Inverse Problems in Geophysics
Part 13: Summary and outlook

2. MGPY+MGIN

Thomas Günther

thomas.guenther@geophysik.tu-freiberg.de

TU BERGAKADEMIE FREIBERG
TUBAF
Die Ressourcenuniversität.
Seit 1765.

What we learned so far

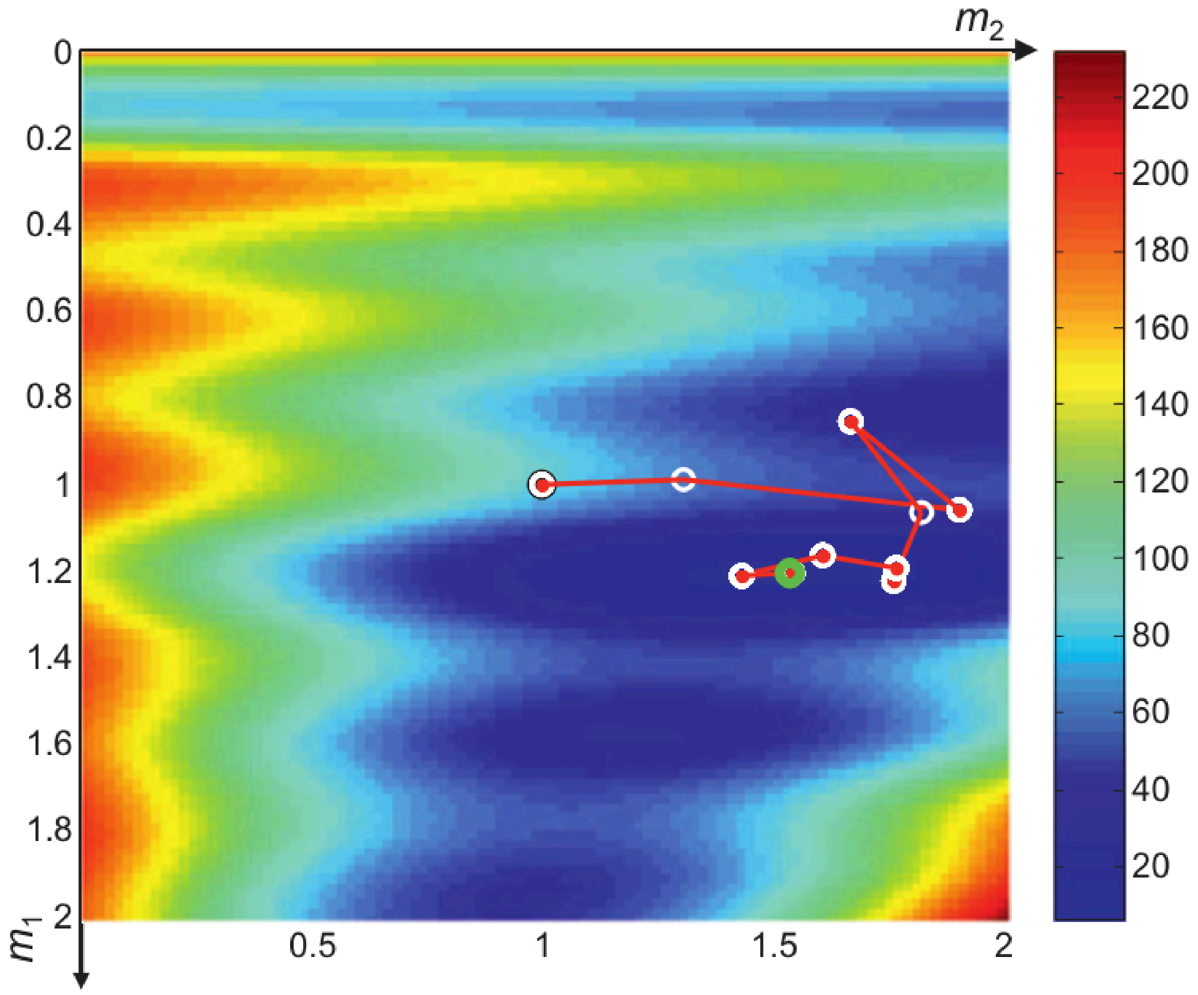
- Linear Inverse problems: Least-squares solution, Minimum-norm solution, (SVD) Pseudoinverse, truncated SVD
- Regularization: Minimum-model, smoothness constraints
- Non-linear problems: linearization by Gauss-Newton method
- Some insight into probability (related to Gaussian statistics)

Monte Carlo methods

- Monte Carlos search: randomly draw solutions from grid
- accept solution if better than old
- Markow chain

$$\{\mathbf{m}^0, \mathbf{m}^1, \ldots\}$$

  - $\mathbf{m}^{k+1}$ depends only on $\mathbf{m}^k$
- Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970)



Monte Carlo method

Recap Bayes Theorem

Bayes theorem

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}$$

- $p(\mathbf{m}|\mathbf{d})$ - posterior probability that model $\mathbf{m}$ is responsible for data
- $p(\mathbf{d}|\mathbf{m})$ - (theoretically) likelihood that $\mathbf{m}$ is capable to generate d
- $p(\mathbf{m})$ - probability that model $\mathbf{m}$ is the actual model
- $p(\mathbf{d})$ - probability of observing $\mathbf{d}$ amongst all possible

Application to constrained inversion

Gaussian likelihood function (data misfit)

$$p(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^P |D^{-1}|} \exp\left(-0.5\delta\mathbf{d}\mathbf{D}^T\mathbf{D}\delta\mathbf{d}\right)$$

Prior probability function (smoothness of the model)

$$p(\mathbf{m}) = \left(\frac{\lambda}{2\pi}\right)^{(N-1)/2} \exp\left(-\frac{\lambda}{2}\mathbf{m}^T\mathbf{C}^T\mathbf{C}\mathbf{m}\right)$$

Metropolis-Hastings algorithm

1. Burn-in phase: random walker is methodically guided toward the region of most probable method
2. Drawing of a candidate $\mathbf{m}'$ from a pdf $q(\mathbf{m}|\mathbf{m}')$ that approximates the posterior pdf (e.g., a Gaussian around the current state $\mathbf{m}$)
3. add candidate $\mathbf{m}'$ to Markov chain with acceptance probability

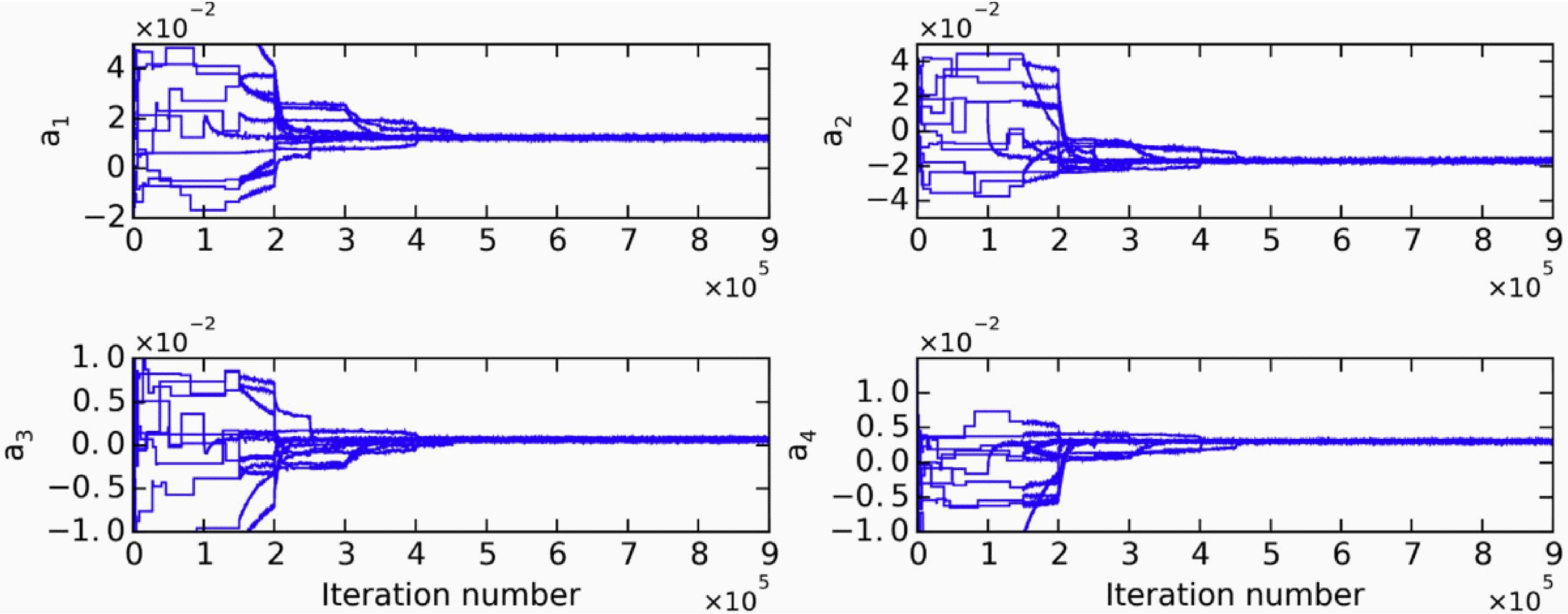
$$\alpha = \min(1, w) \quad \text{with} \quad w = \frac{p(\mathbf{m}') * q(\mathbf{m}|\mathbf{m}')}{p(\mathbf{m}) * q(\mathbf{m}'|\mathbf{m})}$$

