COBB

"Carbon Monoxide (CO) from Biomass Burning (BB):
Estimating the source and variability
using SCIAMACHY and MOPITT measurements"

Pim Hooghiemstra
Maarten Krol

Outline

- Scientific relevance of COBB
- Objectives and how-to-get-there?
- Short introduction to data assimilation
- Current study (Toy model)
- Conlusions & Future research

Scientific Relevance

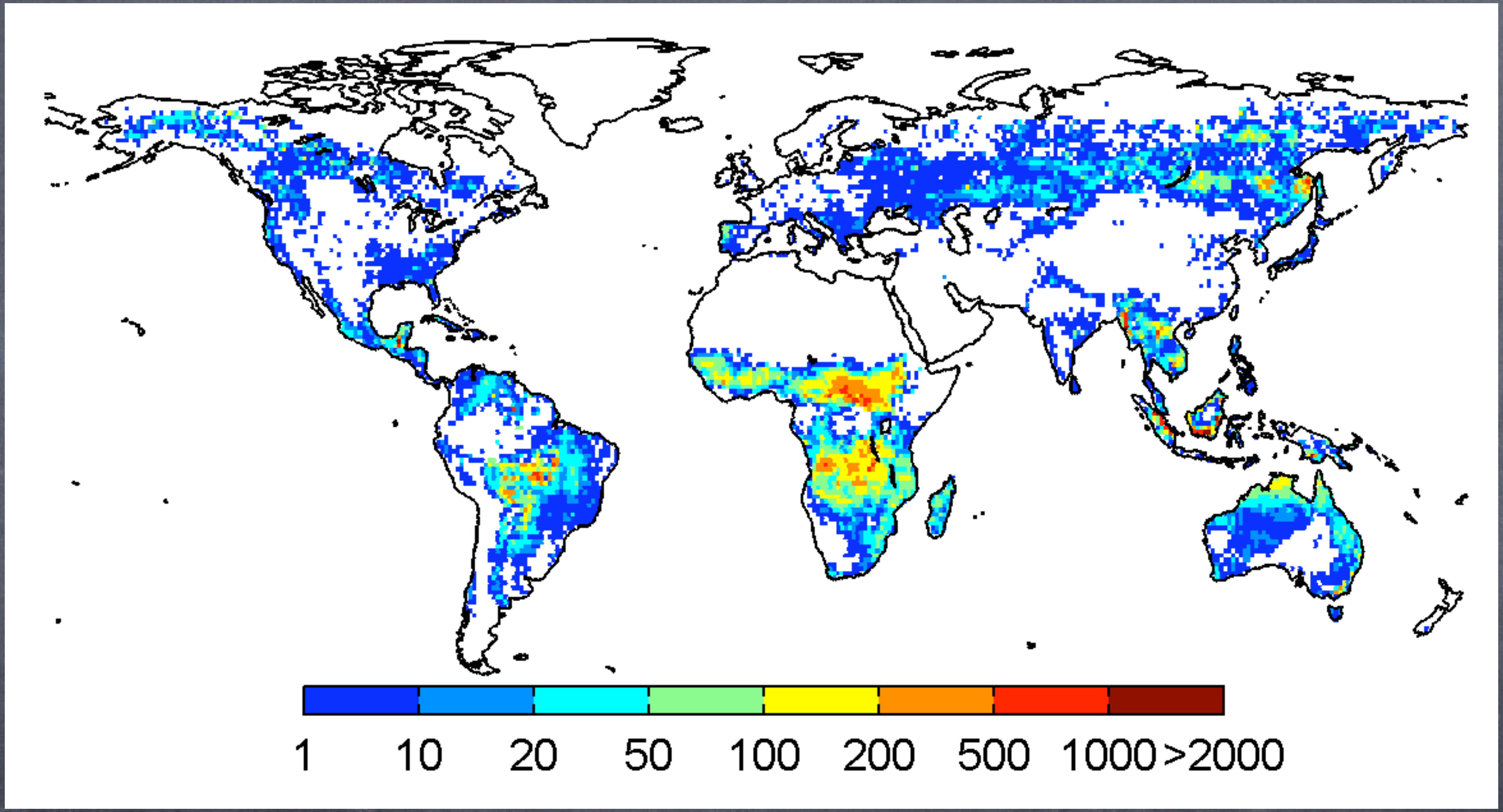
Scientific Relevance

- What is biomass burning?

Scientific Relevance

- What is biomass burning?
- Locating BB from space

Scientific Relevance



g CO/m^2 year

Scientific Relevance

- What is biomass burning?
- Locating BB from space
- BB $\longrightarrow$ trace gas emissions?

Objectives

Estimate the **source** and **variability** of Biomass burning CO

How-to-get-there?

4D-VAR Data Assimilation

background estimate + observations = analysis

How-to-get-there?

