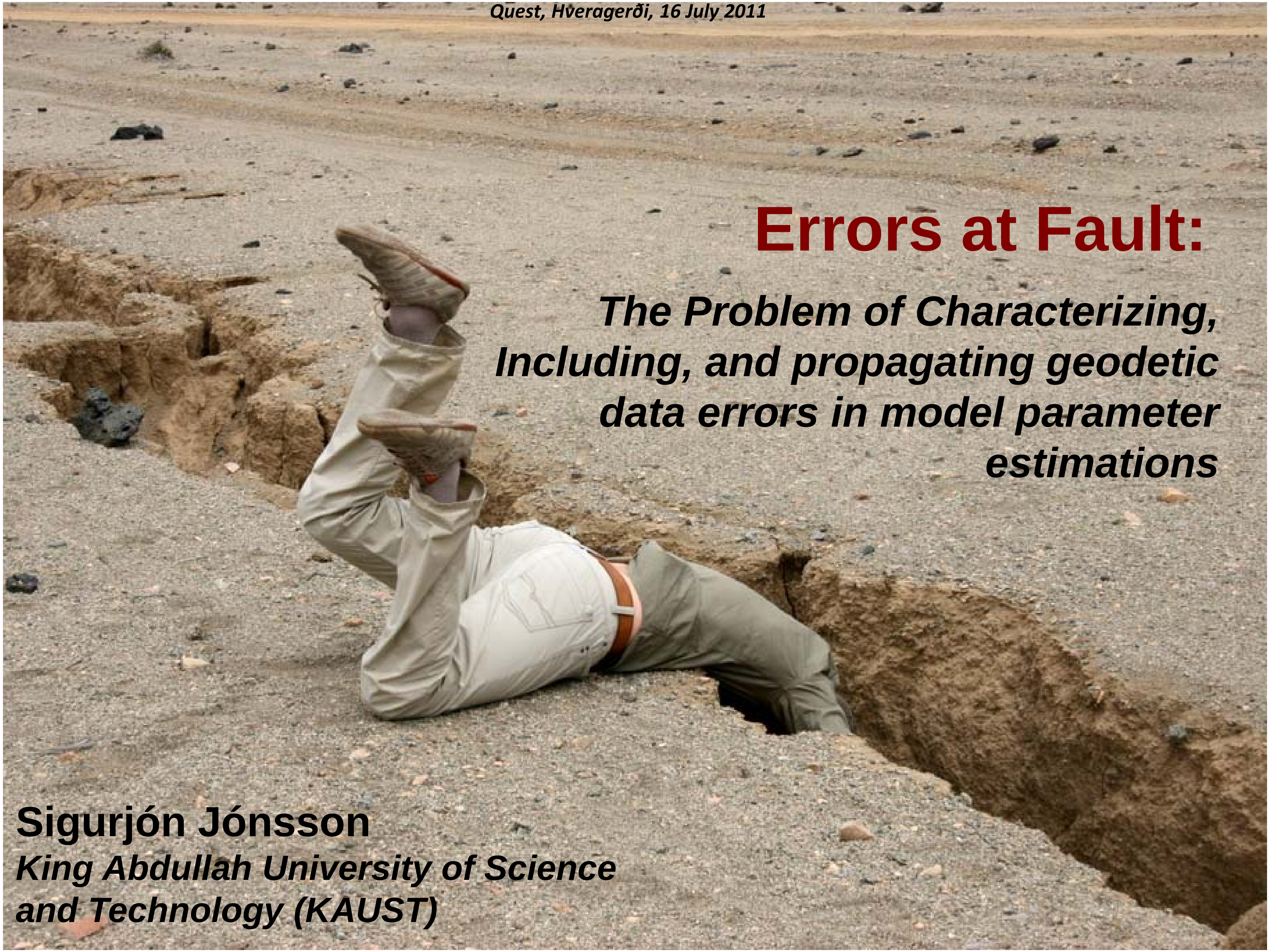


Errors at Fault:

The Problem of Characterizing, Including, and propagating geodetic data errors in model parameter estimations

Sigurjón Jónsson
King Abdullah University of Science and Technology (KAUST)

Motivation

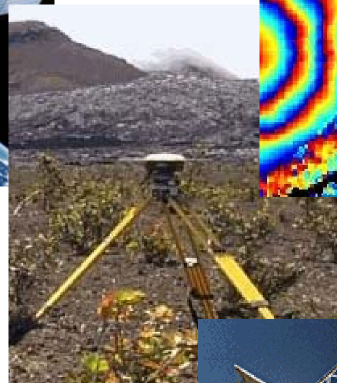
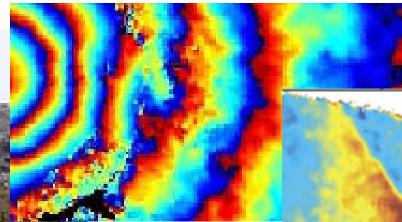
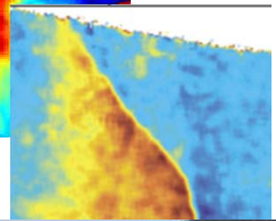
- To tell you a little bit about how many in the geodetic community deal with data errors and estimate model parameter uncertainties

Outline

- Geodetic data
- Estimating data errors
- Making data covariance matrices
- Estimating model parameter uncertainties
- Applications

What are Geodetic Data?

- Global Navigation Satellite Systems: GPS, Glonass, Galileo, Beidou
- InSAR
- Air/satellite photo offsets
- SAR image offsets
- Leveling, Tiltmeters
- Borehole strainmeters
- Creepmeters
- Very Long Baseline Interferometry (VLBI)
- Satellite Laser Ranging (SLR)
- Electronic Distance Measurements (EDM)
- Triangulation
- Underwater acoustic/GPS measurements
- Fault rupture offset measurements
- Coastal uplift, lake-level tilts
- Gravimetric data
- and probably many more…..




- **All Data contain Errors!**
- **Here Vertical Fault Offset**
- **62 cm +/- ???**

King Abdullah University of Science and Technology

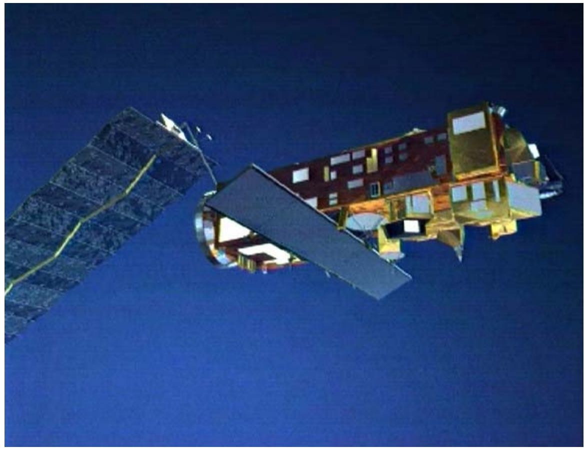
KAUST

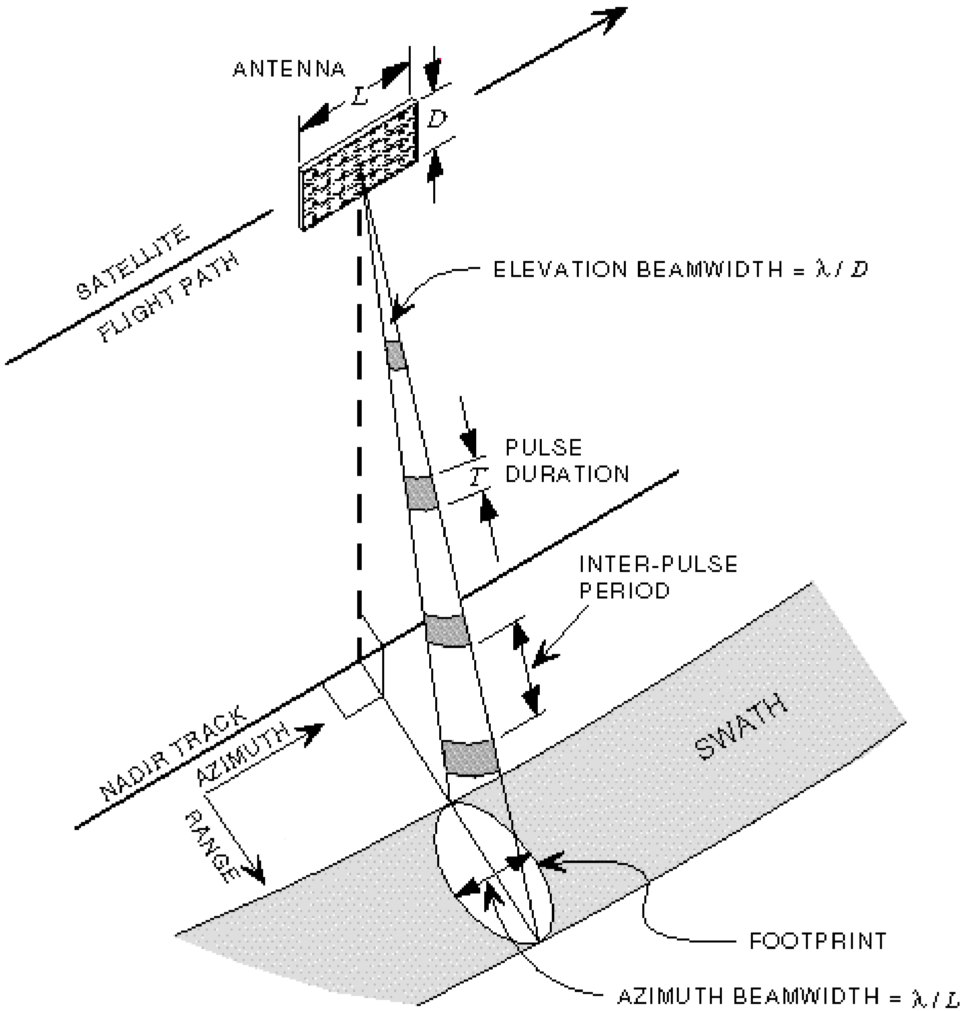
InSAR and InSAR Data Errors

Synthetic Aperture Radar (SAR)

