VERY short intro to double-difference relocation

Felix Waldhauser

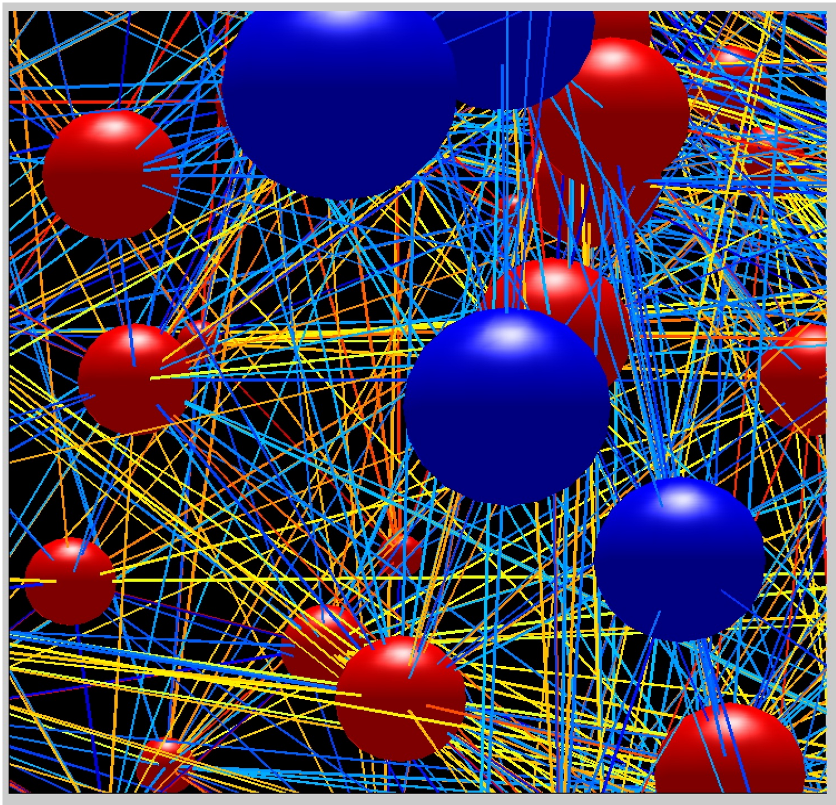
Lamont-Doherty Earth Observatory, Columbia University

HYPODD: A shareware software package for relocating earthquakes using the double-difference algorithm (Waldhauser and Ellsworth 2000; Waldhauser 2001). Relates the residual between the observed and predicted phase travel time difference for pairs of earthquakes observed at common stations to changes in the vector connecting their hypocenters through the partial derivatives of the travel times for each event with respect to the unknown.

ph2dt: computes/optimizes delay time networks, outlier detection, input for HypoDD.

hypoDD: inversion of weighted delay time data for relative locations.

https://github.com/fwaldhauser/hypodd
https://www.ldeo.columbia.edu/~felixw/HYPODD



Reading material

Bulletin of the Seismological Society of America, 90, 6, pp. 1353–1368, December 2000

A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California

by Felix Waldhauser and William L. Ellsworth

Abstract We have developed an efficient method to determine high-resolution hypocenter locations over large distances. The location method incorporates ordinary absolute travel-time measurements and/or cross-correlation *P*-and *S*-wave differential travel-time measurements. Residuals between observed and theoretical travel-time differences (or double-differences) are minimized for pairs of earthquakes at each station while linking together all observed event-station pairs. A least-squares solution is found by iteratively adjusting the vector difference between hypocentral pairs. The double-difference algorithm minimizes errors due to unmodeled velocity structure without the use of station corrections. Because catalog and cross-correlation data are combined into one system of equations, interevent distances within multiplets are determined to the accuracy of the cross-correlation data, while the relative locations between multiplets and uncorrelated events are simultaneously determined to the accuracy of the absolute travel-time data. Statistical resampling methods are used to estimate data accuracy and location errors. Uncertainties in double-difference locations are improved by more than an order of magnitude compared to catalog locations. The algorithm is tested, and its performance is demonstrated on two clusters of earthquakes located on the northern Hayward fault, California. There it collapses the diffuse catalog locations into sharp images of seismicity and reveals horizontal lineations of hypocenters that define the narrow regions on the fault where stress is released by brittle failure.

Introduction

Seismicity analysis for the study of tectonic processes, earthquake recurrence, and earthquake interaction requires knowledge of the precise spatial offset between the earthquake hypocenters. This is particularly the case for crustal faults that are most readily investigated using microseismic activity. The location uncertainty of routinely determined hypocenters is typically many times larger than the source dimension of the events itself, thus putting limits on the study of the fine structure of seismicity.

The accuracy of absolute hypocenter locations is controlled by several factors, including the network geometry, available phases, arrival-time reading accuracy, and knowledge of the crustal structure (Pavlis, 1986; Gomberg *et al.*, 1990). The use of a one-dimensional reference velocity model to locate the earthquakes limits the location accuracy, since three-dimensional velocity variations can introduce systematic biases into the estimated travel times. One can partially account for the velocity variations by including station and/or source terms in the location procedure (e.g., Douglas, 1967; Pujol, 1988; Hurukawa and Imoto, 1992; Shearer, 1997) and/or by jointly inverting the travel-time data for hypocenters and velocity structure (e.g., Crosson, 1976; Ellsworth, 1977; Roecker, 1981; Thurber, 1983; Michael, 1988; Kissling *et al.*, 1994).

The effects of errors in structure can also be effectively minimized by using relative earthquake location methods (Poupinet *et al.*, 1984; Fréchet, 1985; Frémont and Malone, 1987; Got *et al.*, 1994) (for a discussion on relative location errors see Pavlis [1992]). If the hypocentral separation between two earthquakes is small compared to the event-station distance and the scale length of the velocity heterogeneity, then the ray paths between the source region and a common station are similar along almost the entire ray path. In this case, the difference in travel times for two events observed at one station can be attributed to the spatial offset between the events with high accuracy. This is because the absolute errors are of common origin except in the small region where the raypaths differ at the sources.

We can further improve location precision by improving the accuracy of the relative arrival-time readings using waveform cross-correlation methods. Two earthquakes produce similar waveforms at a common station if their source mechanisms are virtually identical and their sources are colocated so that signal scattering due to velocity heterogenei-

July 2014

User Guide to

HYPODD

Version 2.1b – 01/2013

A Computer Program to Compute Double-Difference Earthquake Locations

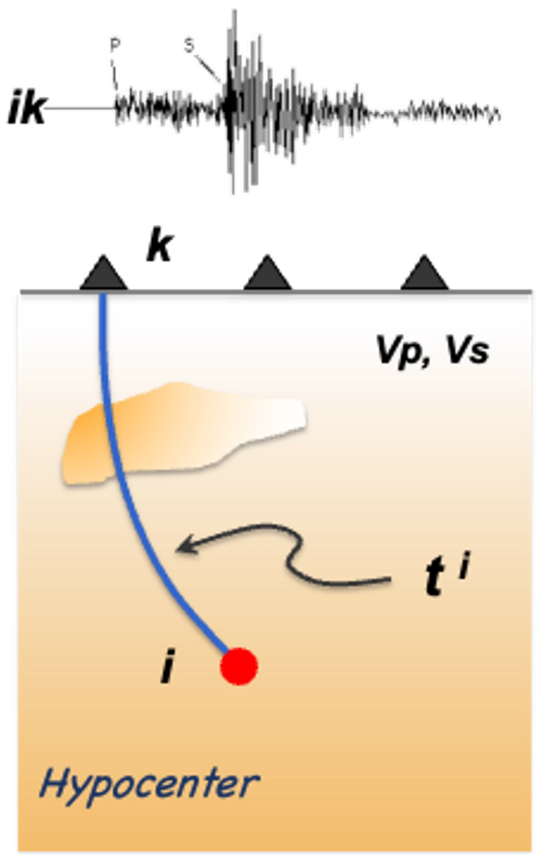
by

Felix Waldhauser

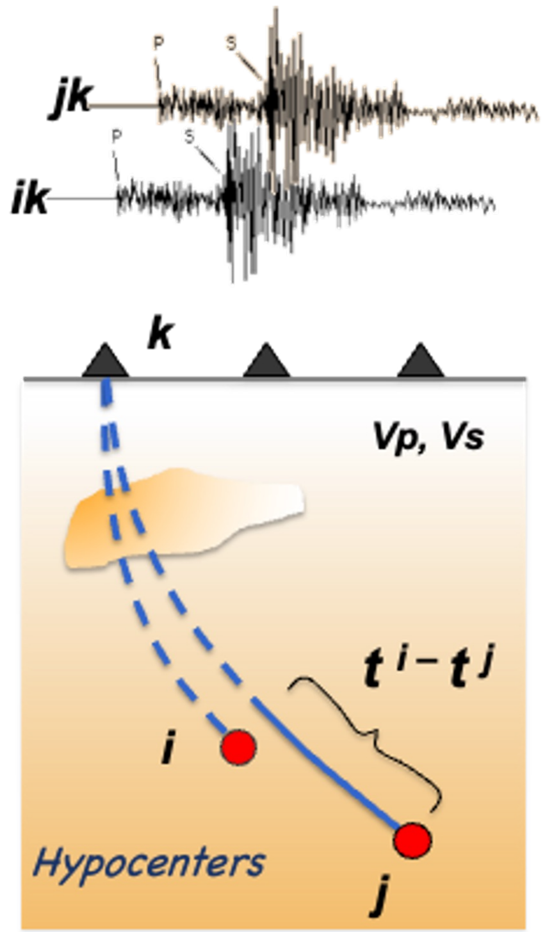
Lamont-Doherty Earth Observatory
Columbia University
felixw@ldeo.columbia.edu

Initial (absolute) location and double-difference relocation

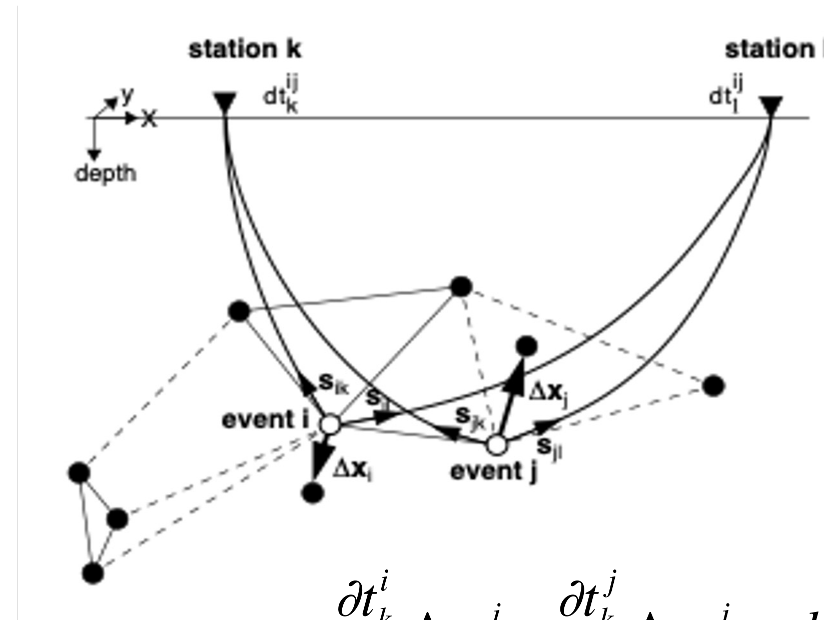
Initial (absolute) location



Relative location



Double-difference relocation



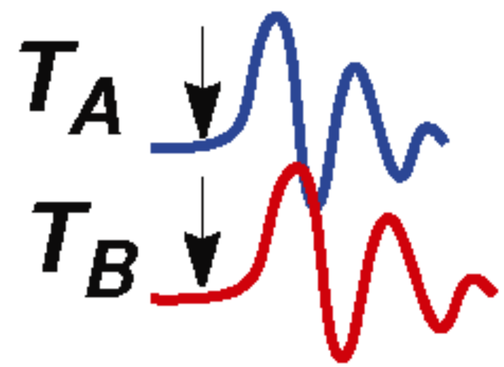
$$\frac{\partial t_k^i}{\partial \mathbf{m}}\Delta\mathbf{m}^i - \frac{\partial t_k^j}{\partial \mathbf{m}}\Delta\mathbf{m}^j = dr_k^{ij}$$

$$dr_k^{ij} = \left(t_k^i - t_k^j\right)^{obs} - \left(t_k^i - t_k^j\right)^{cal}$$

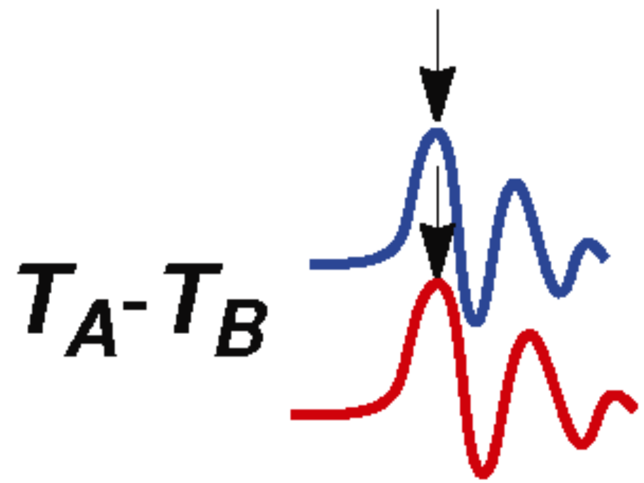
Waldhauser & Ellsworth, 2000 (DD)
Waldhauser 2001 (HypoDD)
Klein, 2002 (HypoInverse)

Arrival time picks vs. correlation delay times

- Automatic or analyst phase onset picks reported in earthquake bulletins of seismic networks. Relates to point of nucleation. Variable accuracy.



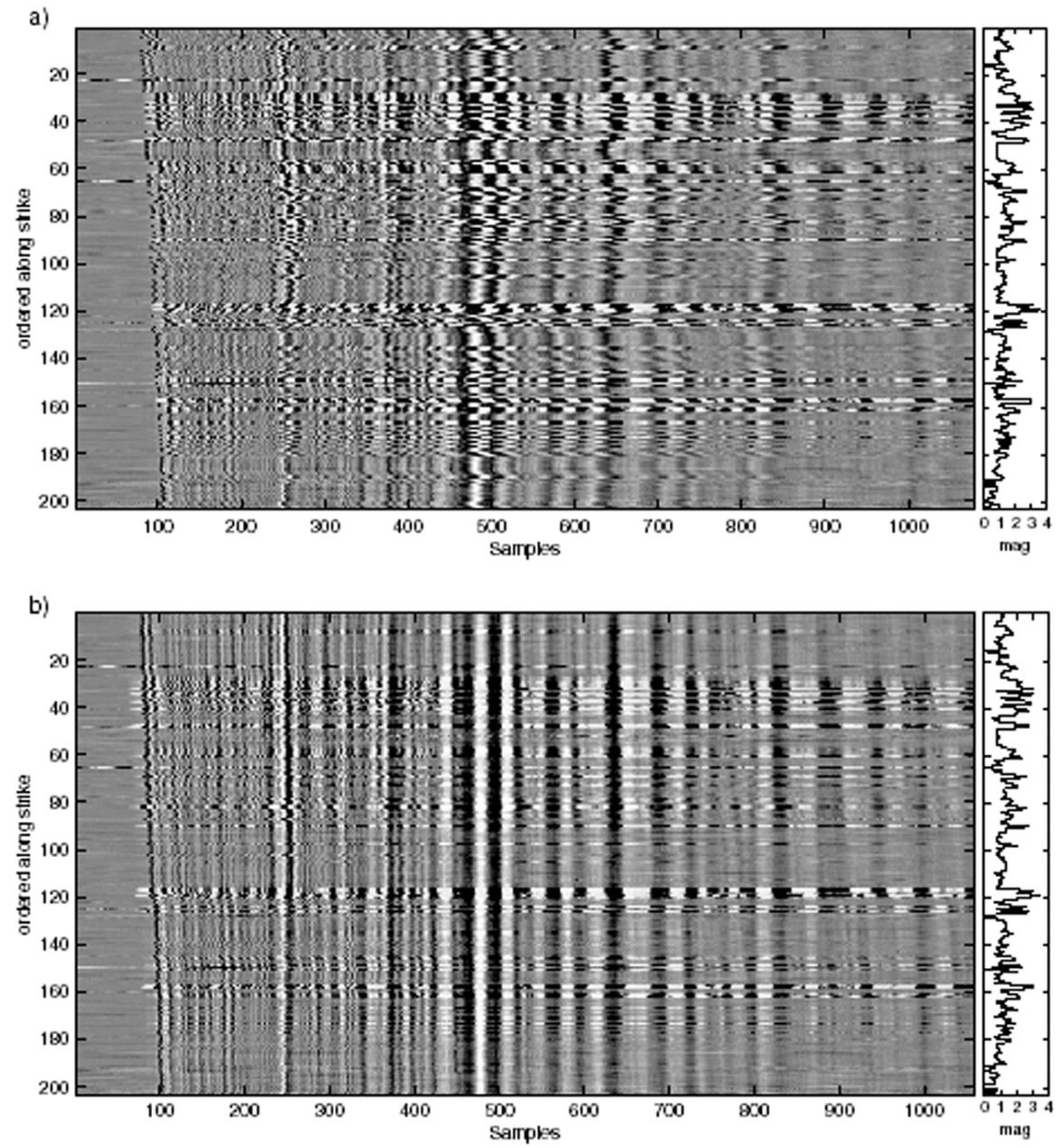
- From waveform cross-correlation of similar seismograms. Relates to point of maximum moment release. Sub-sample precision.



