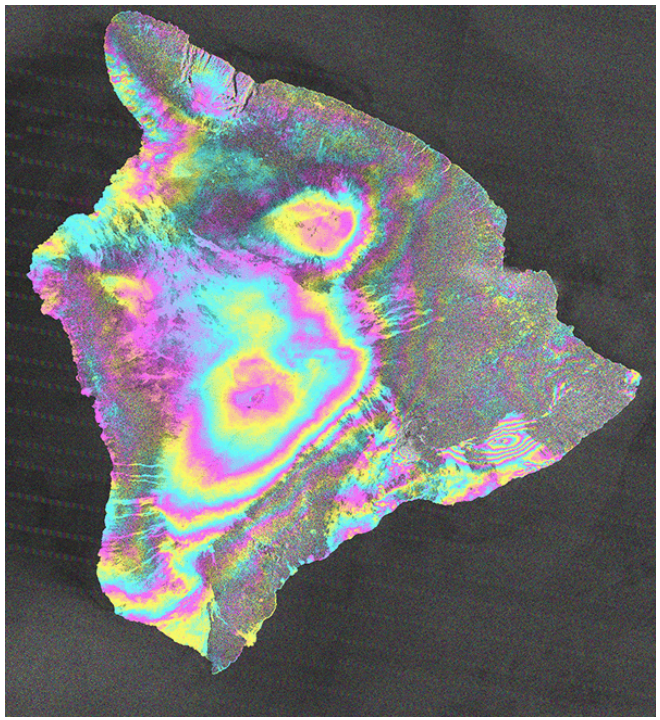
User-friendly products

- Simplified time series view of Kilauea eruption

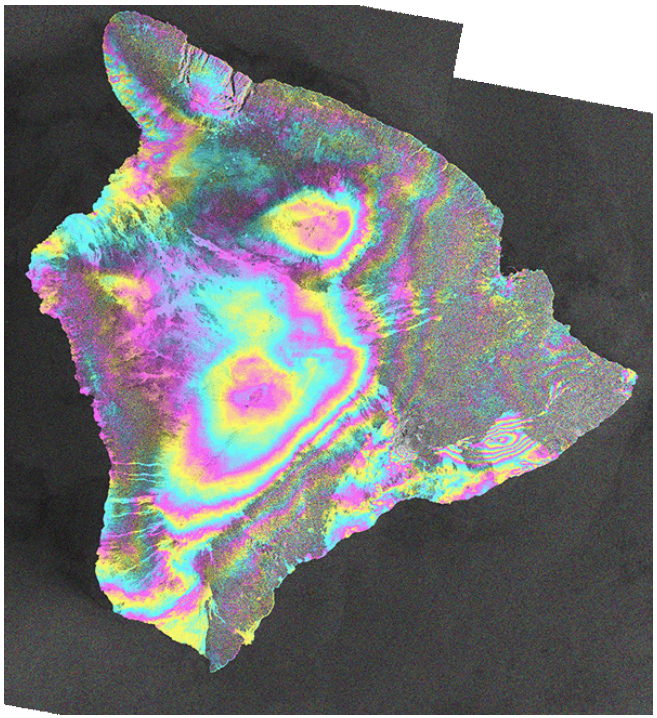


Product comparison

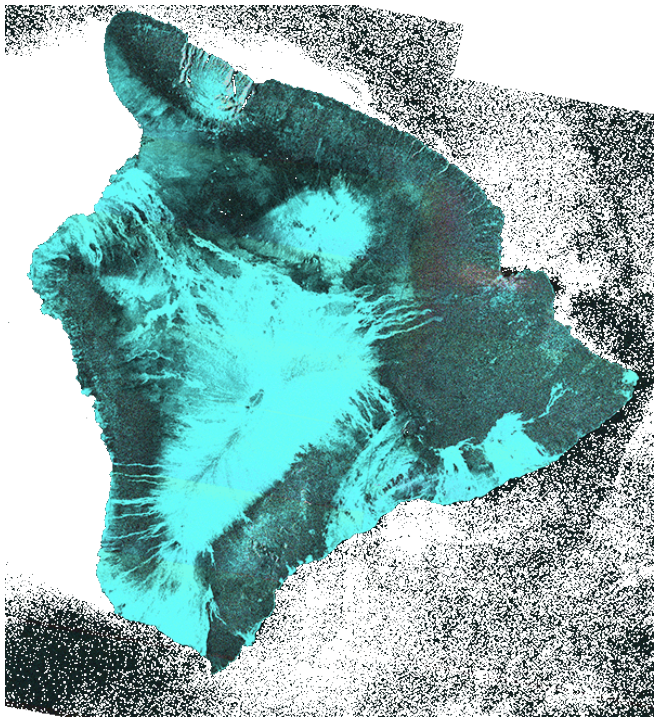
- End products are very similar
- Choice based on efficiency and ease of use