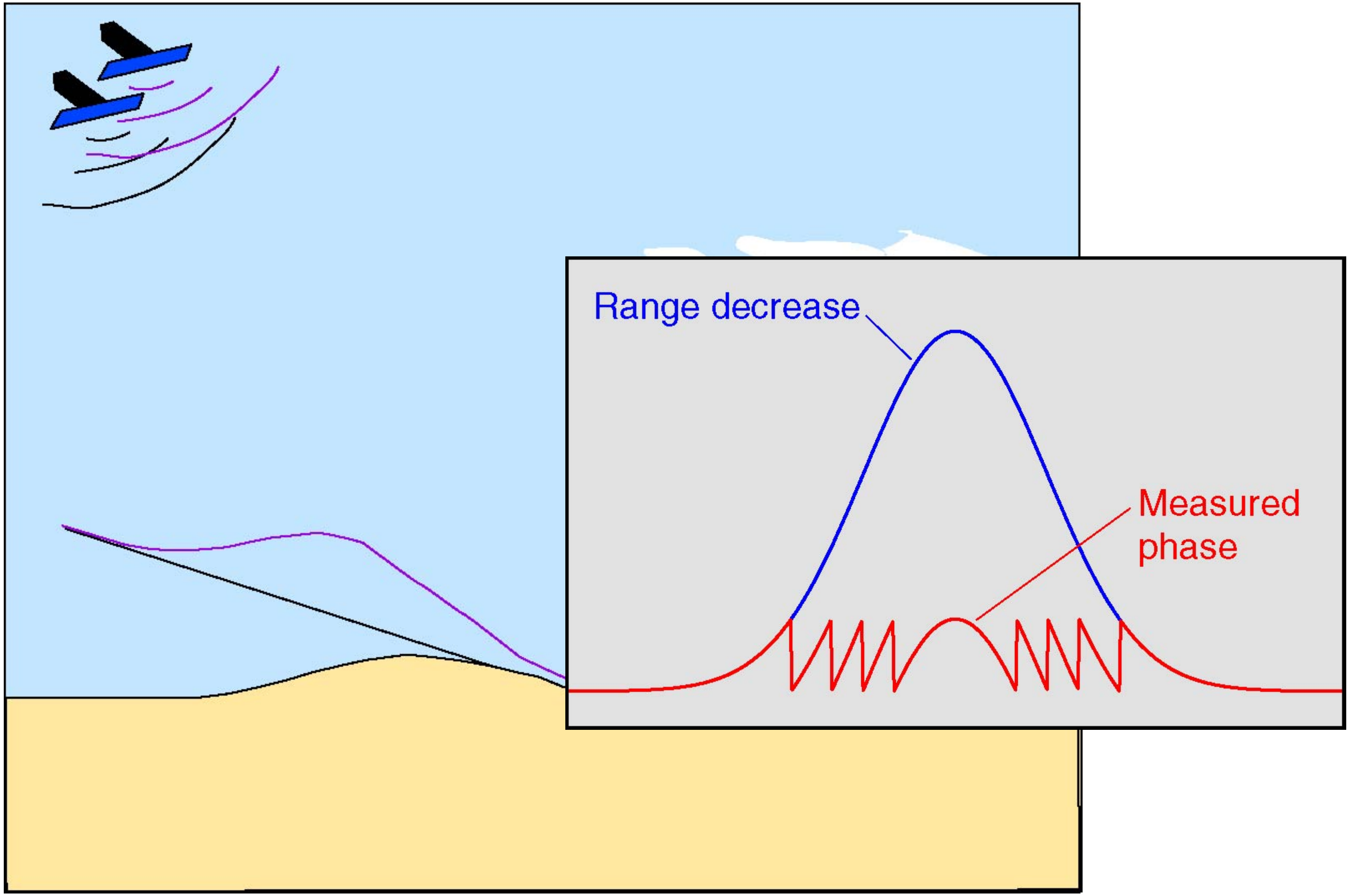
Typical radar satellite configuration:

- Satellite elevation is typically ~7-800 km
- Swath width ~100km
- Antenna 1m x 10m
- Look angle usually 20-50 degrees from vertical



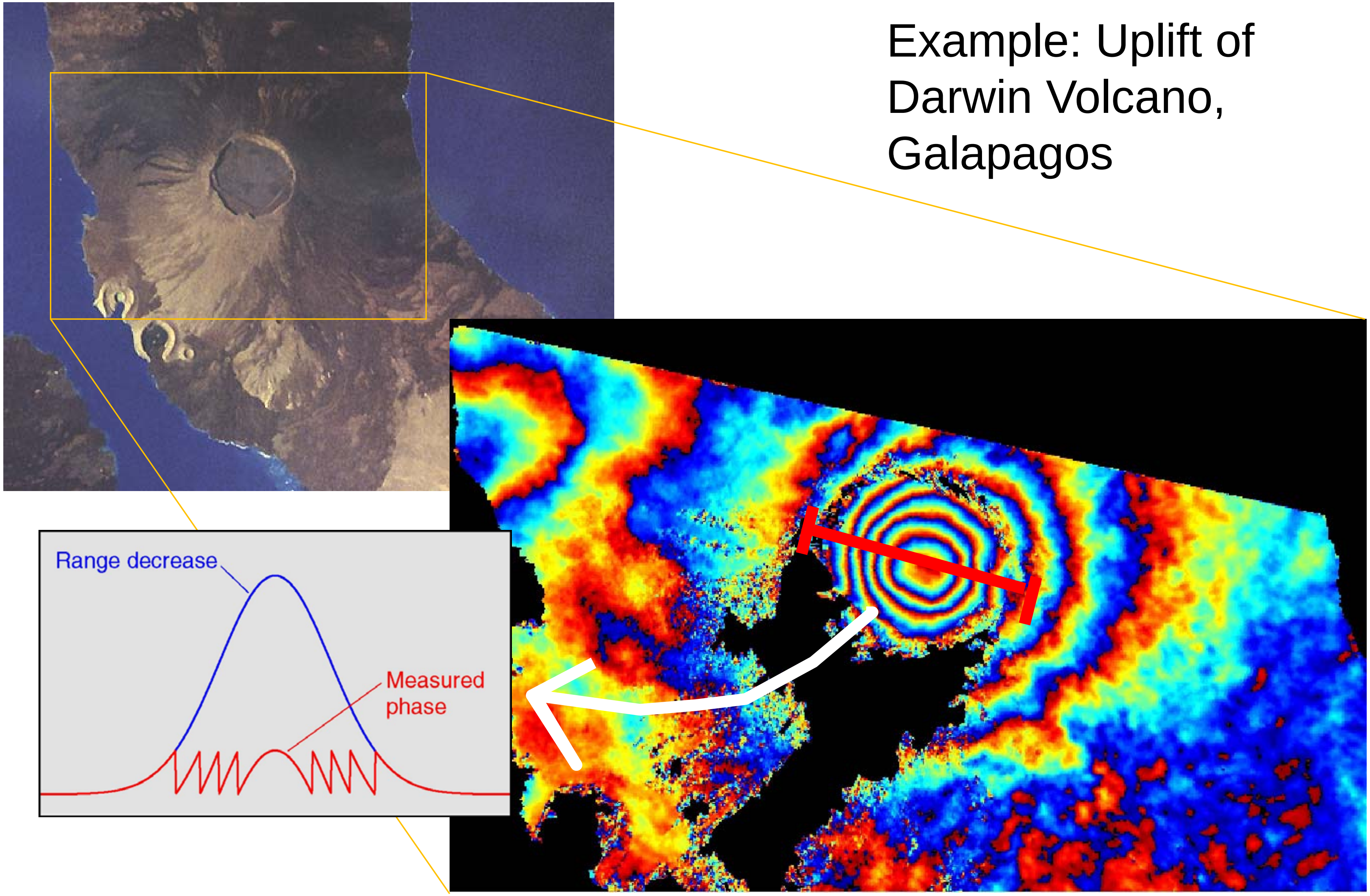


InSAR Phase Difference due to Deformation



InSAR Phase Difference due to Deformation

Example: Uplift of Darwin Volcano, Galapagos



InSAR compared to other geodetic techniques

Advantages:

- Spatial sampling
- Global coverage
- Existing database
- Inexpensive, safe

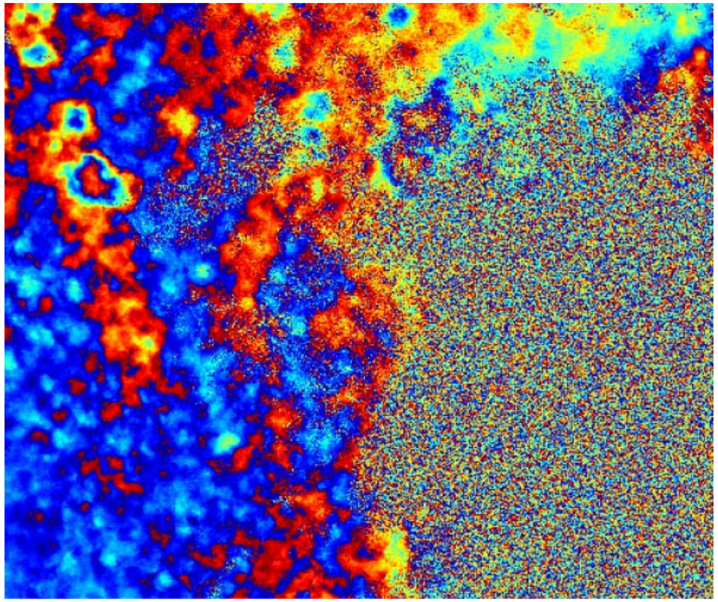
Limitations:

- Poor temporal resolution
- 1D displacements
- Decorrelation

What Causes Errors in InSAR Data?

