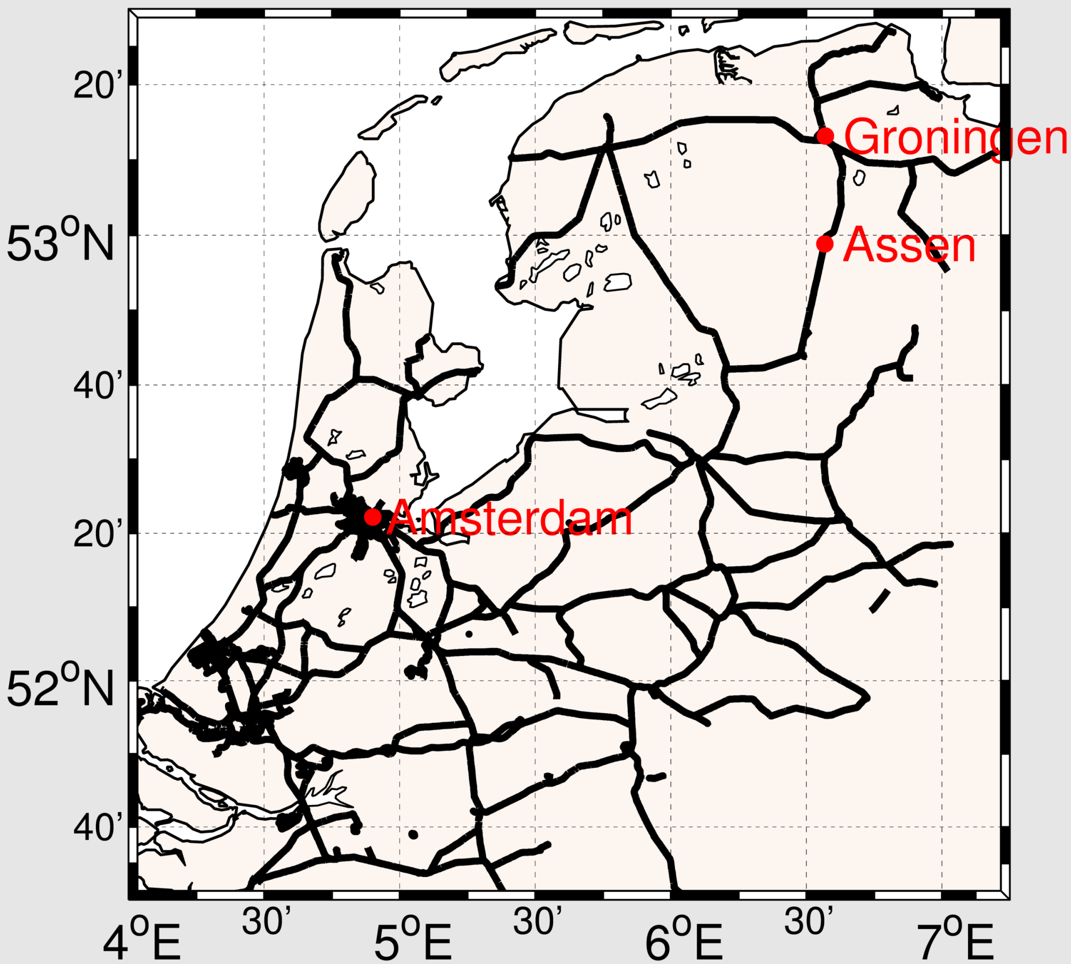
Are DC trains useful for EM exploration?

M. Becken, A. Avdeeva *WWU Münster*
R. Streich *Shell Global Solutions*

Motivation



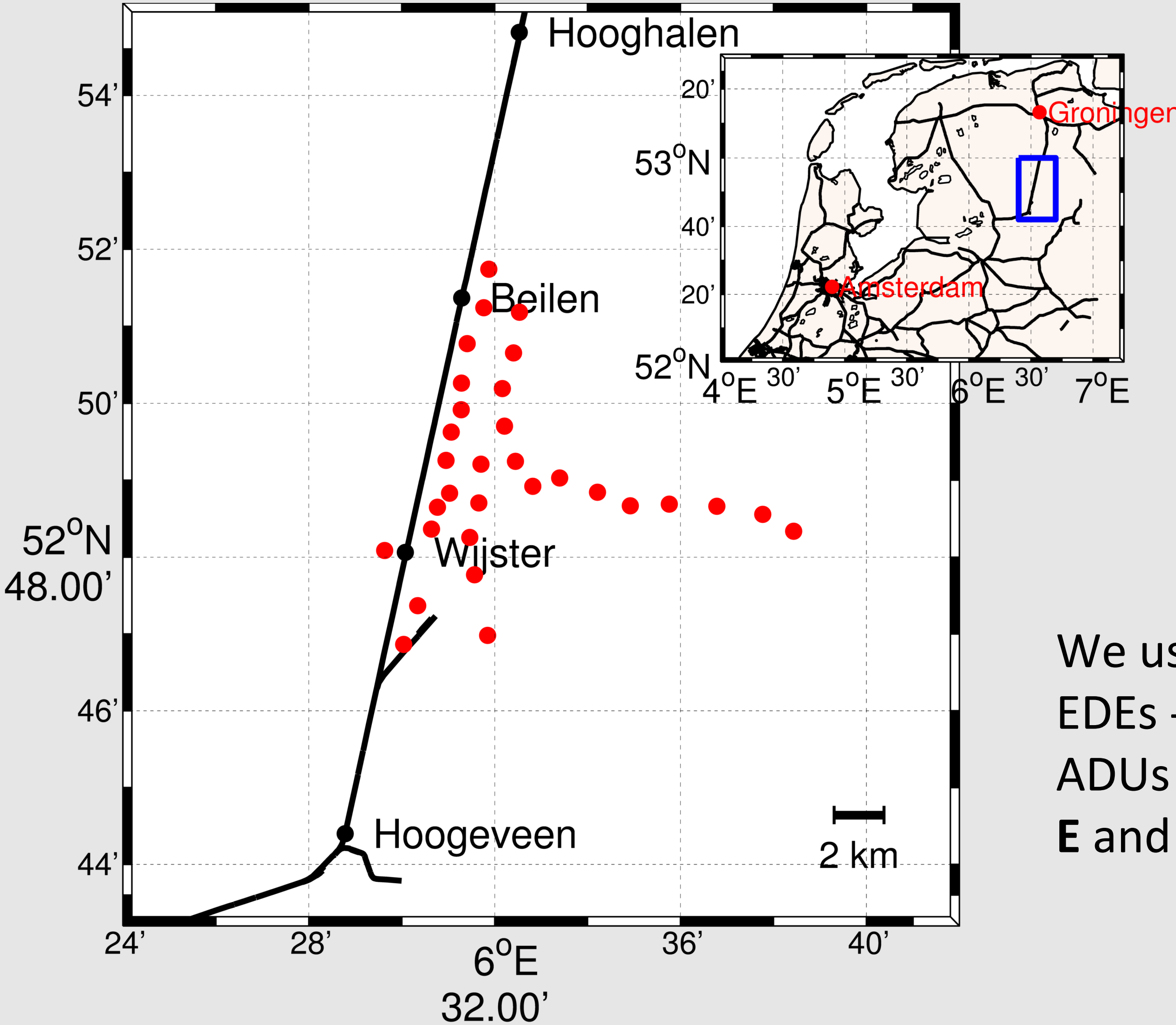
Large amount of (DC) railways

➔ Difficult to conduct standard EM experiments

Some previous studies:

1. Larsen et al., 1995; Egbert et al., 2000 etc.: Viewed signal from electrified railways as noise
2. Tanbo et al., 2003; Neska, 2009: Attempt to use railway signal to image the subsurface
3. Lowes, EPS, 2009: Good overview of DC railways in geomagnetic context

Test site in the Netherlands



31 sites measure **E** field

7 sites measure **B** field

Duration of recordings:

2 to 10 days

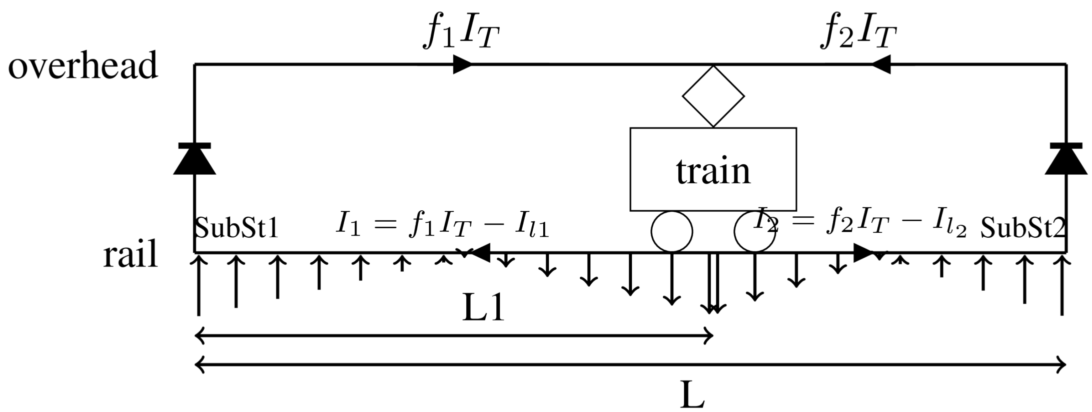
500 Hz sampling rate

We used 2 types of data loggers:

EDEs - WWU Muenster; **E**

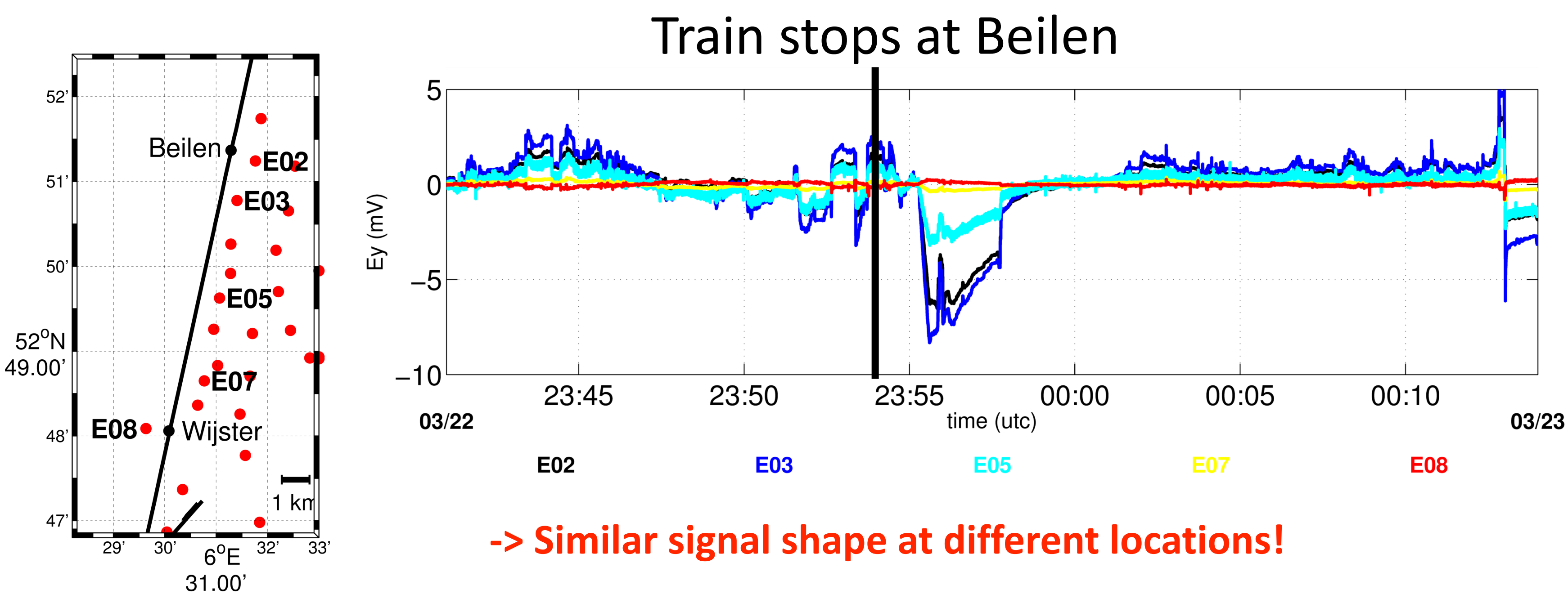
ADUs - WWU Muenster and Shell; **E** and **B**

Theory and Observations



magnetic loop source
current leakage through track
-> dependent on train pos.

after: Lowes et al. (2009)



-> Similar signal shape at different locations!

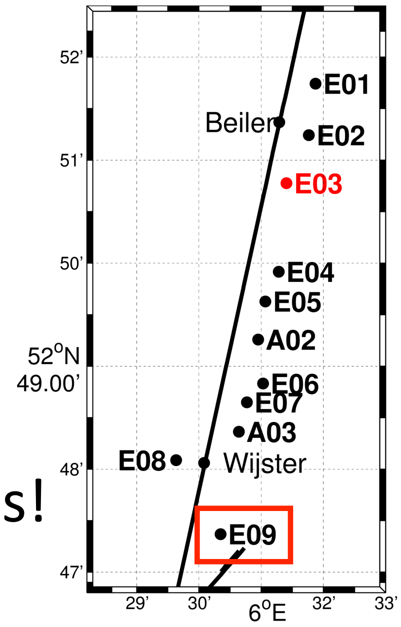
1^{st} approach: interstation transfer functions

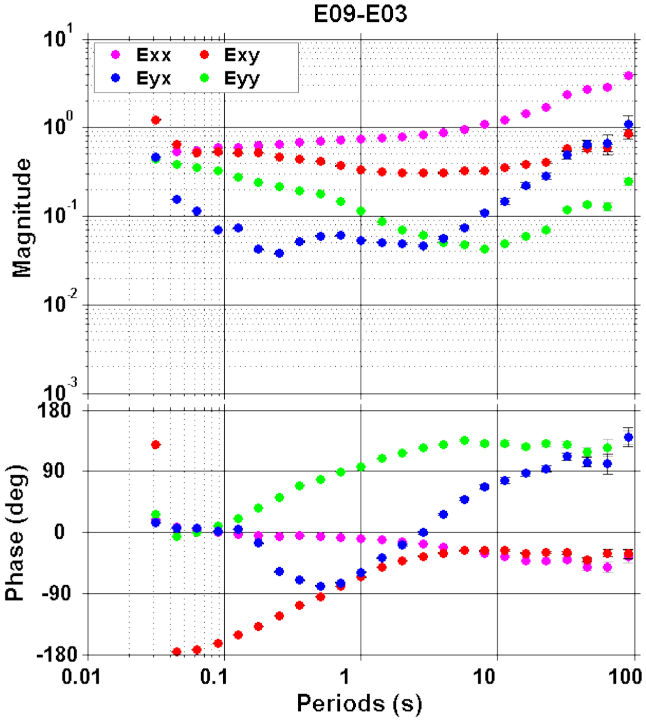
Definition: $\mathbf{E}(\omega, \mathbf{r}_1) = \mathbf{T}(\omega, \mathbf{r}_1, \mathbf{r}_2)\mathbf{E}(\omega, \mathbf{r}_2)$ (bi-variate relation)

Data Example:

- MT-like robust proc.
- 1 day of data

± stable transfer functions!