Arrival time picks vs. correlation delay times

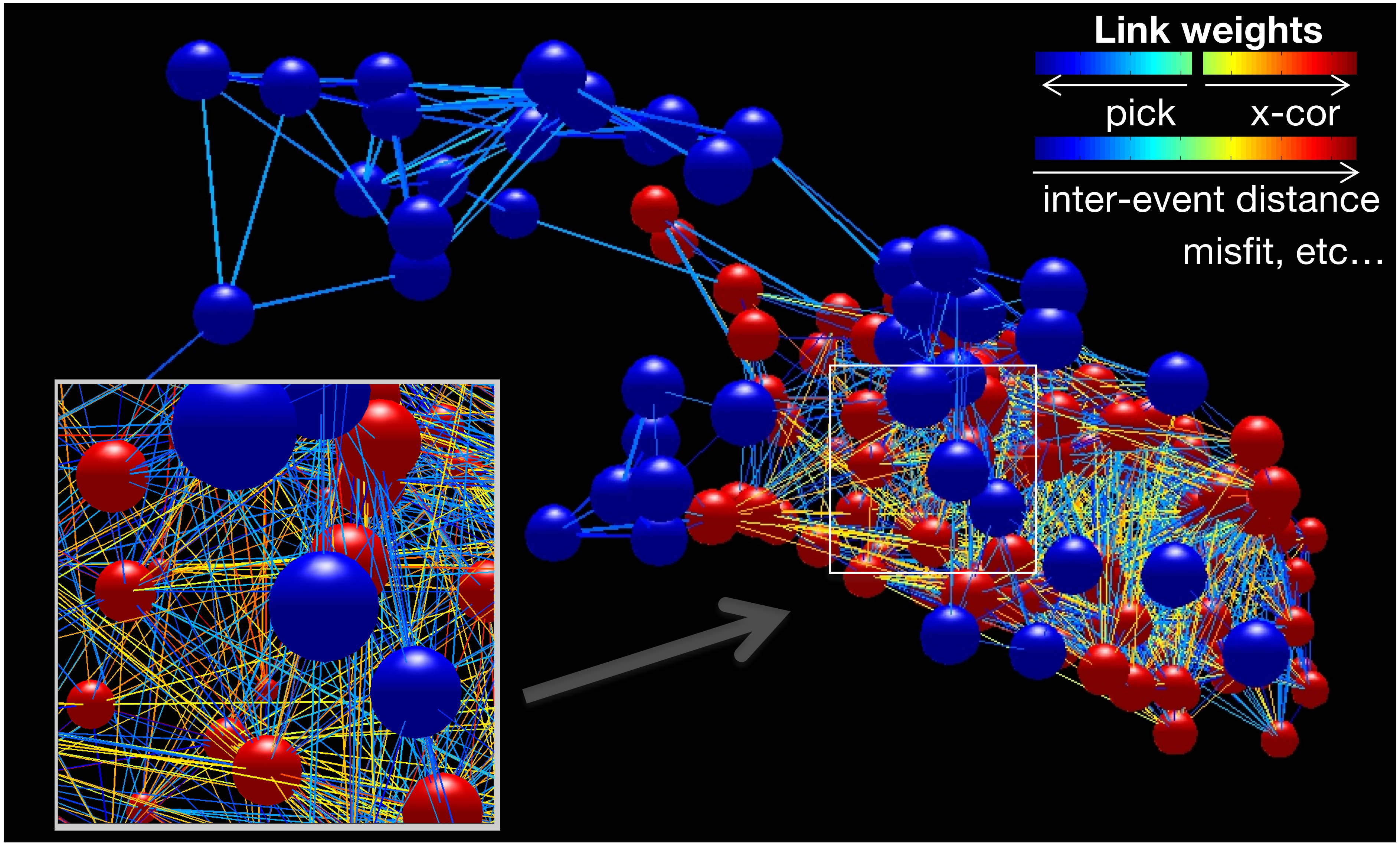
- Aligned on analyst P-wave picks

- Aligned on correlated S-waves



Schaff etal 2005

Network of dynamically weighted delay-time links

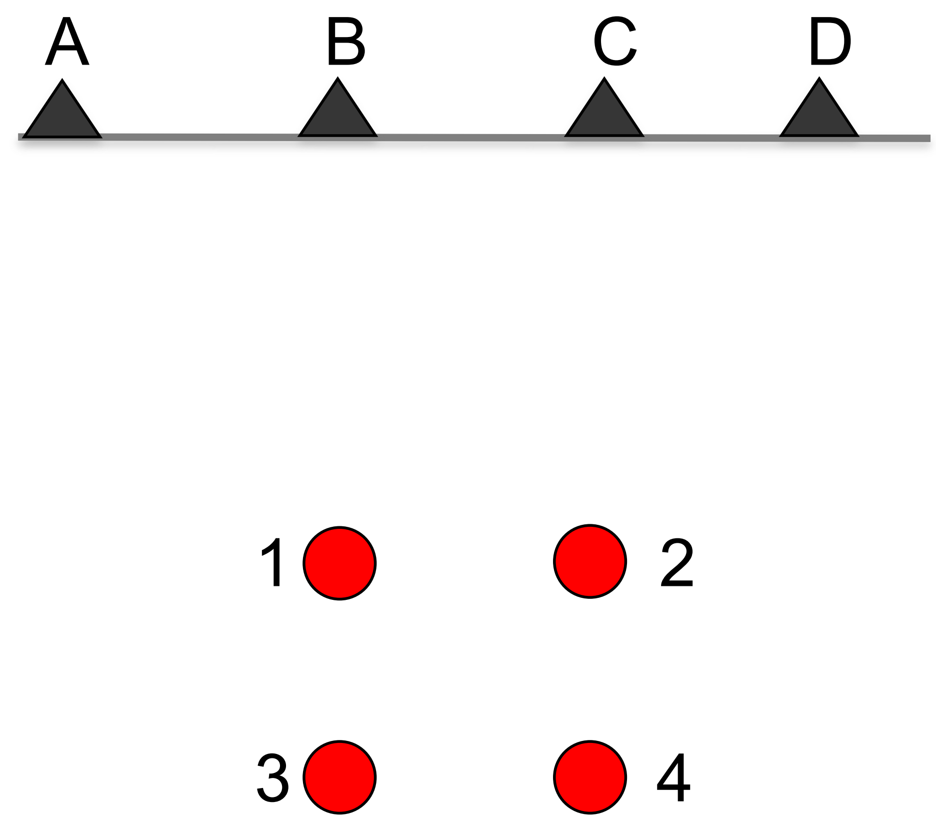


Waldhauser (2001)

Common misconceptions about HypoDD

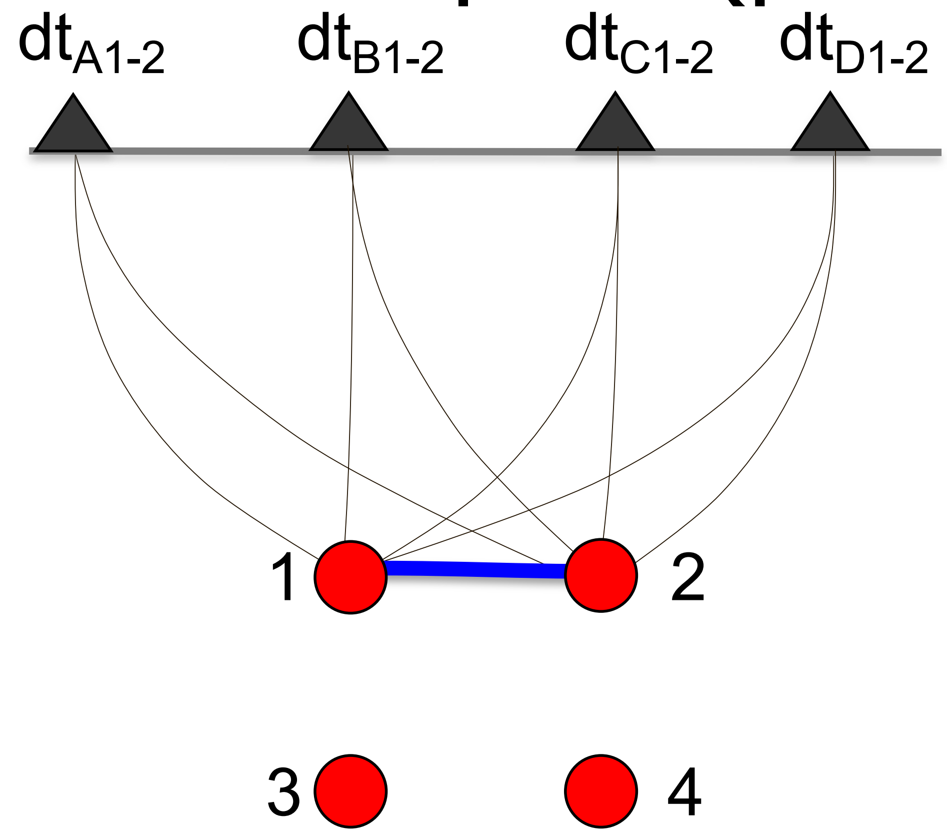
- *HypoDD processes redundant data.*
- *HypoDD removes events*. It does not. It does not relocate events that are ill constrained to begin with. Or spatially isolated single events.
- *Cross-correlation data replace phase picks*. Not in most cases; the two data sets are complementary.
- *It does not accept 3D models.* It accepts homogenous, 1-D layered, and 3D P/S wave velocity models (version 2) [https://github.com/fwaldhauser/HypoDD].

HypoDD does NOT processe redundant data.



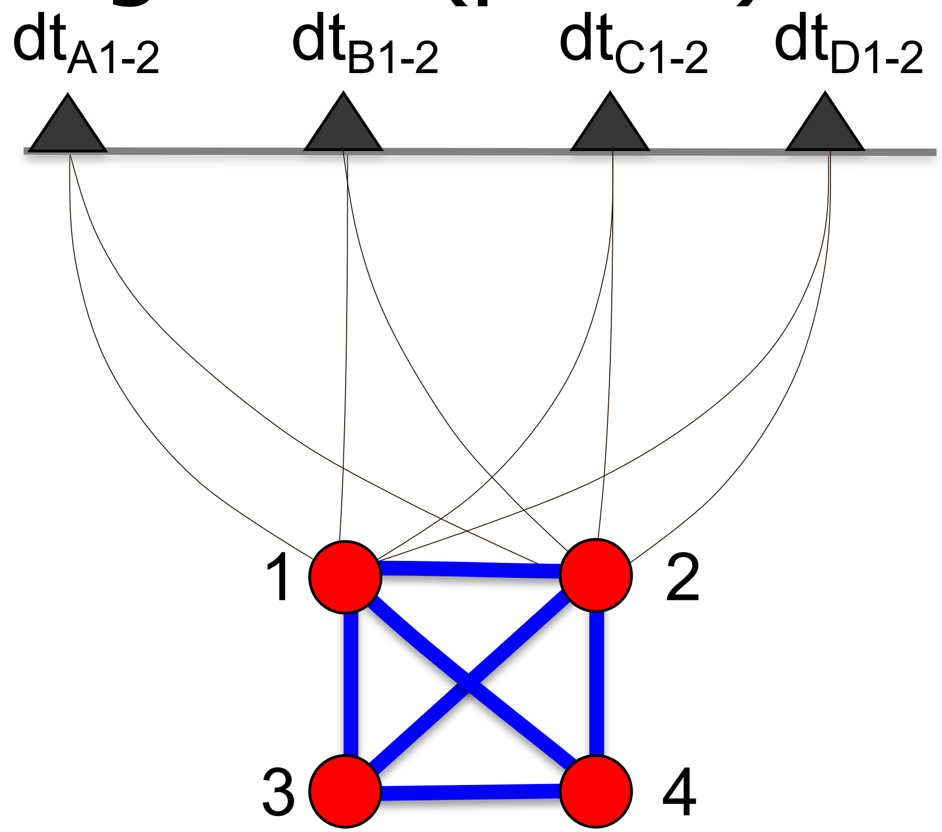
Delay times from phase onset picks (ph2dt)

- Delay times are computed for pairs of nearby events at common stations.



Linking events together (ph2dt)

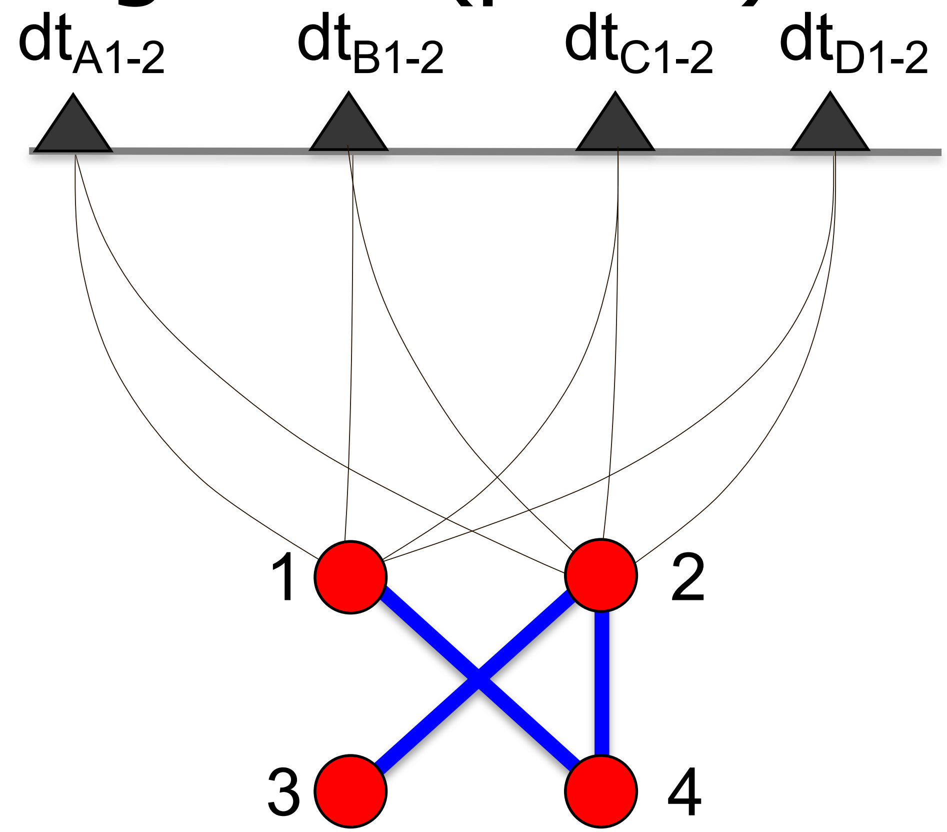
- Delay times are computed for ALL pairs of nearby events at common stations.
- A network of delay time links is computed so that a continuous chain of links connects all earthquakes.