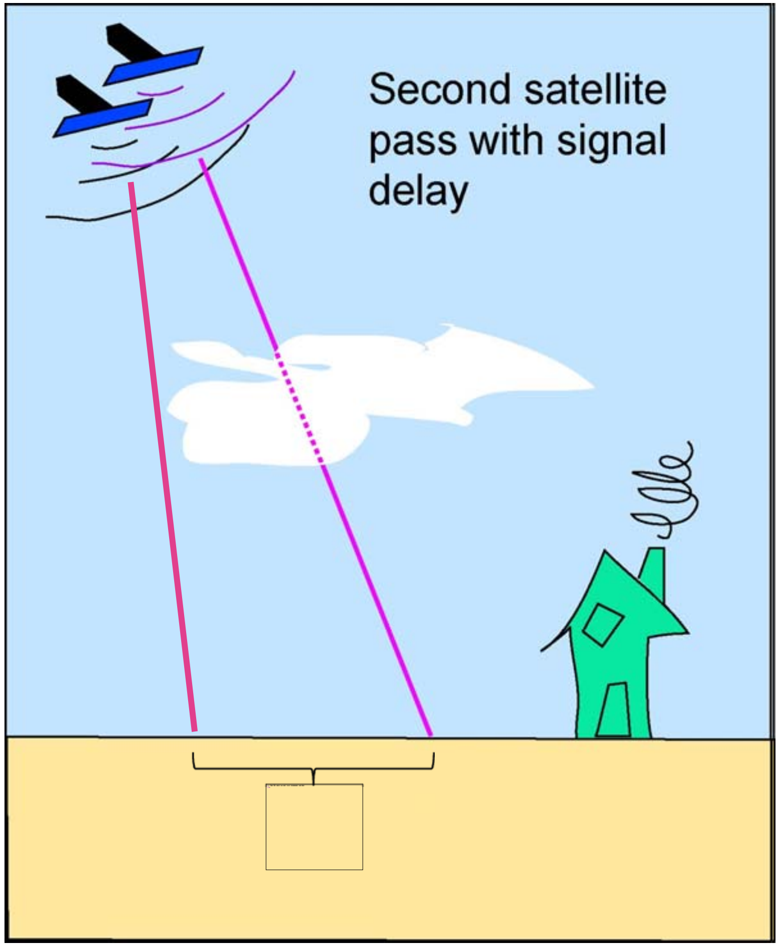
Various errors in Deformation Interferograms:

- DEM errors
- Co-registration errors
- Inaccurate orbit information
- Unwrapping errors
- Geocoding mistakes
- Decorrelation
- Atmospheric artifacts

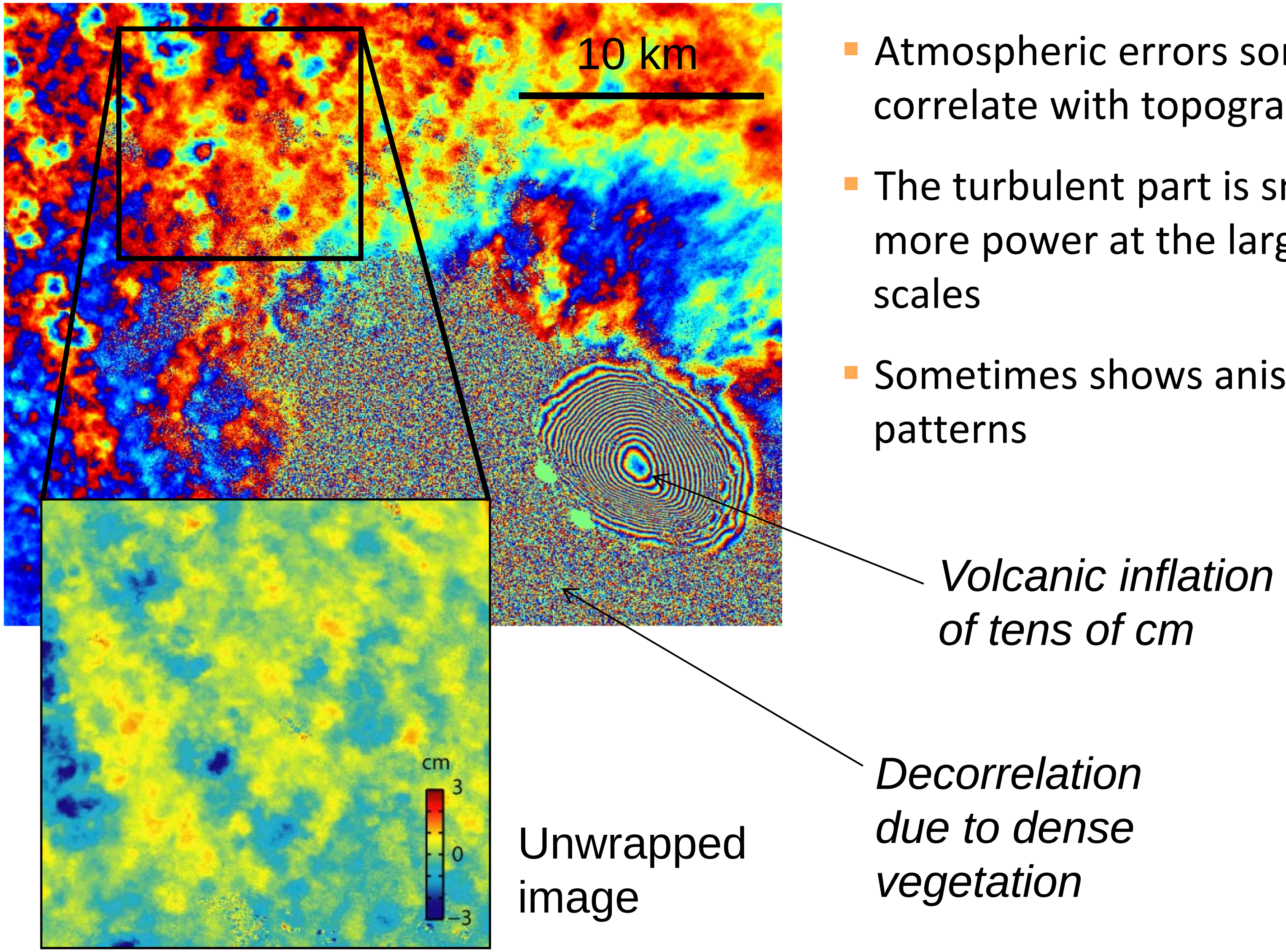


Phase Difference due to the Atmosphere

- Phase difference:
- Where signal delay D_{atm} depends on the refractivity index *n* in the atmosphere: $v=c/n$ ($n=1$ for vacuum)
- Usually refractivity N is used: $N=(n-1)*10^6$
- Propagation delay through the atmosphere (in meters):
- Refractivity composed of:
- and refractivity is clearly:



Example of InSAR Atmospheric Errors

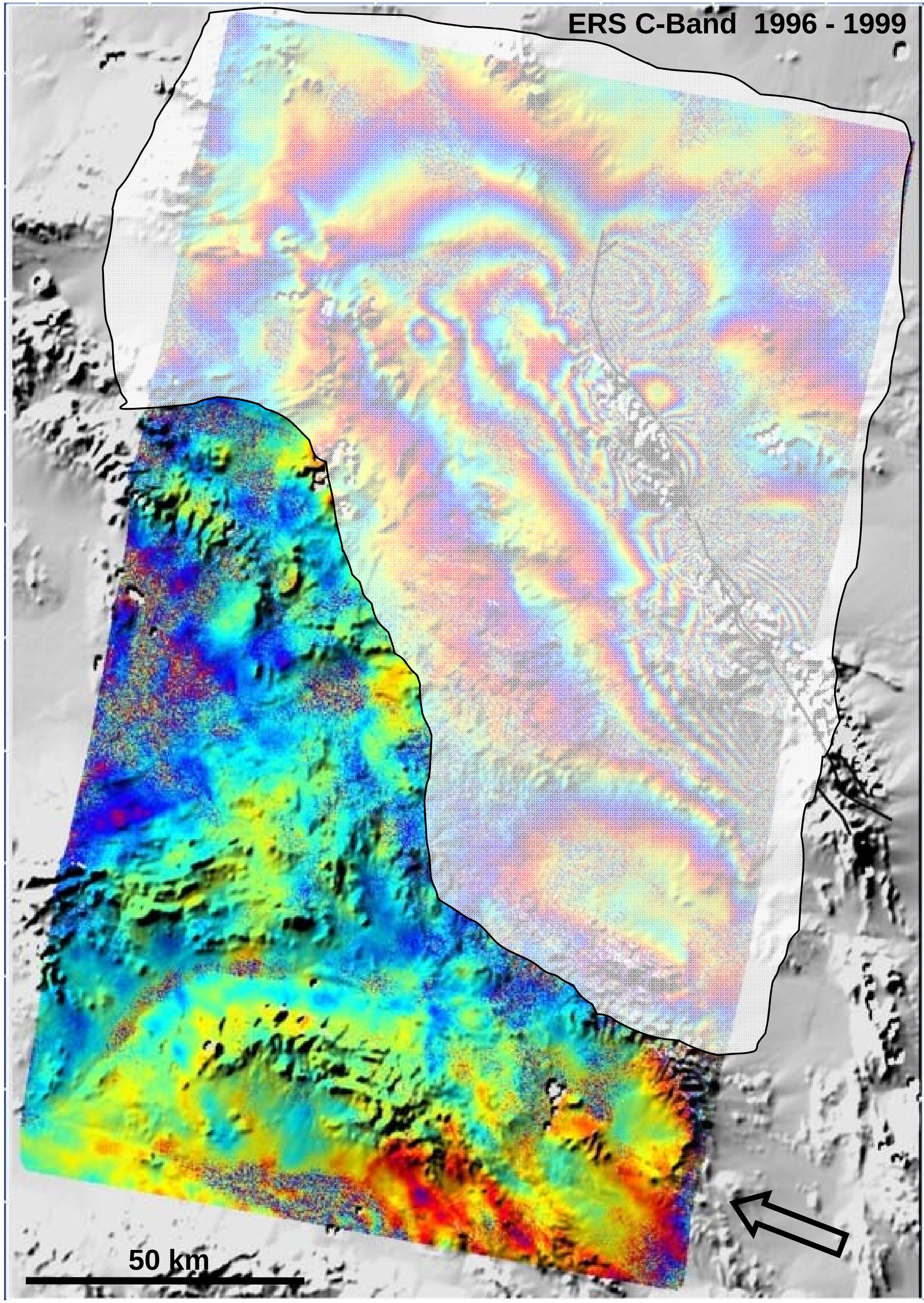


- Atmospheric errors sometimes correlate with topography
- The turbulent part is smooth, has more power at the larger spatial scales
- Sometimes shows anisotropic patterns