Inversion for tfs for sources and conductivity was challenging!

(extreme sensitivity to rail current, independency of two sources, bi-variate approach)

2^{nd} approach: Principal component analysis

Objectives in **EM array processing**:

- determine the number of independent linear combinations
- find a low-dimensional approximation to explain data variance

Collect array observations (at a single frequency):

$$
\begin{array}{cc}
 & \text{time windows} \\
\mathbf{X} = & \left[\begin{array}{ccccc}
E_x^{11} & E_x^{12} & E_x^{13} & \cdots & E_x^{1M} \\
E_y^{11} & E_y^{12} & & & \\
\vdots & & \ddots & & \\
E_x^{N1} & & & E_x^{NM} & \\
E_y^{N1} & & & & E_y^{NM}
\end{array}\right] \text{channels}
\end{array}
\qquad
\begin{array}{cc}
 & \text{channel means} \\
\mathbf{M} = & \left[\begin{array}{cc}
\mu_x^1 & \cdots \\
\mu_y^1 & \\
\vdots & \\
\mu_x^N & \\
\mu_y^N &
\end{array}\right] \text{channels}
\end{array}
$$

SVD of centred data matrix: $P \leq 2N$ non-zero singular values

$$\mathbf{X}_c = \mathbf{X} - \mathbf{M} = \mathbf{USV}^T \qquad \dim(\mathbf{U}) = 2N \times P$$

2^{nd} approach: Principal component analysis

Objectives in **EM array processing**:
- determine the number of independent linear combinations
- find a low-dimensional approximation to explain data variance

The observations are linear combinations in the $\mathbf{U}$ -space

$$\mathbf{X}_c = \mathbf{X} - \mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^T \iff \mathbf{X}_c = \mathbf{U}\mathbf{A} \iff \mathbf{U}^T\mathbf{X}_c = \mathbf{A}$$

observed realizations

$\Rightarrow \mathbf{U}$ are the 'principal components'

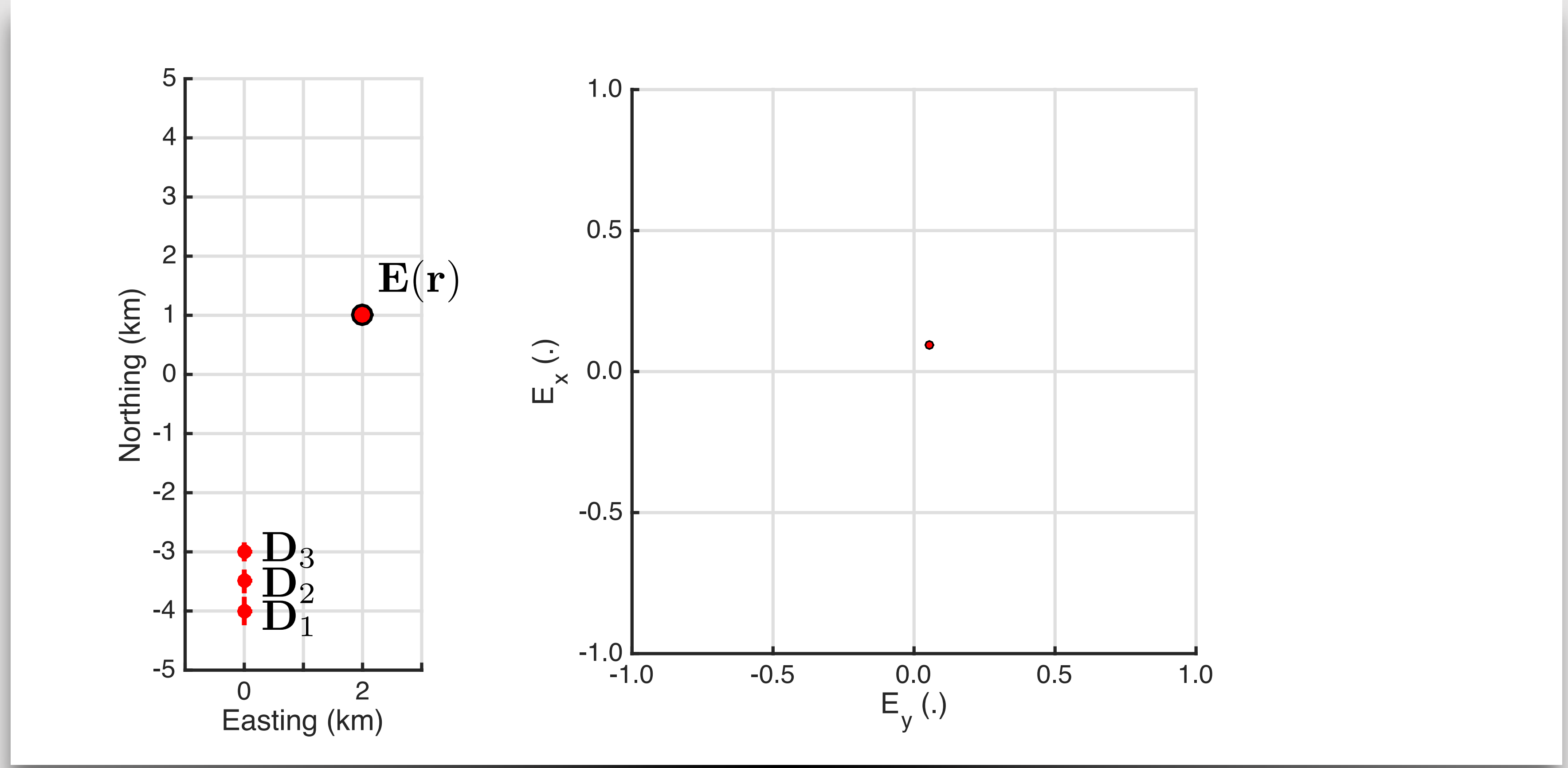
Let $\mathbf{A} = \mathbf{I}$, then $\mathbf{X}_c^U = \mathbf{U}$. $\Rightarrow$ $\mathbf{U}$ can be conceived as electric field

In Practice:
- works for normally distributed random variables
- robust techniques are available (we use `robpca`)

Synthetic example: insulating full-space

Electric field from <u>one</u> set of dipoles: $\mathbf{E}(\mathbf{r}) = \sum_s \mathbf{D}_s \mathbf{E}_s^{unit}(\mathbf{r})$

dipole moments

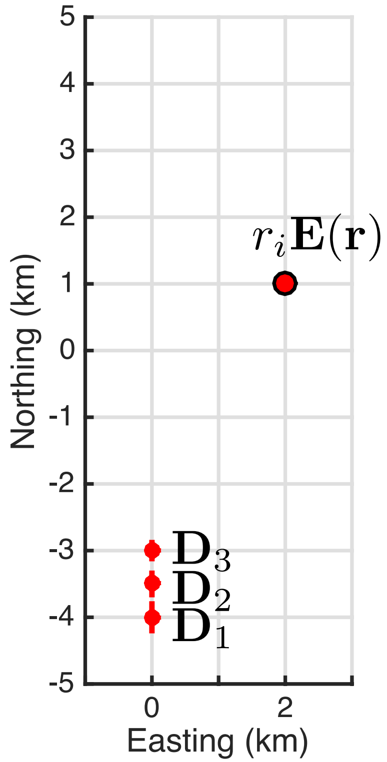


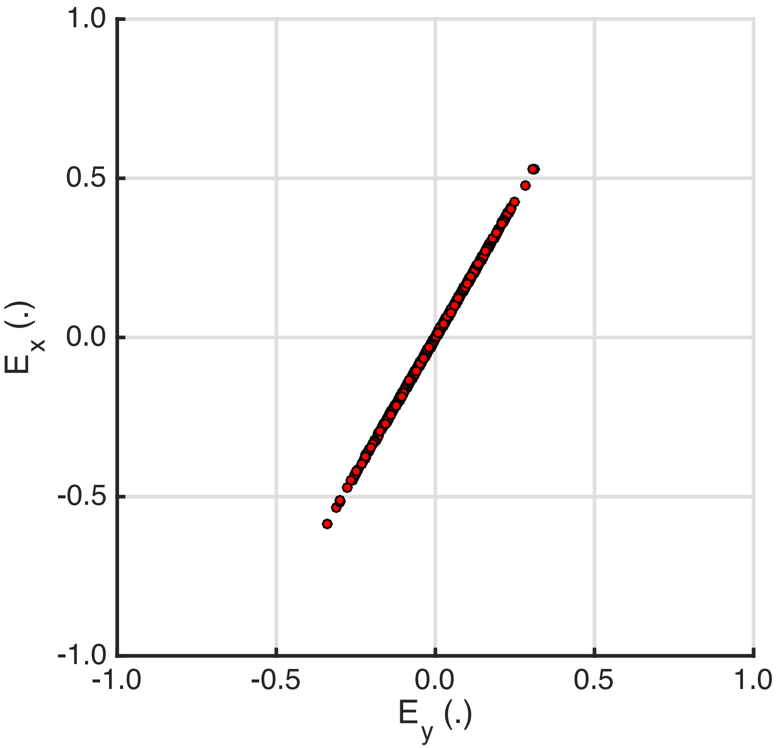
Synthetic example: insulating full-space

Electric field from one set of dipoles: $\mathbf{E}(\mathbf{r}) = \sum_s \mathbf{D}_s \mathbf{E}_s^{unit}(\mathbf{r})$

1000 random realisations: $\mathbf{E}_i(\mathbf{r}) = r_i \mathbf{E}(\mathbf{r})$

random number



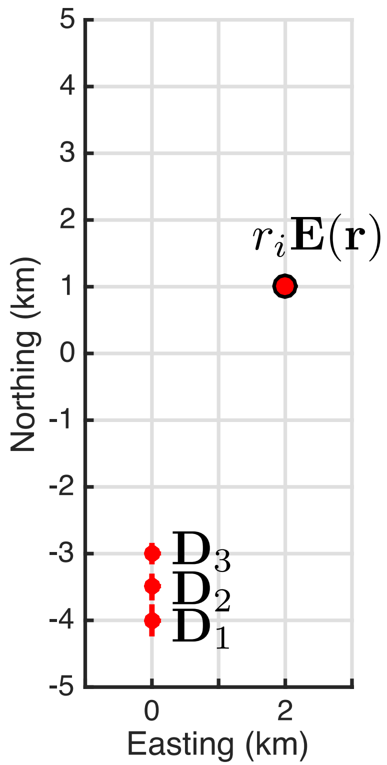


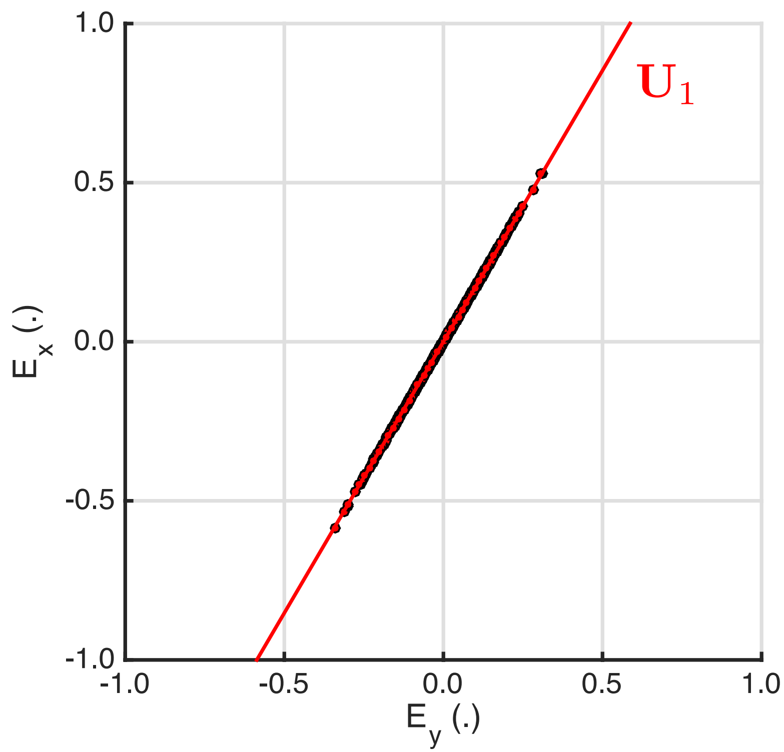
Synthetic example: insulating full-space

Electric field from <u>one</u> set of dipoles: $\mathbf{E}(\mathbf{r}) = \sum_s \mathbf{D}_s \mathbf{E}_s^{unit}(\mathbf{r})$

1000 random realisations: $\mathbf{E}_i(\mathbf{r}) = r_i \mathbf{E}(\mathbf{r})$

<u>One</u> non-zero eigenvalue. $= a_i \mathbf{U}_1$





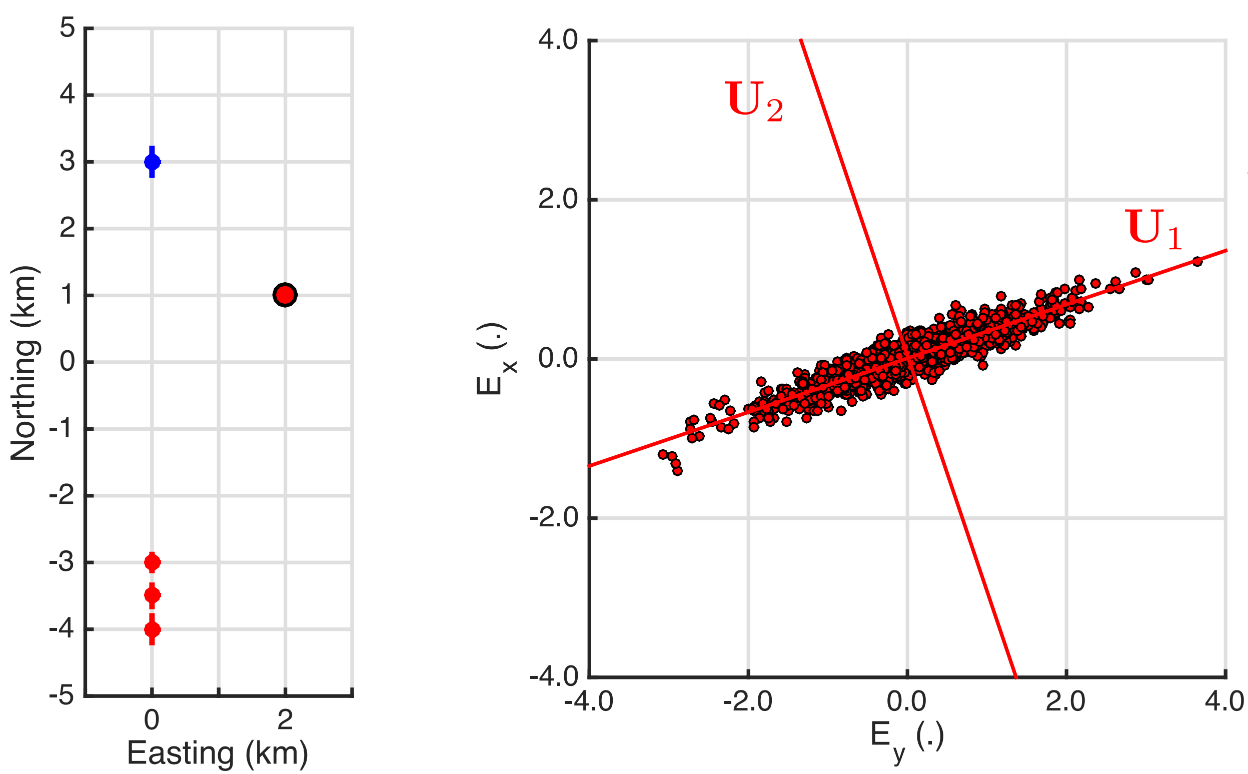
$\mathbf{U}_1$ explains 100 % of the data var.