From L1 SLC



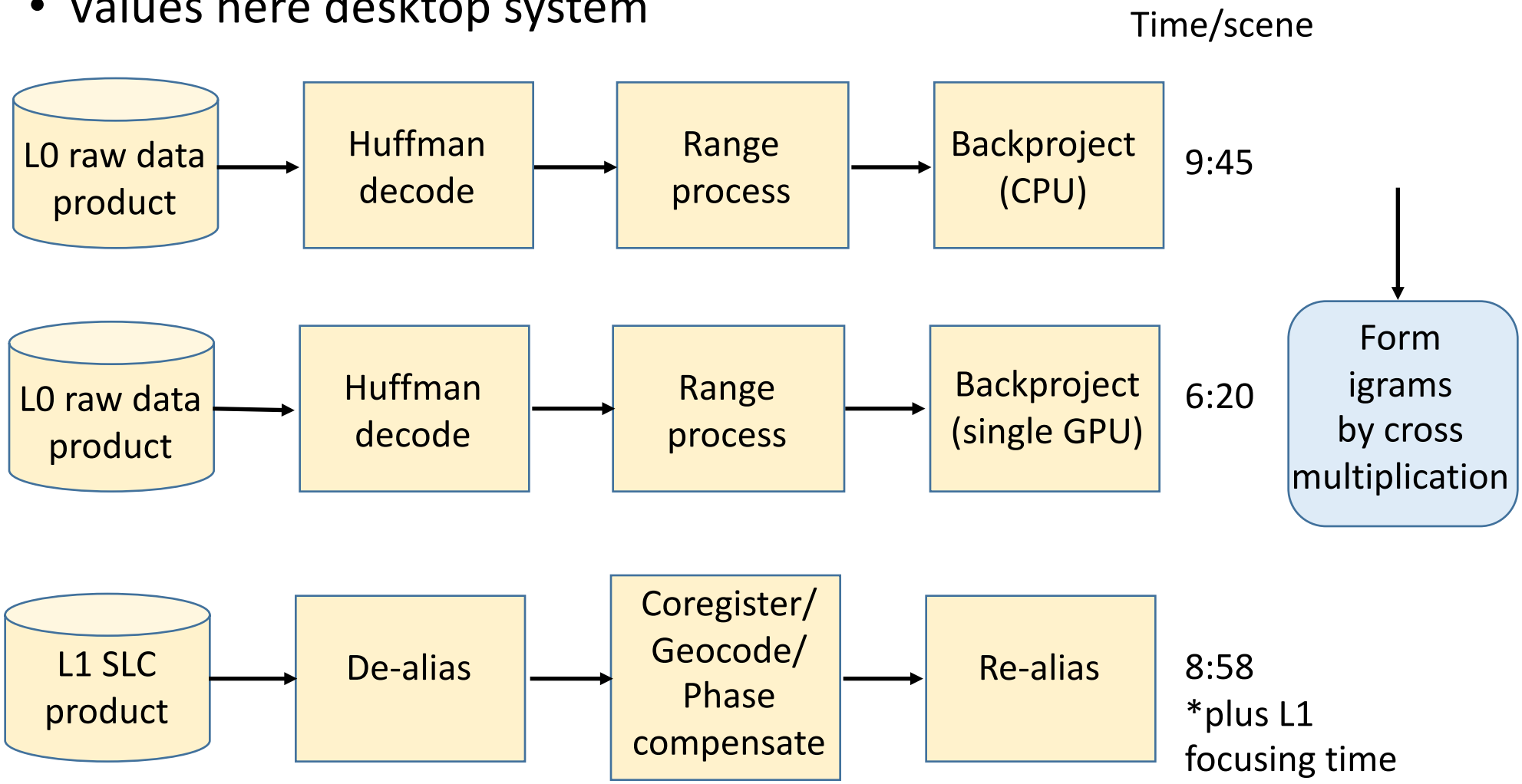
From L0 Raw



Cross-interferogram

Some timing considerations - Sentinel

- These are highly code- and architecture-dependent
- Values here desktop system



Conclusions

- Product generation from range-Doppler SLCs difficult for both producer and end user
- Backprojection
  - Generates user-friendly products directly
  - Simplified SLC processing stream
  - Simplified interferogram generation
  - Computationally efficient with GPU implementation

-> No need to create/store range-Doppler products

-> Feasible on demand delivery system