4D-VAR Data Assimilation

background estimate + observations = analysis

↓

Construct background estimate from inventory from literature

Emission Budget	Yearly Emissions (Tg CO)	Uncertainty
Fossil+Biofuel	571	35%-60%
Tropical Fires	170	70%
Savanna Fires	268	70%
Extratropical Fires	29	70%
Biogenic	160	60%
Oceans	20	-
Oxidation NMVOCs	734	60%
Oxidation Methane	796	-
TOTAL	2748	-

How-to-get-there?

4D-VAR Data Assimilation

background estimate + observations = analysis

↙

- Surface network observations

How-to-get-there?

4D-VAR Data Assimilation

background estimate + observations = analysis

↙

- Surface network observations
- Satellite measurements
  - ⋆ SCIAMACHY
  - ⋆ MOPITT

Data Assimilation

Ingredients:

- Observations $\mathbf{y}$
- Model $\mathbf{H}$
- a priori estimate $\mathbf{x}_b$

Data Assimilation

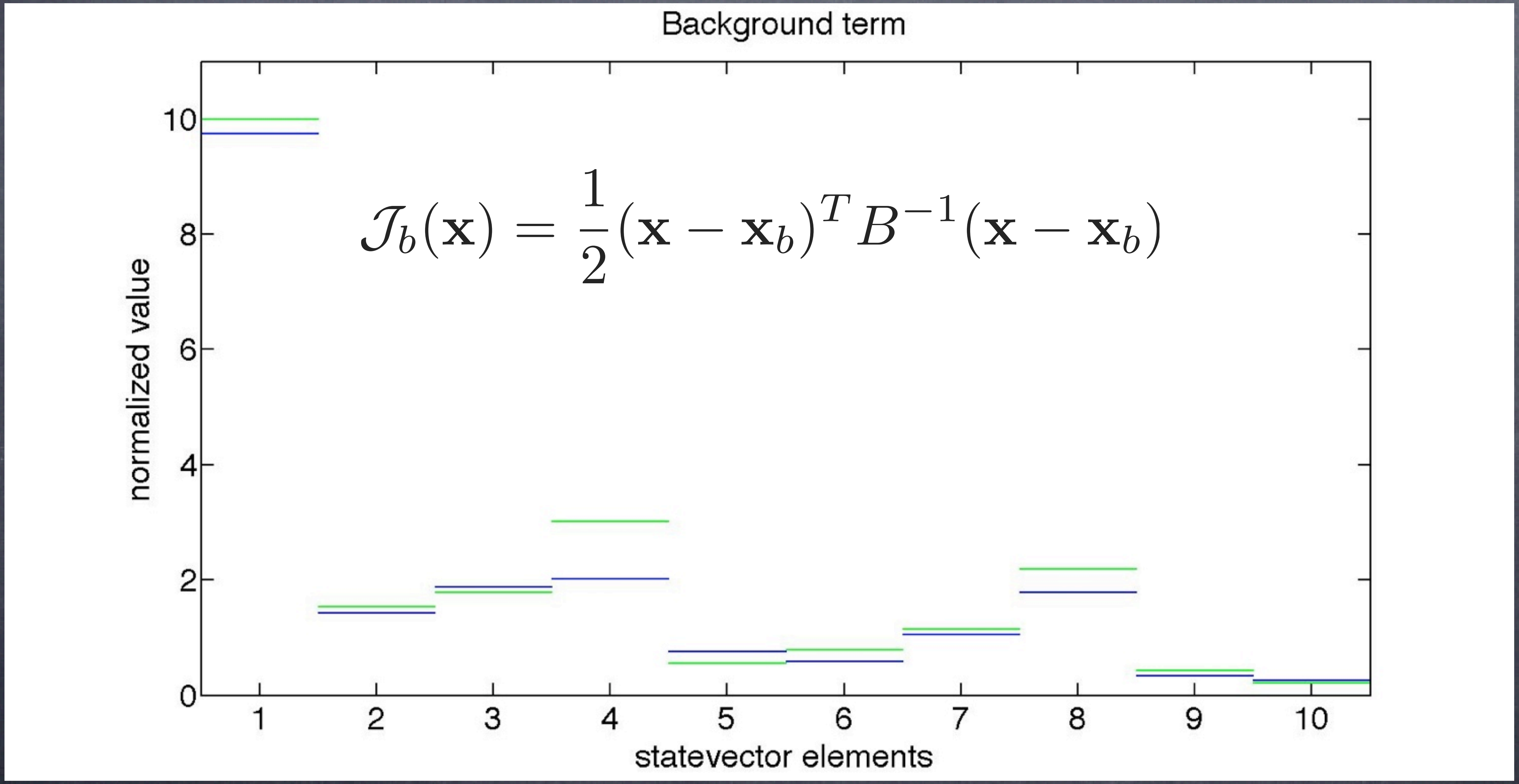
Ingredients:

- Observations **y**
- Model **H**
- a priori estimate $\mathbf{x_b}$

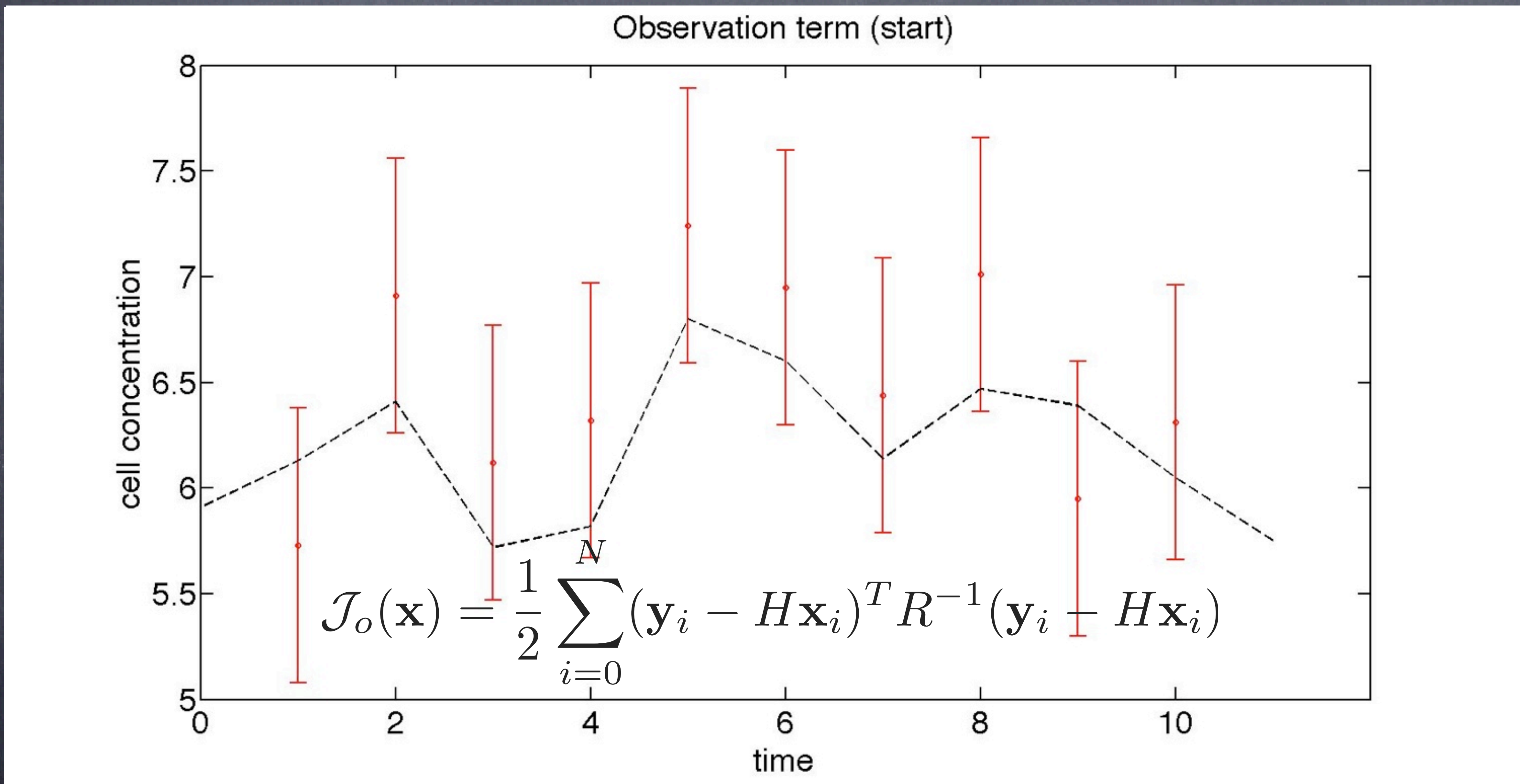
Objective: Find minimum of

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T B^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