Linking events together (ph2dt)

- Delay times are build for pairs of nearby events at common stations.

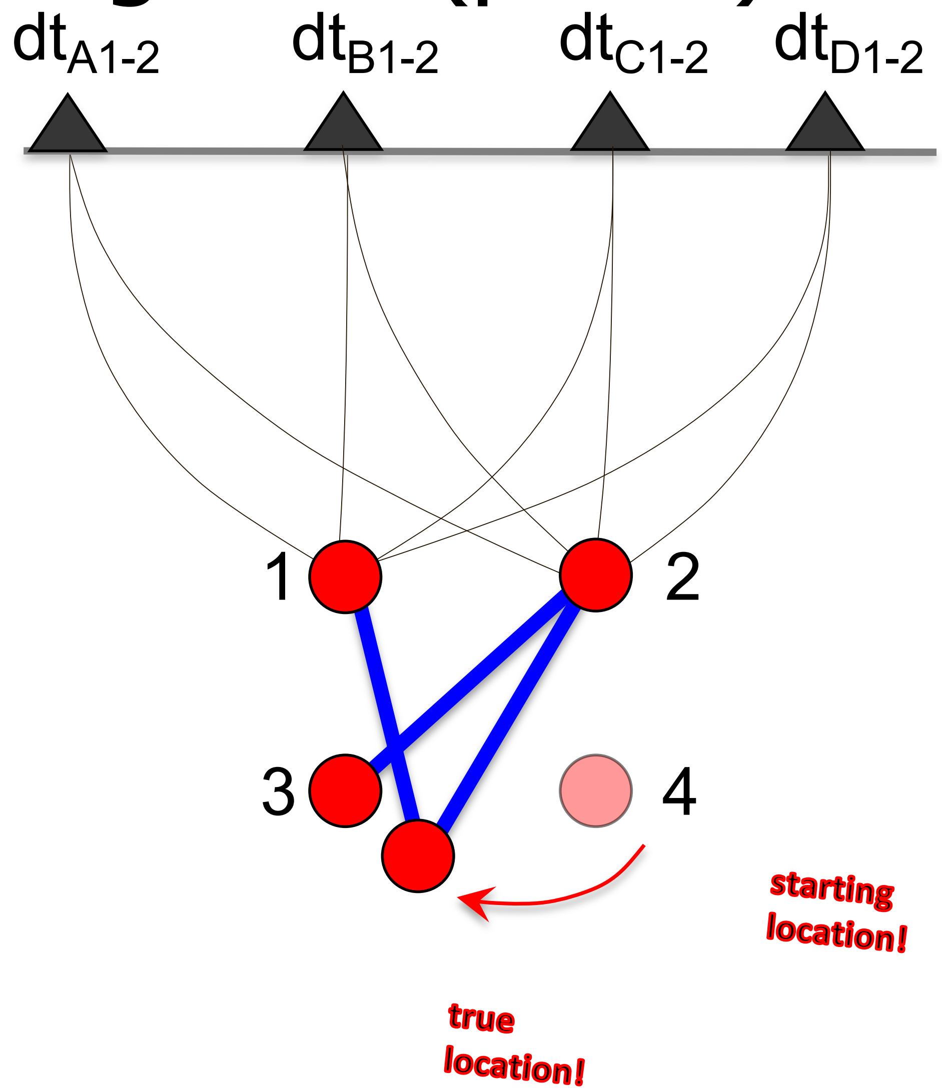
- A network of delay time links is computed so that a continuous chain of links connects all earthquakes.



Linking events together (ph2dt)

- Delay times are build for pairs of nearby events at common stations.

- A network of delay time links is computed so that a continuous chain of links connects all earthquakes.

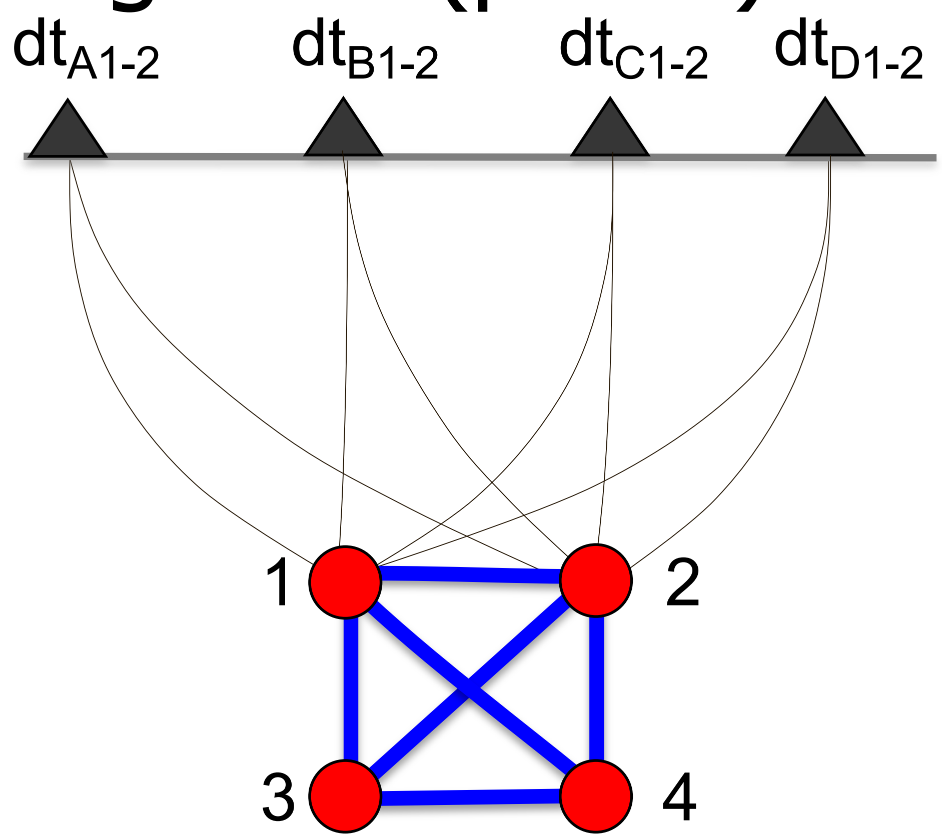


Linking events together (ph2dt)

- Delay times are build for pairs of nearby events at common stations.
- A network of delay time links is computed so that a continuous chain of links connects all earthquakes.

➢ *Data redundancy is wanted and important to find strong and weak links.*

- *mxn G matrix:*
  m = sum[1:Nev-1] × Nsta
  n = Nev x 4

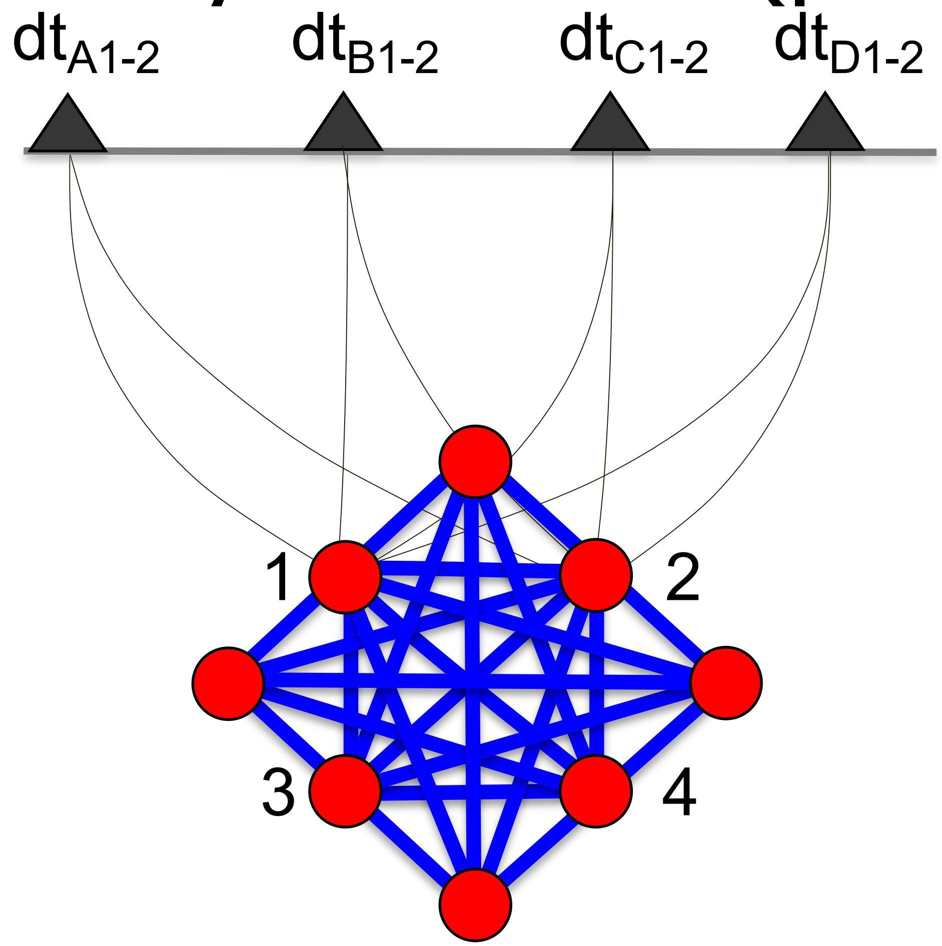


m = 24 n = 16

Building a network of delay time links (ph2dt)

- If we add 4 more events, the number of DD equations leaps to 112.
- For a cluster of
  - 100 events and 10 stations:
    m = 49,500
  - 1000 events and 20 stations:
    m = 10 million

➢ *Need to optimize network!*

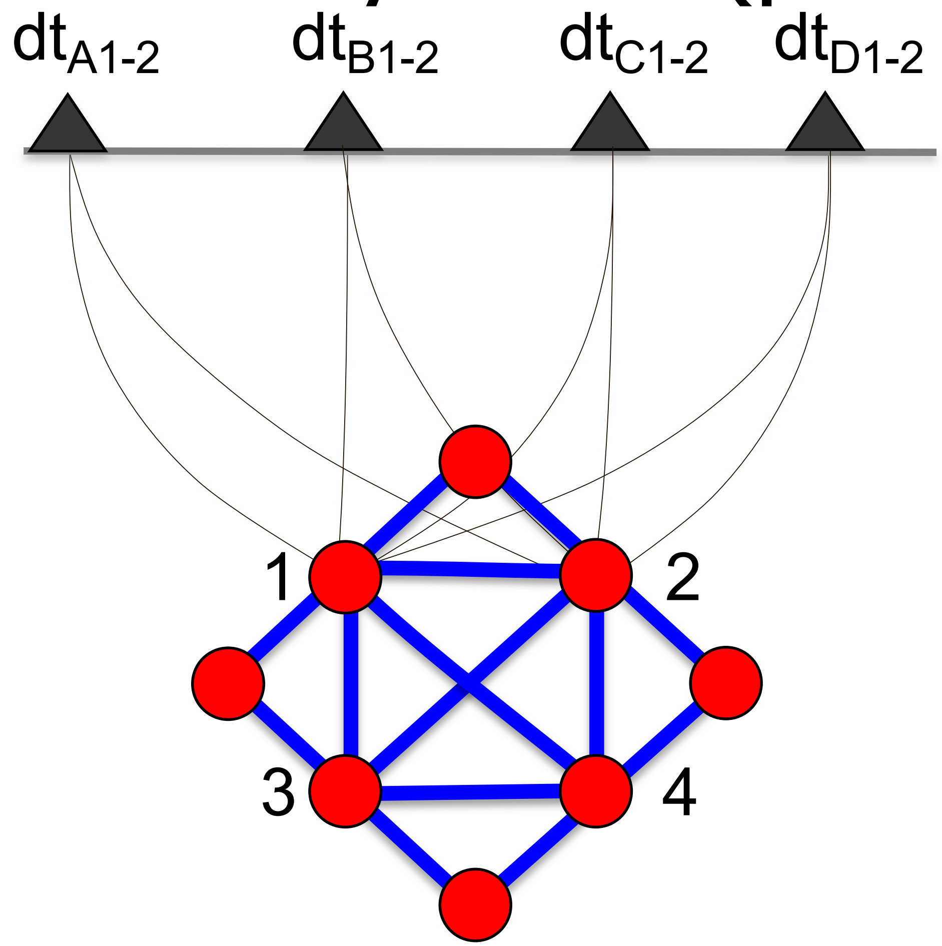


m = 112 n = 32

Optimizing a network of delay times (ph2dt)

- Remove links for hypocenters separated by greater distances (MAXSEP in ph2dt).
- Limit number of nearest neighbors (MAXNGH).
- Remove weak links; i.e. links only constrained by few stations (MINOBS).

➢ *When choosing values consider uncertainty in initial locations!!!*



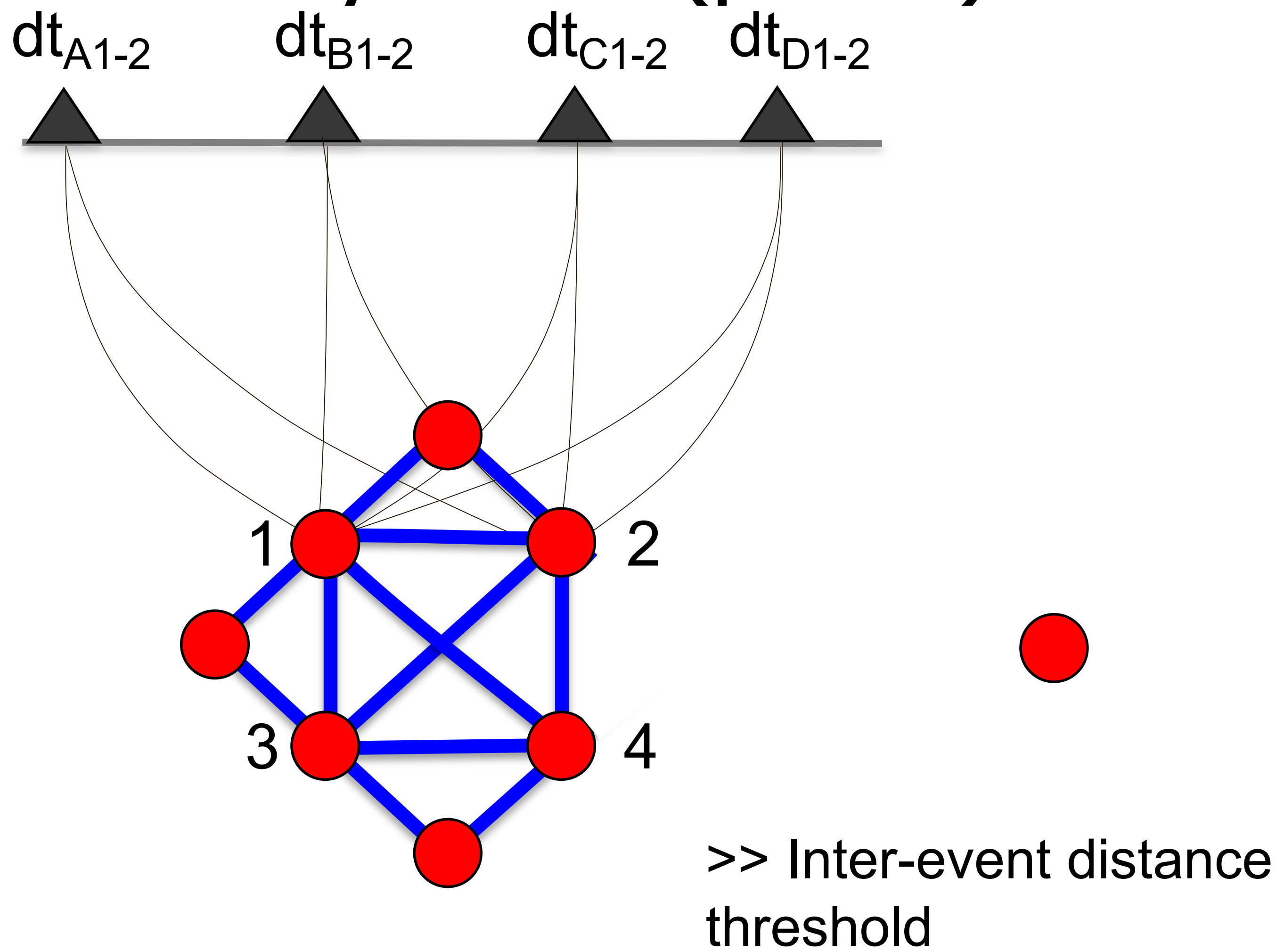
m = 56   n = 32

Optimizing a network of delay times (ph2dt)

- Remove links for hypocenters separated by greater distances (MAXSEP in ph2dt).
- Limit number of nearest neighbors (MAXNGH).
- Remove weak links; i.e. links only constrained by few stations (MINOBS).