using a random variable u in the interval [0, 1] $\Rightarrow$ accept if $u > w$

Key components

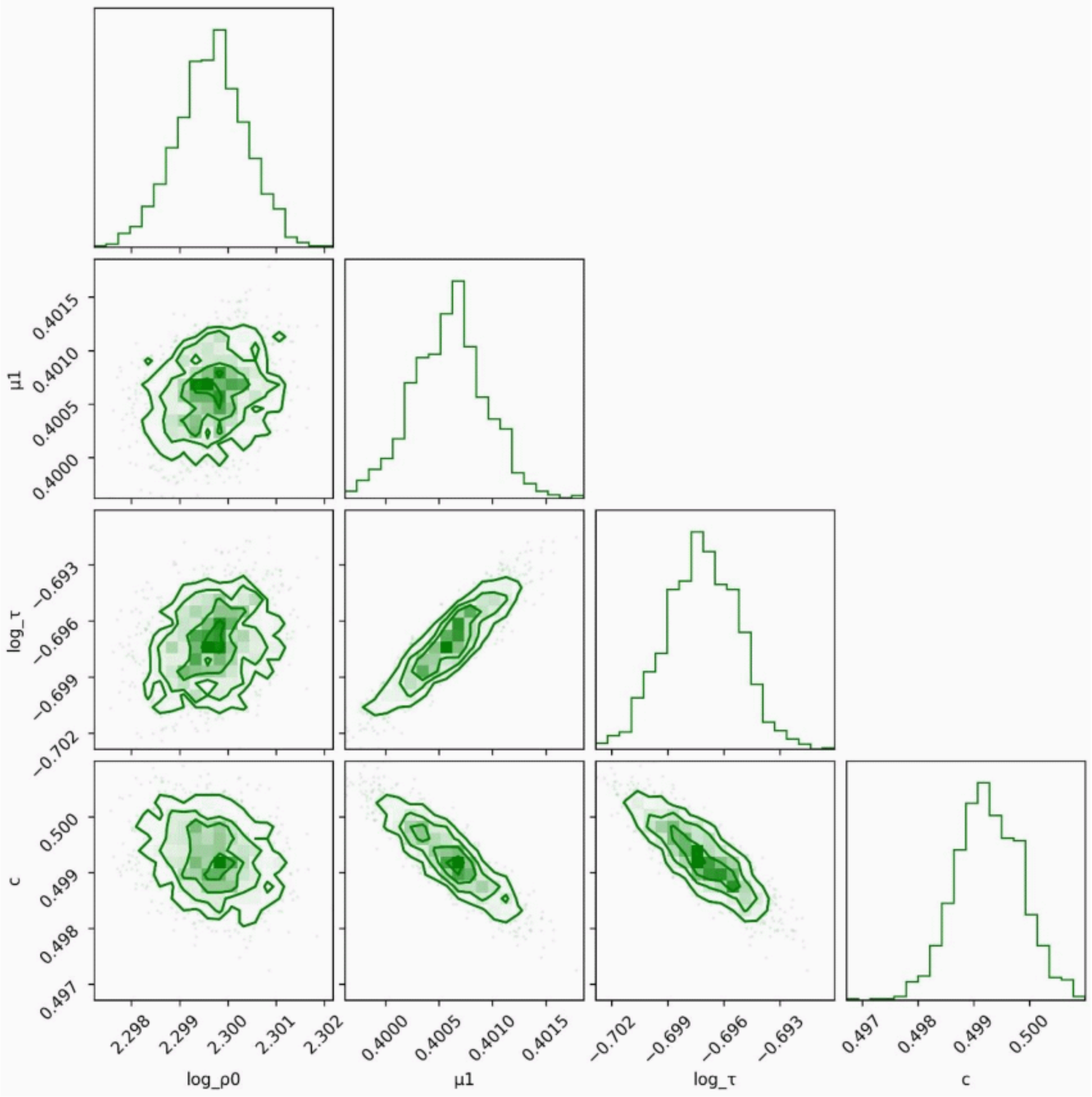
- **Proposal Distribution** (q): Crucial for efficiency. A good proposal distribution explores the space well without being rejected too often.
- **Acceptance Rate**: Good value (20-50%) indicates balance between exploration & exploitation. Too low = slow. Too high = stuck locally.
- **Burn-in Period**: The initial samples often discarded because the chain hasn't converged to the stationary distribution yet.
- **Thinning**: To reduce autocorrelation between samples, you can keep only every nth sample.
- **Convergence Diagnostics**: Check if chain has actually converged to the target distribution. (e.g., Gelman-Rubin statistic, trace plots).

Trace plots



Berube et al. (2017)

Corner plots



(Roudsari et al., in rev.)

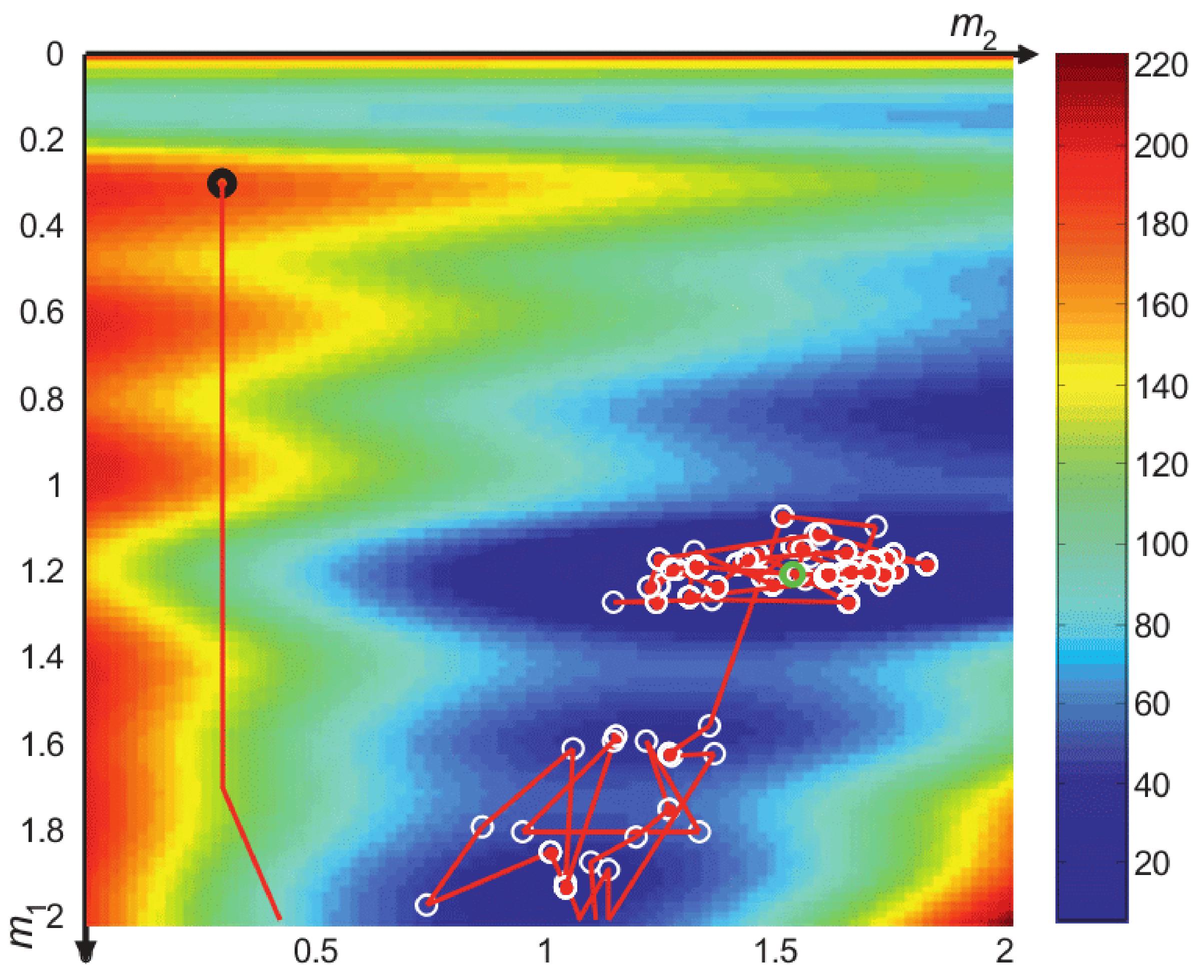
- statistics of models
- correlation of parameters
- uncertainty of the results