Data Assimilation

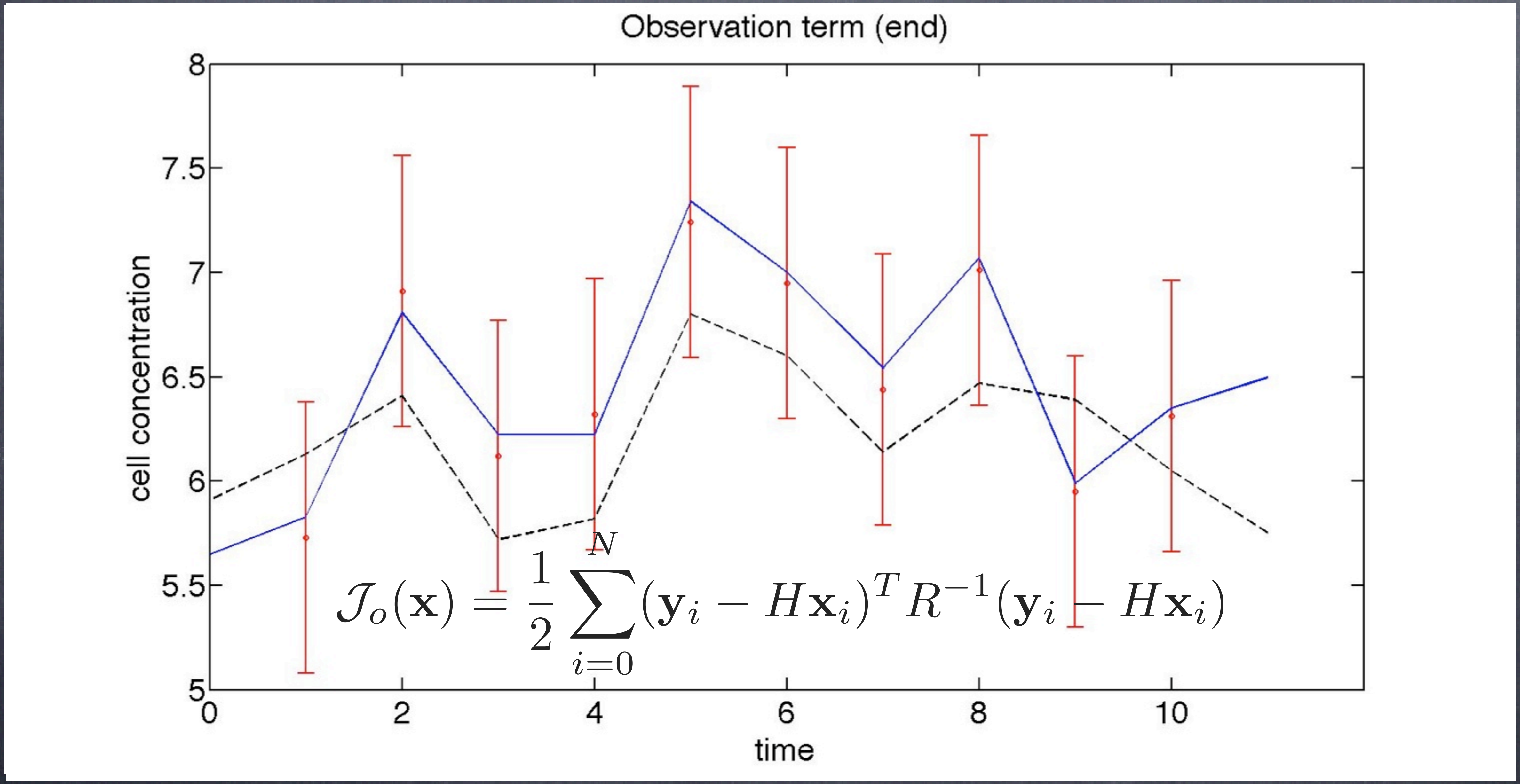


Data Assimilation



$$\mathcal{J}_o(\mathbf{x}) = \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

Data Assimilation



$$\mathcal{J}_o(\mathbf{x}) = \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