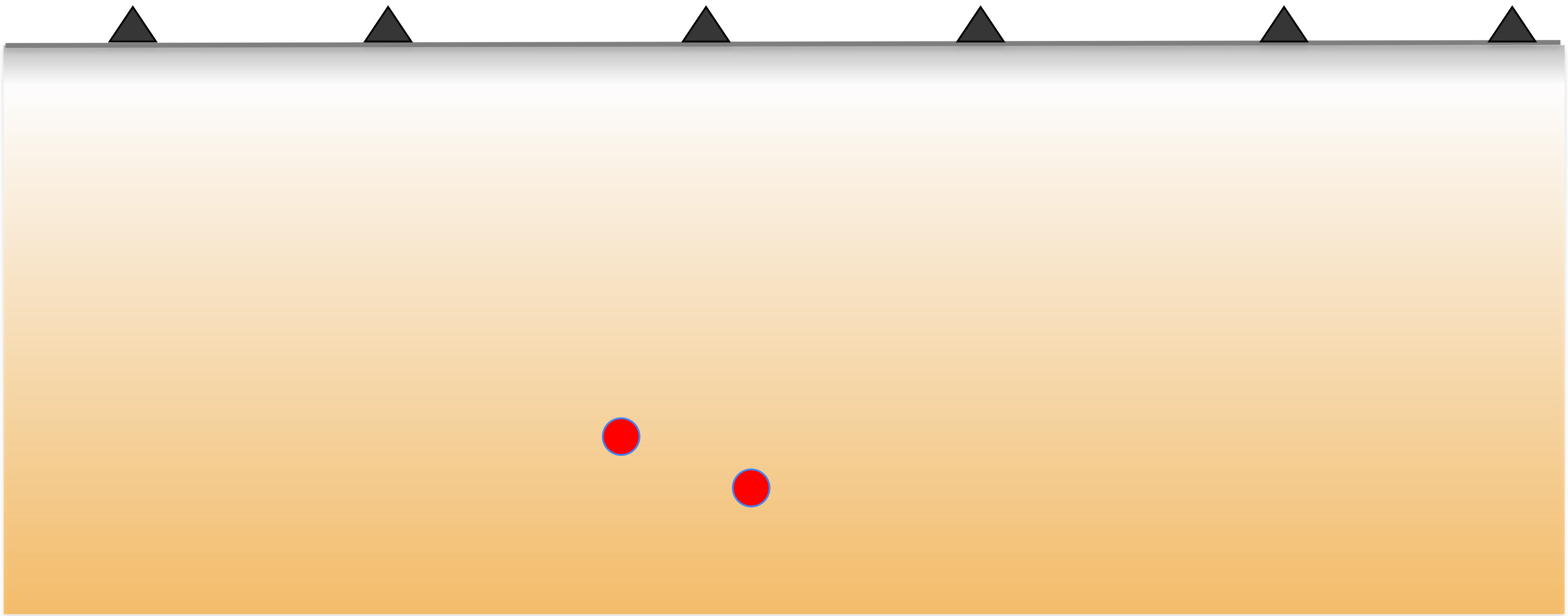
➢ *When choosing values consider uncertainty in initial locations!!!*

➢ *Isolated events removed.*

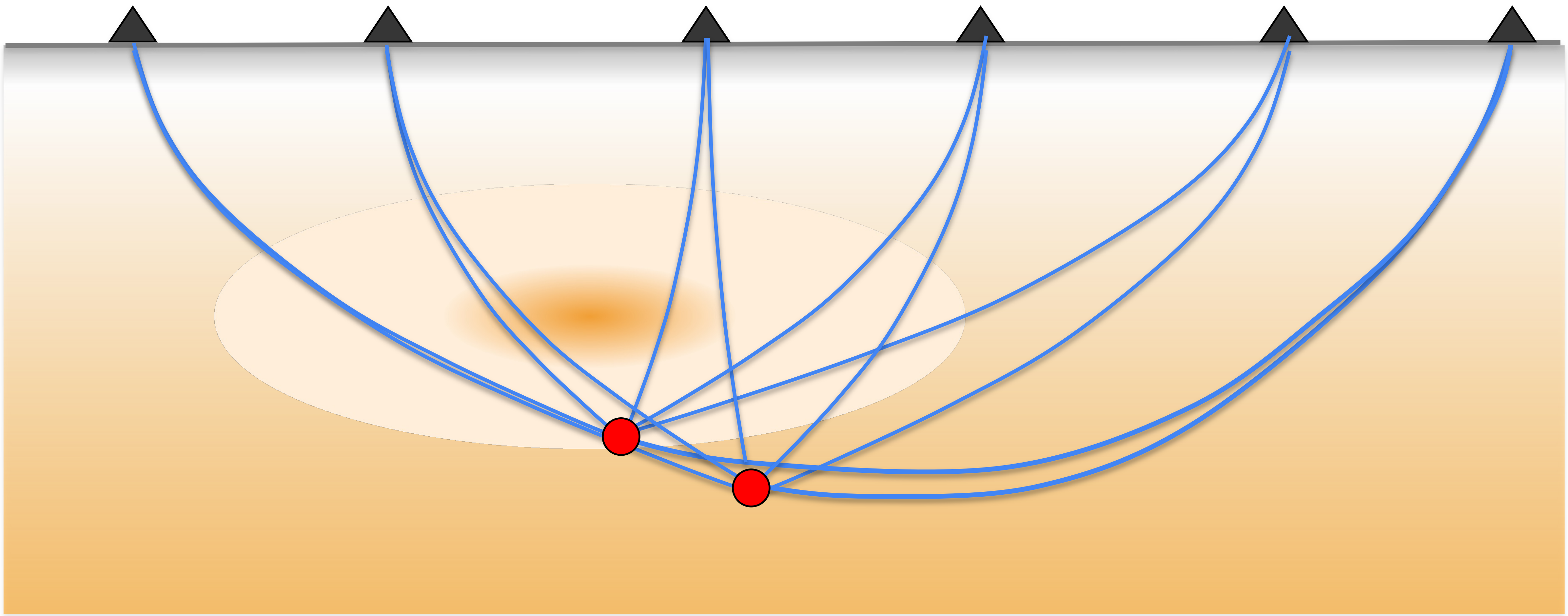


m = 56   n = 32

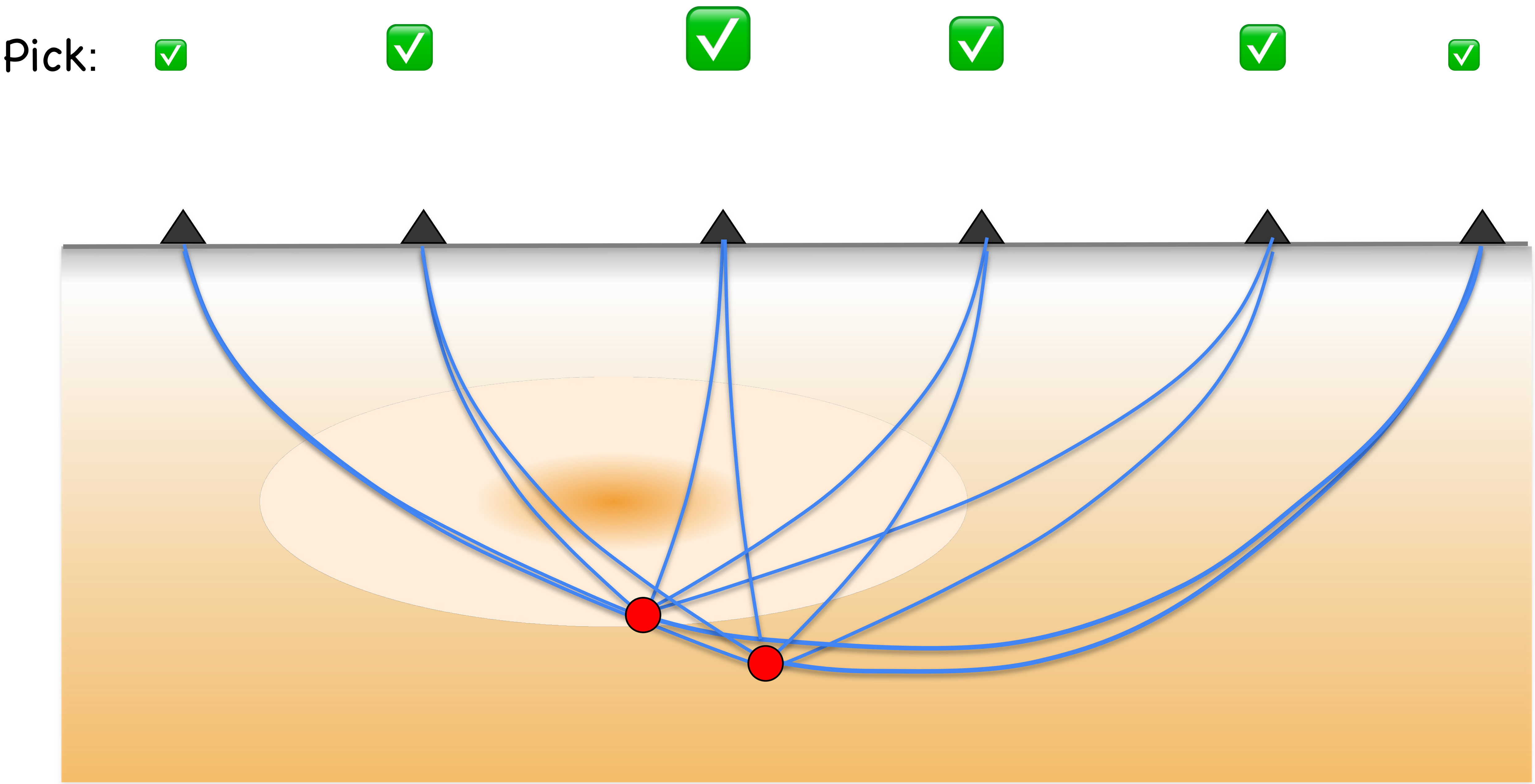
Cross-correlation data CANNOT replace phase picks
(in most cases)



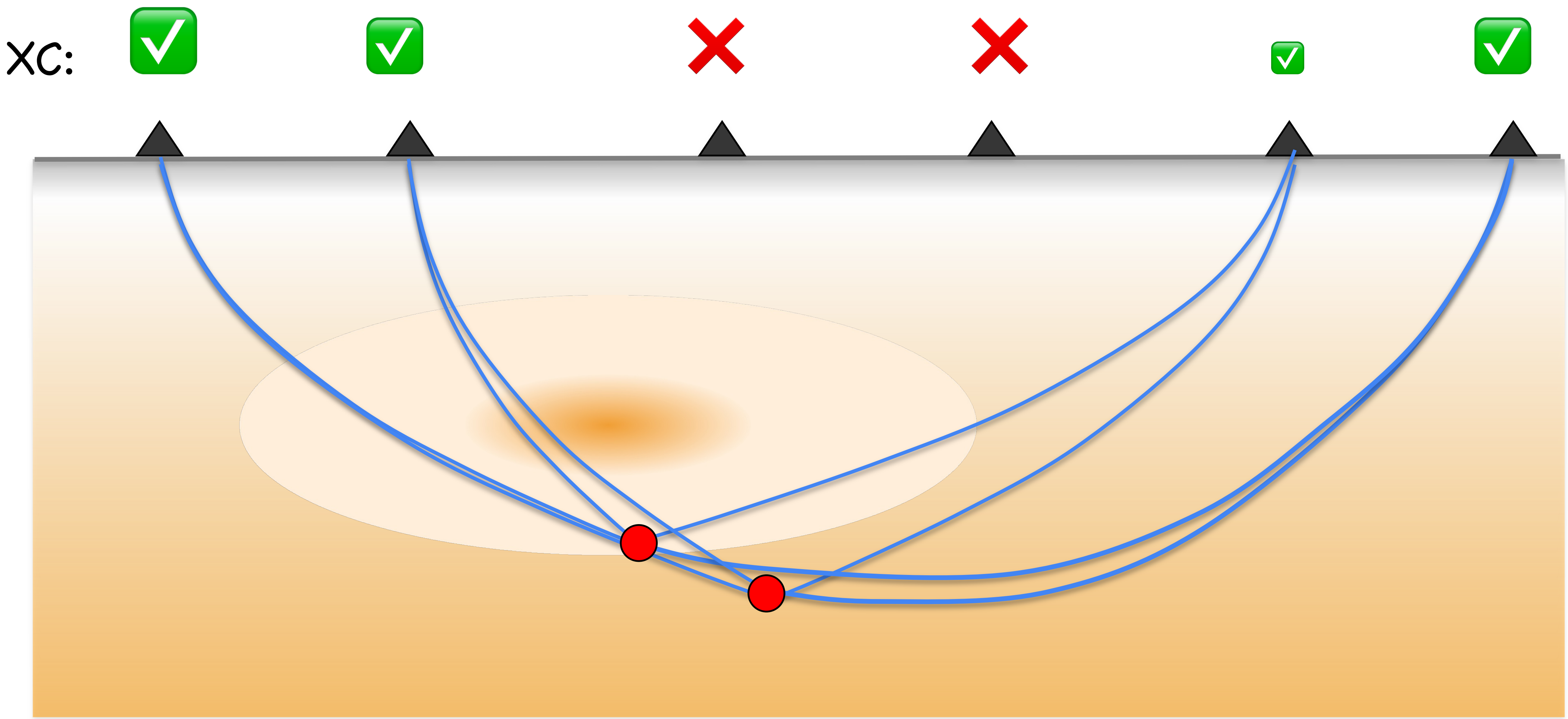
Cross-correlation data CANNOT replace phase picks
(in most cases)