King Abdullah University of Science and Technology

KAUST

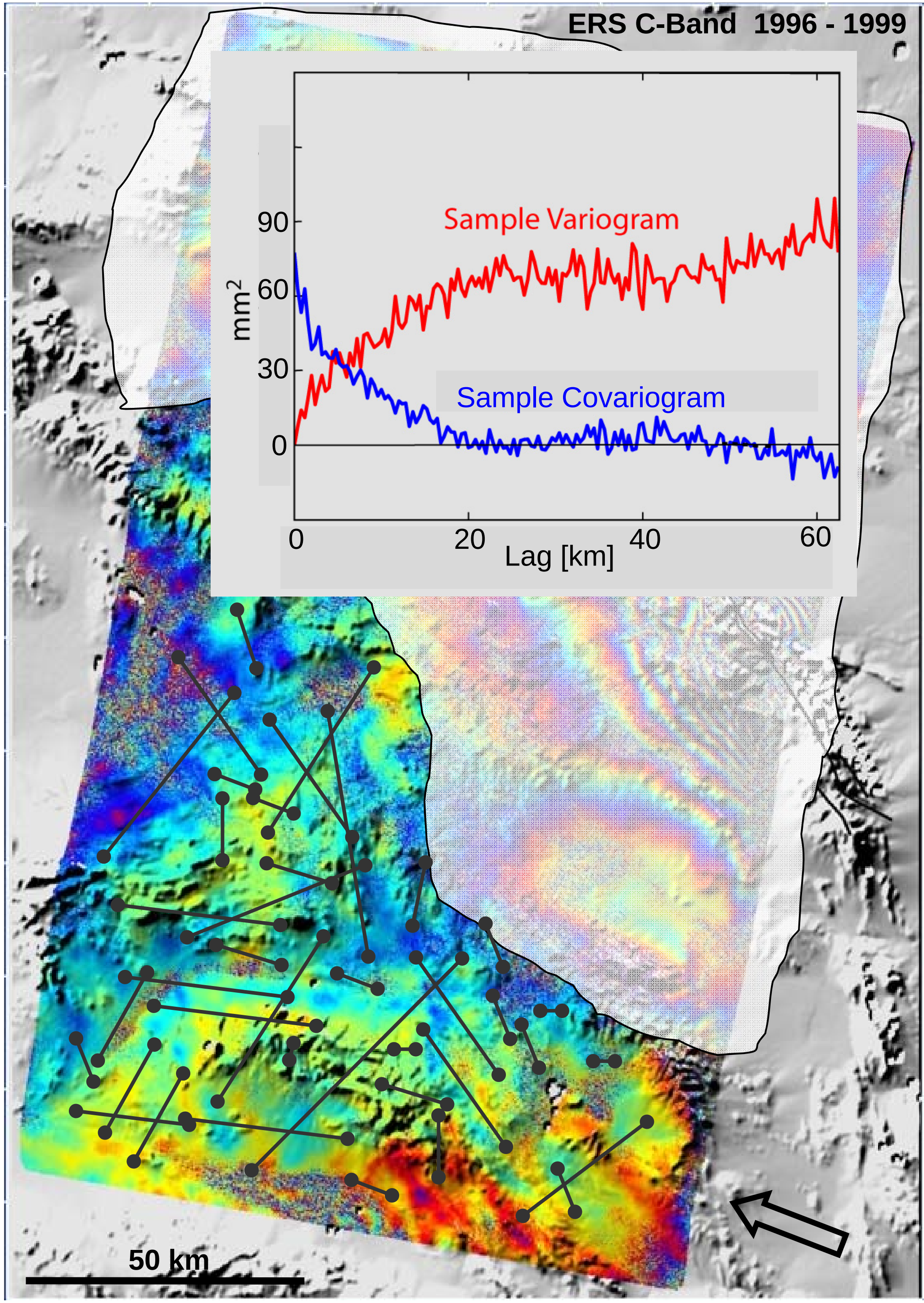
How do we Estimate Errors in InSAR Data?



Zirkuh, Iran, M_w7.2

- The interferogram contains both deformation and atmospheric signals
- Mask out deformation area
- Estimate error statistics in the non-deforming region
- Assume stationarity
- Start by building an empirical isotropic semi-variogram

Sudhaus and Jonsson, GJI, 2011

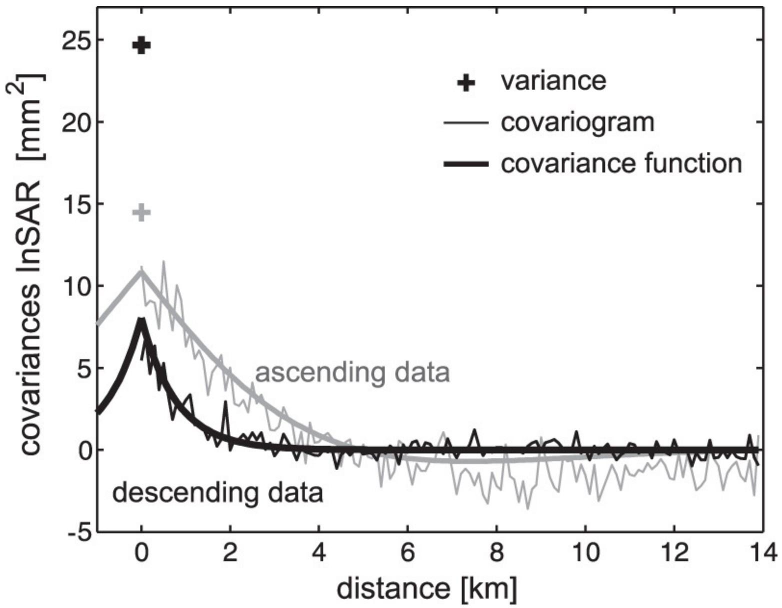


Zirkuh, Iran, M_w7.2

- For each distance-bin, select random point-pairs and calculate the semi-variogram value
- At certain distance the variogram has a plateau
- For second-order stationary conditions, we have a relationship between semi-variograms and covariograms

Covariance Function Fitting

Sudhaus and Jónsson, GJI. 2009

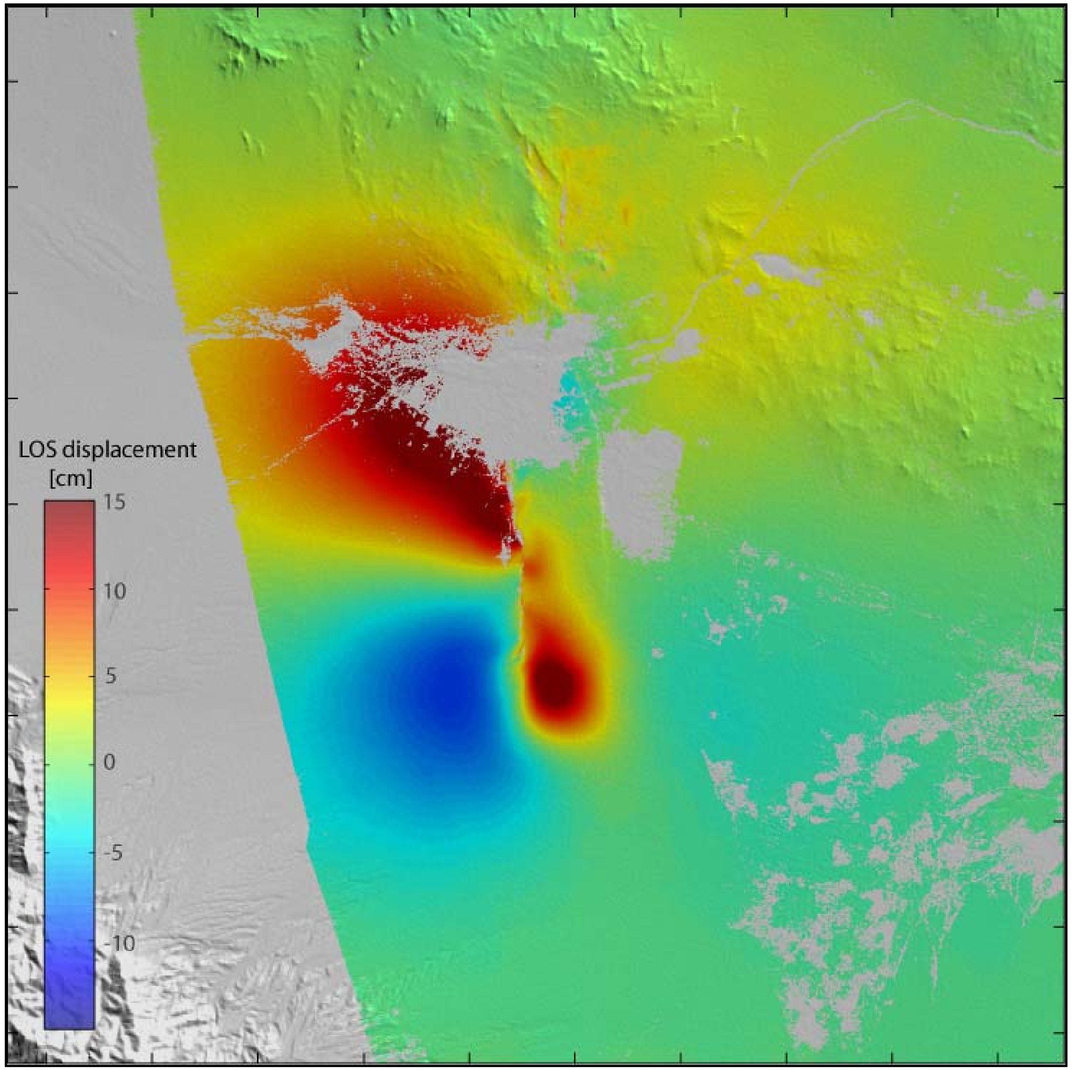


- To obtain a covariance value for any distance, we fit a covariance function to the covariogram, common function types are

- and a function allowing for anti-correlation:

- With a covariance function, we can form the data covariance matrix

Large Covariance Matrices!