Minimization routine

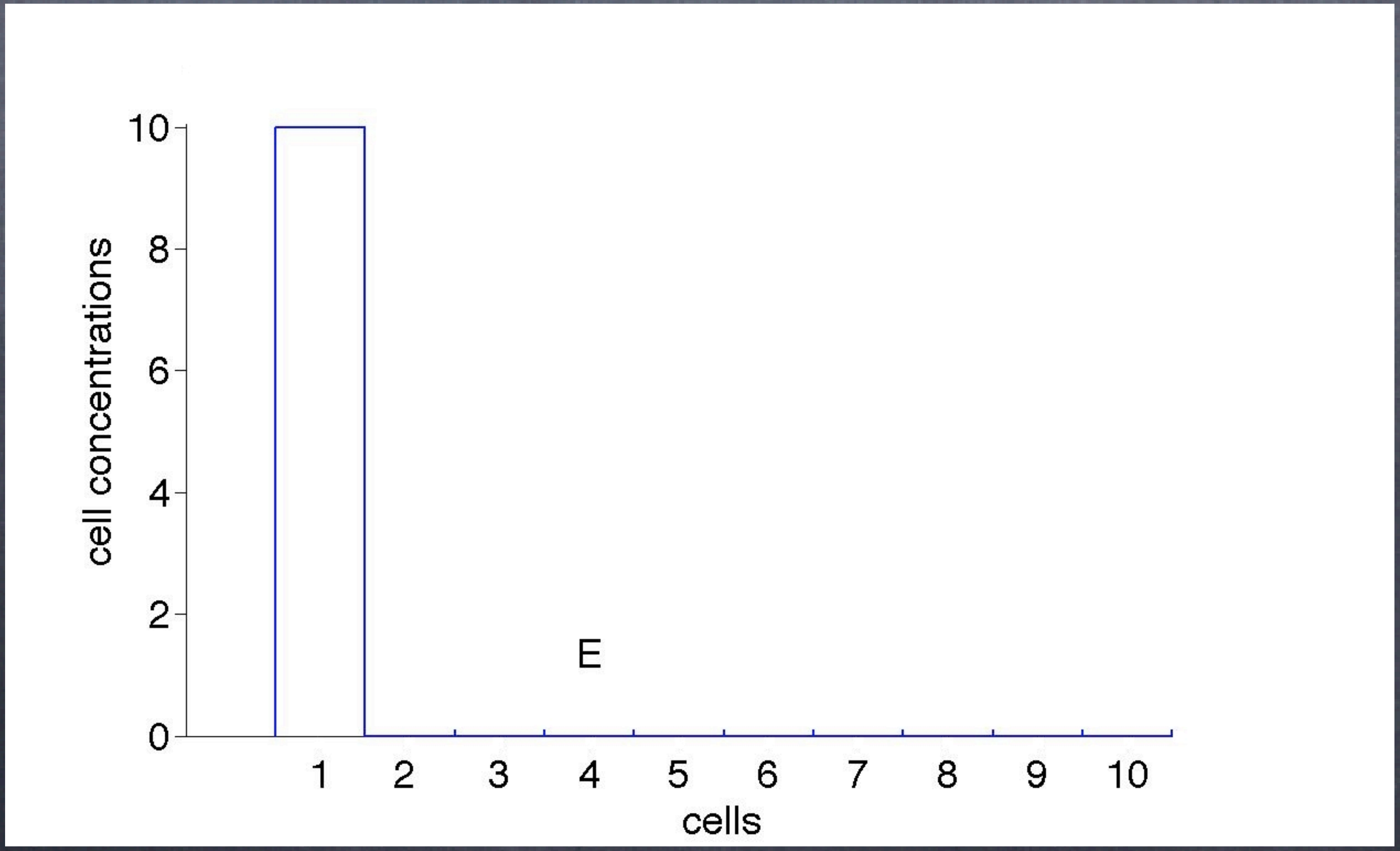
Conjugate Gradient method (CONGRAD):

- Based on Lanczos algorithm for approximating the leading eigenpairs of Hessian of cost function **J**.

- Simultaneously minimization of **J** and estimation of leading eigenpairs of Hessian.
- A posteriori error covariance matrix $\mathbf{A}_a$ derived from: $\mathbf{A}_a \approx \sum_{k=0}^{M} \frac{1}{\lambda_k} \mathbf{v}_k \mathbf{v}_k^T$

Toy model (1)

- 1D domain with 10 cells
- Initial tracer concentration
- Emission in a predefined cell
- constant windspeed from left to right



Toy model (2)

- Underlying equation of model is the advection-emission equation:

$$\frac{\partial q}{\partial t} + \bar{u}\frac{\partial q}{\partial x} = E$$

- Use finite volume method with upwind discretization of fluxes

Toy model (2)

- Underlying equation of model is the advection-emission equation:

$$\frac{\partial q}{\partial t} + \bar{u}\frac{\partial q}{\partial x} = E$$

- Use finite volume method with upwind discretization of fluxes

- Simulation movie!

Toy model (3)