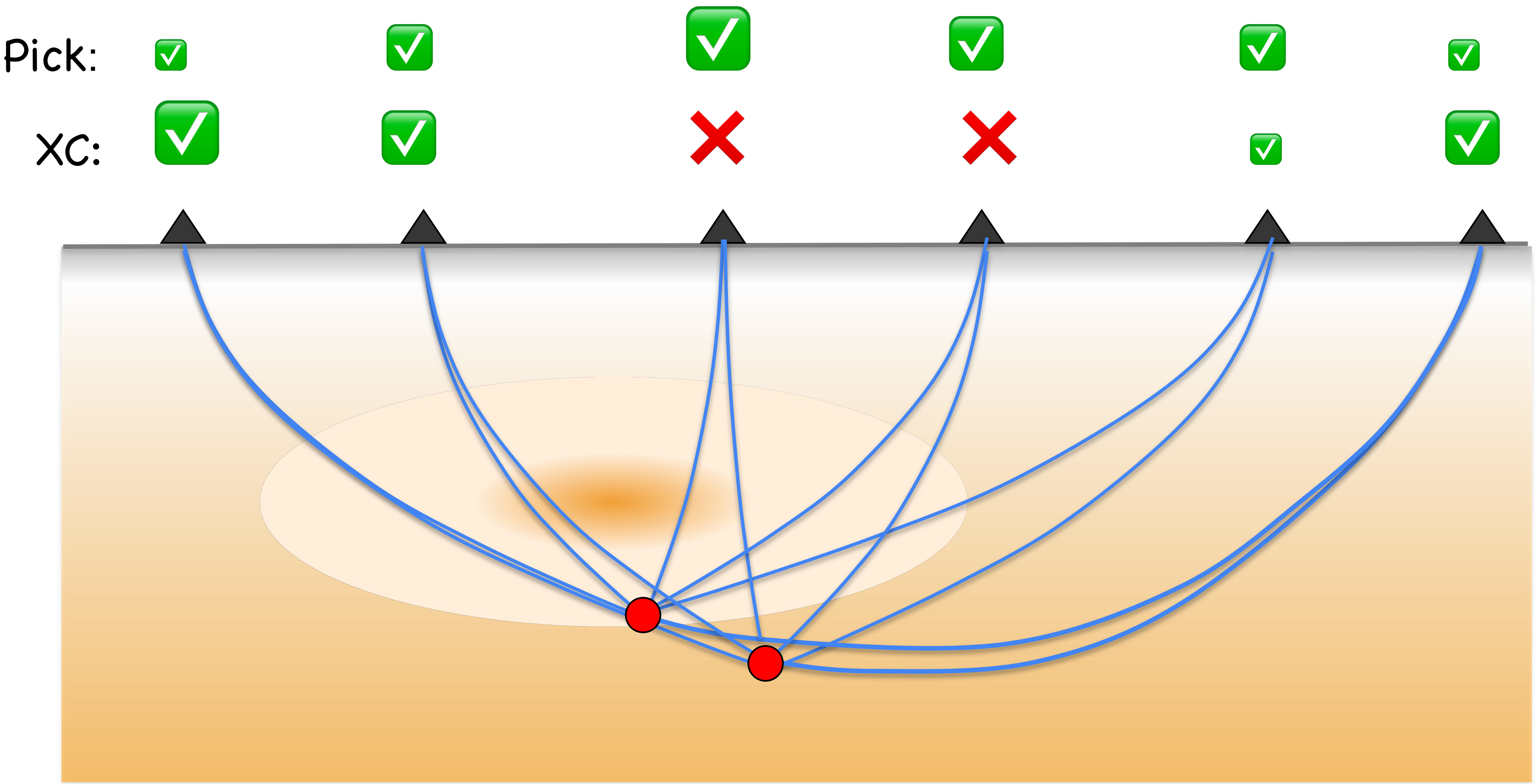
Phase picks vs. correlation times



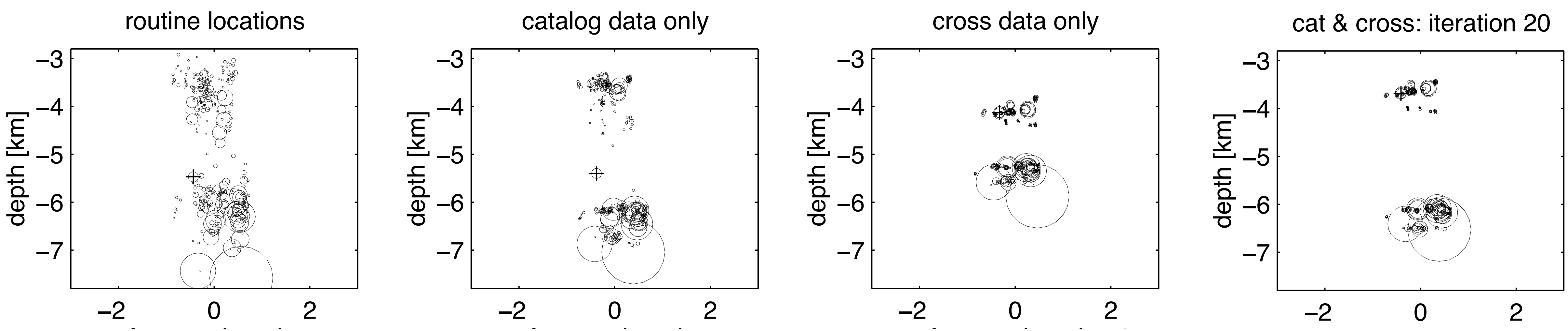
Phase picks vs. correlation times



Phase picks vs. correlation times

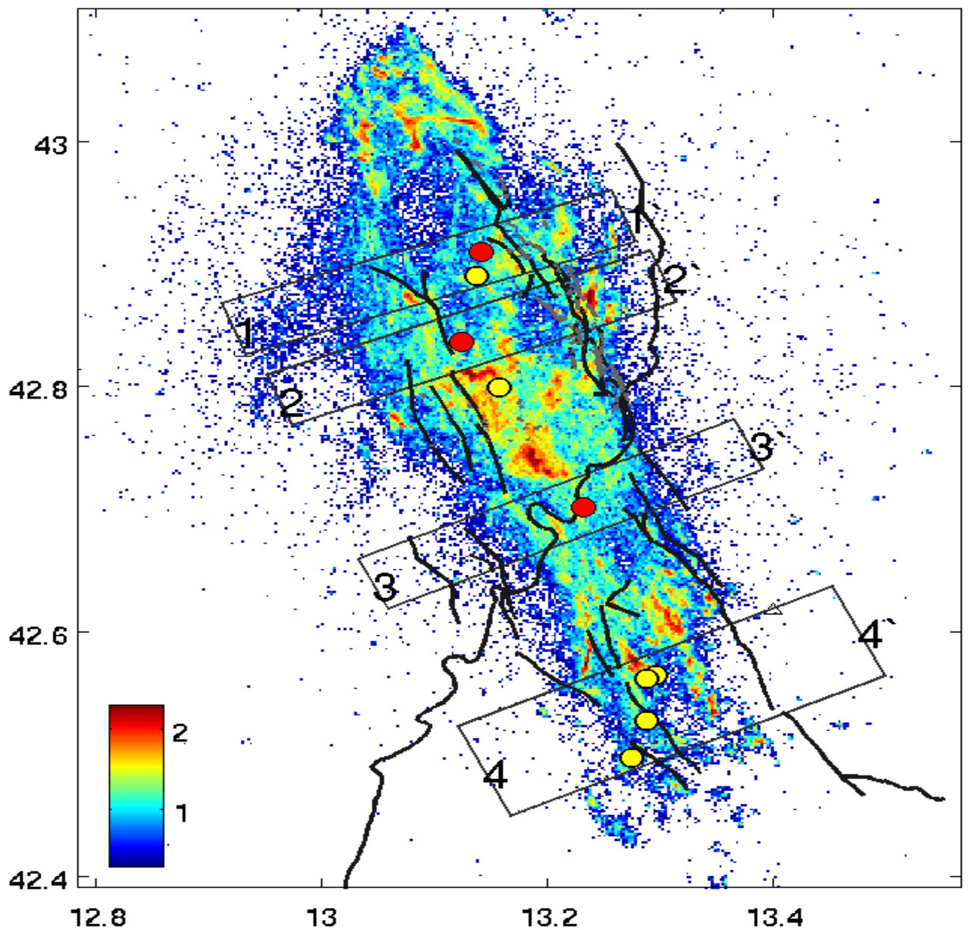


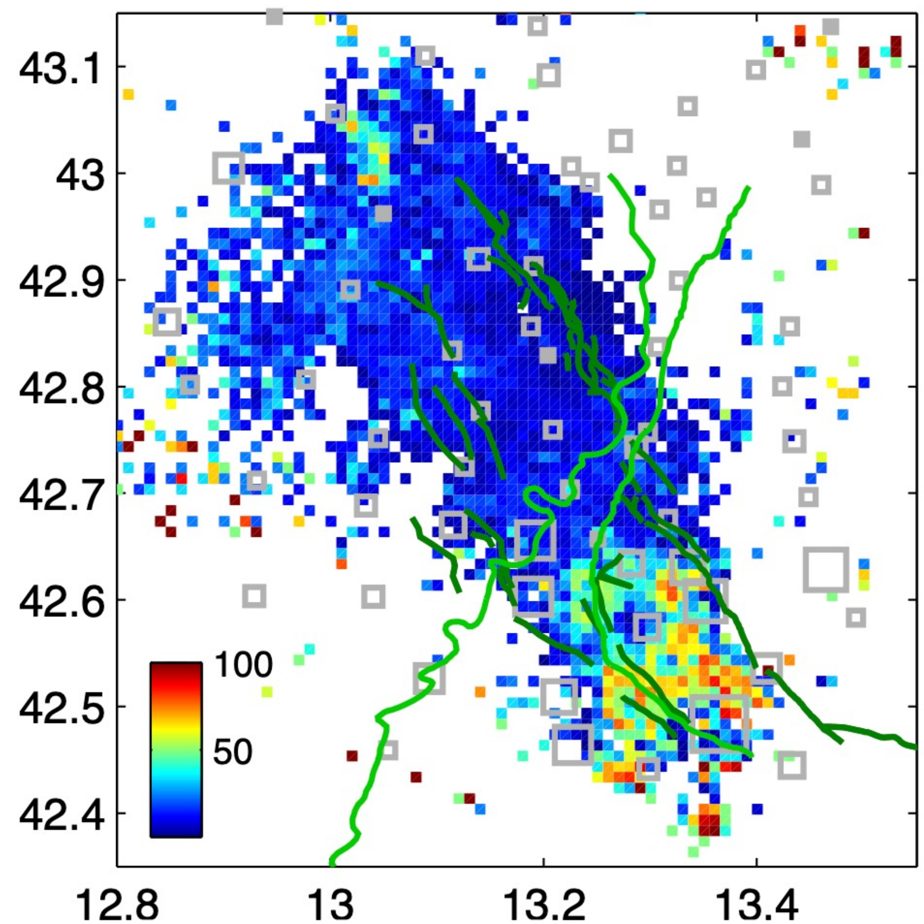
Phase picks vs. correlation times



Waldhauser 2001

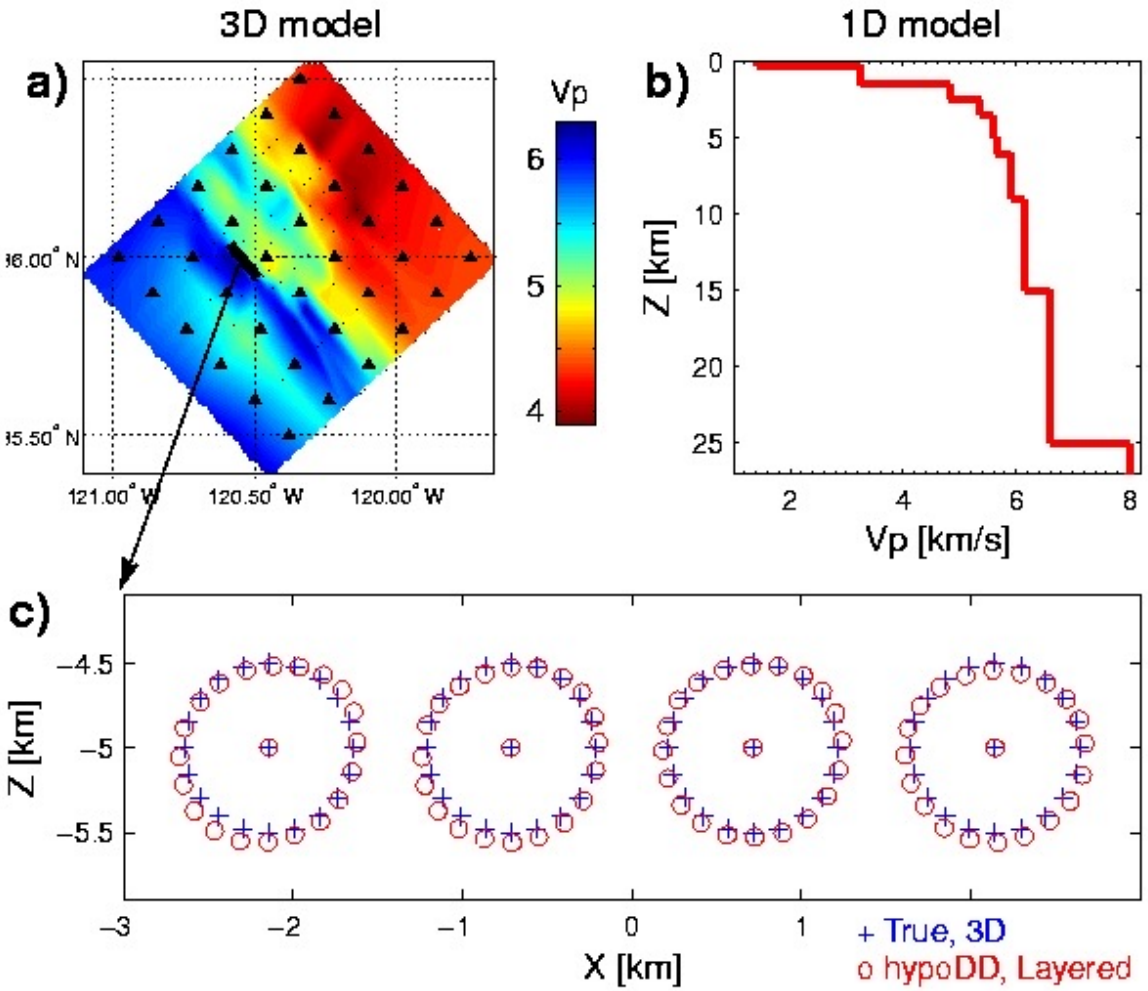
Fault/source complexity controls waveforms similarity (Amatrice sequence)





Waldhauser et al (2021)

It DOES accept 3D models (HypoDD 2.1)



Build layered models using:

- *Velest* to compute minimum 1D model
- Active source data

3D models from:

- 3D/DD tomography
- Active source data

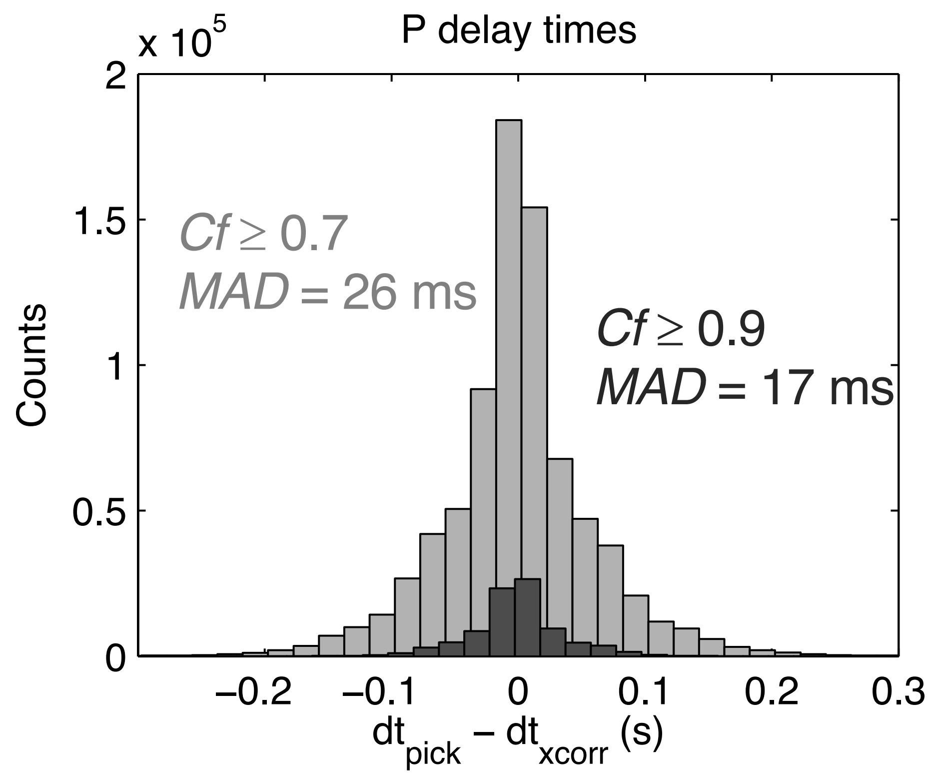
Do’s and don’ts

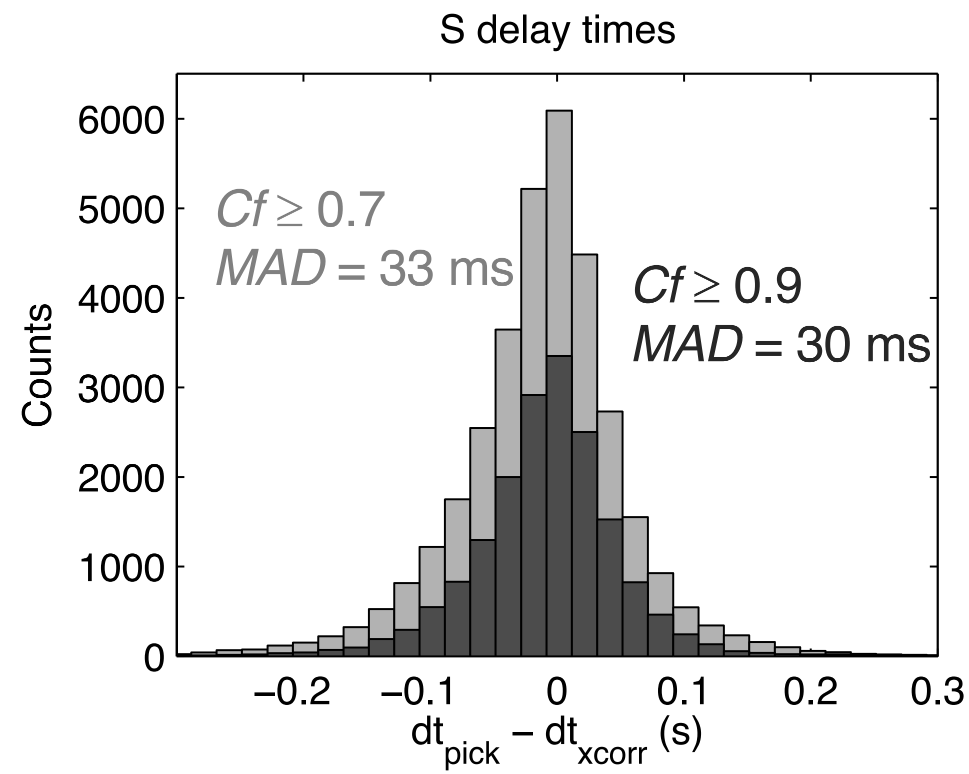
Do’s and don’ts

- Always look at the data.