Synthetic example: insulating full-space

Electric field from <u>two</u> sets of dipoles: $\mathbf{E}^j(\mathbf{r}) = \sum_s \mathbf{D}_s^j \mathbf{E}_s^{unit}(\mathbf{r})$

1000 random realisations: $\mathbf{E}_i(\mathbf{r}) = r_i^1 \mathbf{E}^1(\mathbf{r}) + r_i^2 \mathbf{E}^2(\mathbf{r})$

<u>Two</u> non-zero eigenvalues: $= a_{1i}\mathbf{U}_1 + a_{2i}\mathbf{U}_2$



$\mathbf{U}_1$ explains 98.5 % of the data var.

$\mathbf{U}_2$ explains 1.5 %

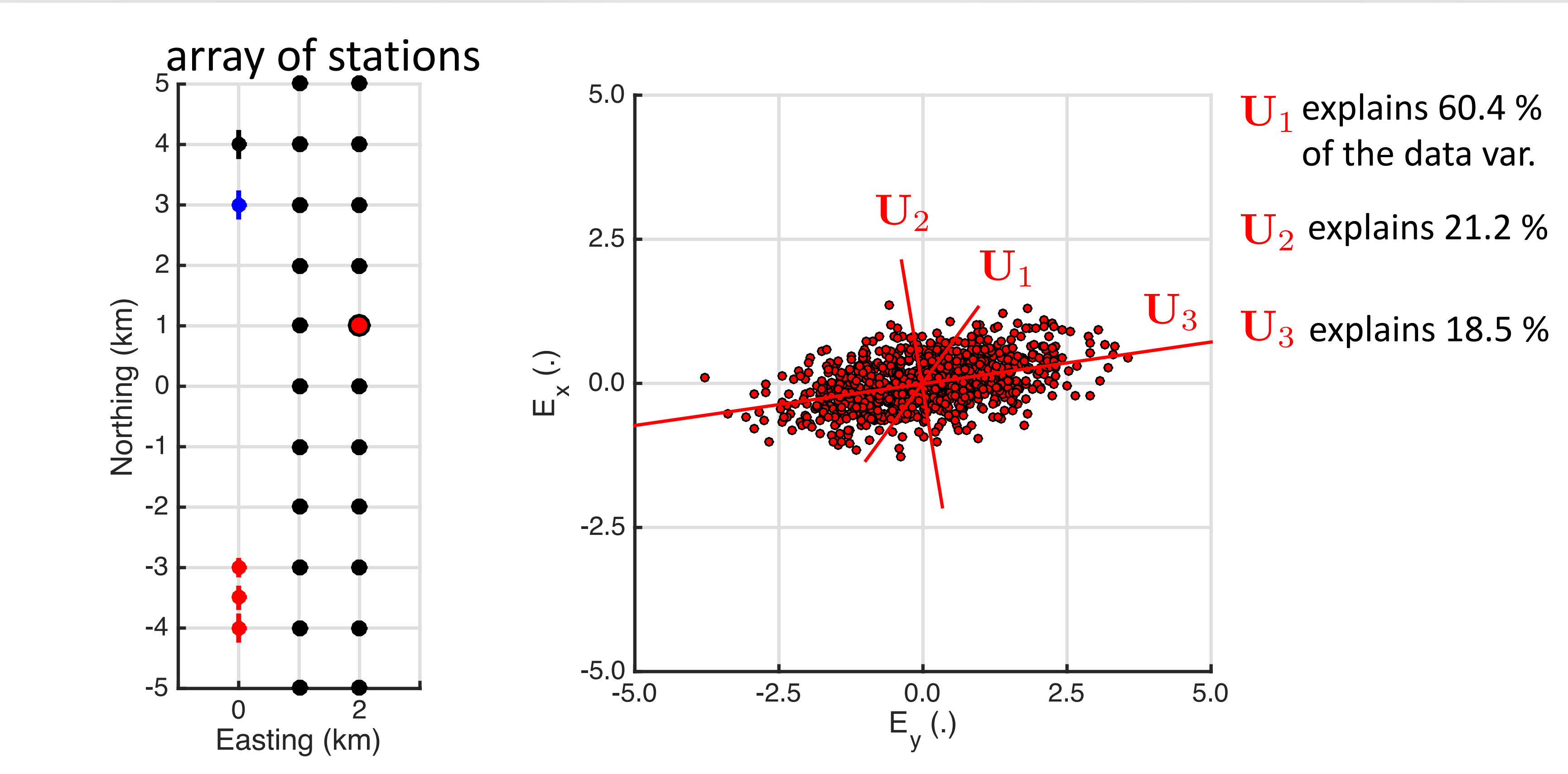
<u>*Note:*</u> PCs relate to some linear comb. of independent sources.

Synthetic example: insulating full-space

Electric fields from <u>three</u> sets of dipoles: $\mathbf{E}^j(\mathbf{r}) = \sum_s \mathbf{D}_s^j \mathbf{E}_s^{unit}(\mathbf{r})$

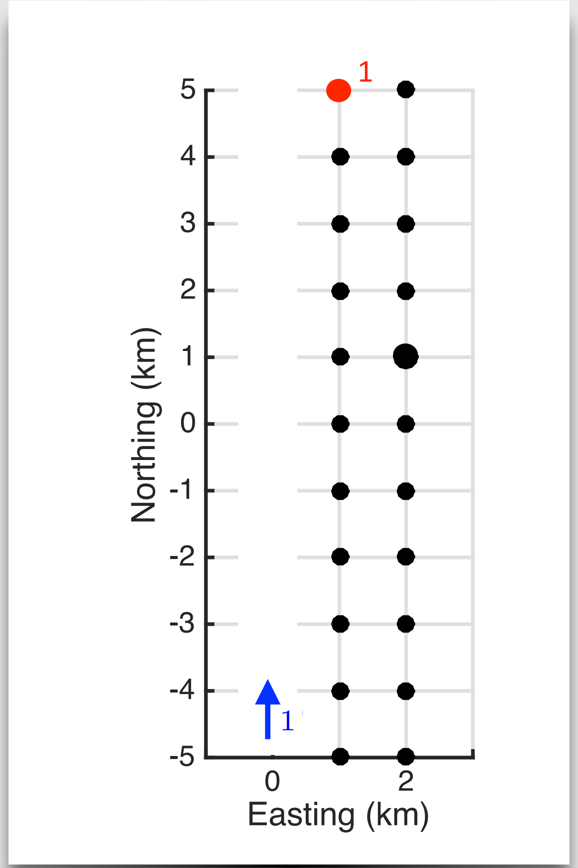
1000 random realisations: $\mathbf{E}_i(\mathbf{r}) = r_i^1 \mathbf{E}^1(\mathbf{r}) + r_i^2 \mathbf{E}^2(\mathbf{r}) + r_i^3 \mathbf{E}^3(\mathbf{r})$

<u>Three</u> non-zero eigenvalues: $= a_{i1}\mathbf{U}_1 + a_{i2}\mathbf{U}_2 + a_{i3}\mathbf{U}_3$



Inversion of principal components for sources

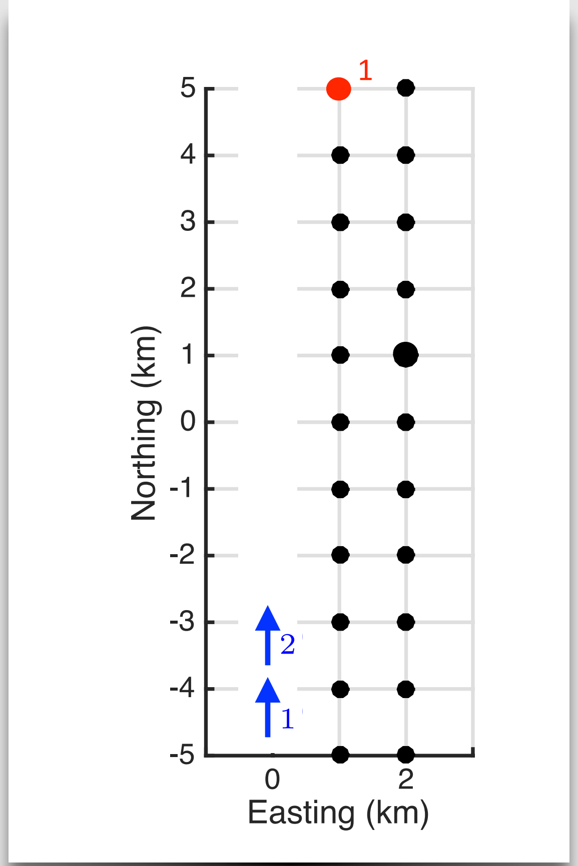
For each principal component $\Rightarrow$ linear inverse problem



$\mathbf{E}_1^{unit,1}$

Inversion of principal components for sources

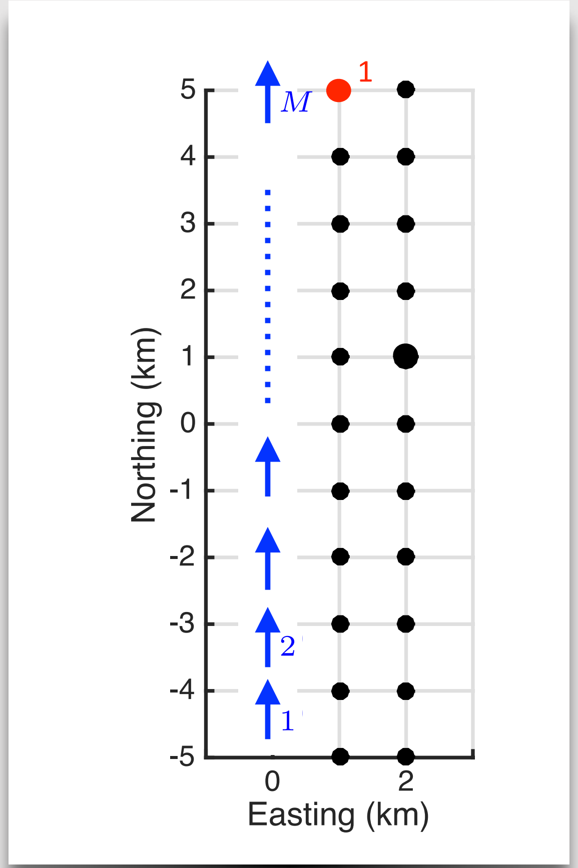
For each principal component $\Rightarrow$ linear inverse problem



$\mathbf{E}_1^{unit,1}, \mathbf{E}_1^{unit,2}$

Inversion of principal components for sources

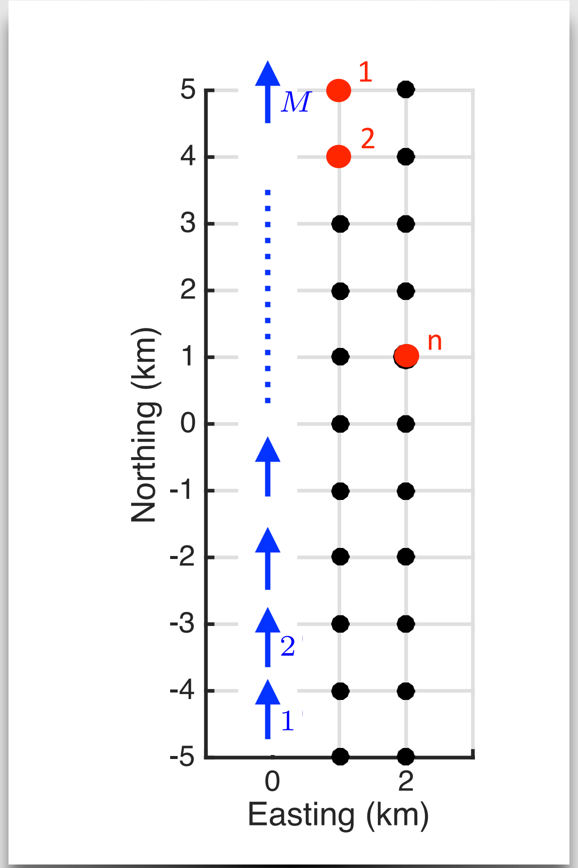
For each principal component $\Rightarrow$ linear inverse problem



$$\mathbf{E}_1^{unit,1}, \mathbf{E}_1^{unit,2}, ..., \mathbf{E}_1^{unit,M}$$

Inversion of principal components for sources

For each principal component $\Rightarrow$ linear inverse problem



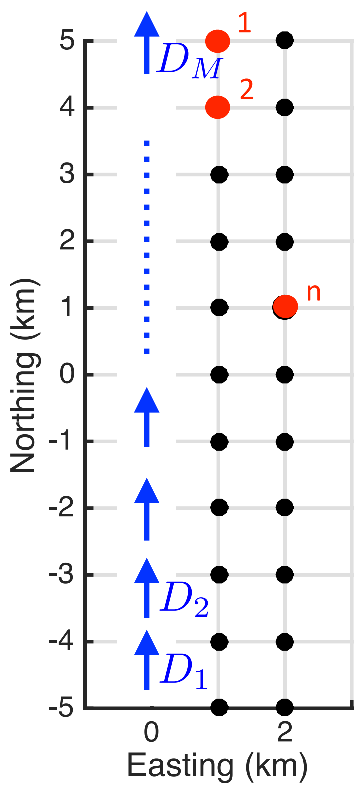
$$\left.\begin{array}{l} \mathbf{E}_1^{unit,1}, \mathbf{E}_1^{unit,2}, \ldots, \mathbf{E}_1^{unit,M} \\ \mathbf{E}_2^{unit,1}, \mathbf{E}_2^{unit,2}, \ldots, \mathbf{E}_2^{unit,M} \end{array}\right\}$$ kernel $\mathbf{X}^{unit}$, precomputed