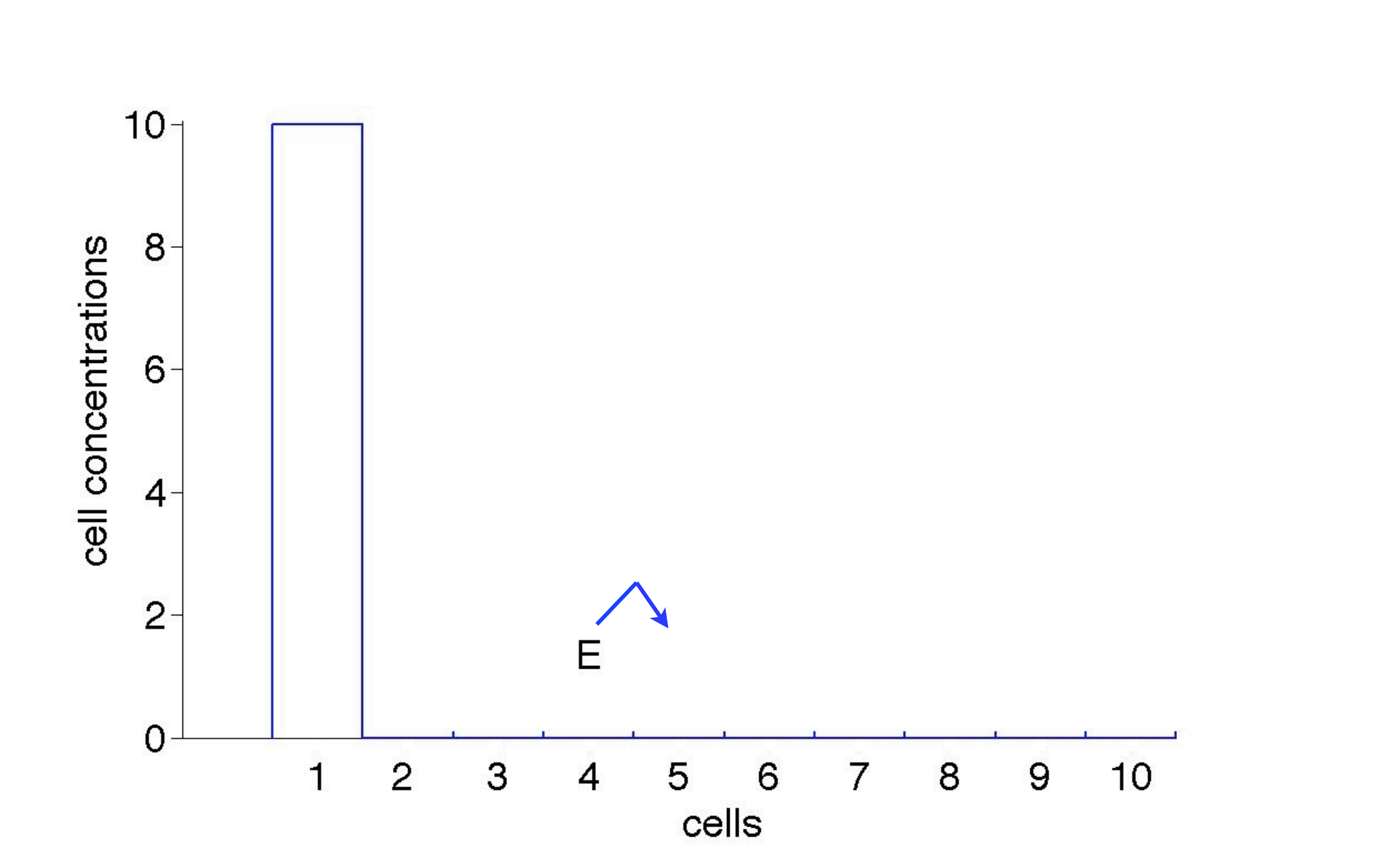
Question: Assume that **initial tracer concentration** distribution and the **emissions** are unknown. Using the model and observations in an (inverse) data assimilation scheme, how good is the approximation of the initial state?

We use the **BLUE** analysis as data assimilation method.
This is equivalent to minimization of cost function J.

Choices

- # measurement times
- # measurement stations
- background estimate $\mathbf{x}_b$
- error covariance matrices **B** and **R**

Choosing background



concentration profile equal to initial, background emission is placed in cell 5

Many stations & observations

- # measurement times = 101
- # measurement stations = 10
- backgr. conc. error = 100%
- backgr. emis. error = 200%
- R diag. matrix, entries = 0.1

Cost function value

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T B^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