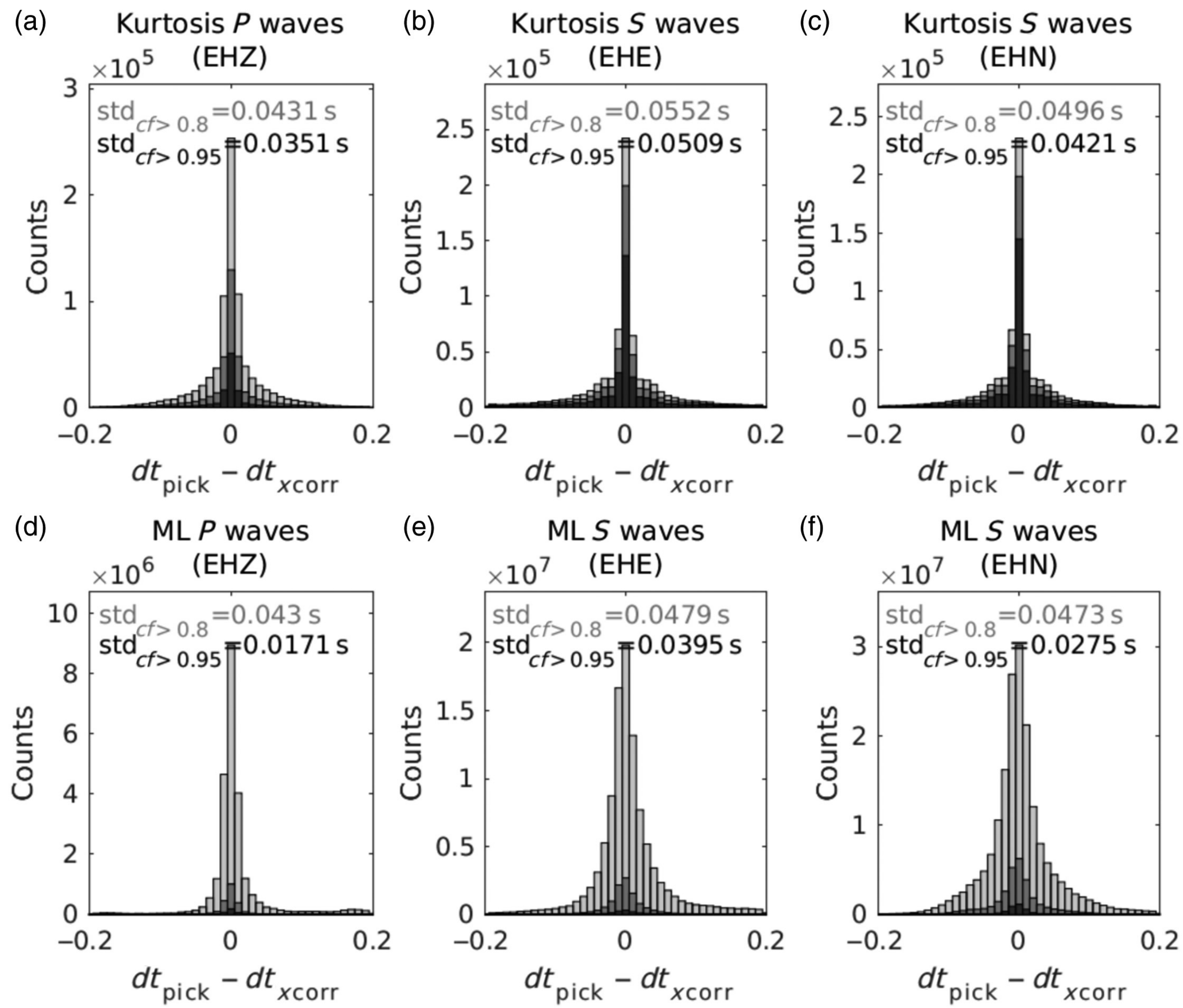
Test/Evaluate Phase Picks

- Look at differences between pick and corresponding cross-correlation delay times:



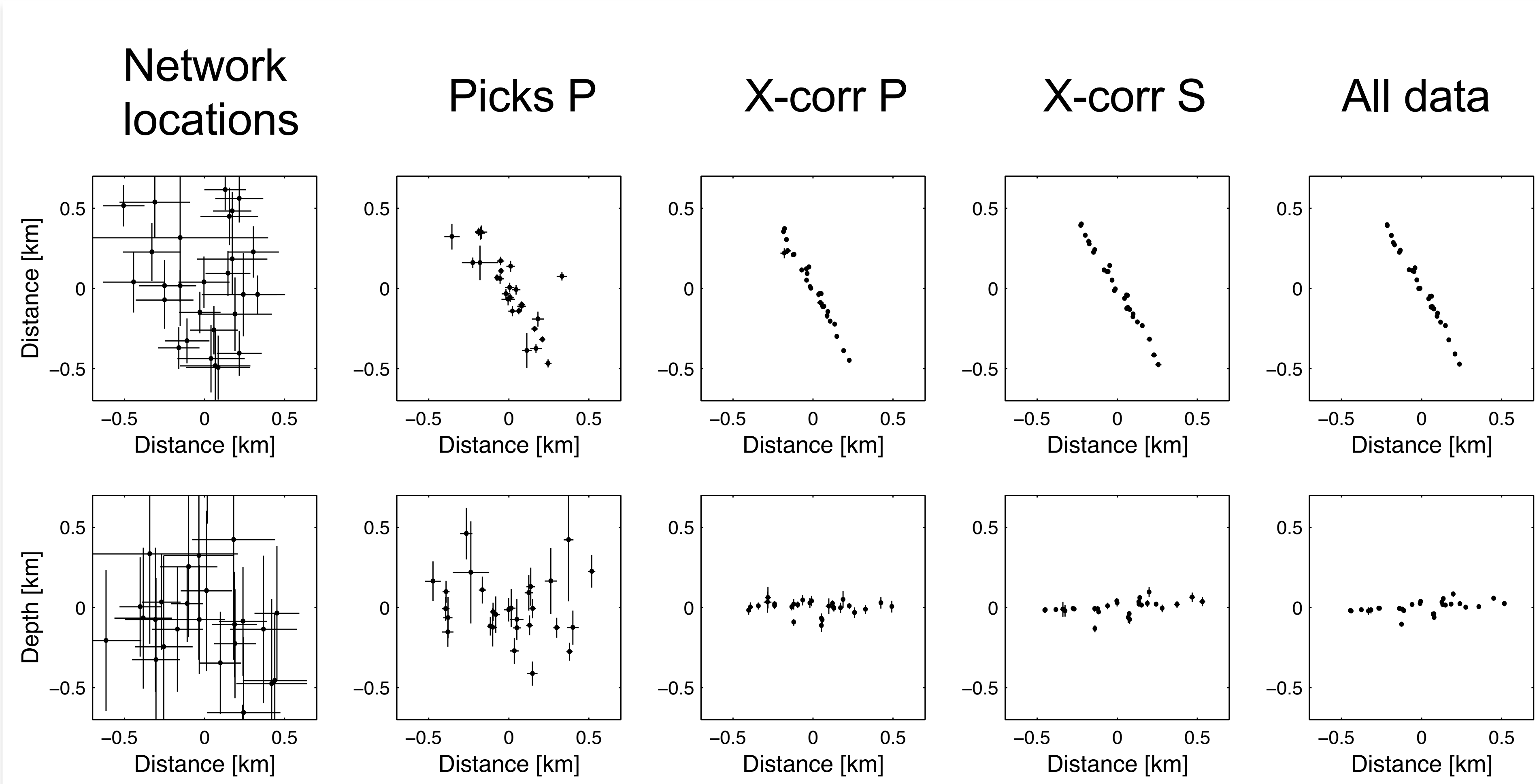


$dt_{pick} – dt_{xcorr}$ at Axial Seamount (OBS)



Always look at the data first

- Relocate earthquake with each data set individually:

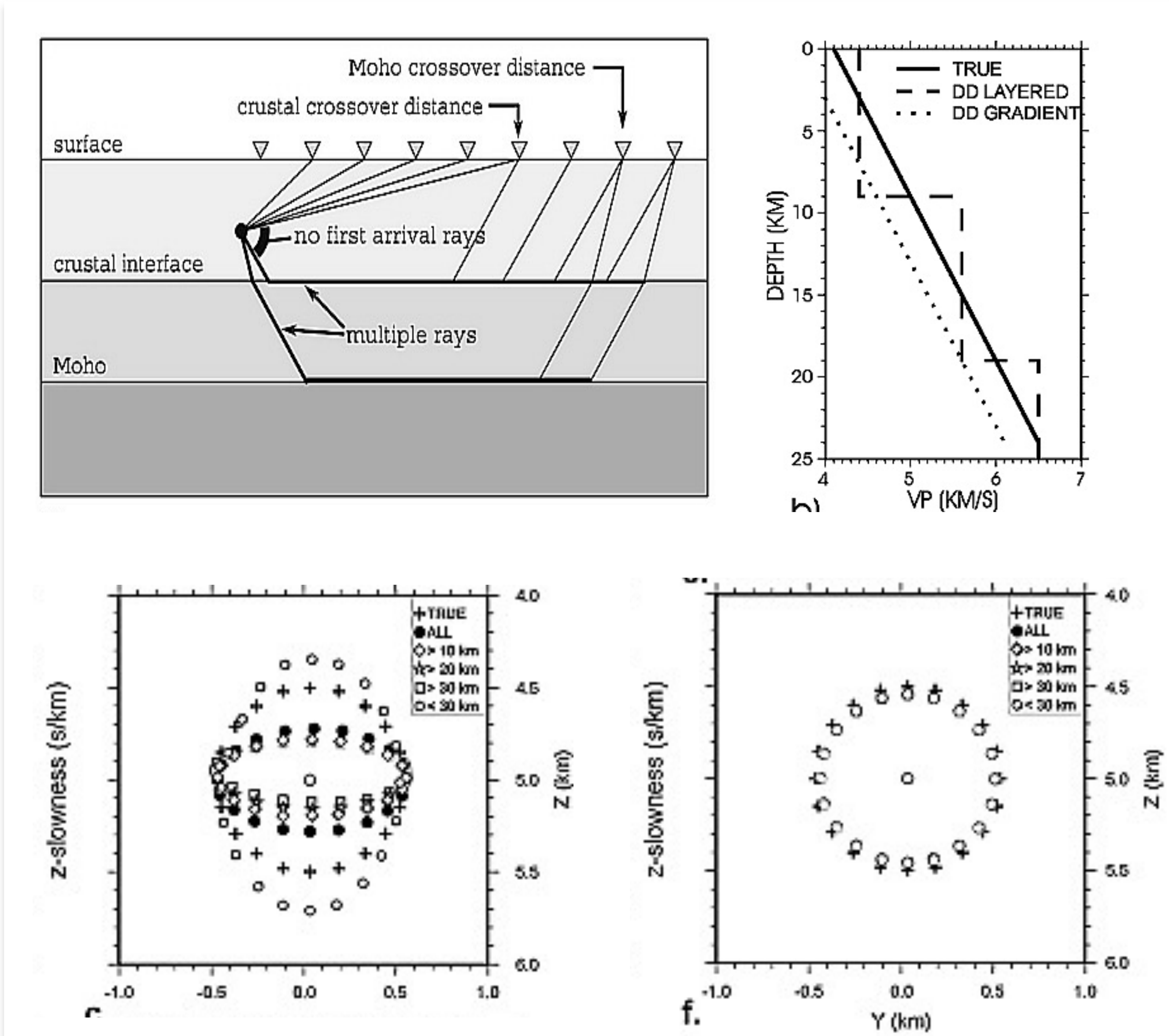


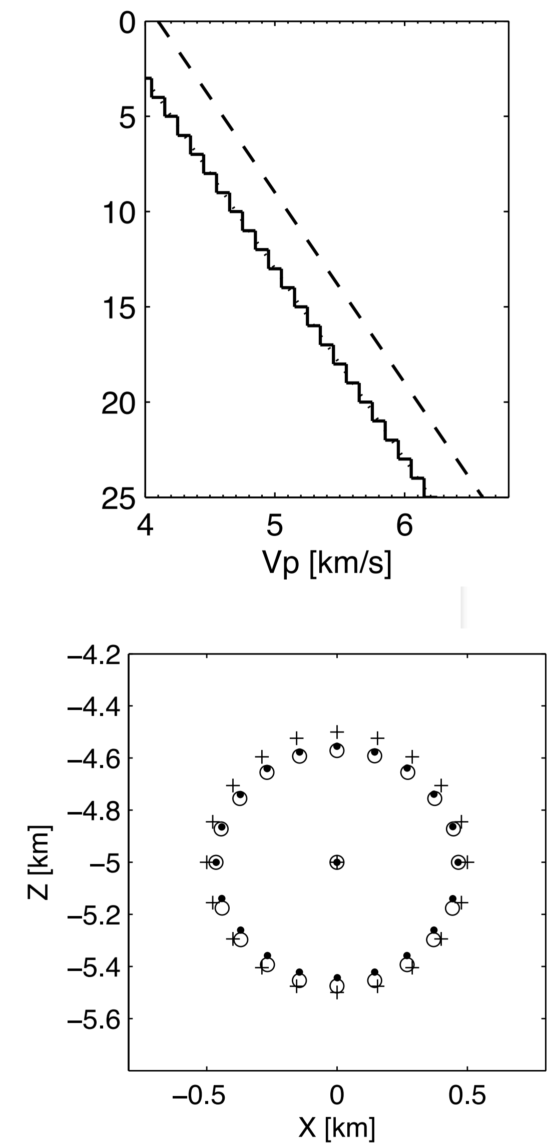
Do’s and don’ts

- Always look at the data.