Simulated Annealing

Gibbs (Boltzmann) sampling

$$P(E) = e^{-\frac{E}{k_B T}} = e^{-(\Phi(\mathbf{m})-\Phi(\mathbf{m}^p))/T}$$

- draw models neighboring best fit
- large movements if far from target fit
- small if approaching target

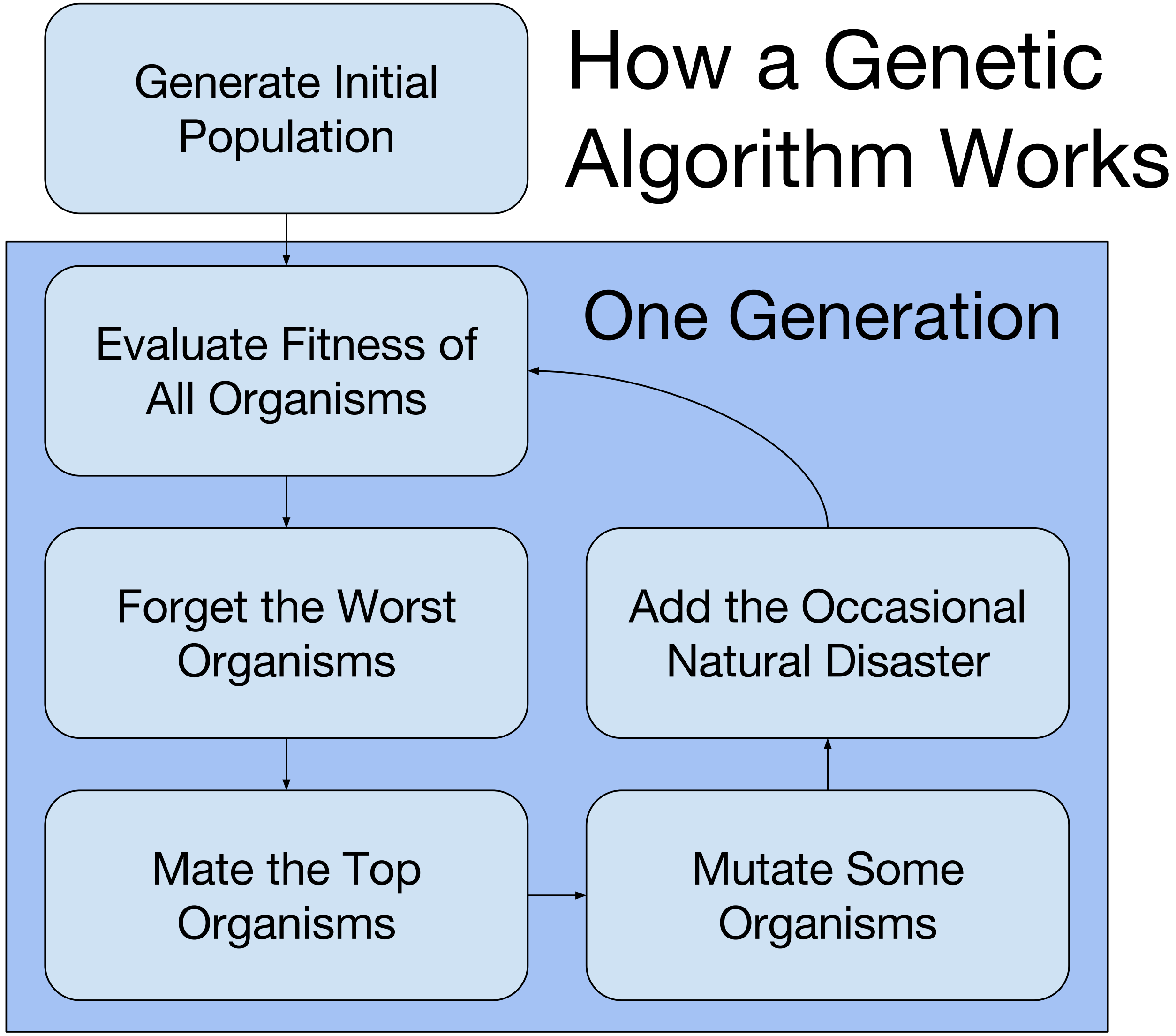


Simulated Annealing

Bio-inspired global search methods

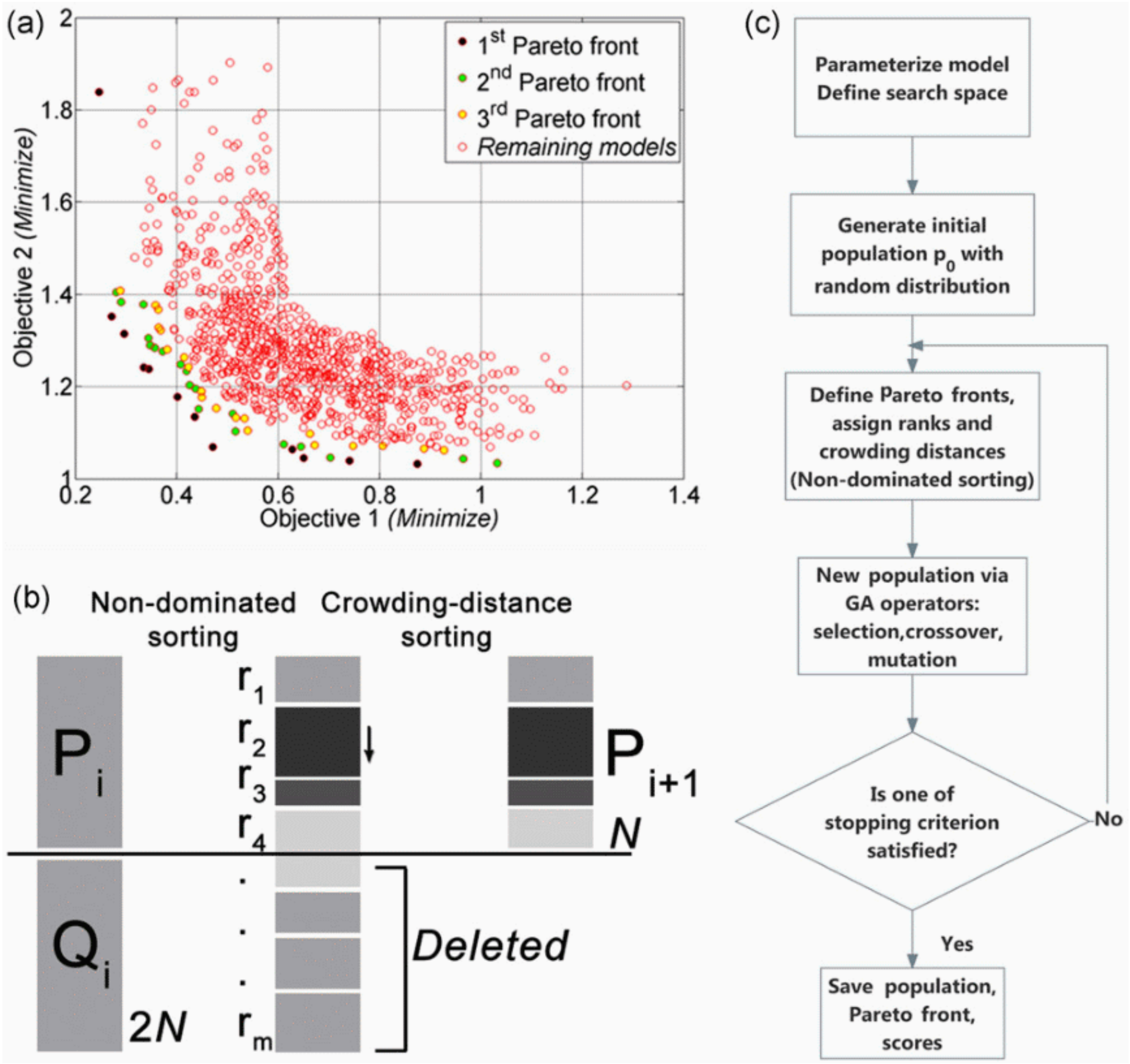
- Genetic Algorithm
- Evolution Strategy
- Differential Evolution Algorithm
- Estimation of Distribution Algorithm
- Particle Swarm Optimization
- Ant Colony Optimization
- Pareto Archived Evolution Strategy (PAES)
- Nondominated Sorting Genetic Algorithm (NSGA-II)