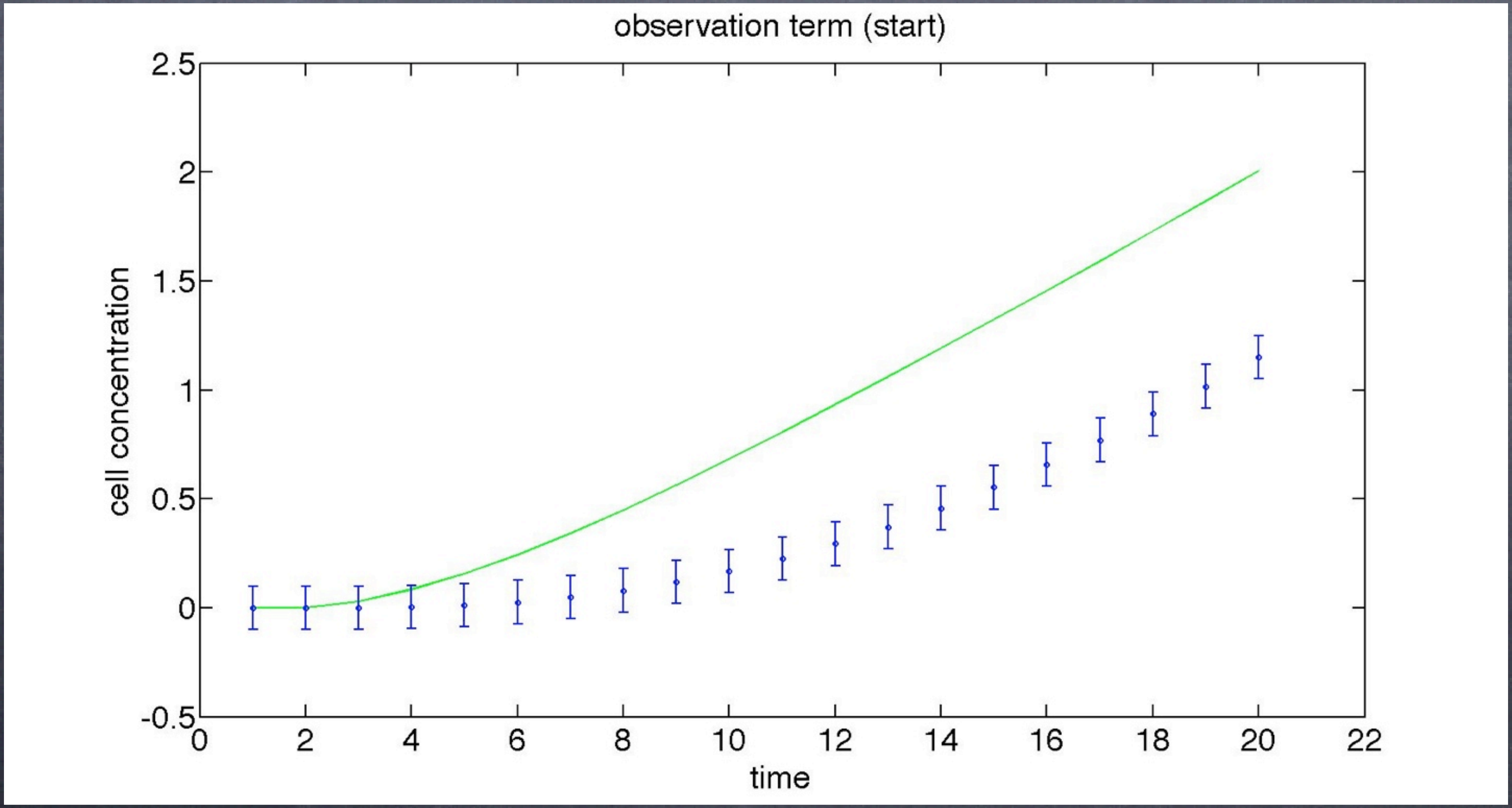
Cost function value

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T B^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