$\ldots\ldots$

Inversion of principal components for sources

For each principal component $\Rightarrow$ linear inverse problem

model vector $\mathbf{D}$



$$\mathbf{E}_1^{unit,1}, \mathbf{E}_1^{unit,2}, \ldots, \mathbf{E}_1^{unit,M}$$
$$\mathbf{E}_2^{unit,1}, \mathbf{E}_2^{unit,2}, \ldots, \mathbf{E}_2^{unit,M}$$
....

kernel $\mathbf{X}^{unit}$, precomputed

Pose linear inverse problem:

$$\mathbf{U} = \mathbf{X}^{unit}\mathbf{D}$$

$$\mathbf{D}^{est} = \left[\mathbf{X}^{unit,T}\mathbf{X}^{unit} + \lambda \mathbf{W}_m\right]^{-1} \mathbf{X}^{unit,T}\mathbf{U}^{obs}$$

regularization term

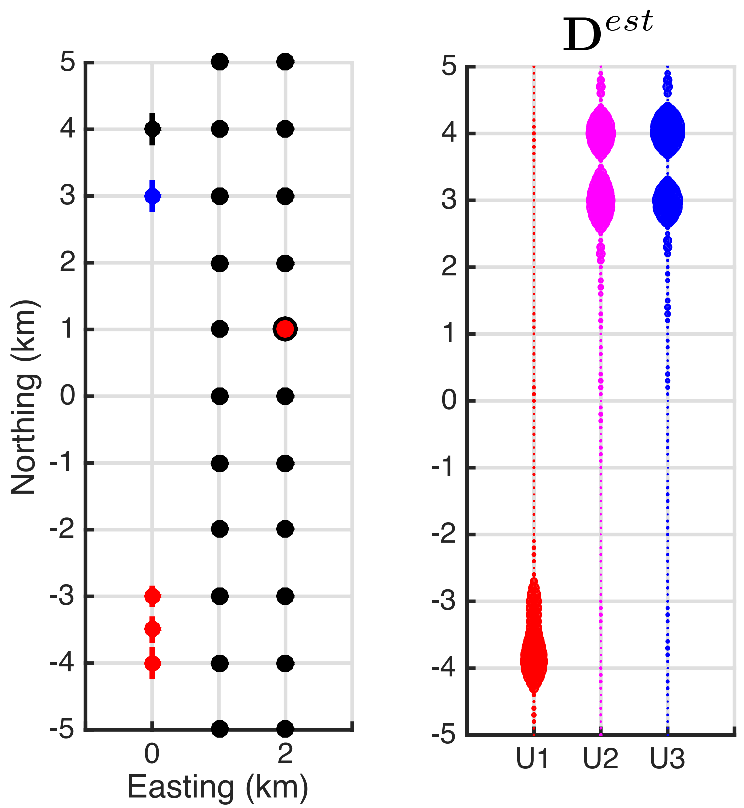
Assumptions required:
source geometry, regularization

Inversion of principal components for sources

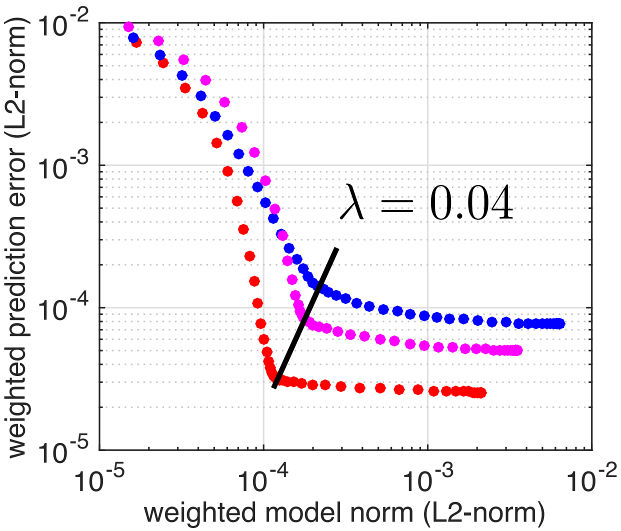
For each principal component $\Rightarrow$ linear inverse problem:

Source estimate is particular linear combination of the three sources

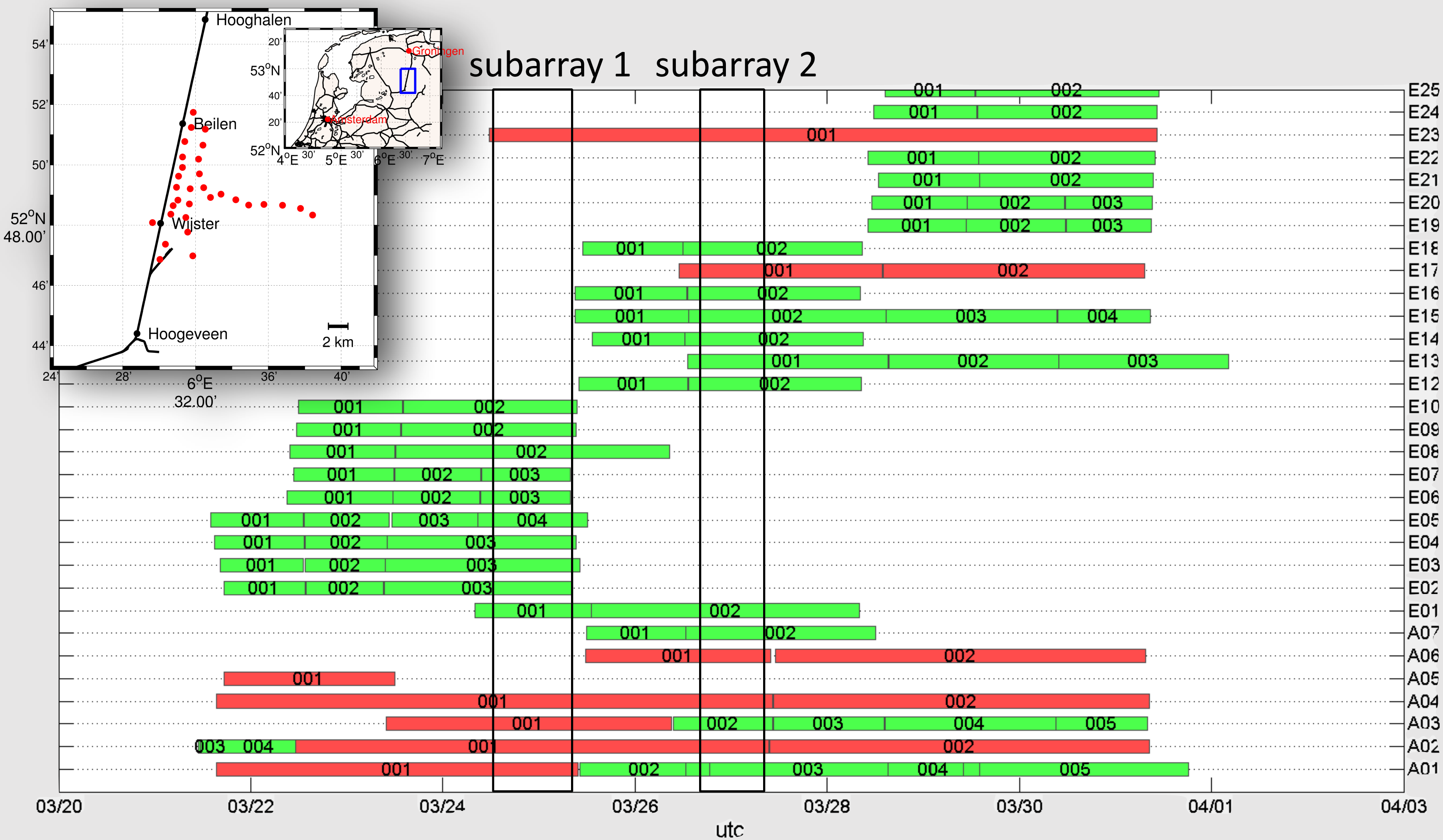
(for the inversion: 100 m dipole spacing)



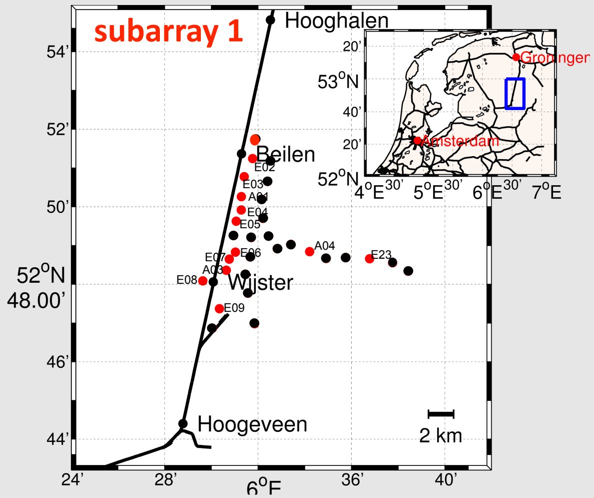
- noise was added to electric field
- data fit perfect (not shown)
- source estimate: linear combination of true sources



Application to field data - runtimes



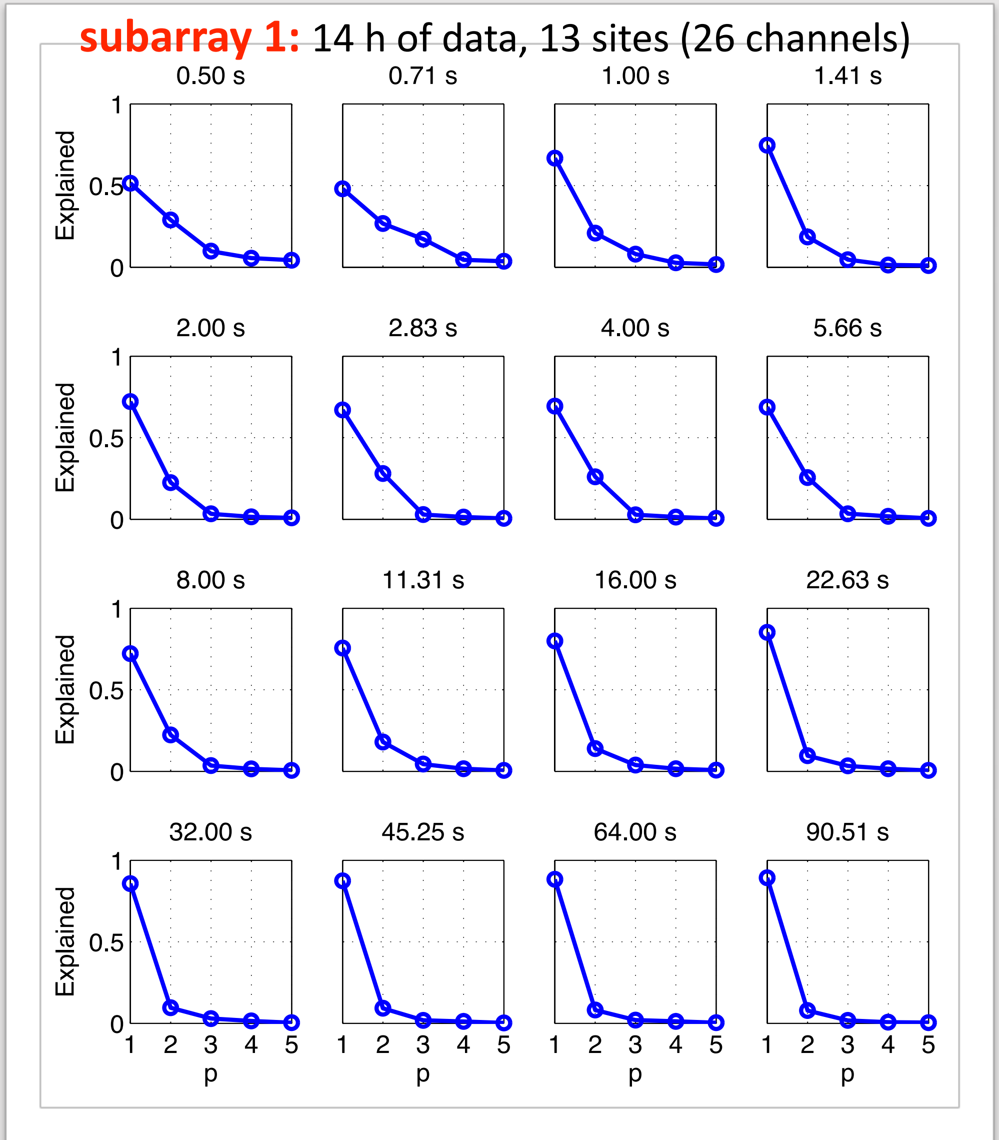
PCA of field data - eigenvalues of subarray 1



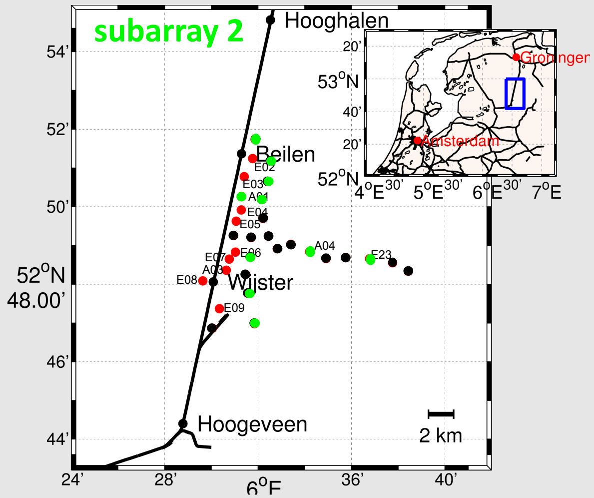
For each period, collect fourier coefs. for all E-channels in matrix $\mathbf{X}$:

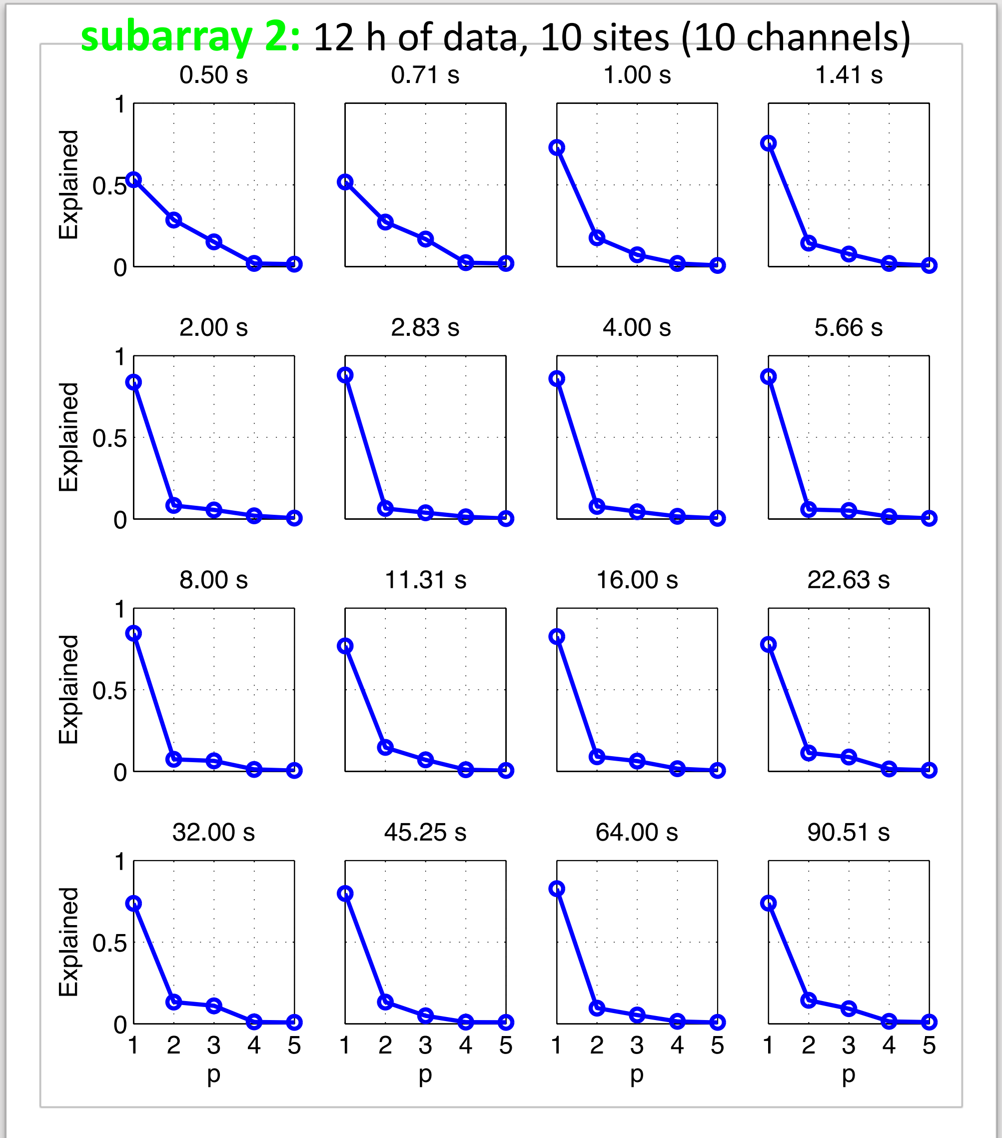
$\Rightarrow$ PCA: $\mathbf{X}_c = \mathbf{X} - \mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^T$
$= \mathbf{U}\mathbf{A}$

$$\text{Explained} = \frac{\lambda_p}{\sum_p \lambda_p}$$



PCA of field data - eigenvalues of subarray 2



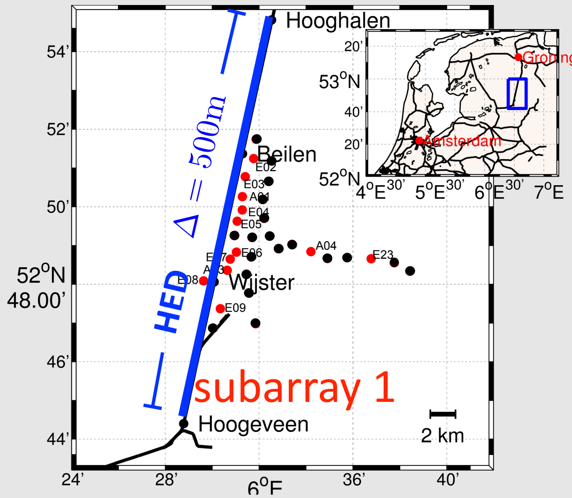


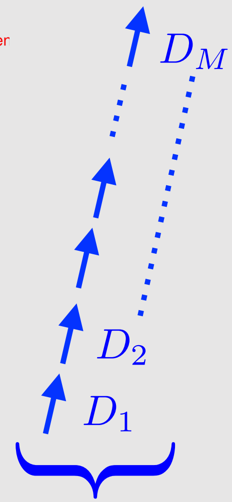
For each period, collect fourier coefs. for all E-channels in matrix $\mathbf{X}$:

$$\Rightarrow \text{PCA: } \mathbf{X}_c = \mathbf{X} - \mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U}\mathbf{A}$$

$$\text{Explained} = \frac{\lambda_p}{\sum_p \lambda_p}$$

Inversion of PCs for the sources





model vector $\mathbf{D}$

- $\mathbf{E}_1^{unit,1}, \mathbf{E}_1^{unit,2}, \ldots, \mathbf{E}_1^{unit,M}$
- $\mathbf{E}_2^{unit,1}, \mathbf{E}_2^{unit,2}, \ldots, \mathbf{E}_2^{unit,M}$
- ....

kernel $\mathbf{X}^{unit}$, precomputed for a priori resistivity model

Pose linear inverse problem:

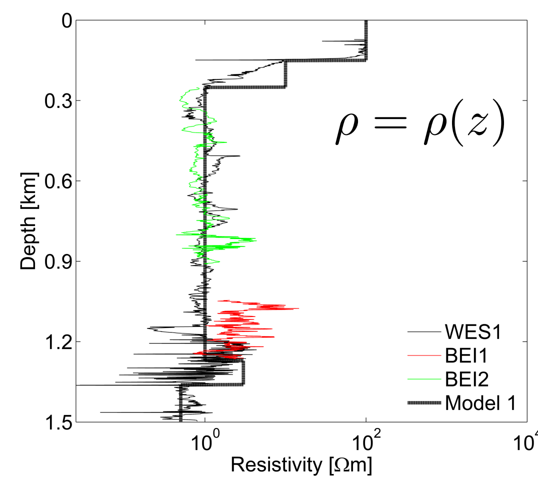
$$\mathbf{U} = \mathbf{X}^{unit}\mathbf{D}$$

regularization term

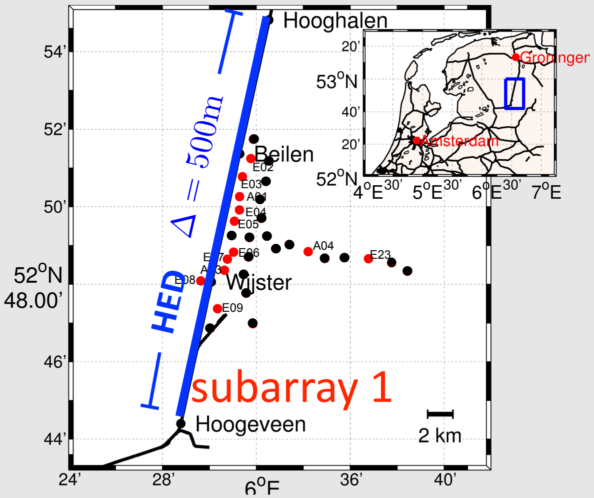
data weighting

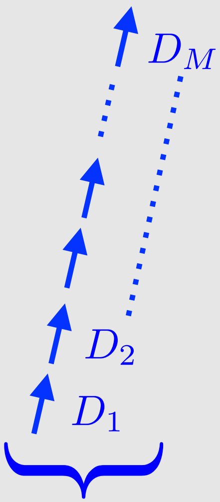
$$\mathbf{D}^{est} = \left[\mathbf{X}^{unit,T}\mathbf{W}_e\mathbf{X}^{unit} + \lambda\mathbf{W}_m\right]^{-1}\mathbf{W}_e\mathbf{X}^{unit,T}\mathbf{U}^{obs}$$

$\boxed{\text{Assumptions required: source geometry, a priori resistivity model, regularization, data weighting}}$

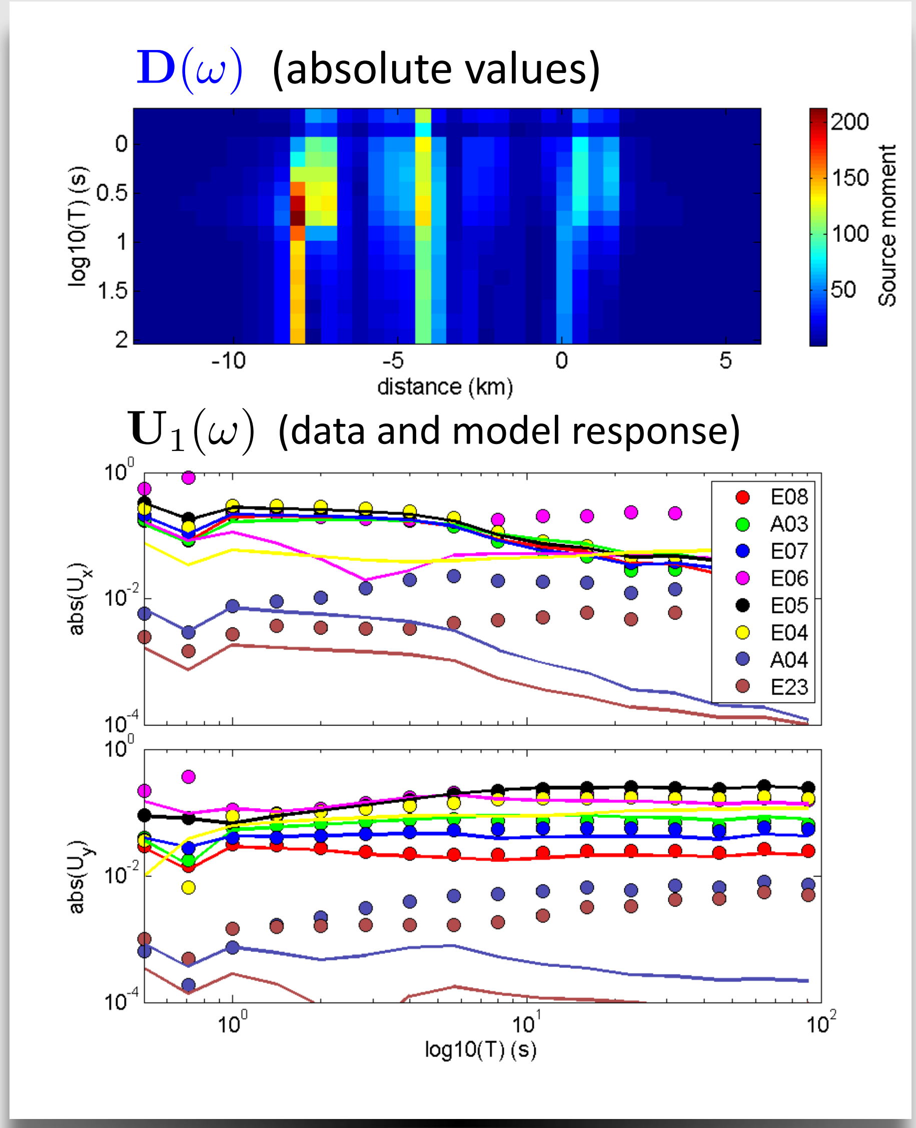


Testing inversion of subarray 1 for 1st source





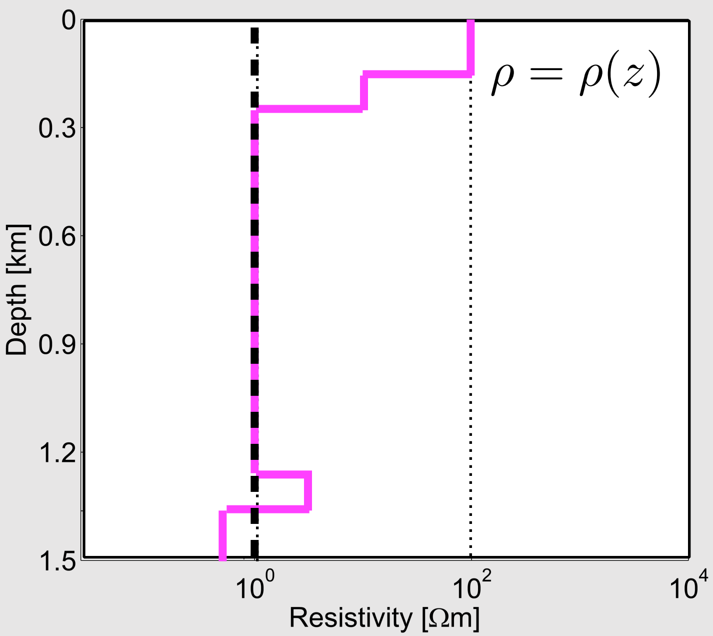
model vector $\mathbf{D}$

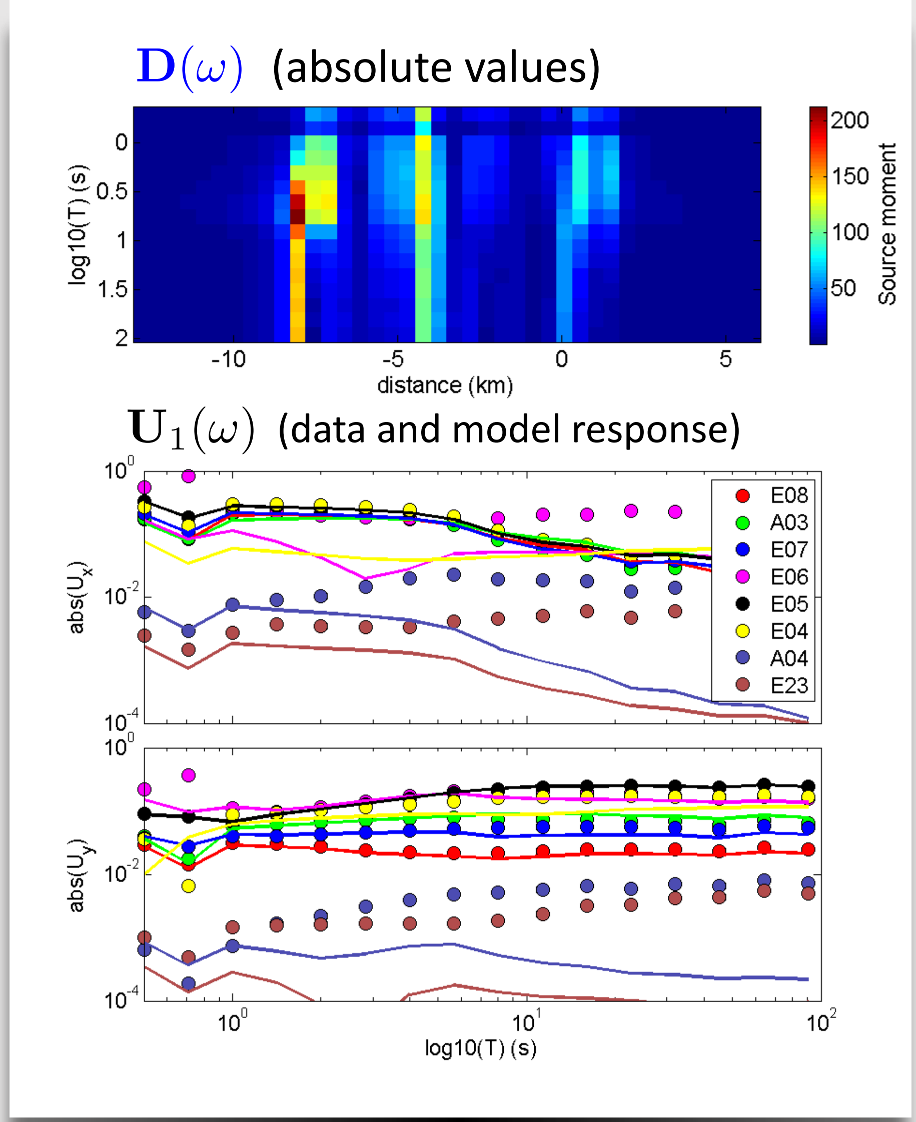


Testing inversion of subarray 1 for 1^{st} source

Testing resistivity models:

(i) 'borehole' model



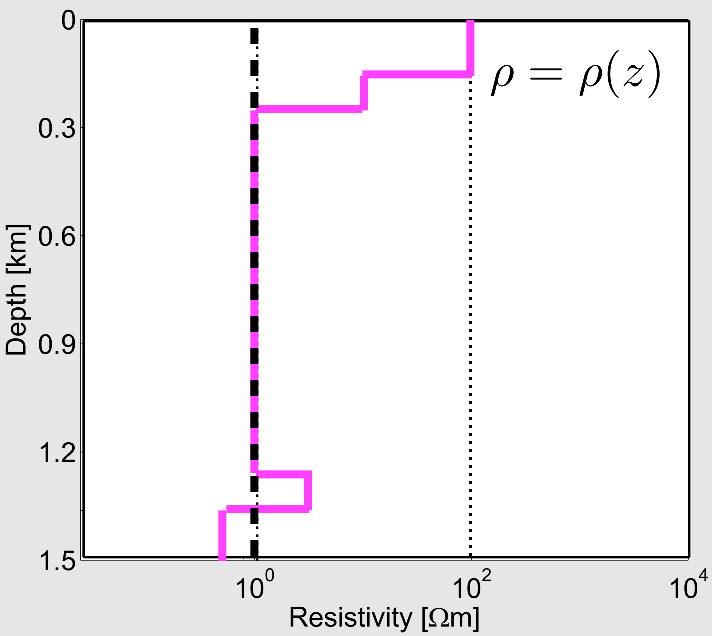


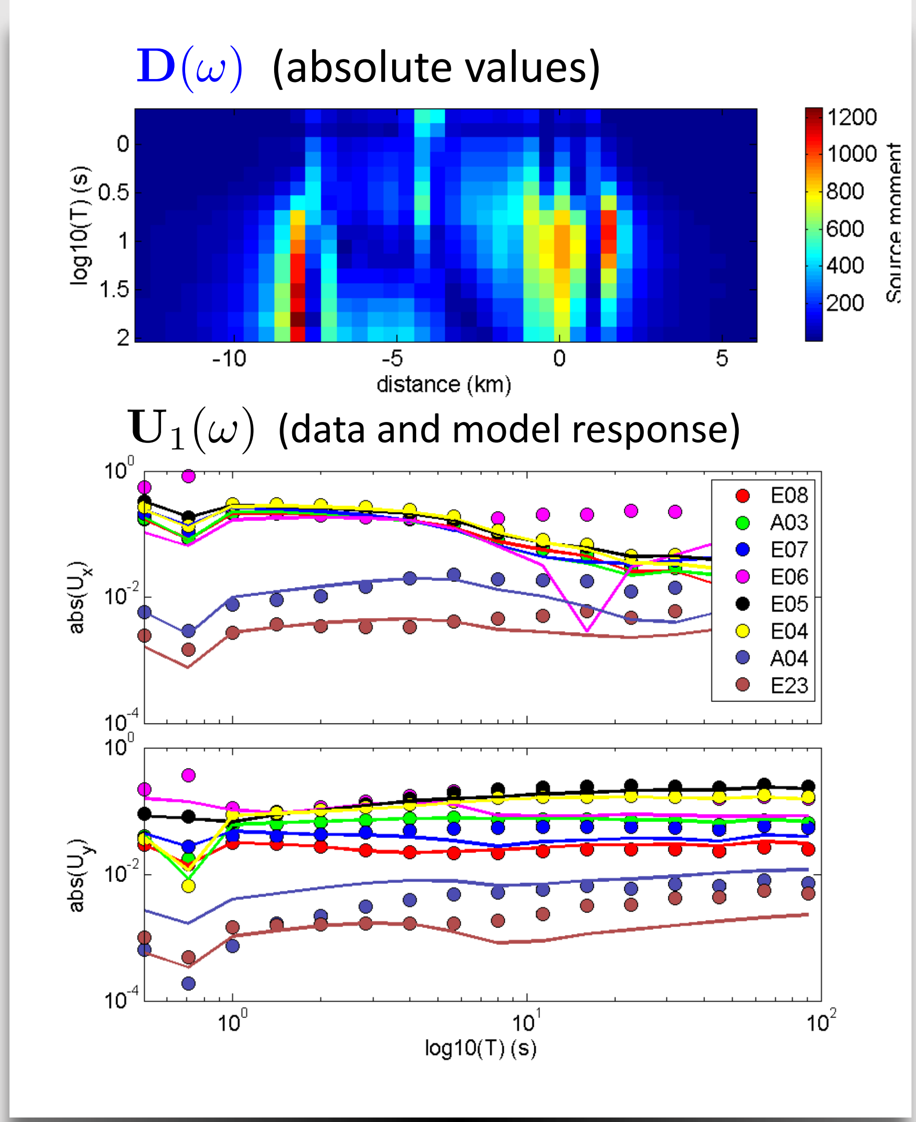
Testing inversion of subarray 1 for 1st source

Testing resistivity models:

(i) 'borehole' model

(ii) 1 Ohm-m half-space

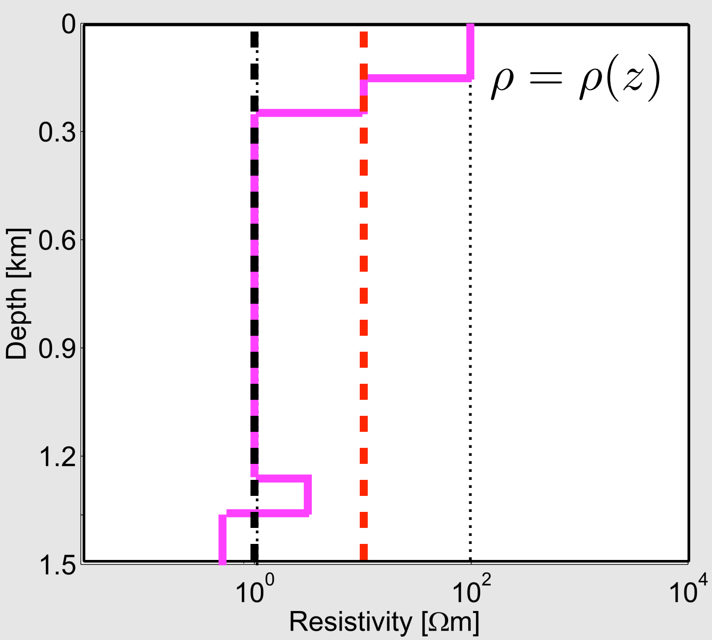


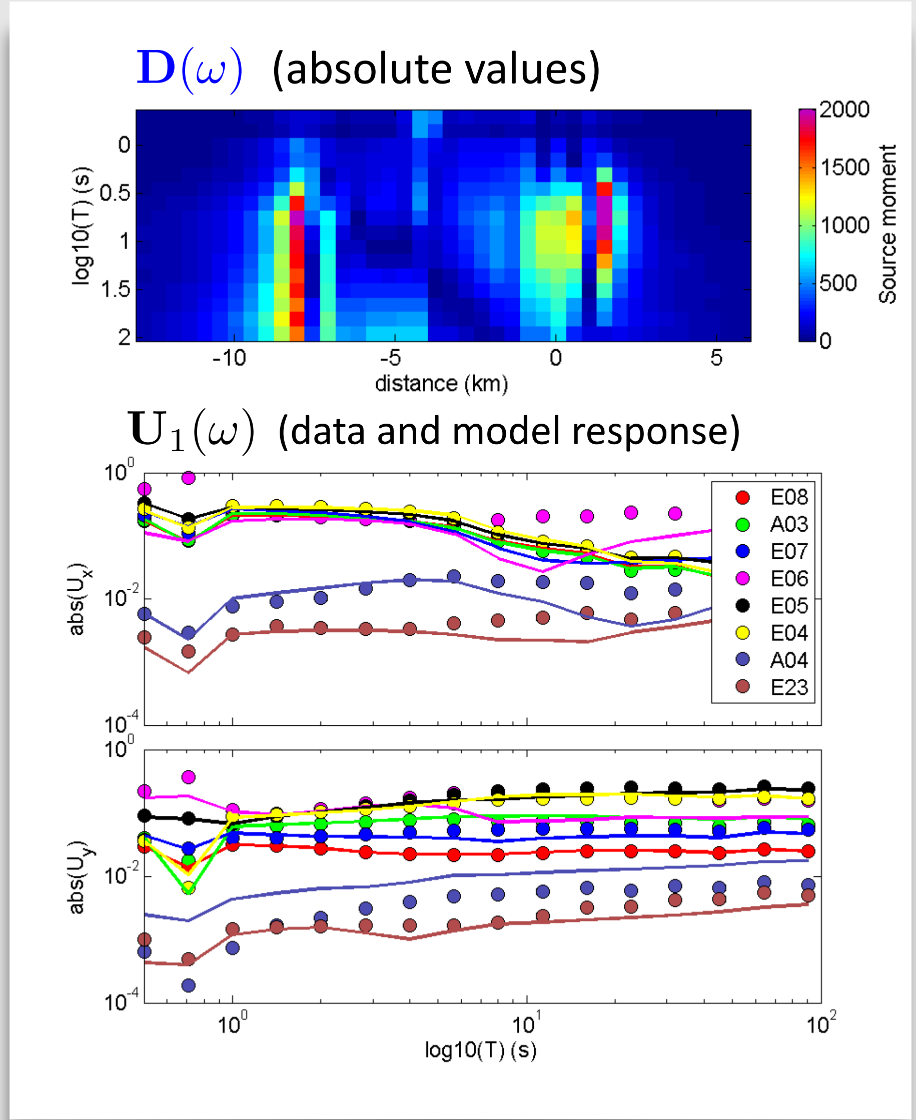


Testing inversion of subarray 1 for 1st source

Testing resistivity models:

(i) 'borehole' model
(ii) 1 Ohm-m half-space
(iii) 10 Ohm-m half-space

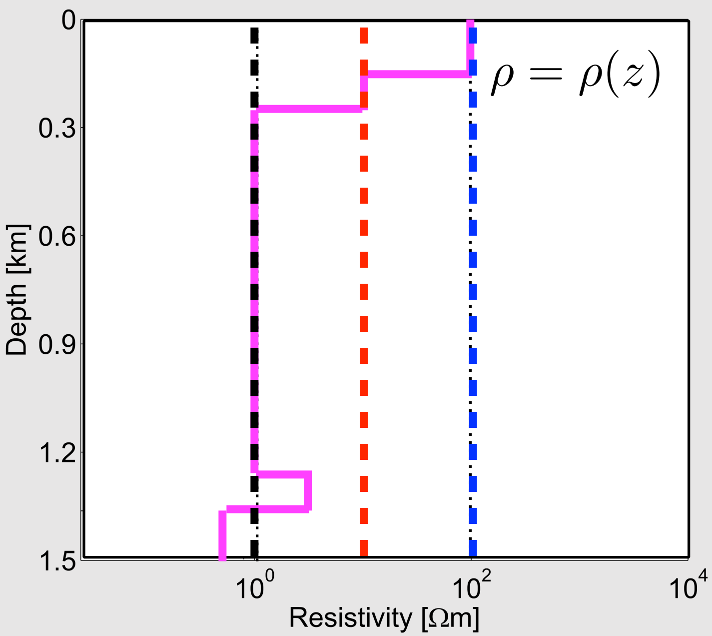


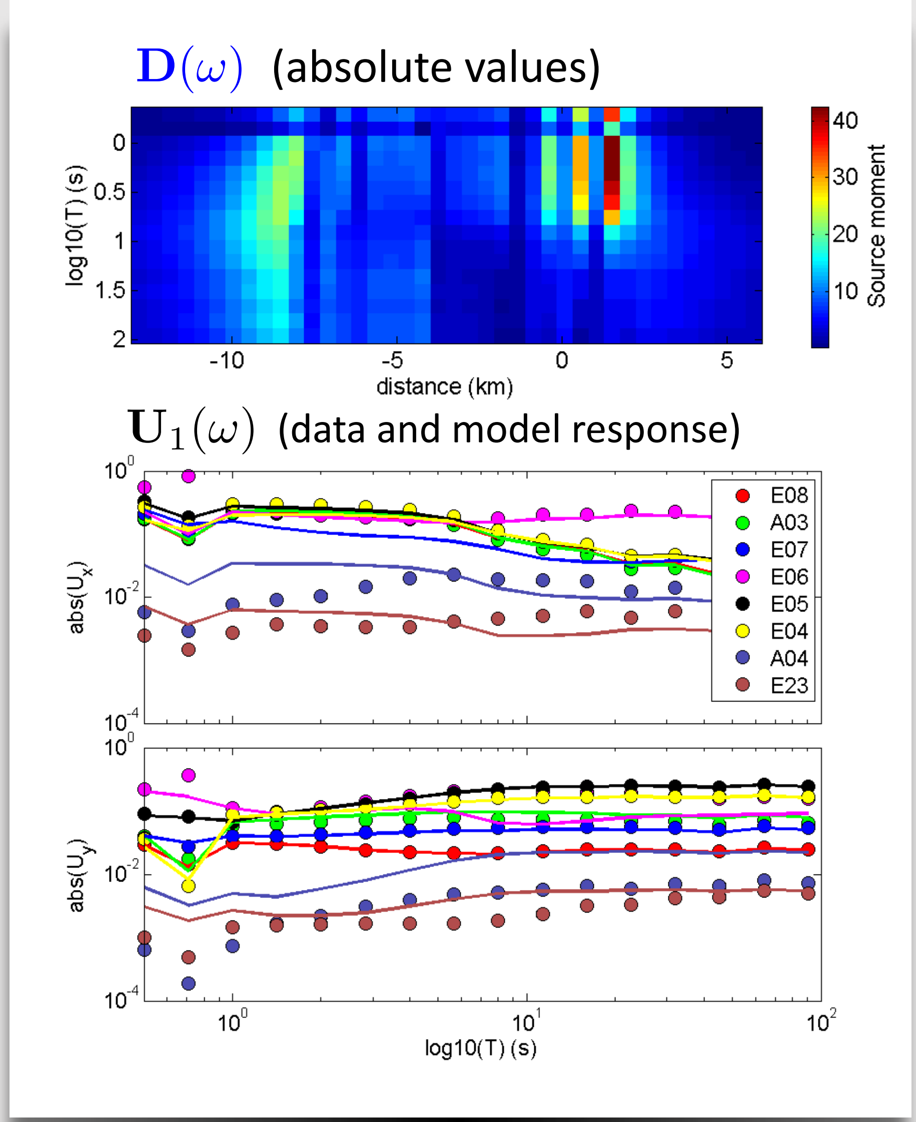


Testing inversion of subarray 1 for 1st source

Testing resistivity models:

(i) 'borehole' model
(ii) 1 Ohm-m half-space
(iii) 10 Ohm-m half-space
(iv) 100 Ohm-m half-space

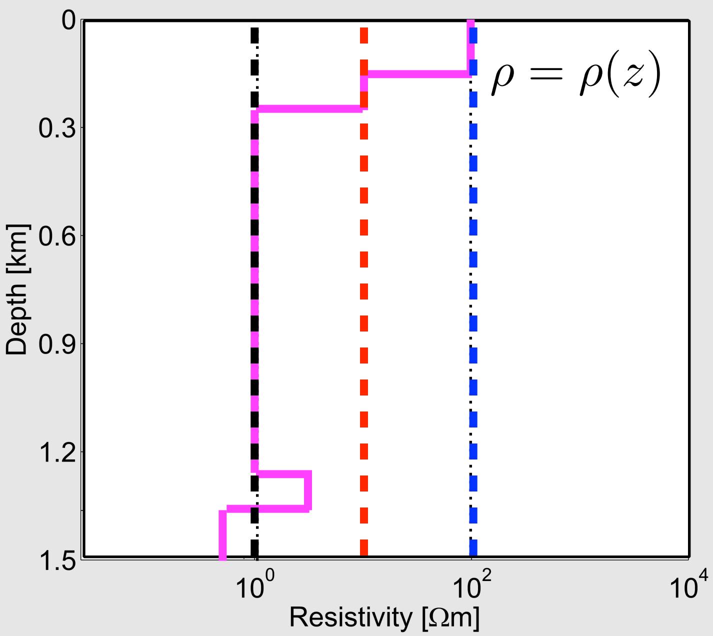




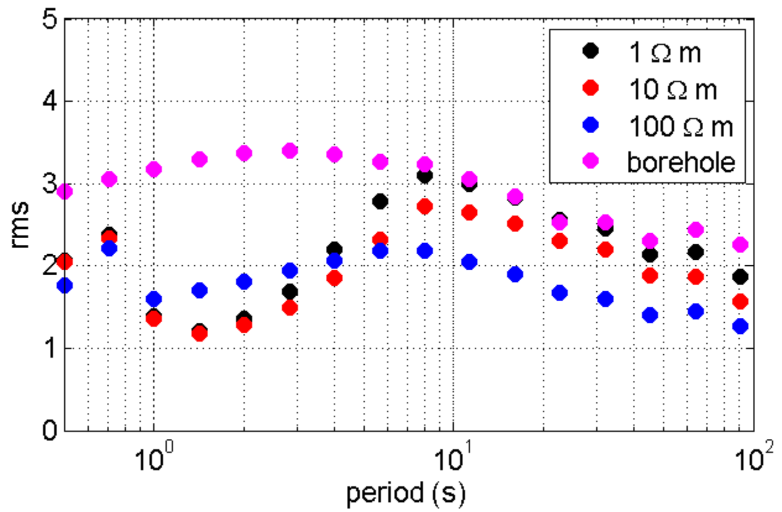
Testing inversion of subarray 1 for 1^{st} source

Testing resistivity models:

(i) 'borehole' model
(ii) 1 Ohm-m half-space
(iii) 10 Ohm-m half-space
(iv) 100 Ohm-m half-space



Misfit for different resistivity models:



$\Rightarrow$ source estimates and global misfit are sensitive to the a priori resistivity model

(for the particular dipole geometry used)

Merging subarrays 1 and 2

A spatially extended array may better capture the source modes.
Can we merge the subarrays?
Problem: PCA not unique, e.g. $\mathbf{X}_c = \mathbf{UA} = \mathbf{UR}^T\mathbf{RA}$

$\Rightarrow$ determine $\mathbf{R}$ from overlapping channels

For 2 PCs

subarray 1: $\mathbf{X}_c^1 = \mathbf{U}^1\mathbf{A}^1$

subarray 2: $\mathbf{X}_c^2 = \mathbf{U}^2\mathbf{A}^2$

pull-out overlapping channels:

$$\mathbf{U}^{1,overlap} = \begin{bmatrix} U_{a2}^1 & U_{a2}^1 \\ U_{b2}^1 & U_{b2}^1 \\ \vdots & \\ U_{n2}^1 & U_{n2}^1 \end{bmatrix} \qquad \mathbf{U}^{2,overlap} = \begin{bmatrix} U_{a2}^2 & U_{a2}^2 \\ U_{b2}^2 & U_{b2}^2 \\ \vdots & \\ U_{n2}^2 & U_{n2}^2 \end{bmatrix}$$

find orthogonal rotation matrix $\mathbf{R}$ from:

$$\left\|\mathbf{U}^{2,overlap} - c\mathbf{U}^{1,overlap}\mathbf{R}\right\| \to min$$

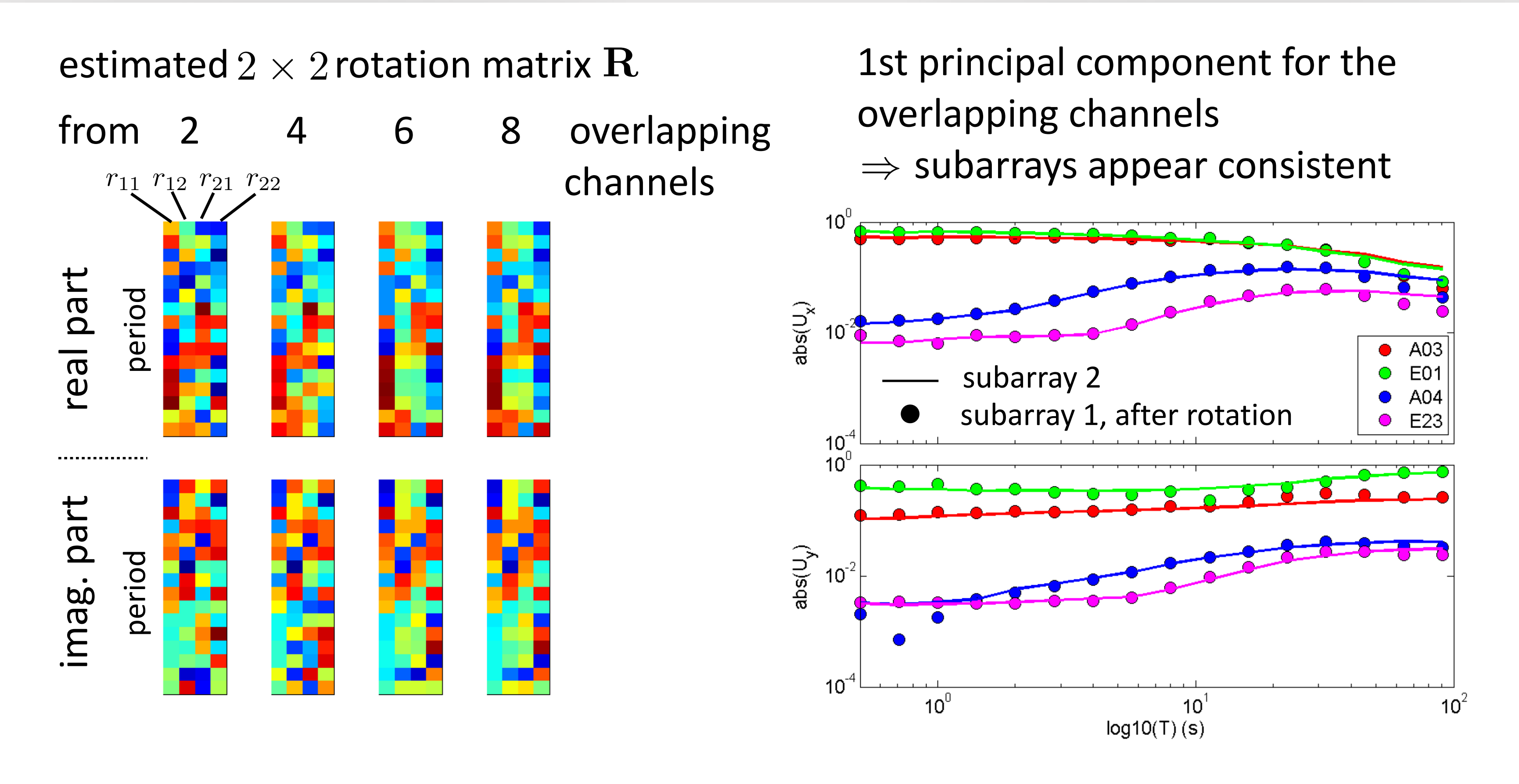
(we use matlab function `rotatefactors` and lsq estimate of scaling factor *c*)

form principal component of merged array:

$$\mathbf{U}^{merged} = \begin{bmatrix} c\mathbf{U}^1\mathbf{R} \\ \mathbf{U}^2 \end{bmatrix}$$

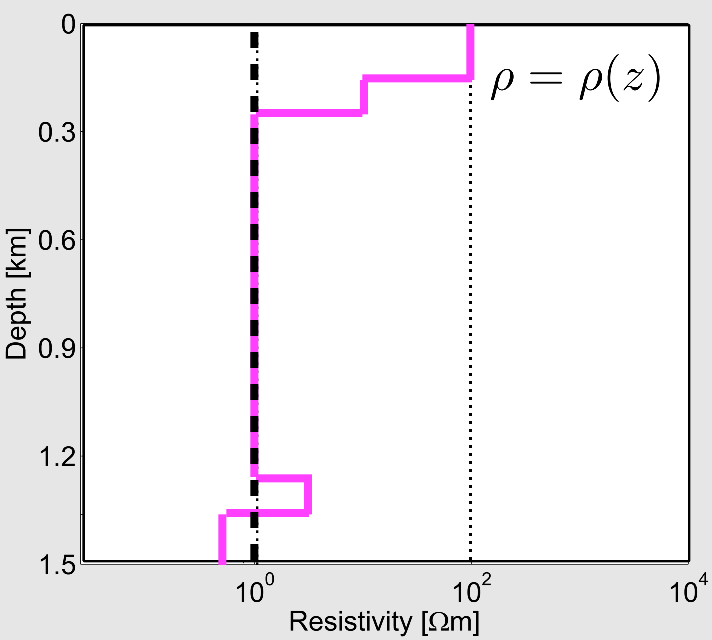
Merging subarrays 1 and 2

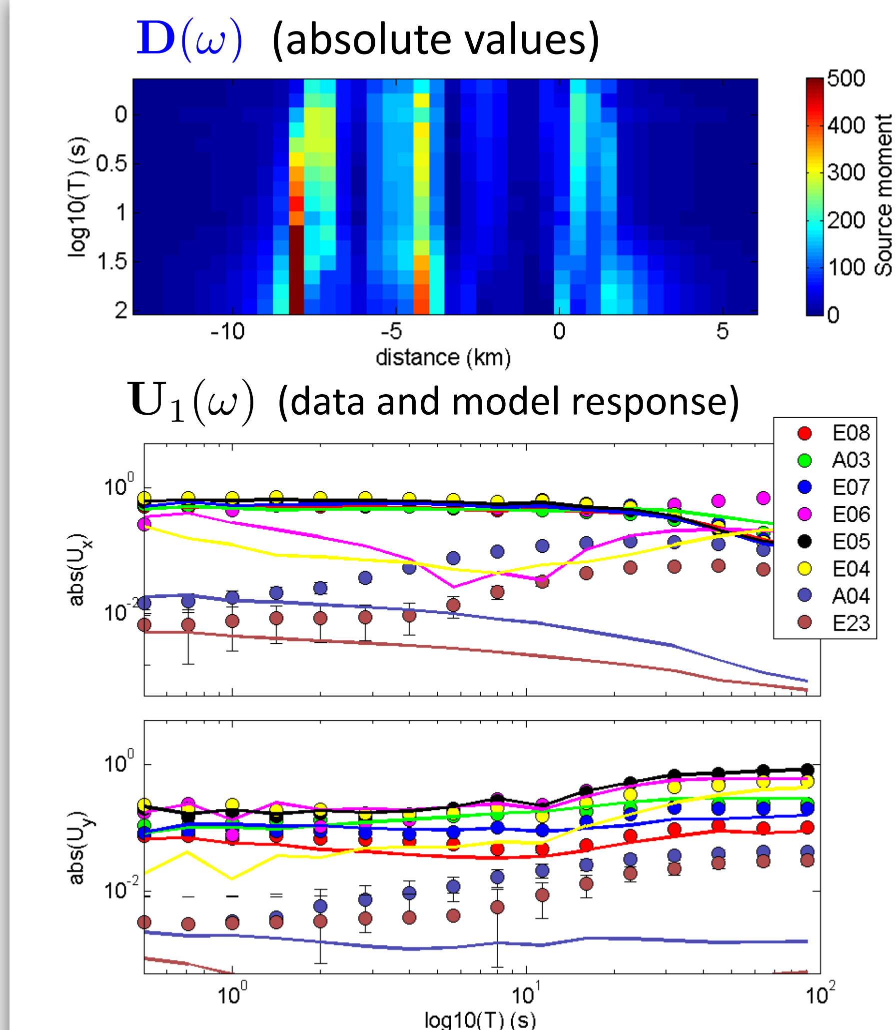
We use the first two principal components from our data.