Genetic algorithms



- inspired by evolution of species (*survival of the fittest*)
- code models in a binary (DNA) sequence
- randomly generate initial population
- let the fittest (data fit!) survive & mate and the unfittest die
- mutation of the DNA to create variability

Nondominated Sorting Genetic Algorithm

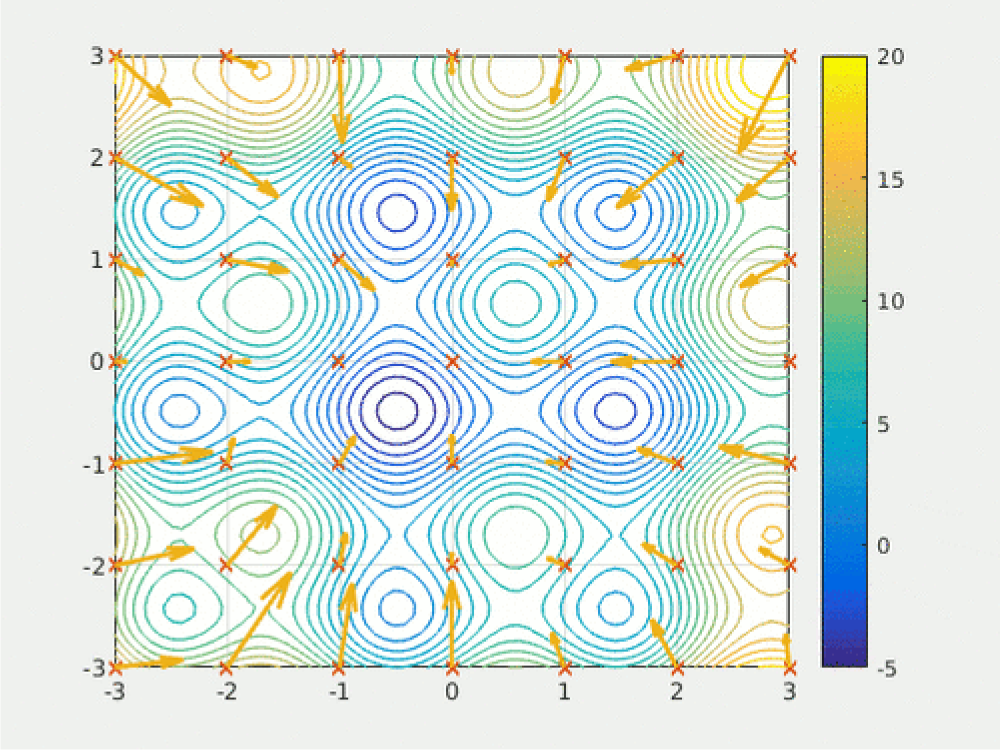


- minimization of two objective functions
  - two geophysical methods
  - data misfit and model roughness
- Pareto front: all non-dominated individuals
- sort Pareto rank and play god

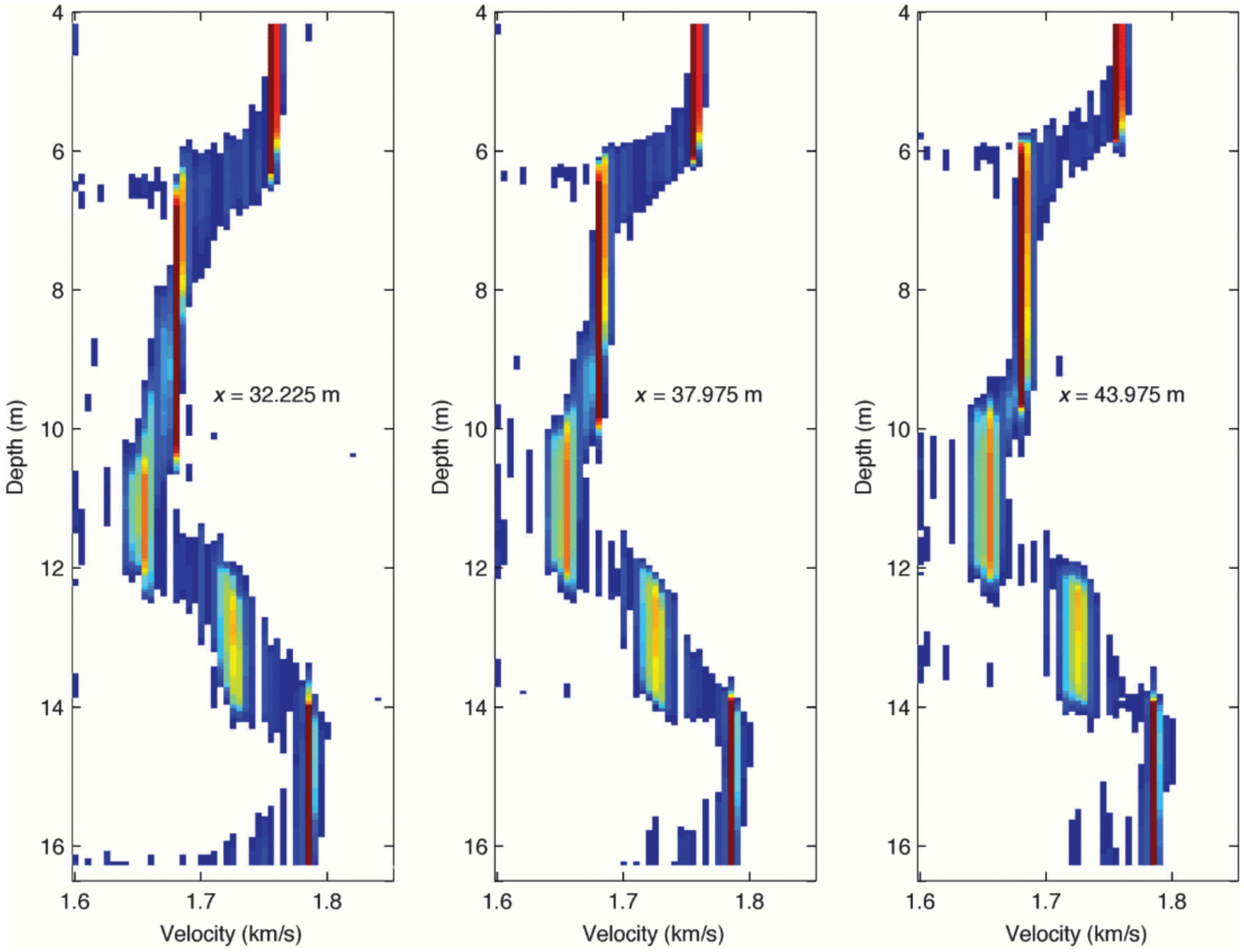
Particle swarm optimization

- movements of individuals in swarm (impulse & attraction)
- individual position ($\mathbf{m}$) & velocity

- distribute population over range
- acceleration from attraction
- move with velocity