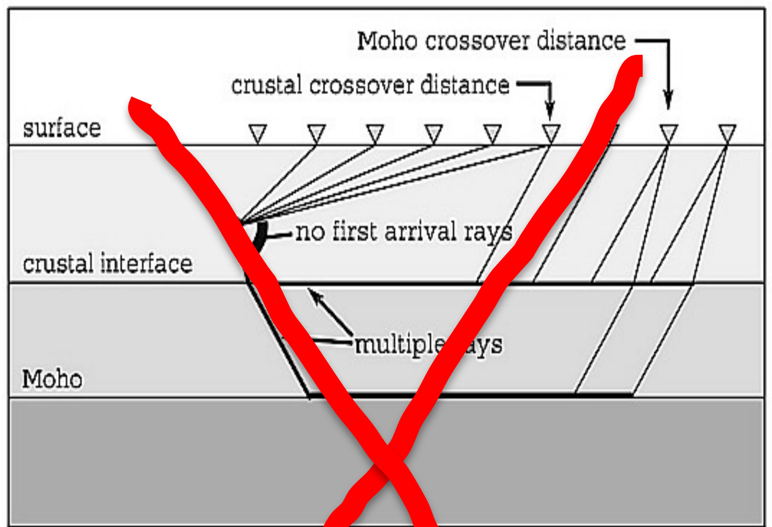
➢ Choose an appropriate model.

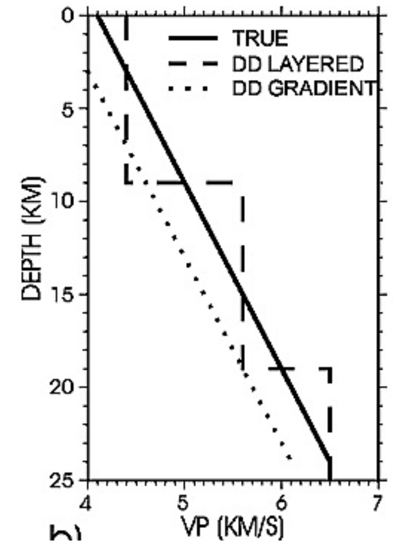
The velocity model IS important

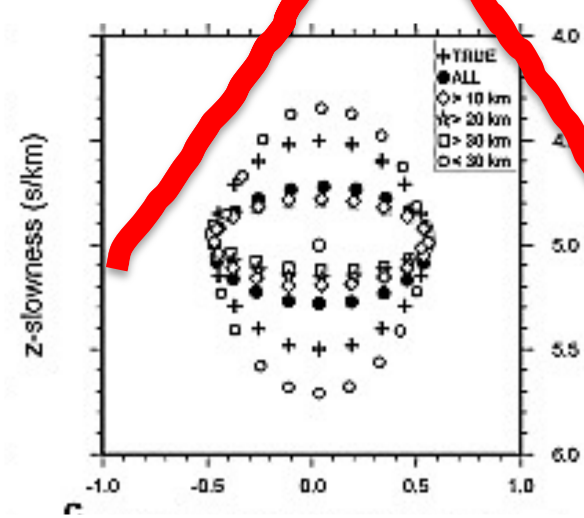


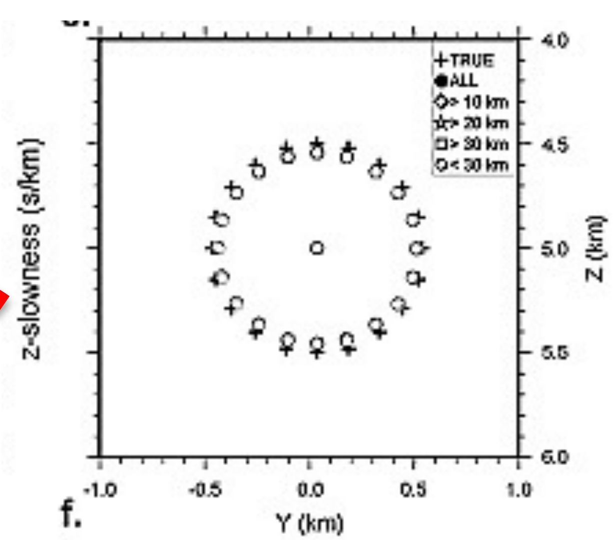


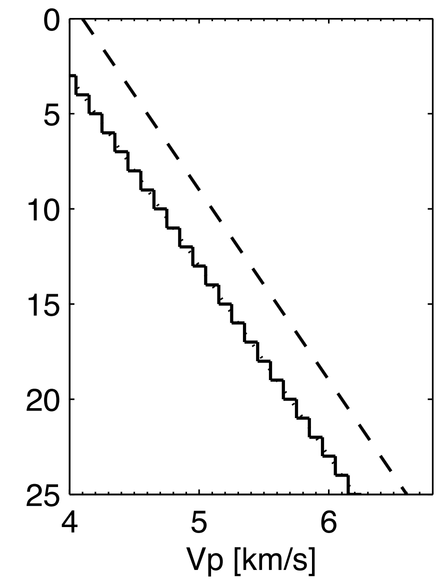
The velocity model IS important

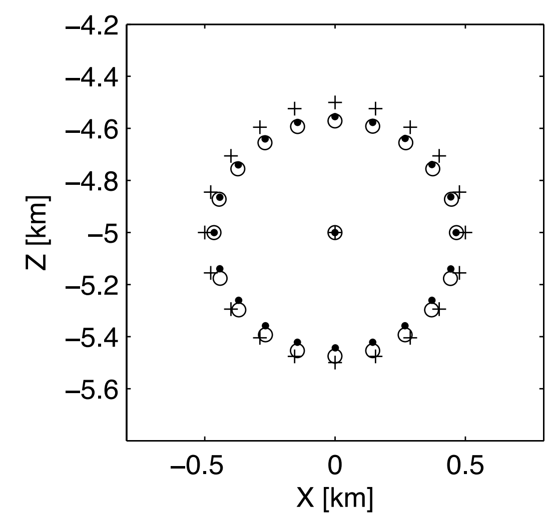












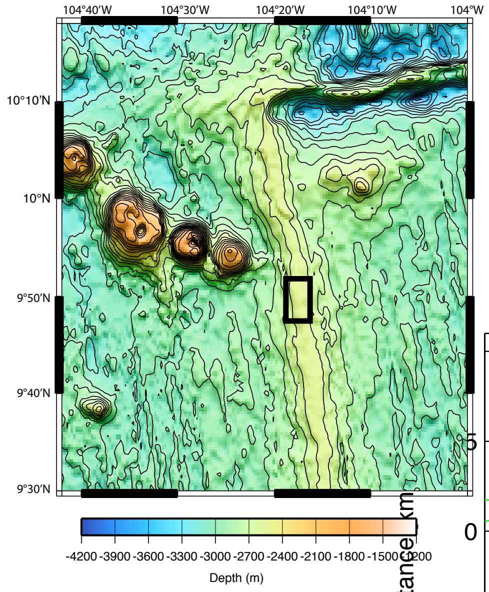
Do's and don'ts

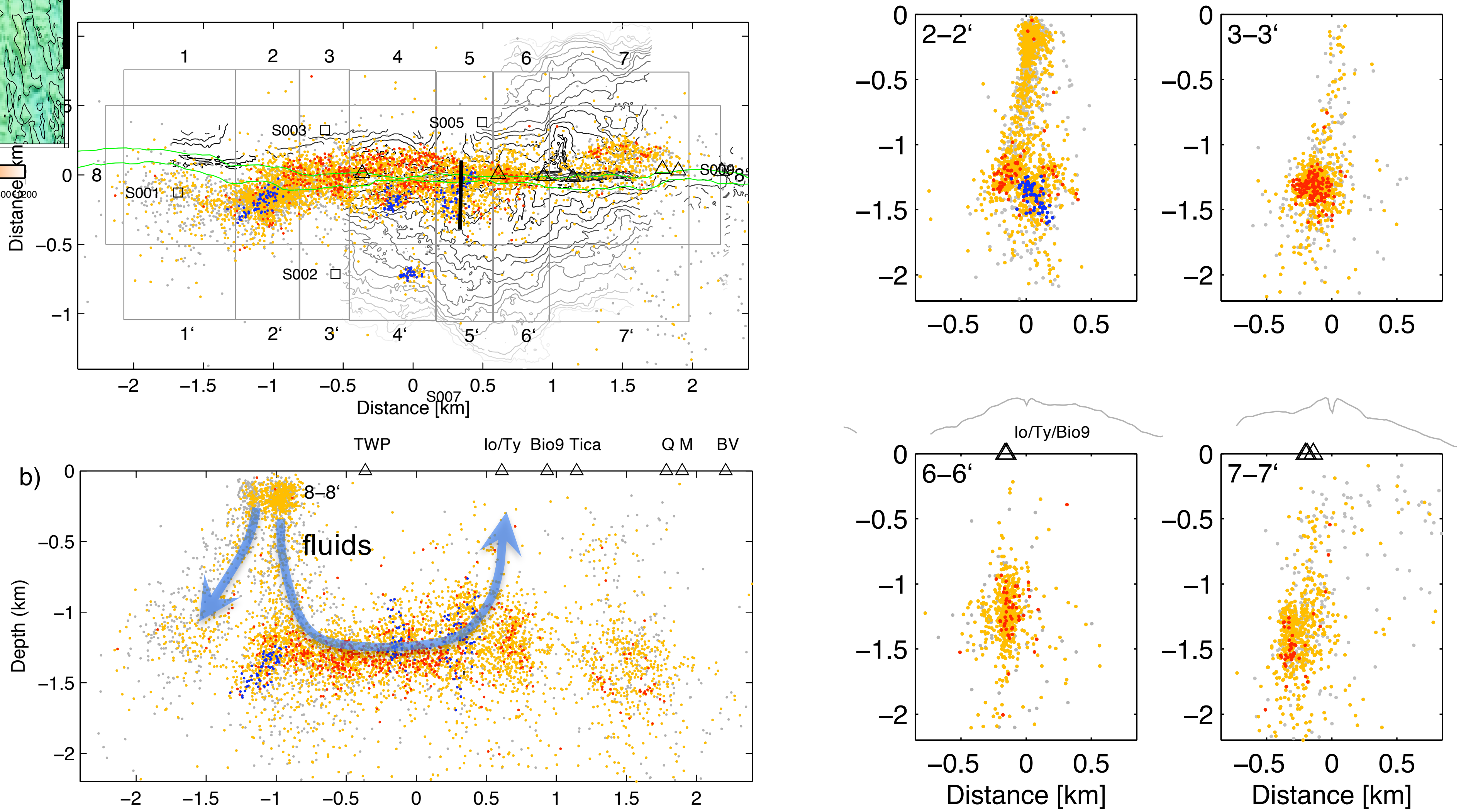
- Always look at the data.

- Choose an appropriate model.

➢ Evaluate results.

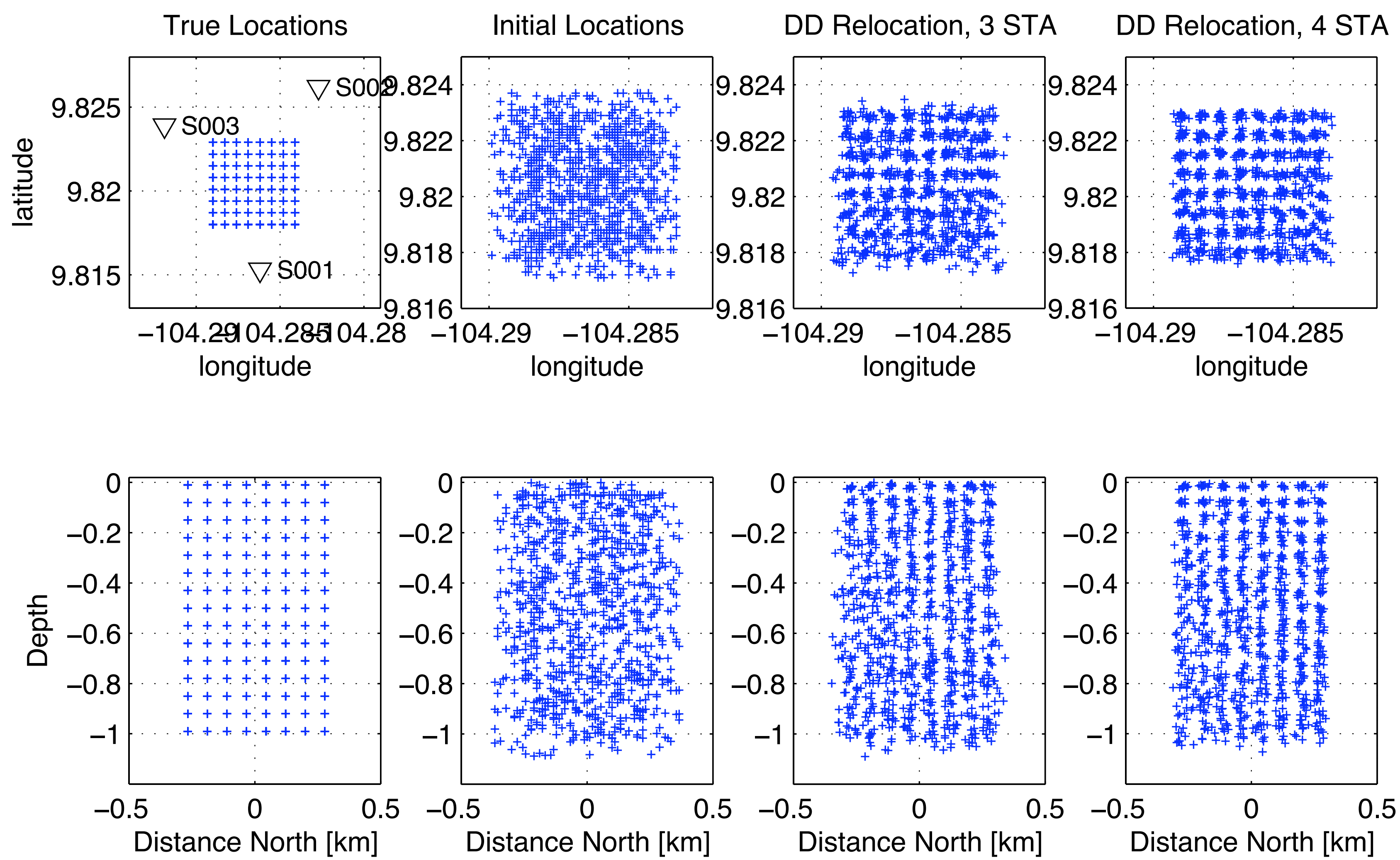
Evaluate results: Example - EPR





Waldhauser and Tolstoy, 2011

Can we resolve the down-flow pipe?



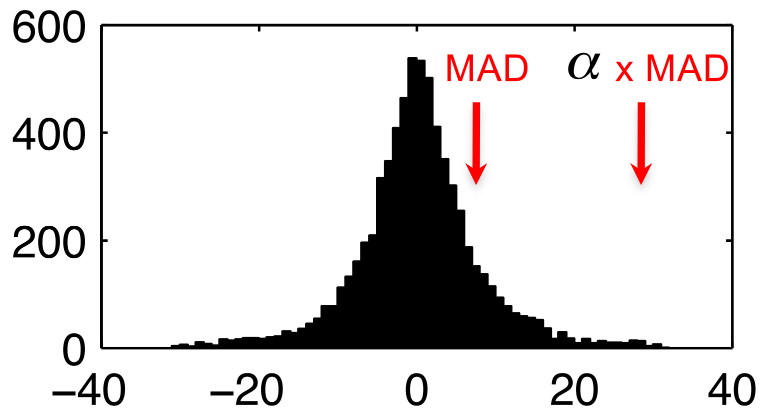
➢ Synthetic resolution tests for the down-flow pipe

A priori data weighting

- **Data type (W = 0-1):**
  - P-, S-waves
  - Phase onset catalog picks
  - Cross-correlation data
  - These weights are iteration dependent!