Inversion of merged subarrays for 1st source

(i) ‘borehole’ model



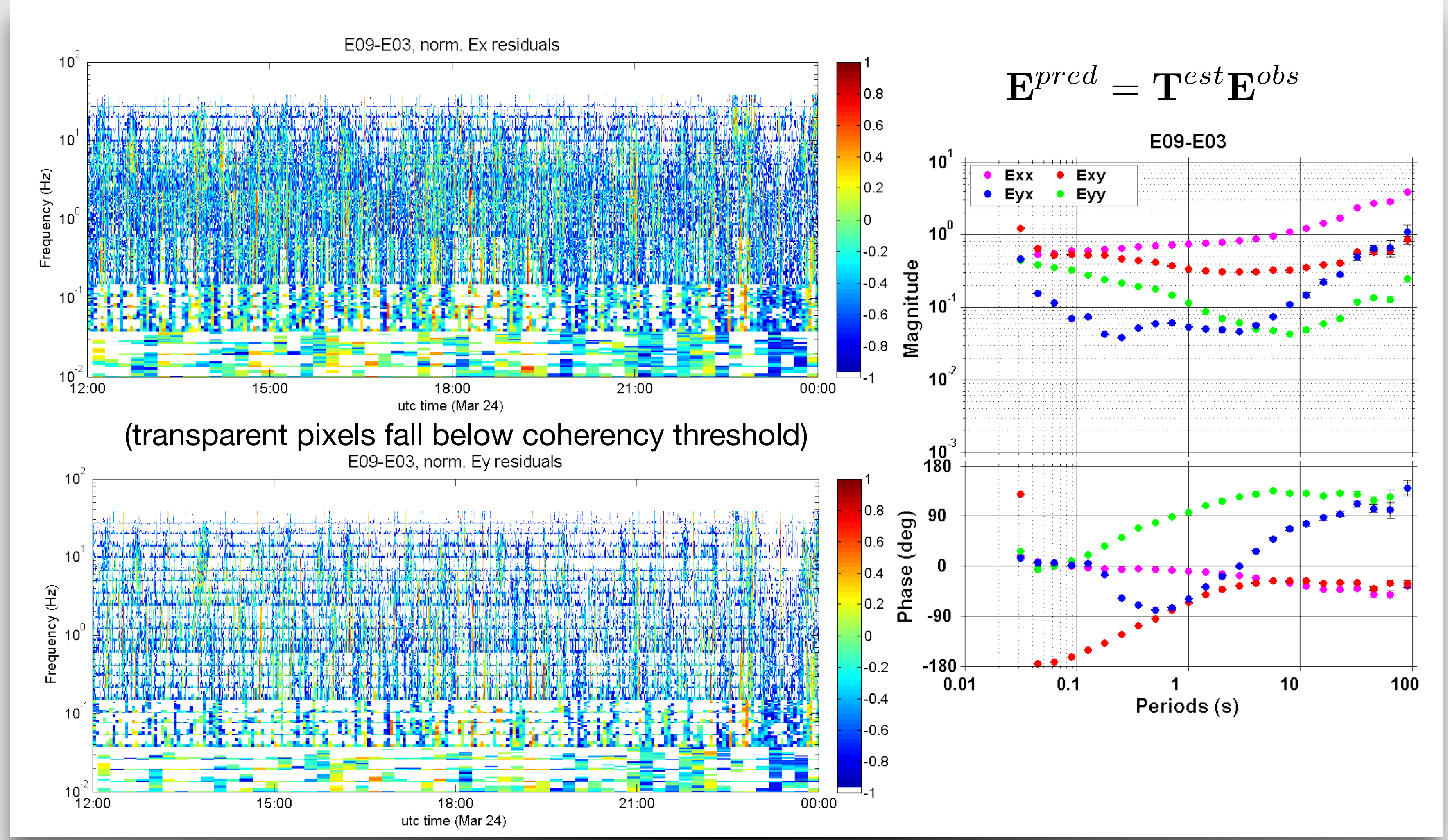


Are DC trains useful for EM exploration?

- PCA of EM array recordings yields *P=2-5* independent sources for the railway segment near Beilen
- Merging of the two subarrays appears feasible
- Inversion for the source assumes
  - a priori resistivity model
  - a priori geometry (distribution, orientation, type)
  - data weighting
- Sequential inversion (alternating updates of source and conductivity model) has been attempted on synthetic data
  - experienced convergence problems (local minima?)

$\Rightarrow$ No final answer yet!

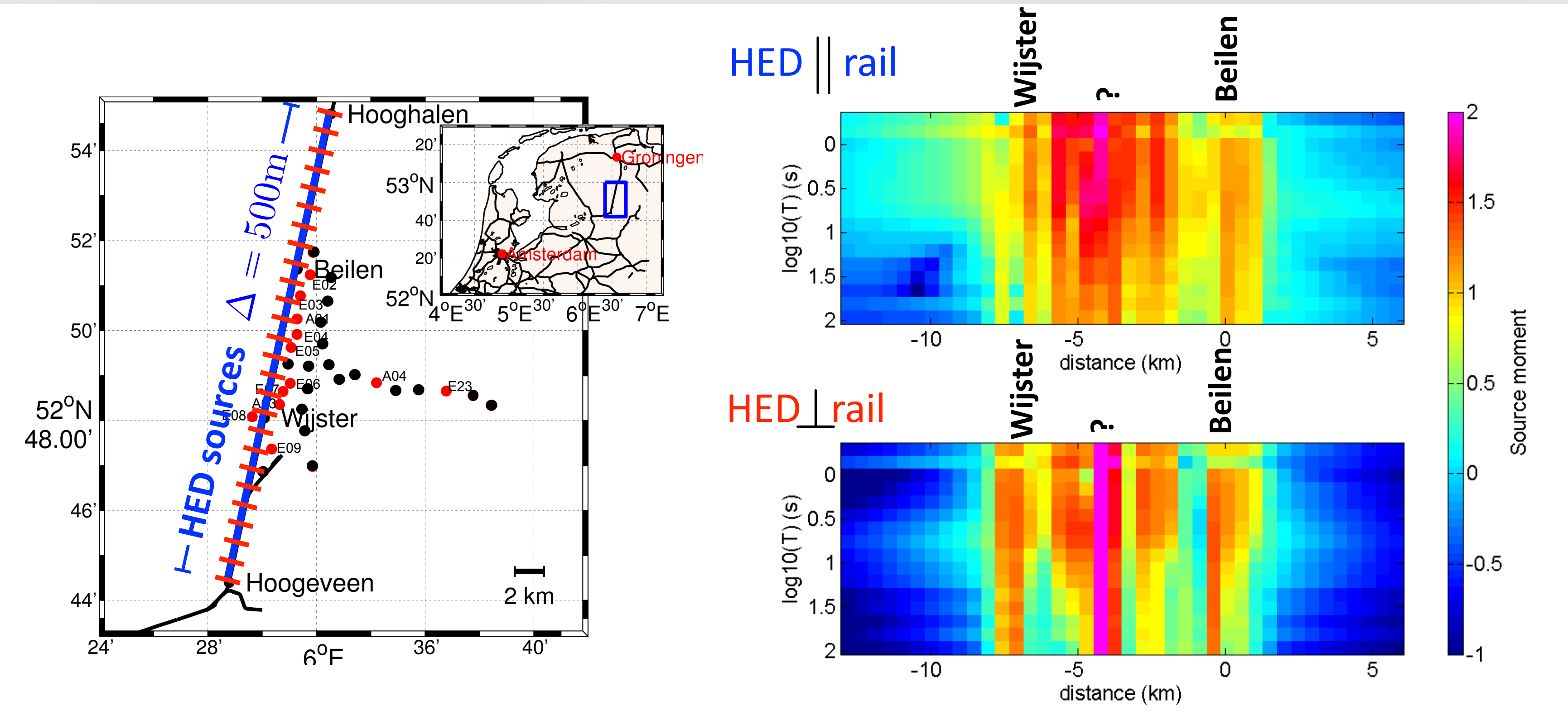
Residuals for bi-variate approach



Inversion for the sources

Source model obtained from inversion of the first PC for the source
and with 40 dipoles added ⊥ to the rail

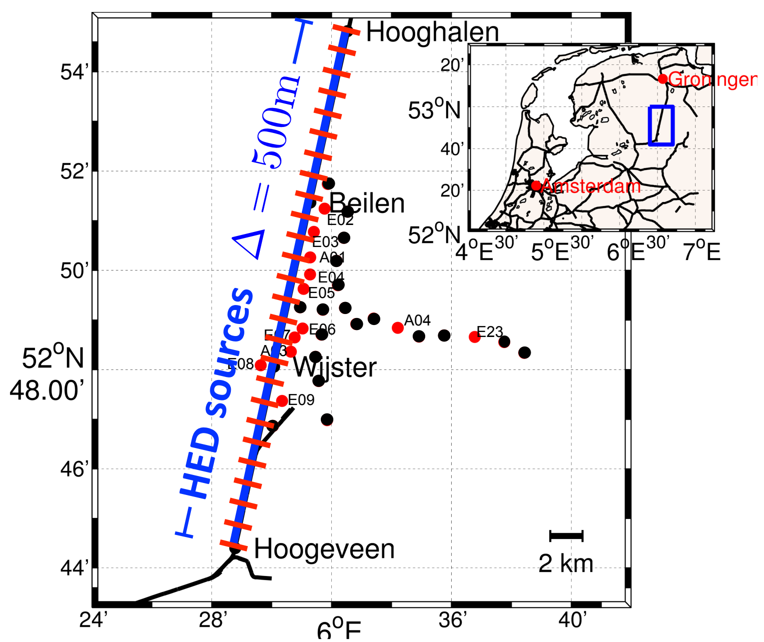
(only absolute values are shown; units of source strength are not physical.)

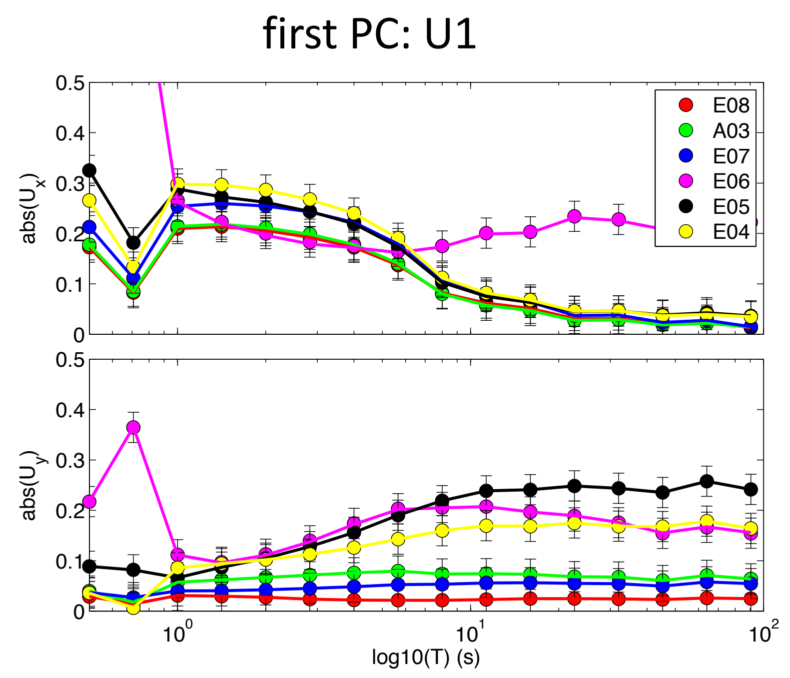


Inversion for the sources

Fit of the principal electric fields Ux and Uy for the first PC (selected stations, absolute values are shown).

$\Rightarrow$ perfect data fit (means that we have lost conductivity infos)



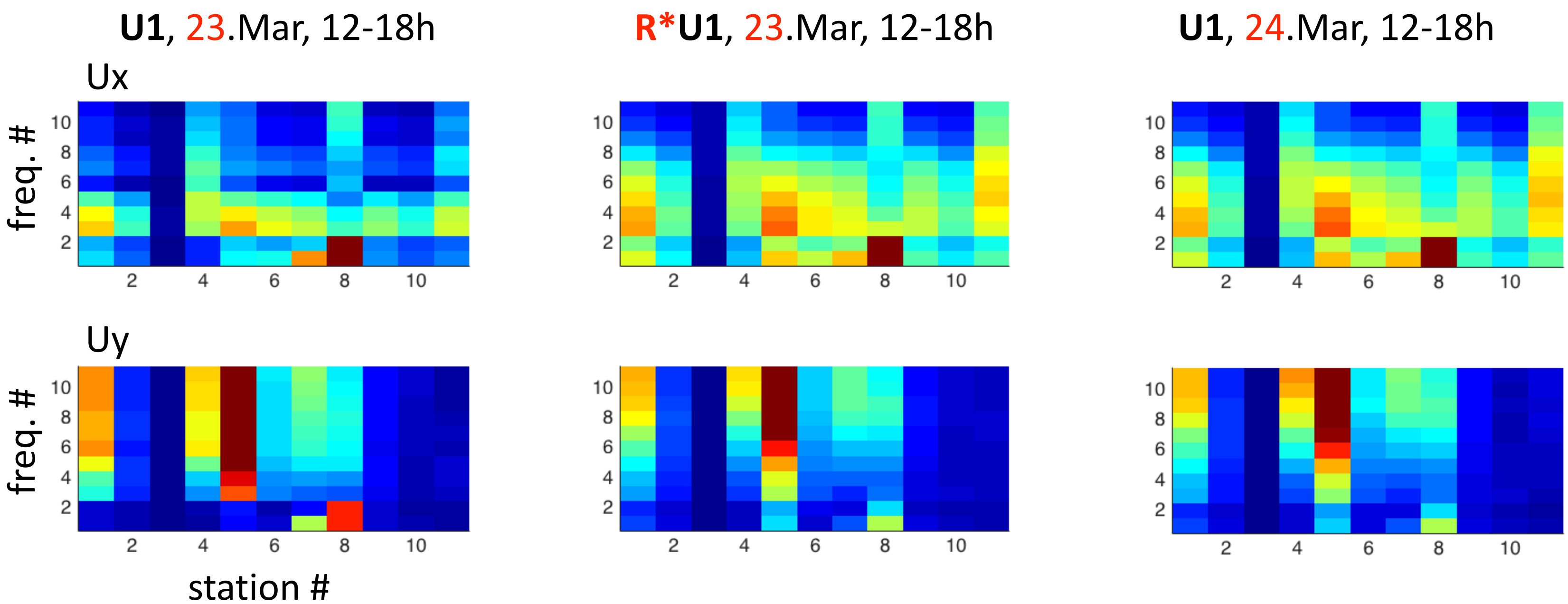


Merging subarrays

PCA is ambiguous, since for any invertible $P \times P$ matrix B, we have $\mathbf{X} = \mathbf{UA} = \mathbf{UB}^{-1}\mathbf{BA}$.

Rotation of subarrays required.

<u>Example:</u> Same array but two different days



To determine the $P \times P$ rotation matrix, <u>at least</u> P^2 overlapping channels are required!

-> not feasible for our data.