Model variance



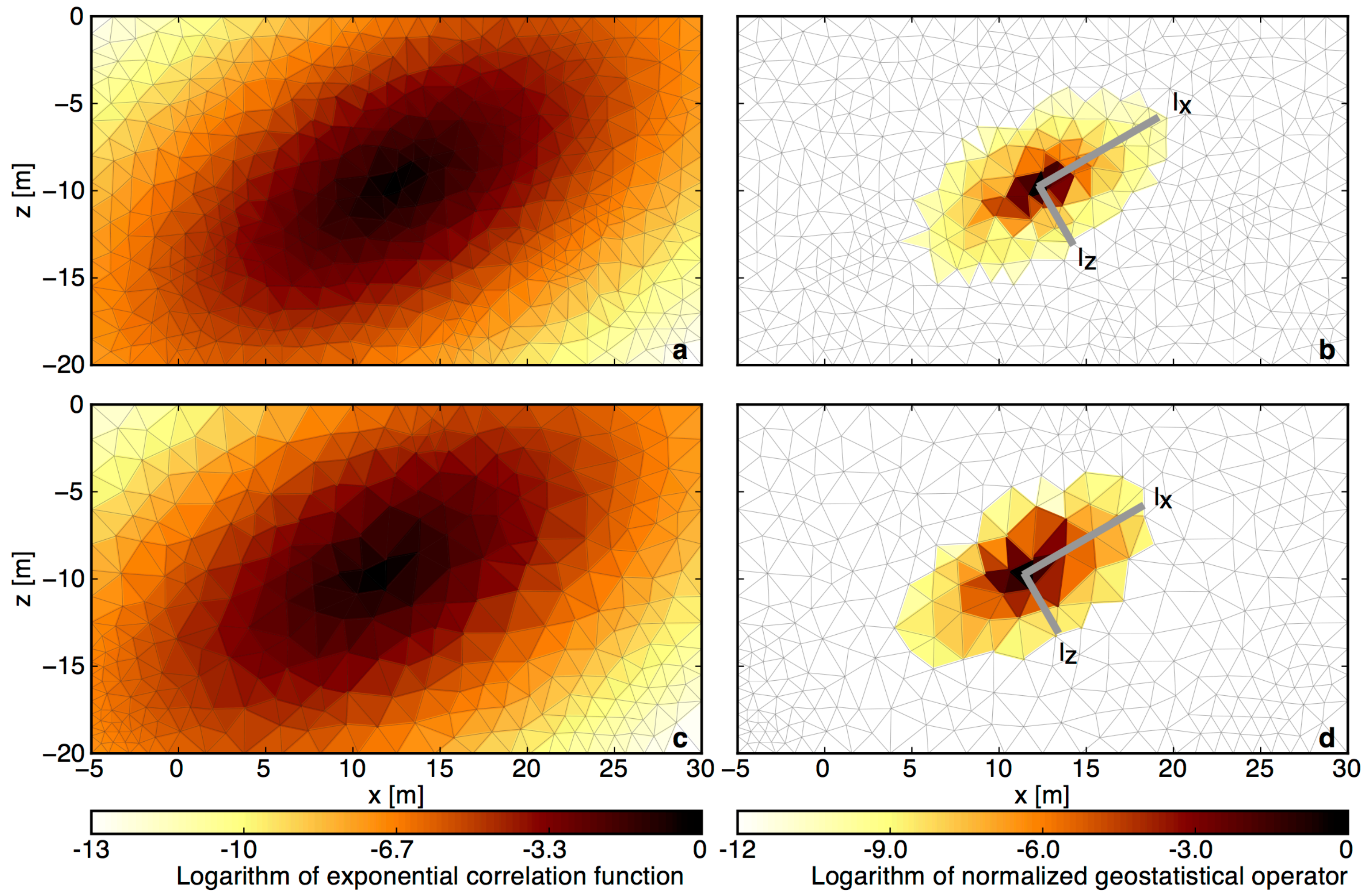
- derivation of mean values and standard deviations
- statistic simulations

(Tronicke et al., 2012)

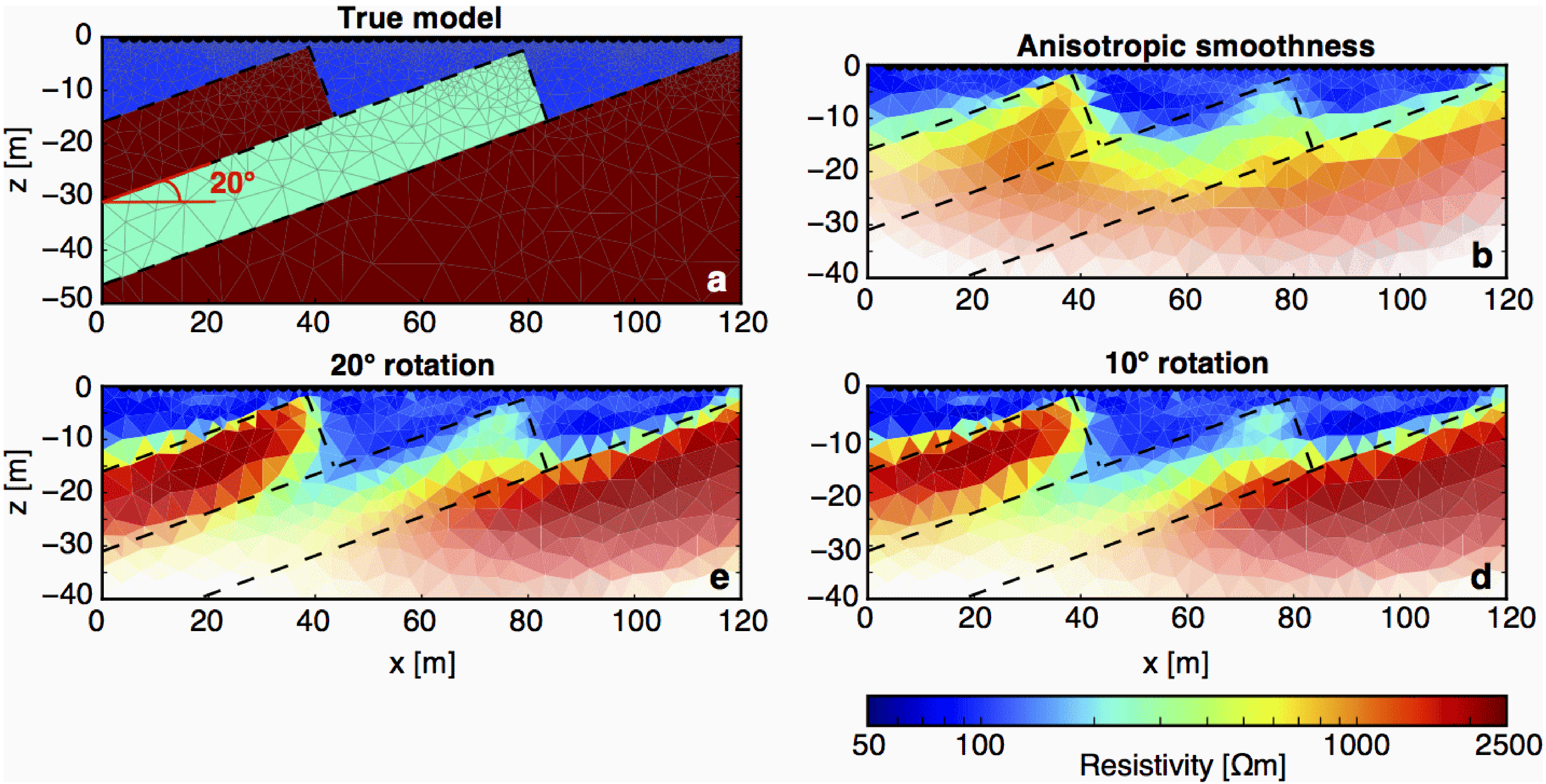
Modern regularization

- L1 minimization & minimum (gradient) support
- geostatistical regularization
- region-wise regularization
- prior and structural constraints
- structurally coupled cooperative inversion
- cross gradients joint inversion