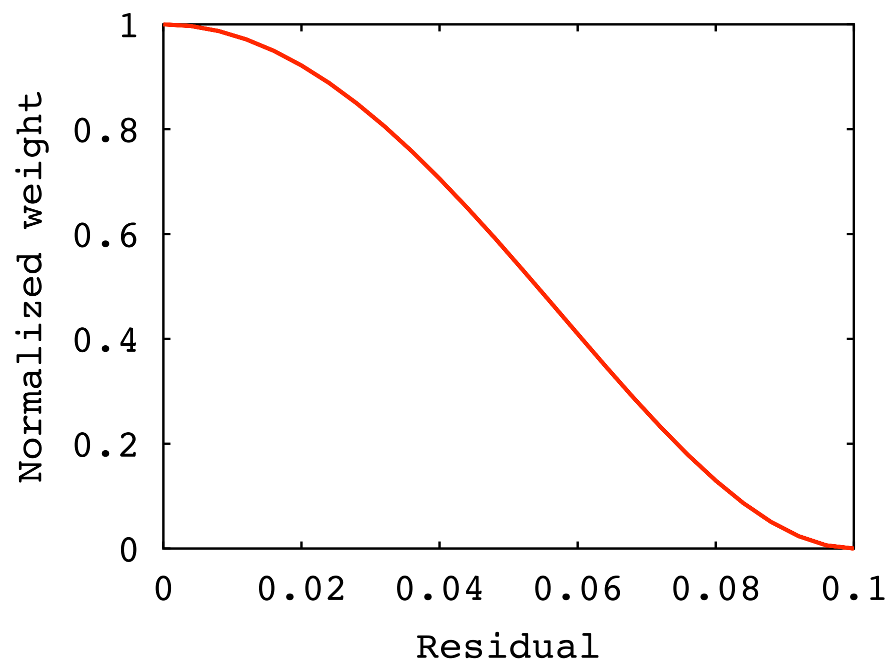
- **Data quality (W = 0-1):**
  - Pick uncertainty:
    W = *f* (pick error [s]; pick quality [0,1,2…]; onset characteristic [I,E])
  - Cross-correlation coefficient: W = $(Cf)^2$

Residual weighting

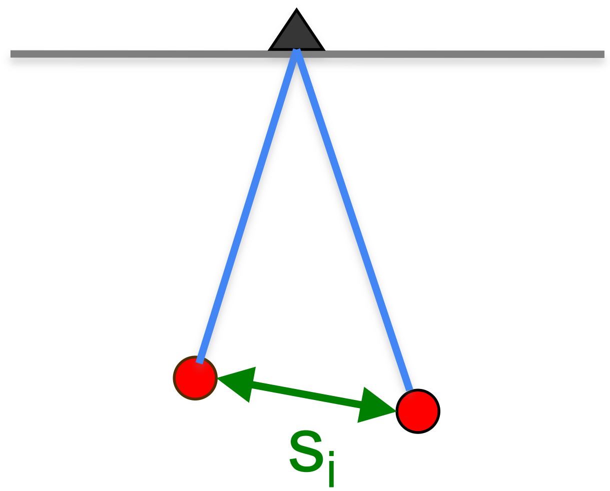


$$W_i = \max^2\left(0, 1 - \left(\frac{dr_i}{\alpha\frac{\mathbf{dr}_{\mathrm{MAD}}}{\sigma_{\mathrm{MAD}}}}\right)^2\right)$$

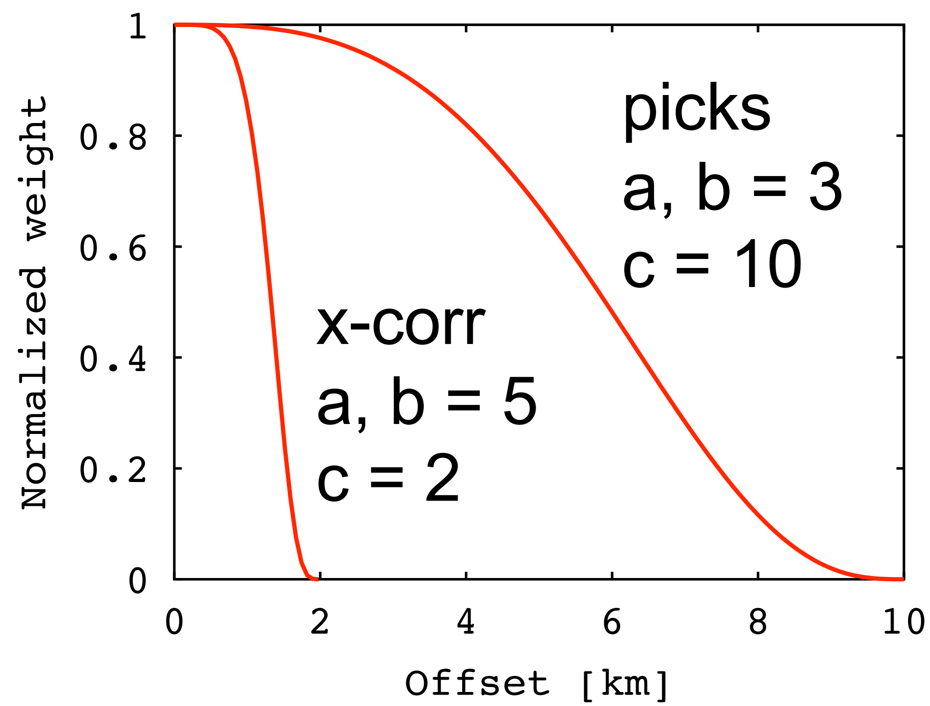
- α = WR in ***hypoDD***
- Weight applied to residuum computed after each iteration



Inter-event distance weighting

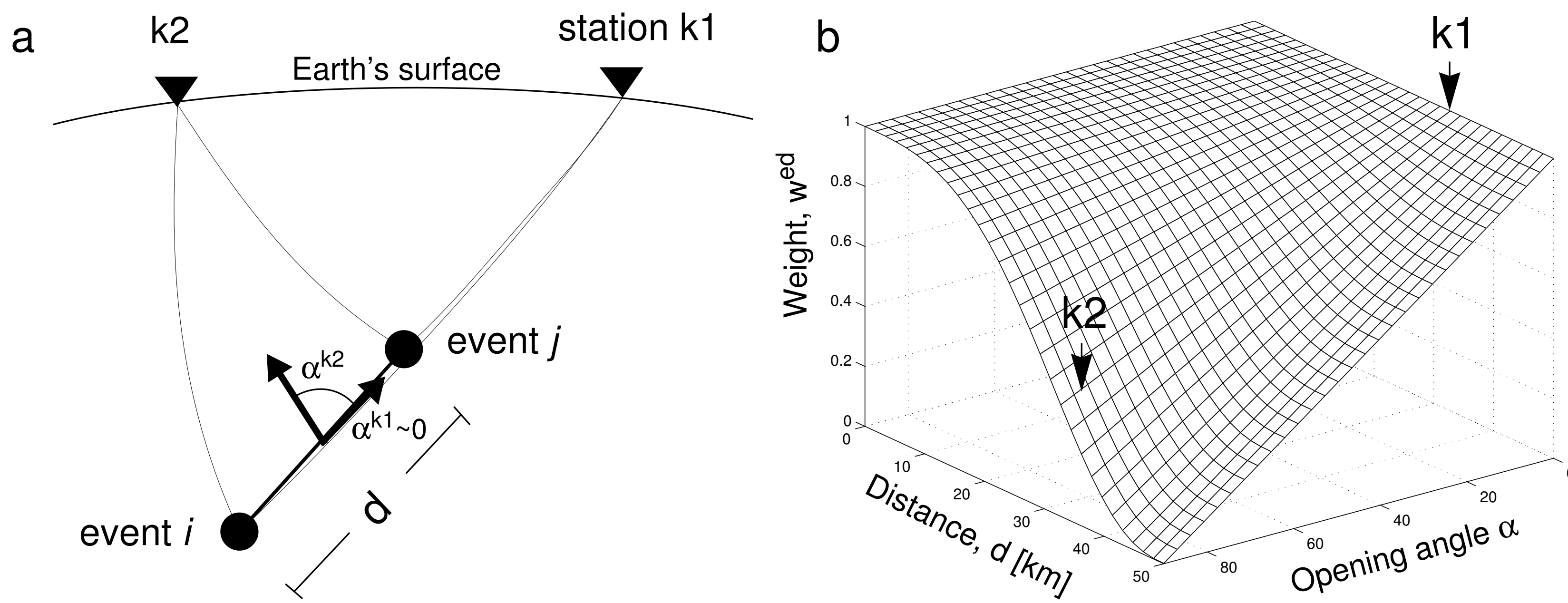


$$W_i = \max{}^{b}\left(0, 1 - \left(\frac{s_i}{c}\right)^{a}\right)$$



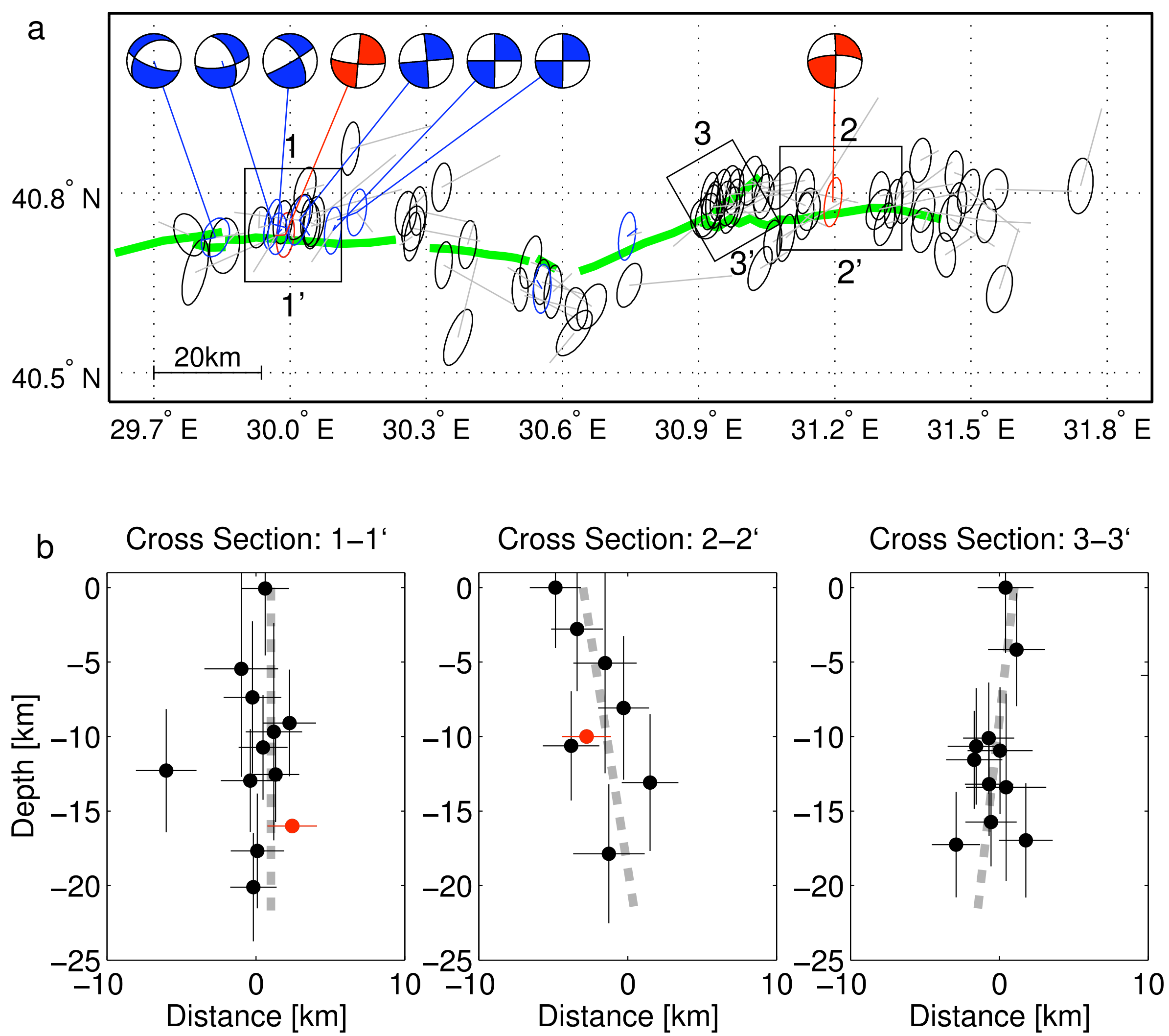
- c = parameter WD in ***hypoDD*** (distance cutoff)
- Weight applied to residuum computed after each iteration

Direction dependent inter-event distance weighting

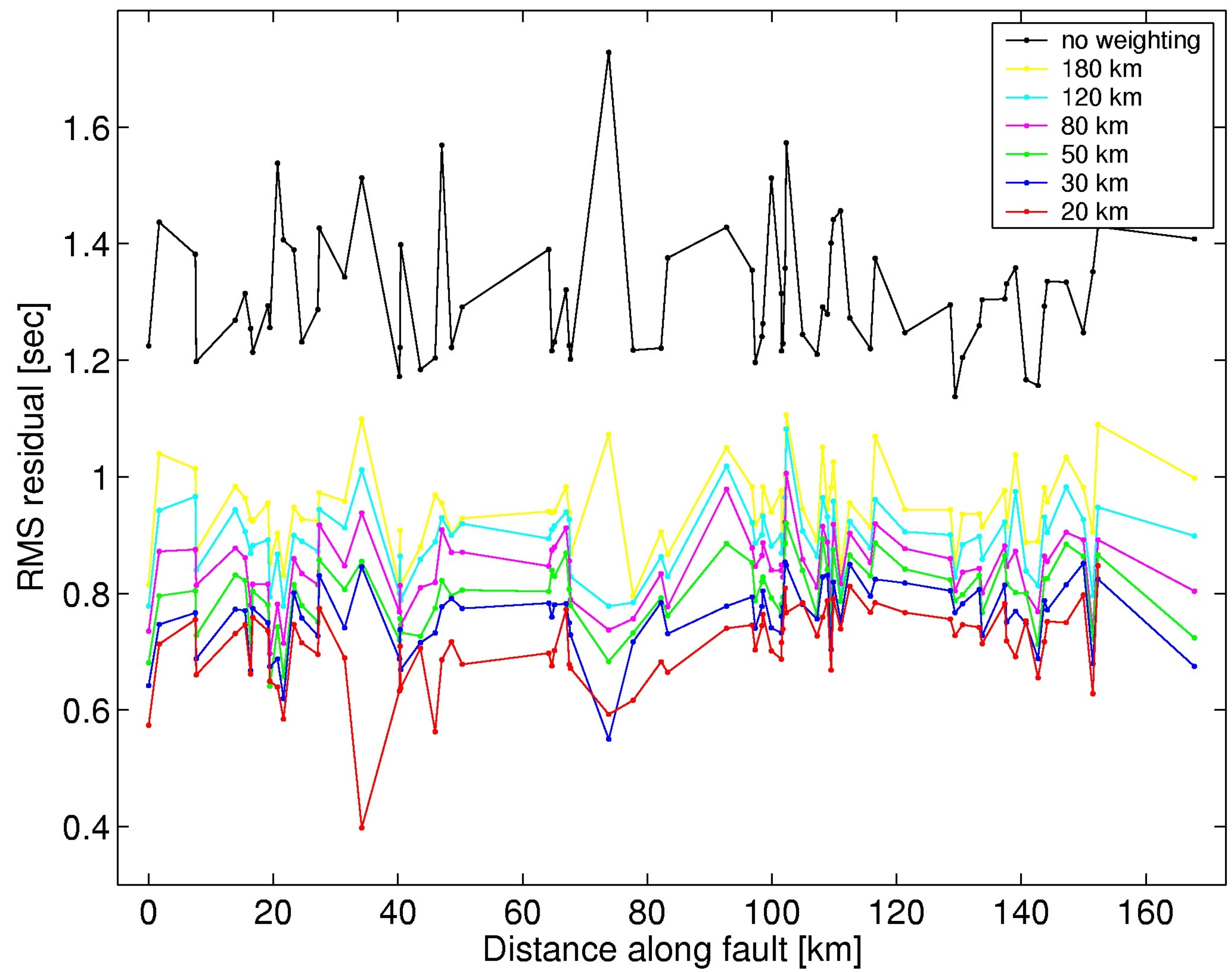


Waldhauser and Schaff, 2007

Example: Izmit/Düzce on North Anatolian Fault



Example: effect of distance weighting on RMS residuals



Waldhauser and Schaff, 2007

Example: effect of angle weighting on RMS residual