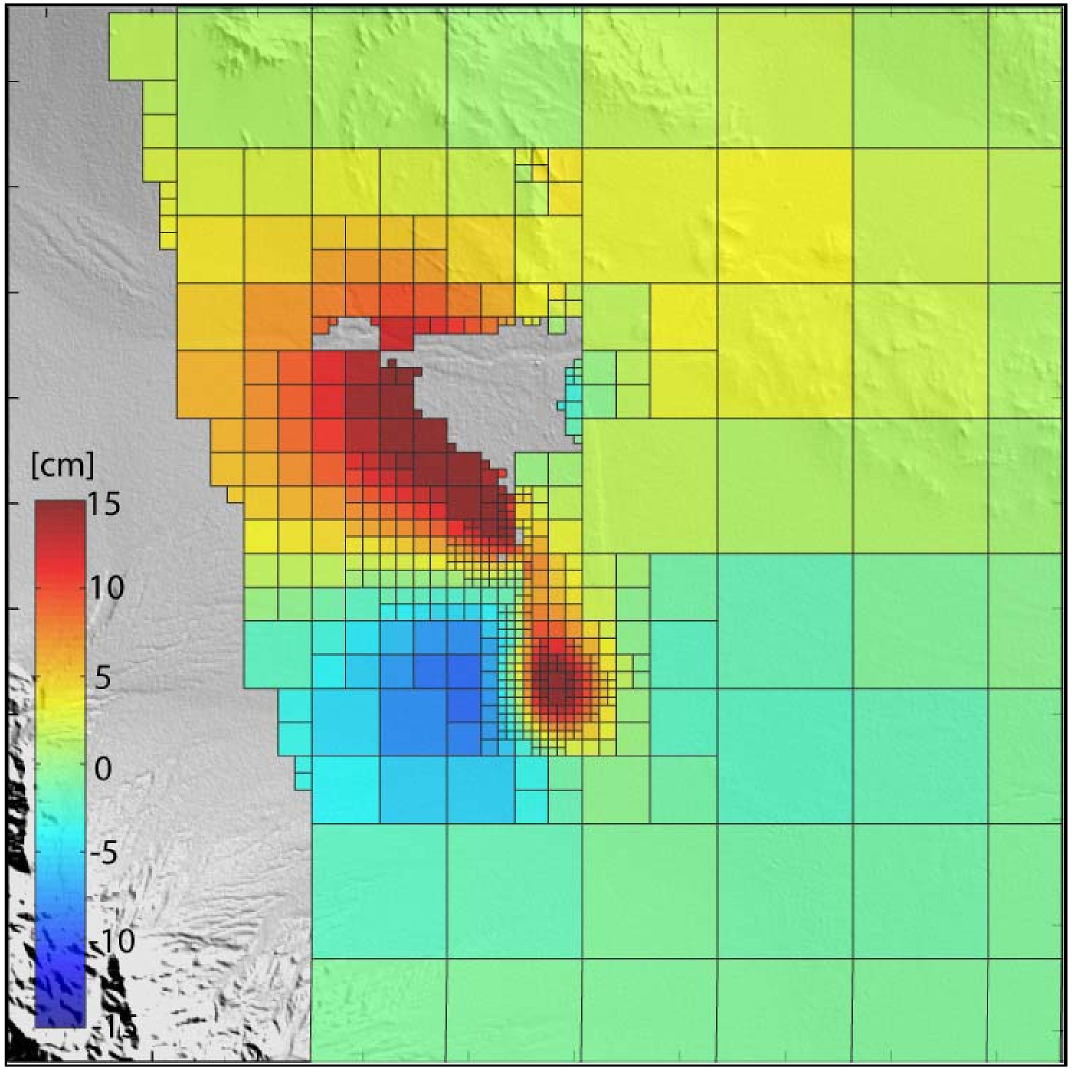


2003 Bam (Iran) Earthquake

- First step is phase unwrapping
- This interferogram has 6 million data points
- The corresponding data covariance matrix Σ_d includes 3.6×10^{13} elements! (~150TB)
- What to do?

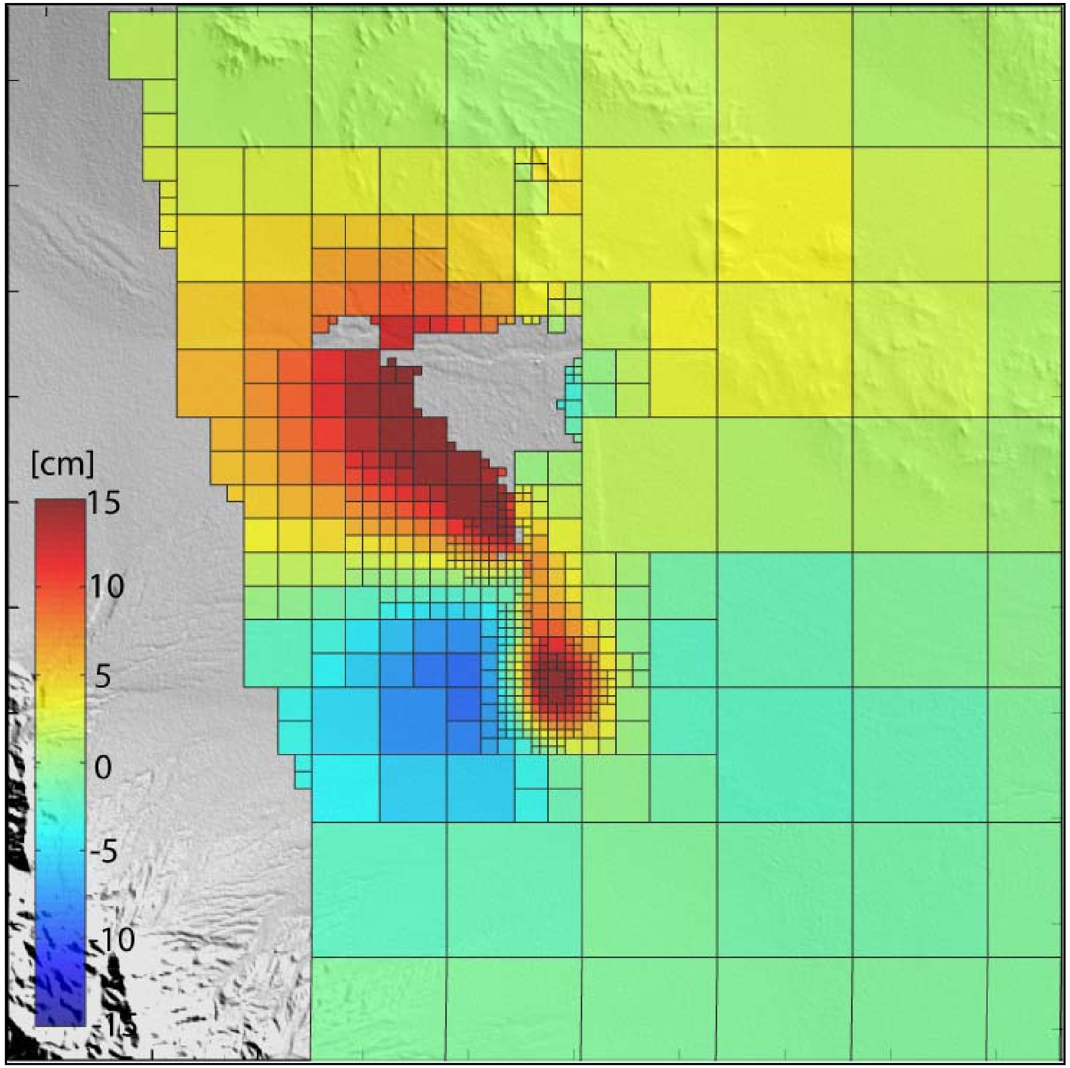
–> Data reduction

InSAR Data Reduction



- Quadtree partitioning or Subsampling
- Assign e.g. the mean or median of displacement each quadtree square to the center of the square
- Here the mean, can use linear subsampling:
- From 6 million data points to 600, in this case

Error Propagation



- How do we get an error estimate for each of the down-sampled points?
- When using a linear subsampling:
  we can simply propagate the full data covariance matrix:
- One problem is the size of Σ_d (~150TB)
- Still possible to calculate, e.g. by storing only single lines of Σ_d at once
- Σ_{sub} only 600×600 matrix

King Abdullah University of Science and Technology KAUST

Fault Modeling using InSAR Data

Fault Modeling

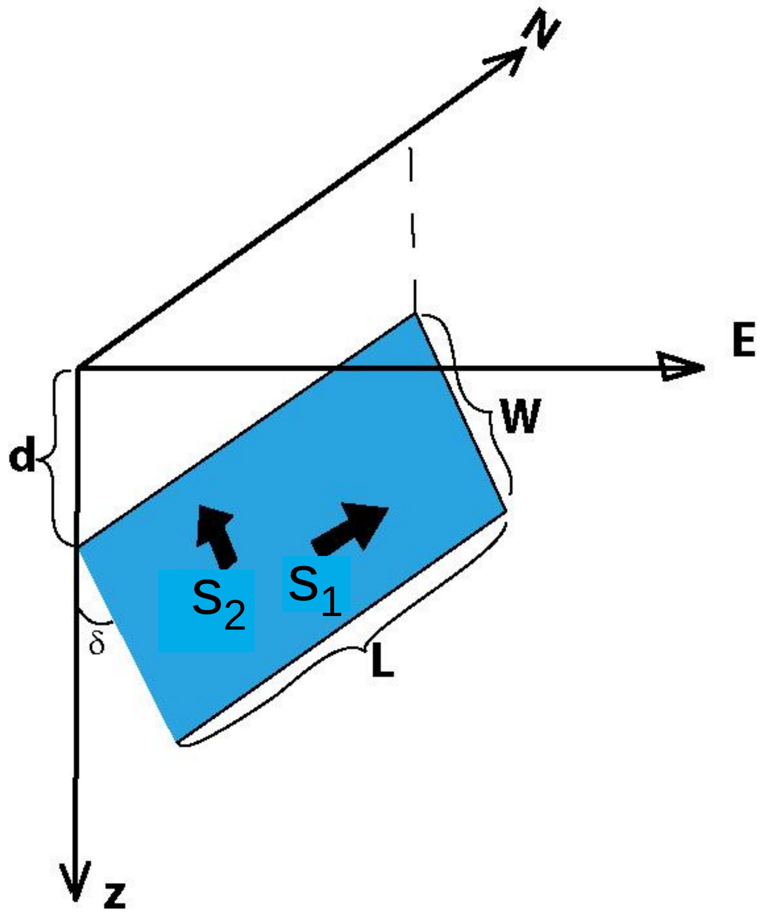
First let's assume the data vector $\boldsymbol{d}_{sub}$ can be described by a finite number of model parameters **m**:

where *g* relates slip on a fault (described by **m**) to surface displacements. Then we want to minimize

Where weighting matrix W is found by decomposition (Cholesky) of the inverse of the covariance matrix:

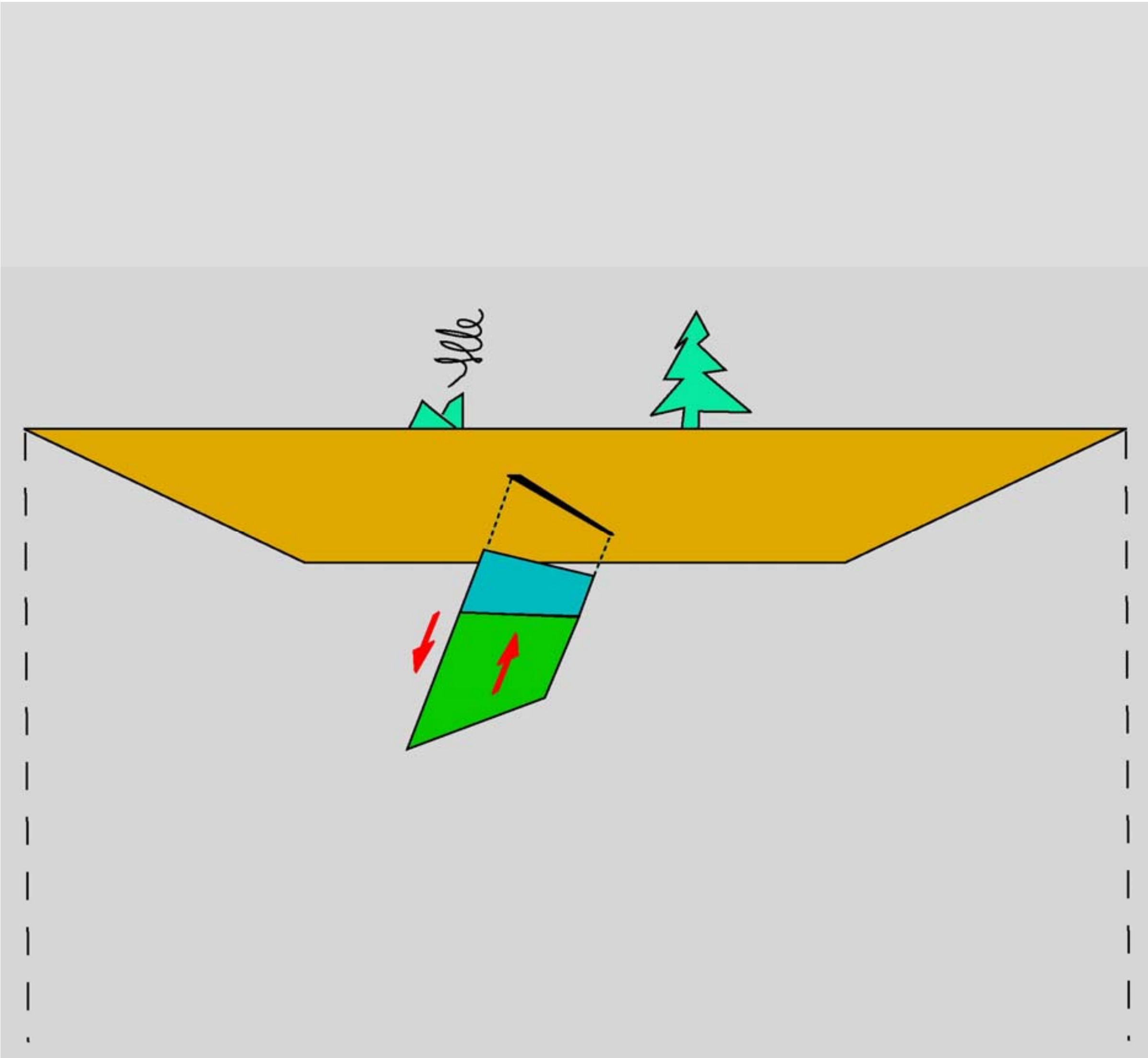
The Green's functions can be from elastic homogeneous or layered half-space, or some other earth model. Use 10 model parameters that describe the geometry and displacement across a rectangular surface:

$$\mathbf{m}^T=[L,W,d,\delta,\phi,x_N,x_E,s_1,s_2,s_3]$$

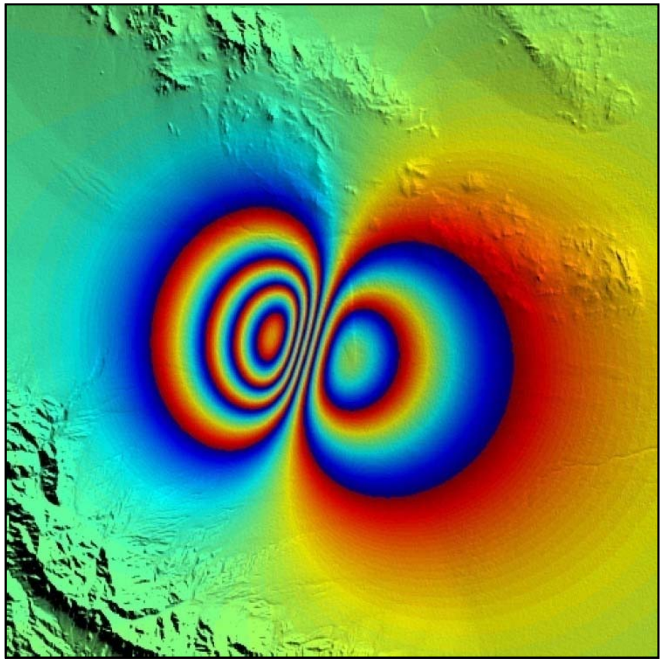


Thus, the problem is to find the set of model parameters **m** that best match our observed displacements **d**

Fault Modeling

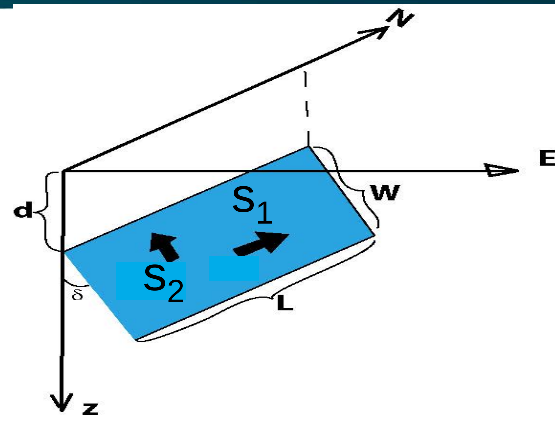


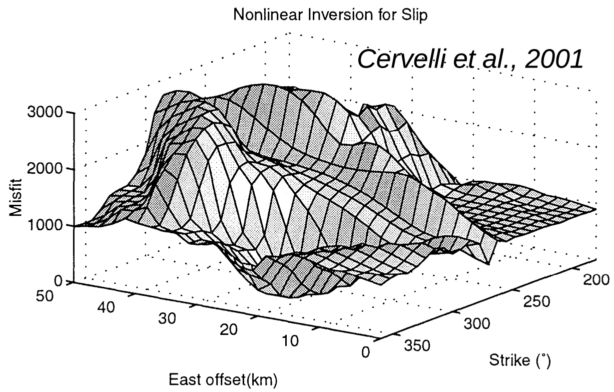
We search for the set of model parameters **m** that minimizes:



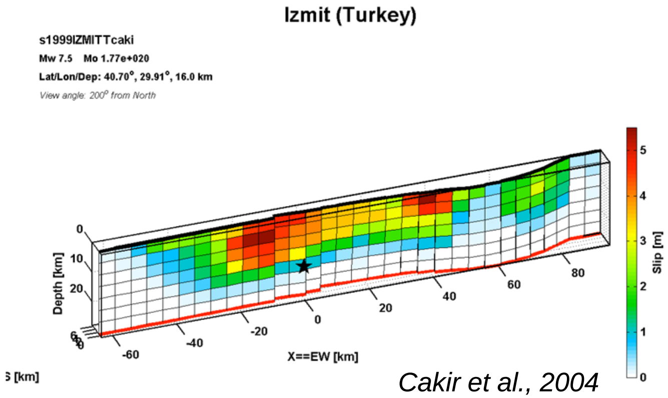
Fault Modeling

- For moderate-sized earthquake, do the fault parameter estimation in two steps:
- First, we search for the for the best set of parameters of one or more rectangular faults using non-linear optimization scheme, like simulated annealing, followed by a derivative based method (repeat a few times to ensure stability)
- In the second step, we use the location, strike, and dip of the fault plane found in step 1, expand the fault's length and width, discretize the fault plane into many fault patches, and invert for fault spatially variable slip using a certain degree of smoothening





Cervelli et al., 2001

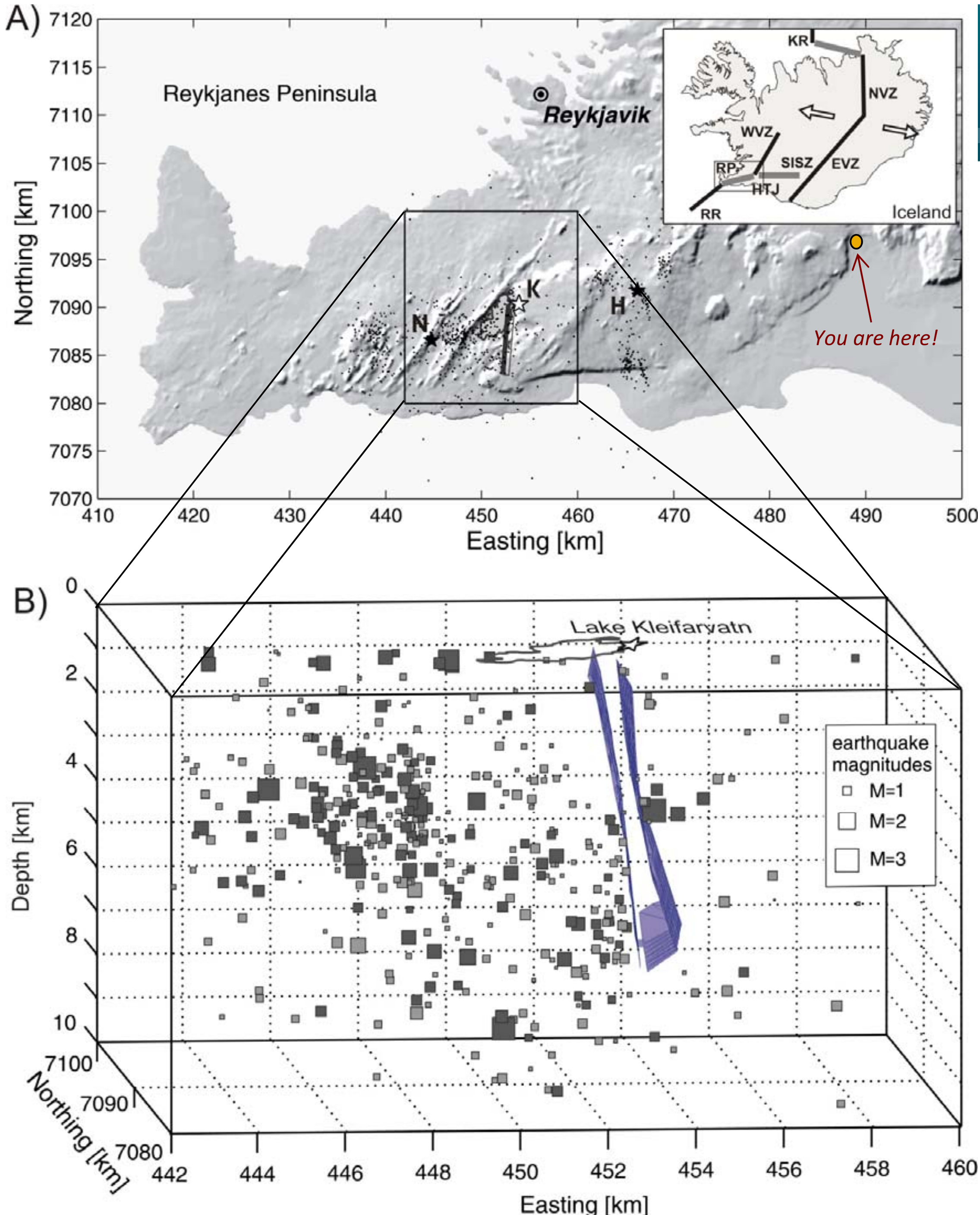


Cakir et al., 2004

King Abdullah University of Science and Technology

KAUST

How can we estimate the Model Parameter Uncertainties?