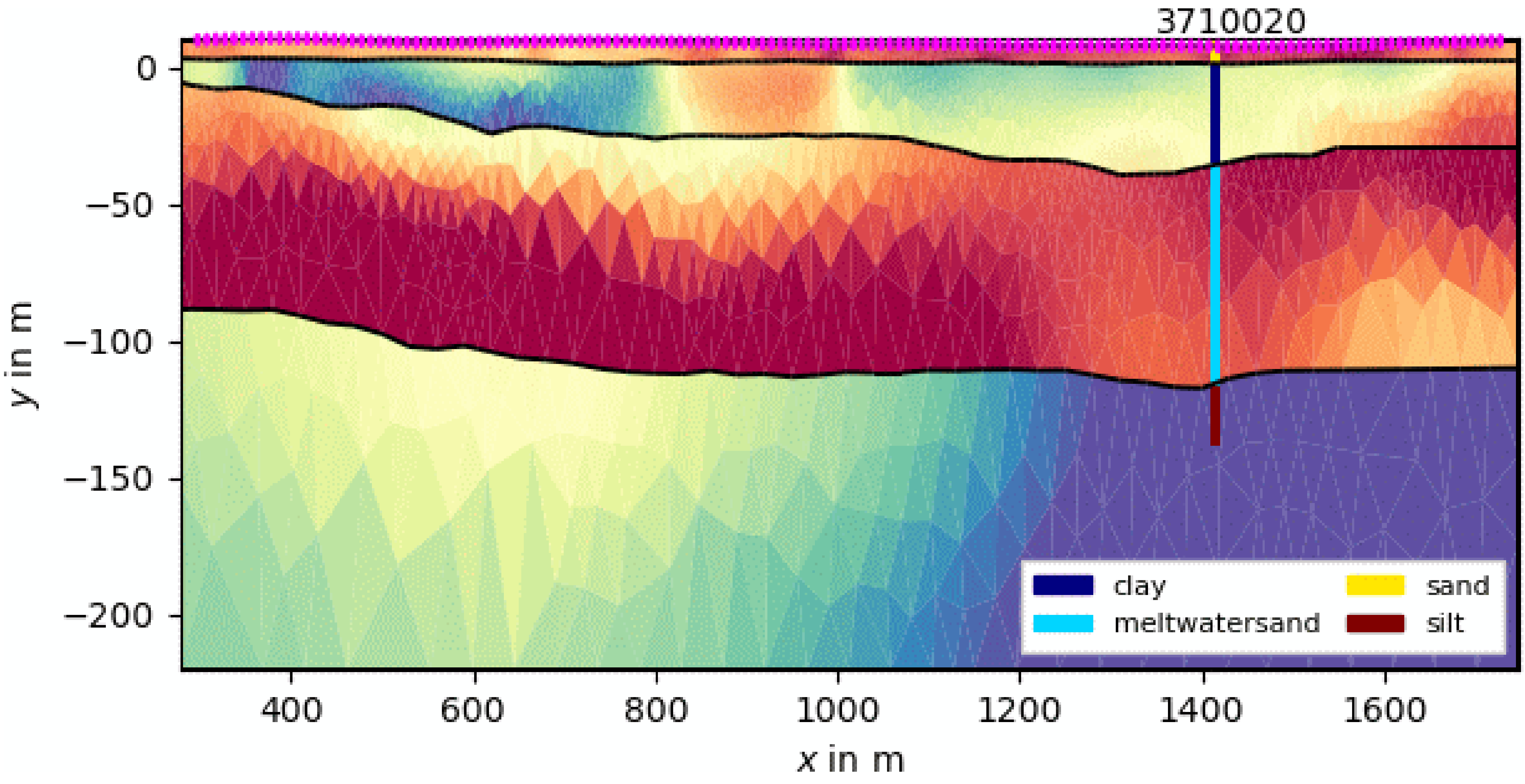
Geostatistical regularization (Jordi et al. 2018)



Geostatistical regularization (Jordi et al. 2018)



Region-wise settings



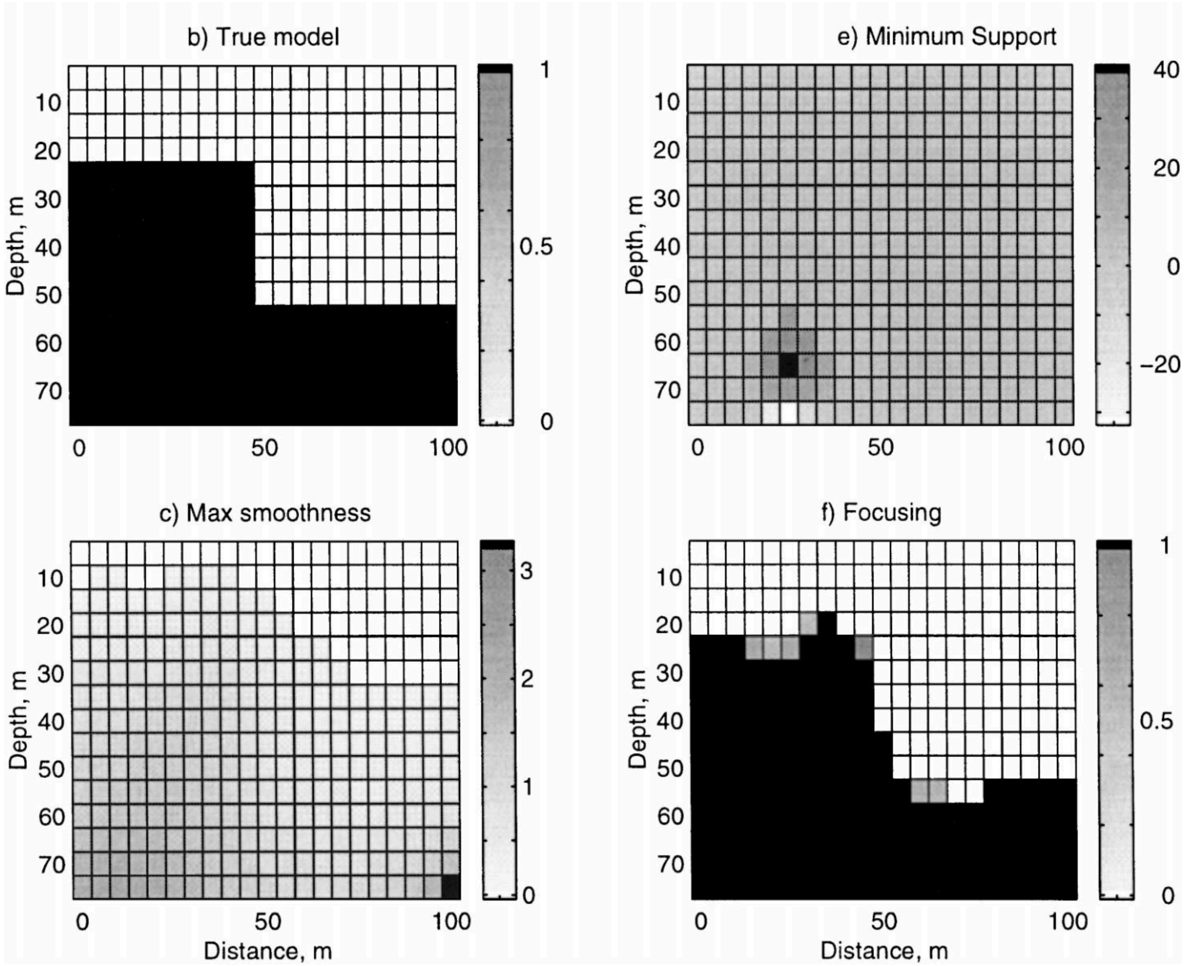
Alei et al. (2025)

Focusing inversion

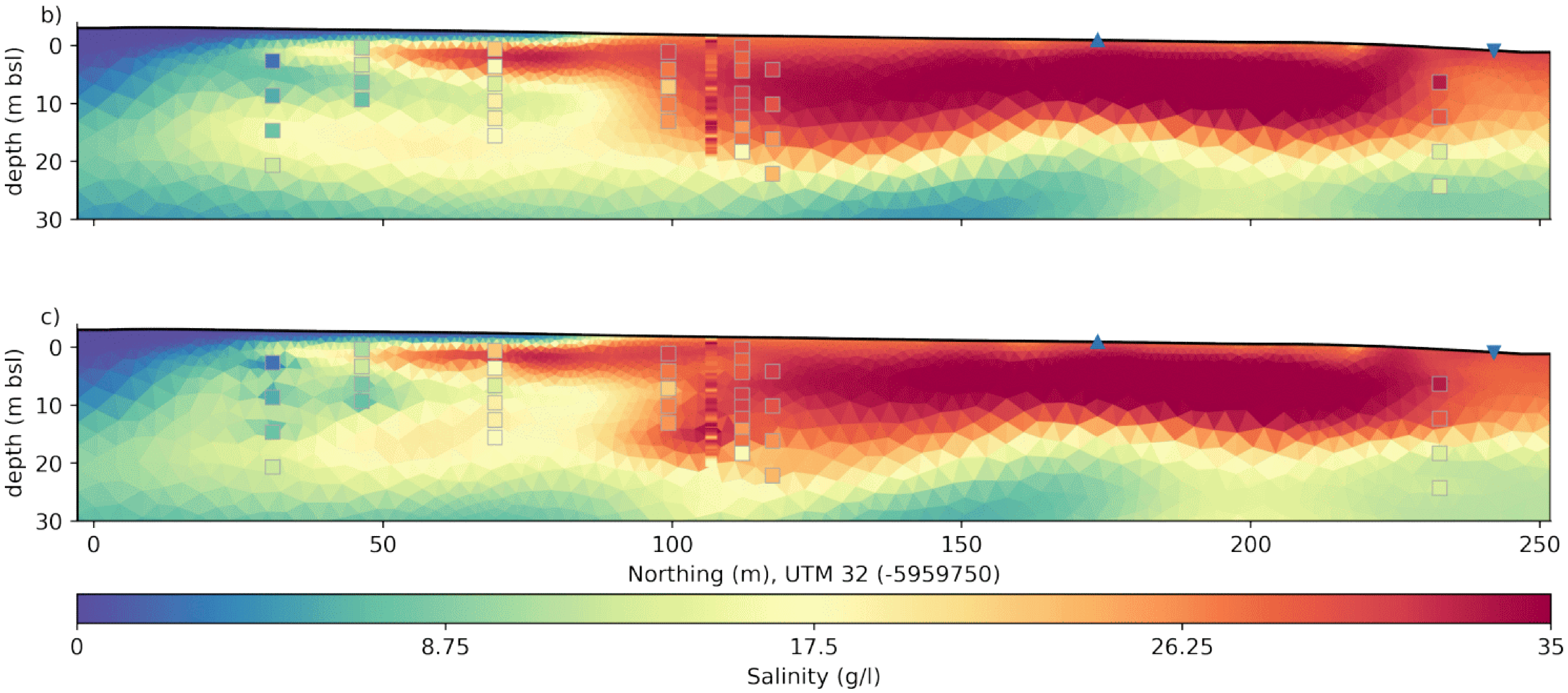
- L_1 norm (robust) minimization of the model roughness: iterative reweighting leads to sharper contrasts
- minimum gradient support:

$$\Phi_m = \int \frac{\nabla m \cdot \nabla m}{\nabla m \cdot \nabla m + \beta^2}$$

number of contrasts get penalized

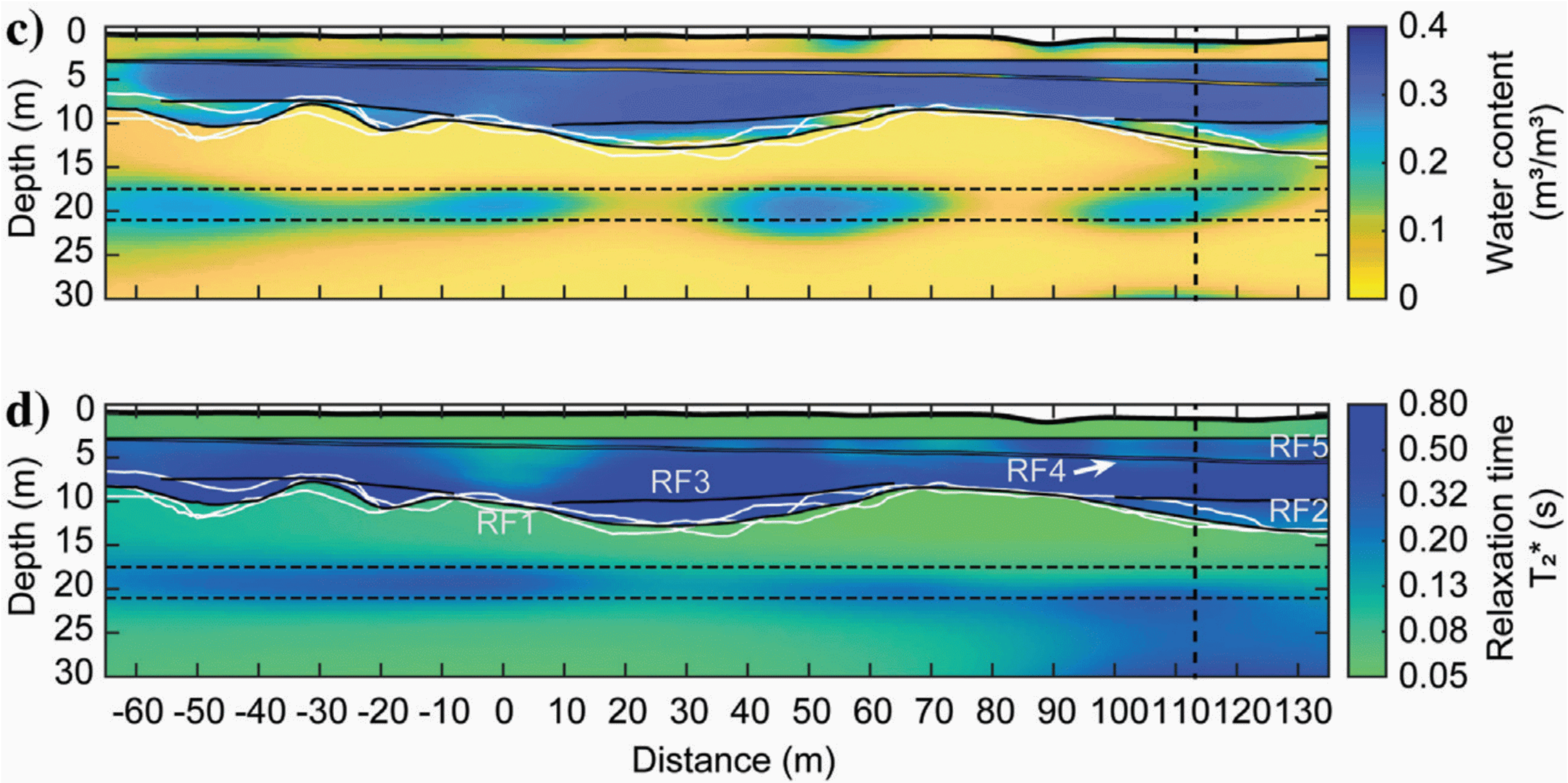


Inversion with prior data

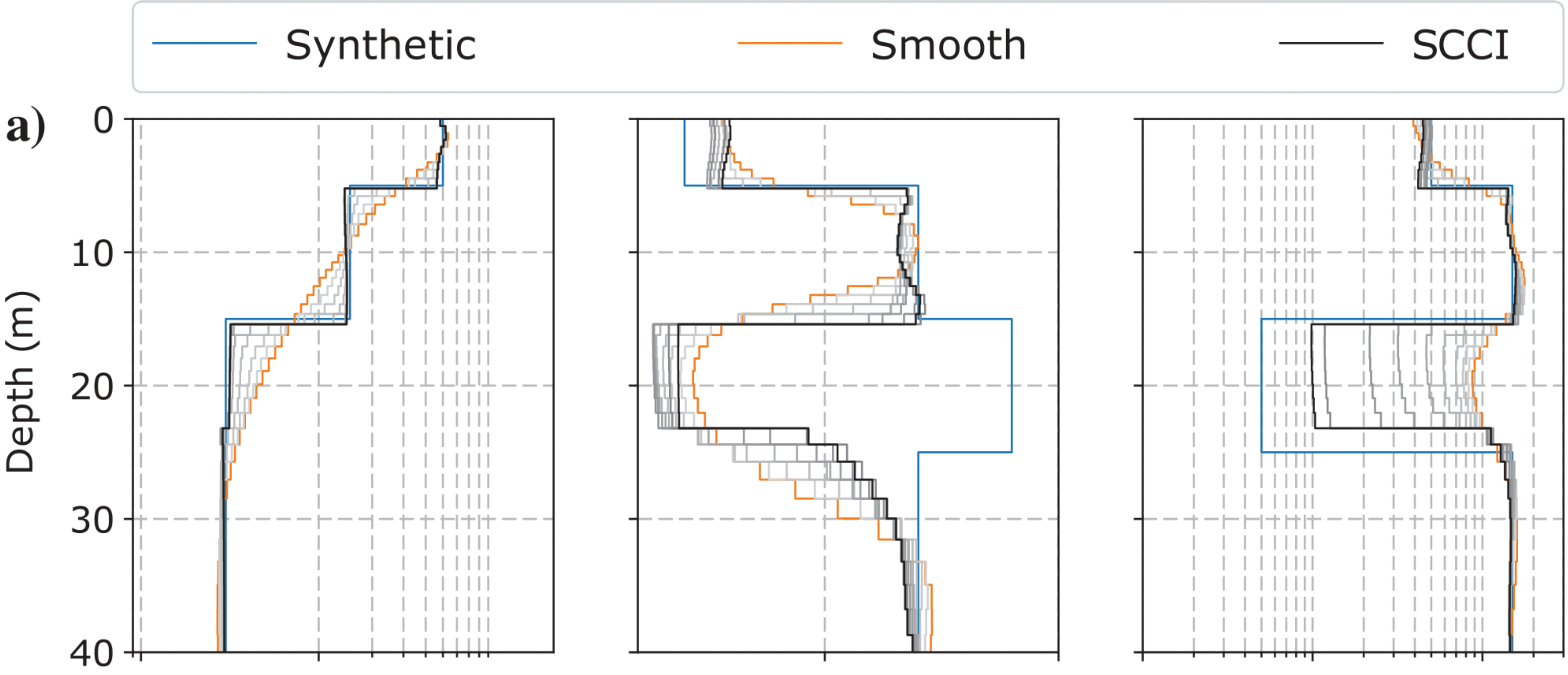


Structurally constrained inversion

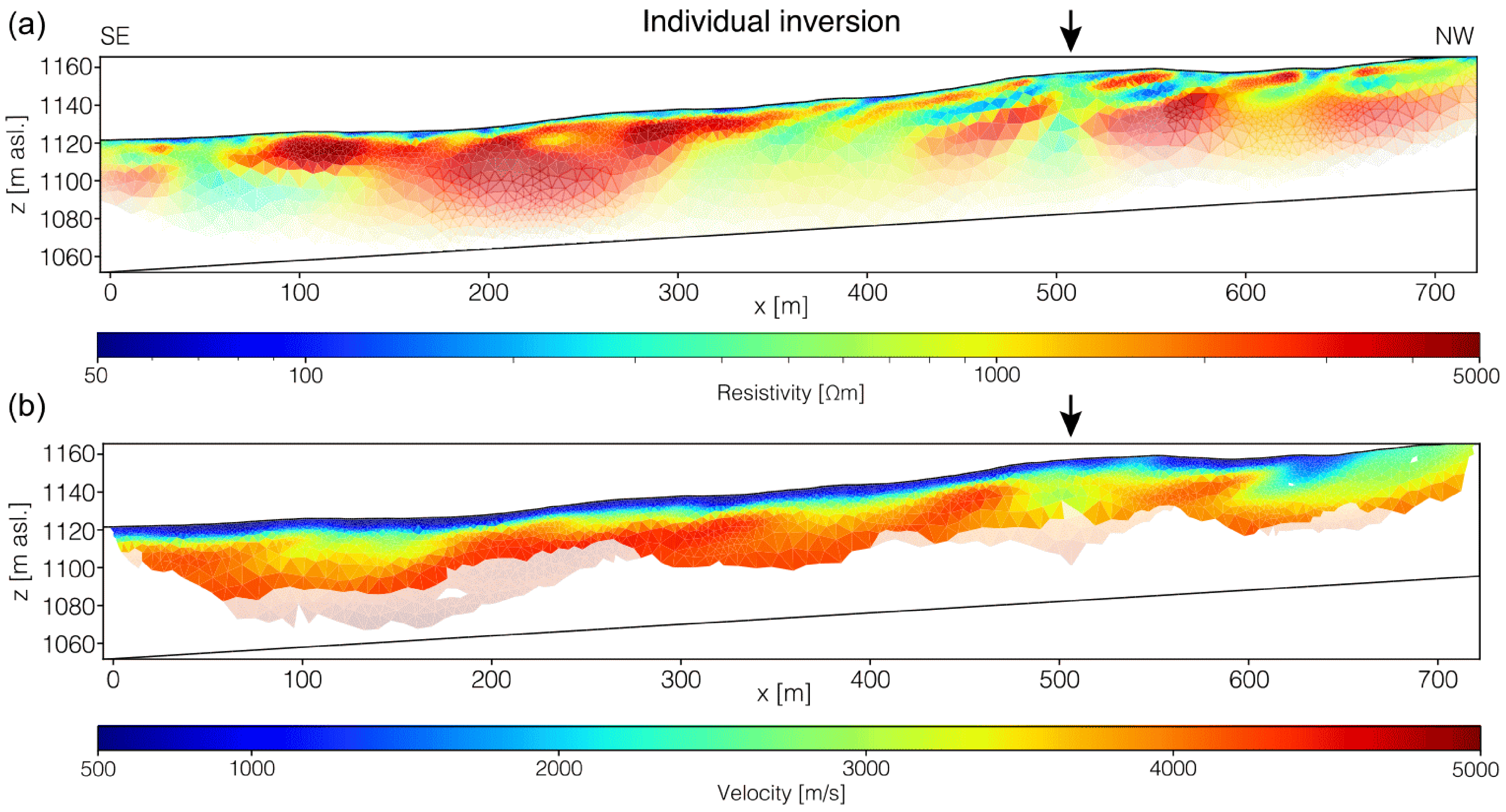