Woessner et al. in prep. 2011

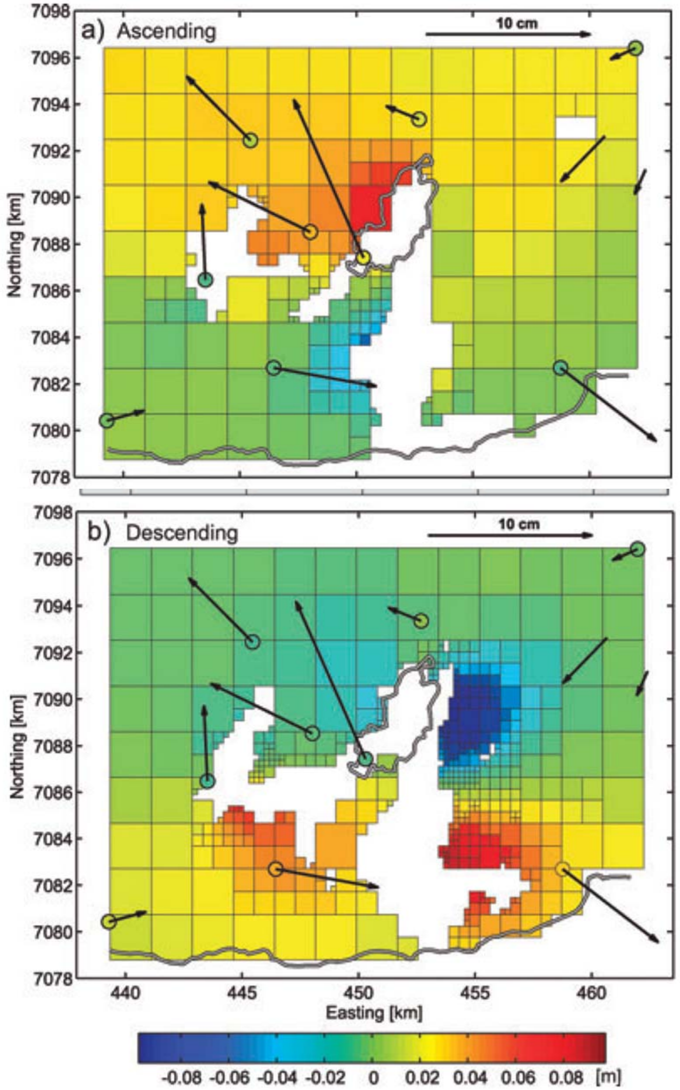
Kleifarvatn EQ

Kleifarvatn Earthquake M_w5.9

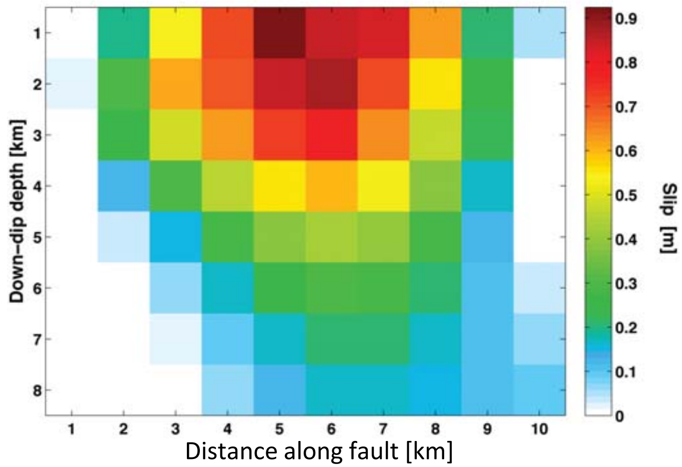
- Occurred in June 2000, 15 km south of Reykjavik
- North-south, near-vertical, right-lateral strike-slip earthquake
- 3D volume shows aftershocks and the "envelope" shows the range of possible fault geometries as estimated from GPS and InSAR

Kleifarvatn InSAR and GPS Data

Sudhaus and Jónsson, GJI. 2009

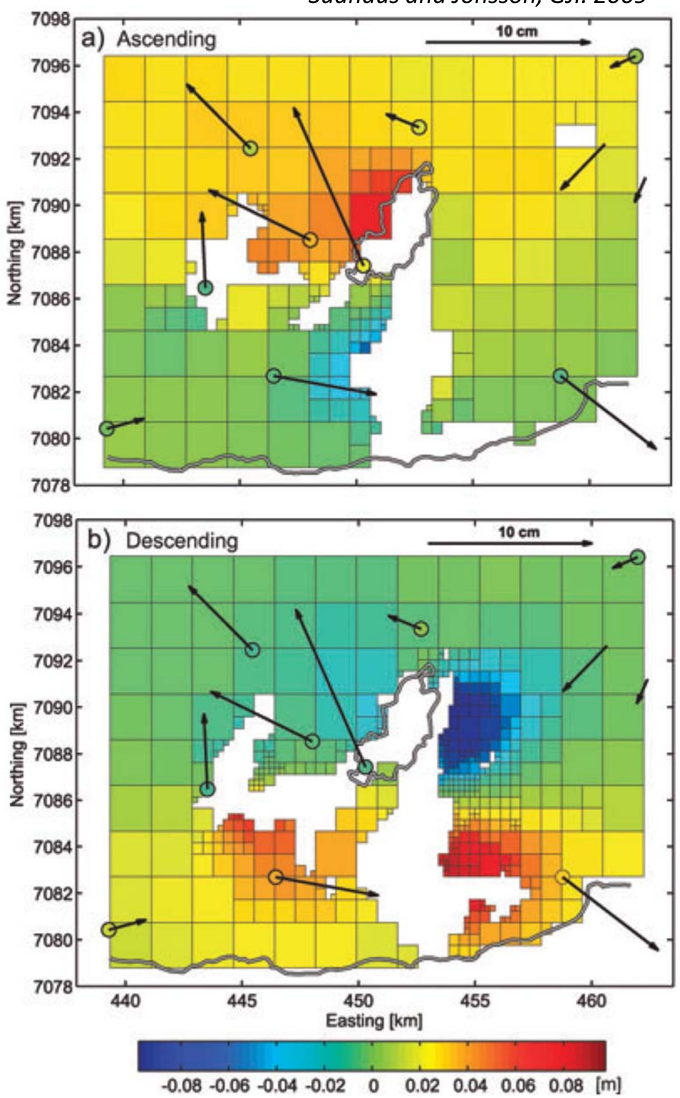


- Here we use all the tricks described:
  - Unwrap the InSAR data
  - Generate sample covariograms
  - Fit-covariance functions
  - Subsample InSAR data
  - Build covariance matrices
  - Combine data sets
- Estimate optimal model geometry and slip distribution

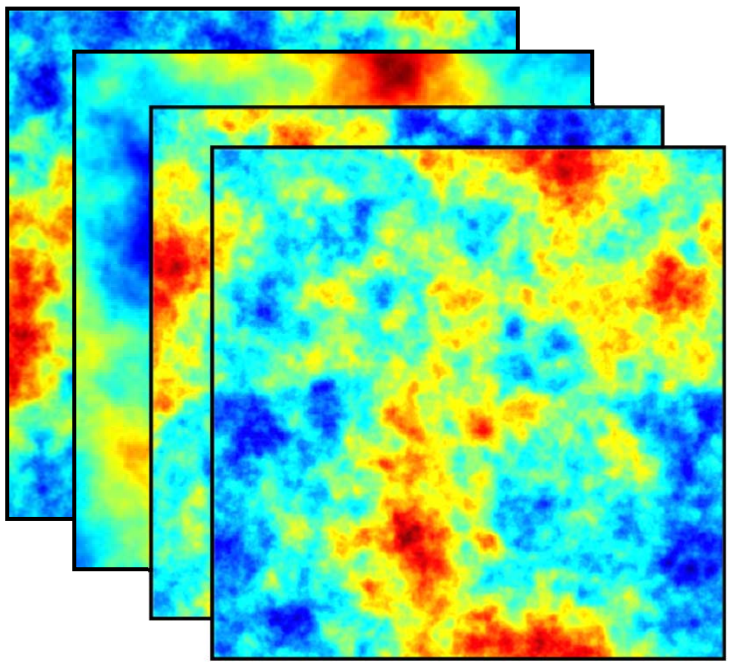


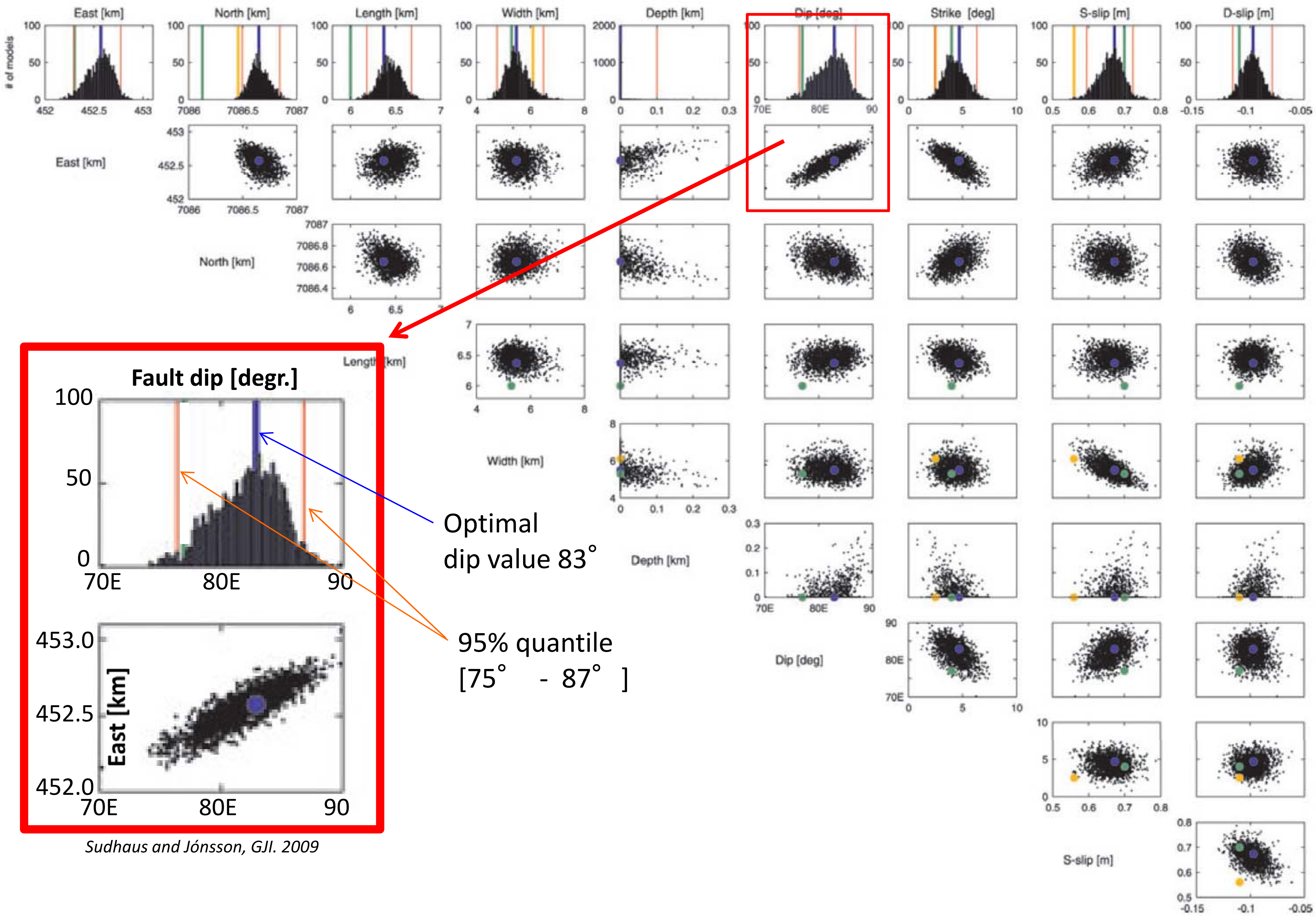
Kleifarvatn Fault Model Uncertainty Estimation

Sudhaus and Jónsson, GJI. 2009



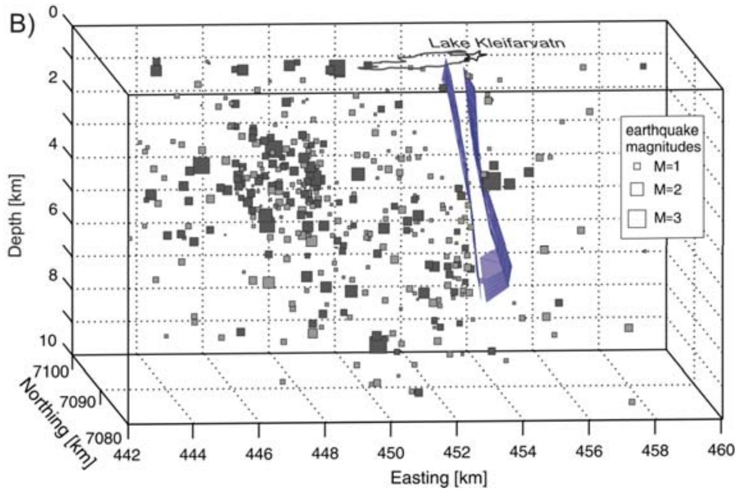
- Generate 2500 synthetic data errors
- Add each error realization to the data and estimate fault geometry and slip distribution using biased data
- Results in 2500 different source models

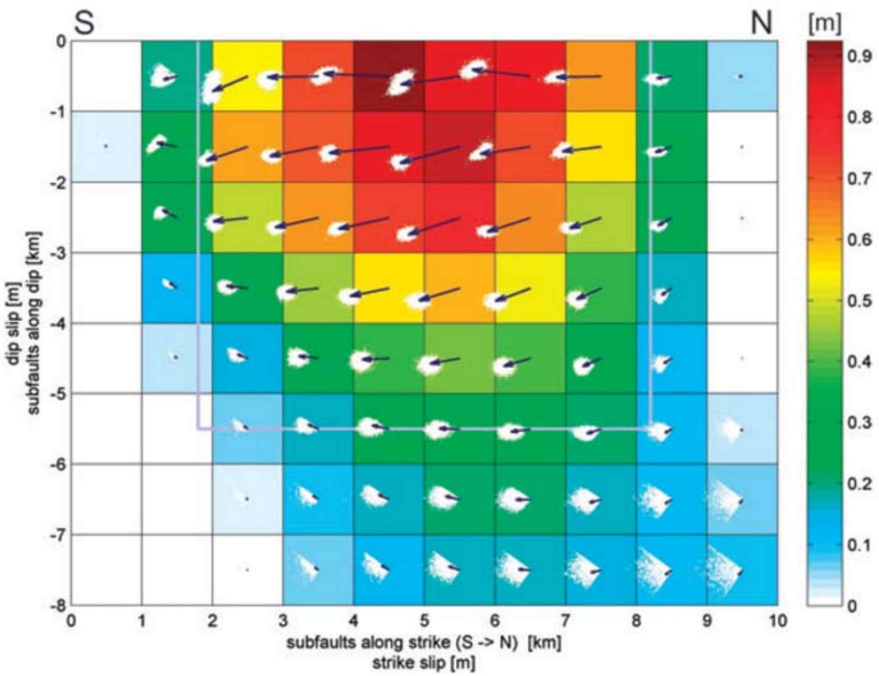




Sudhaus and Jónsson, GJI. 2009

Kleifarvatn Fault Model and Slip Distribution



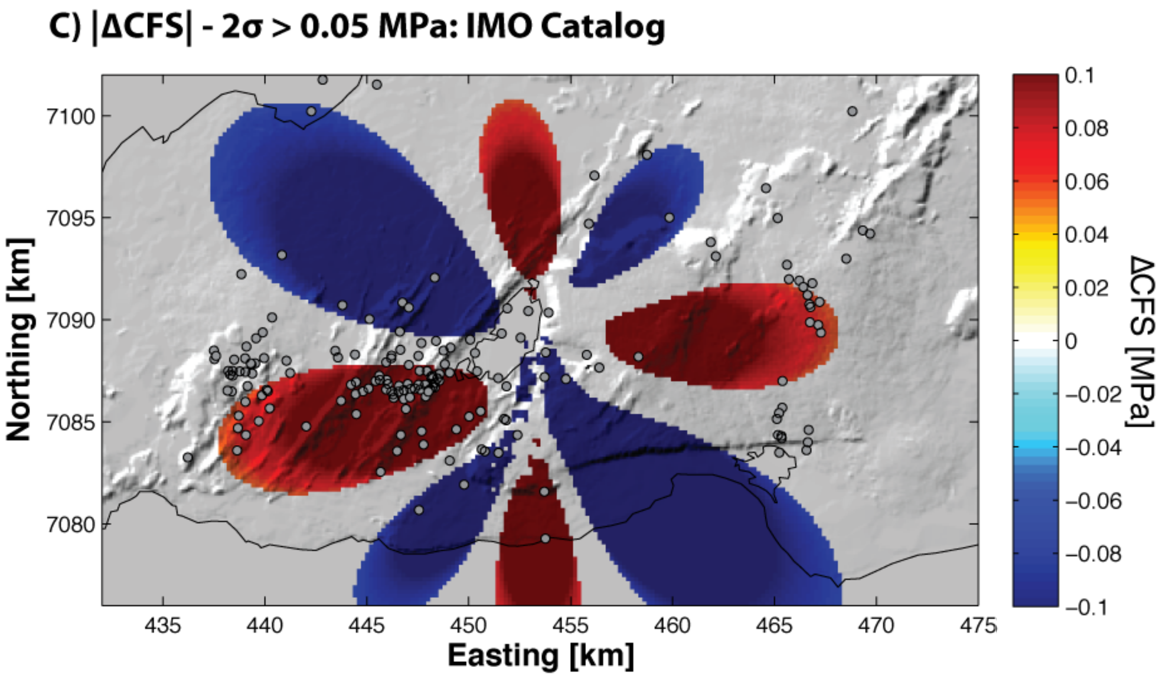


Sudhaus and Jónsson, GJI. 2009

- 3D volume shows an envelope indicating the variability of the resulting 2500 fault geometries
- Slip distribution is simple, shows the slip of the optimal model
- Slip values of the 2500 models are shown as white dots
- Slip magnitude and rake well constrained near the surface
- Poor resolution with depth

What can we do with Model Parameter Errors?

Applications of Model Parameter Uncertainties



Woessner et al. "almost" submitted, 2011

- Coulomb Failure Stress changes (ΔCFS) near the Kleifarvatn fault
- We have estimated uncertainties for the fault geometry and slip distribution
- Propagate fault errors to obtain ΔCFS errors
- Reduces the size of "reliable" CFS changes
- Islands of reliable and strongly positive/negative ΔCFS remain

Conclusions

- Important to remember that all data are uncertain
- Geodetic data uncertainties are fairly well understood, in many cases, but can be quite variable and tricky to estimate
- Correlations are important, need full data covariance matrices
- Covariance matrices help in combining different data in model parameter estimations, e.g. multiple interferograms. However, different types of data can still be complicated to combine
- Here we only discussed the influence of estimated data errors on model parameter uncertainties. When SNR is high, model errors dominate (e.g. homogenous/layered halfspace simplifications)
- Fault model uncertainties are more important than a single “best-fit” candidate fault model, also for fault model applications, e.g. when comparing CFS changes with aftershock locations