decouple smoothness at boundaries (seismics, GPR, boreholes)



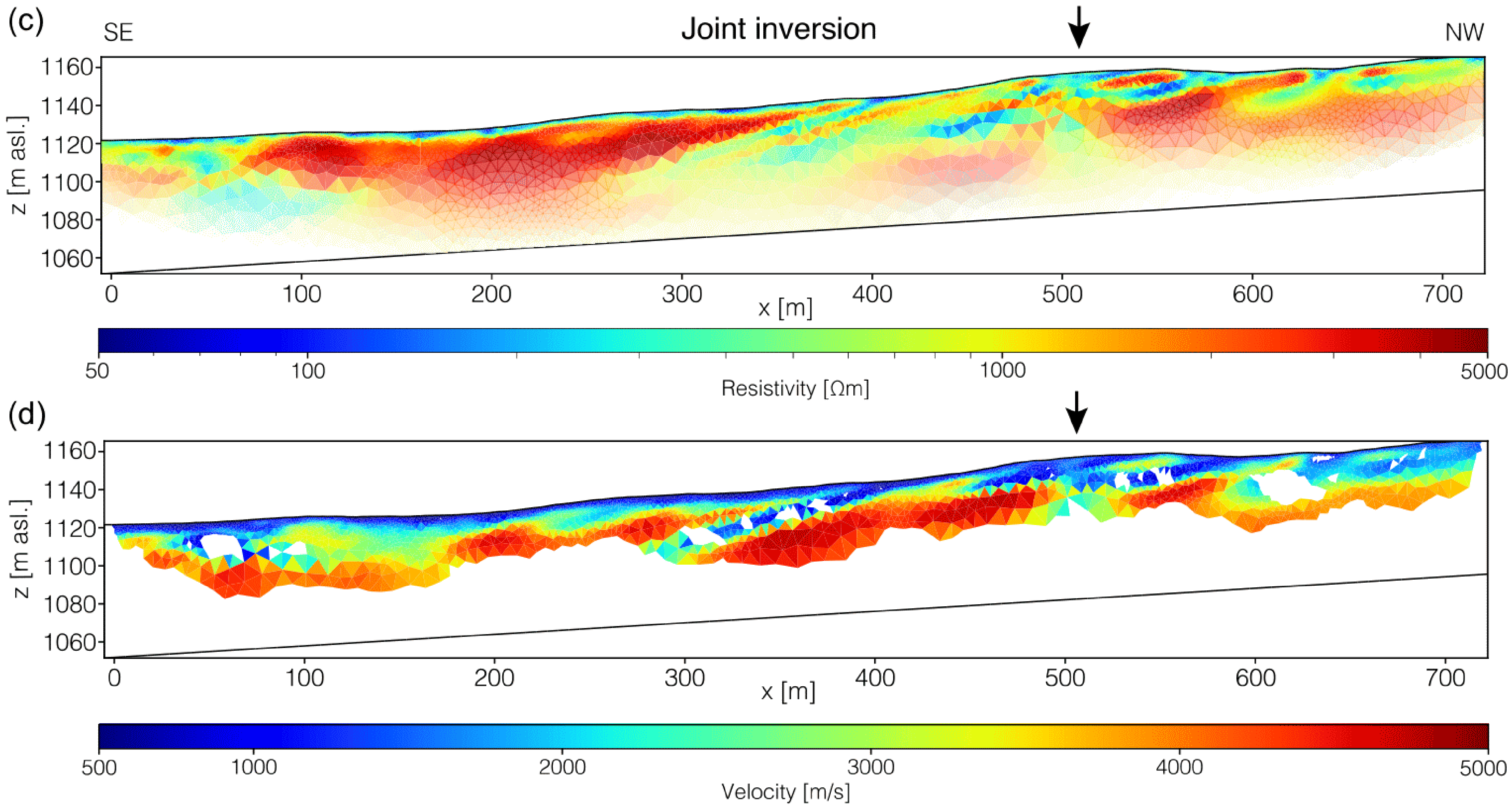
Structurally coupled cooperative inversion



Joint inversion with cross gradients



Joint inversion with cross gradients



Summary

Regularization is key to subsurface imaging

Add as much information as you have to inversion