Show movie

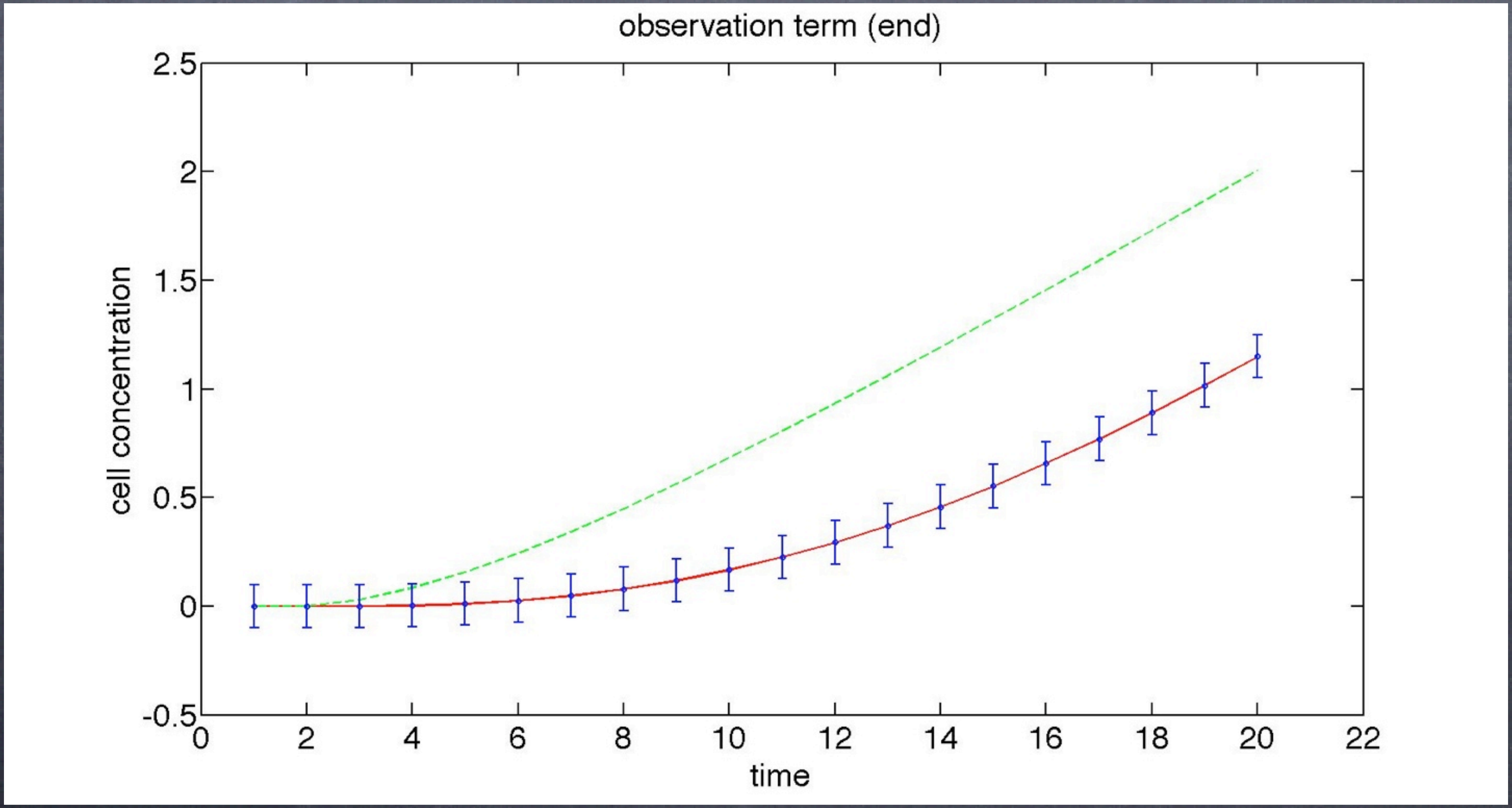
Cost function value

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T B^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$



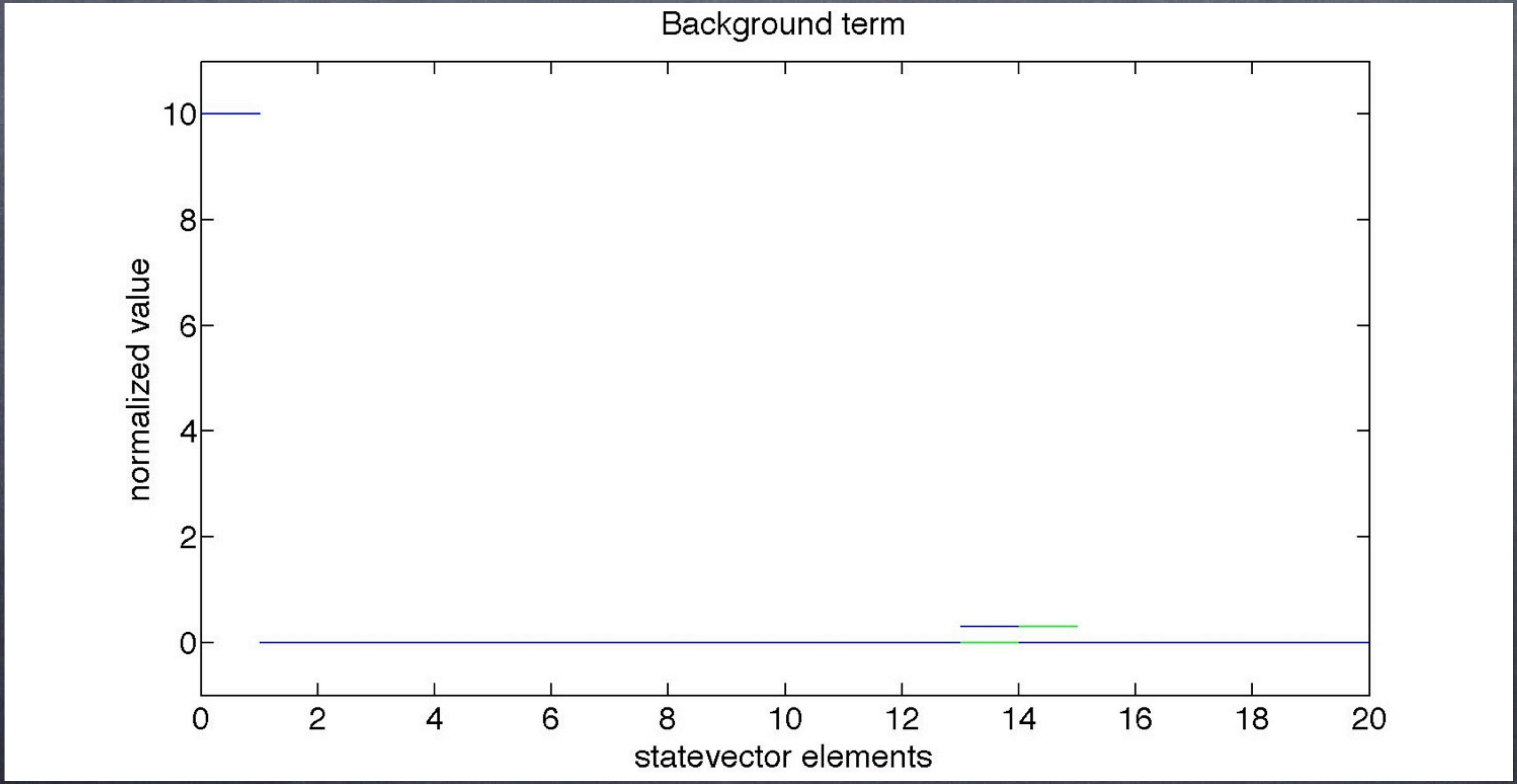
Cost function value

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T B^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

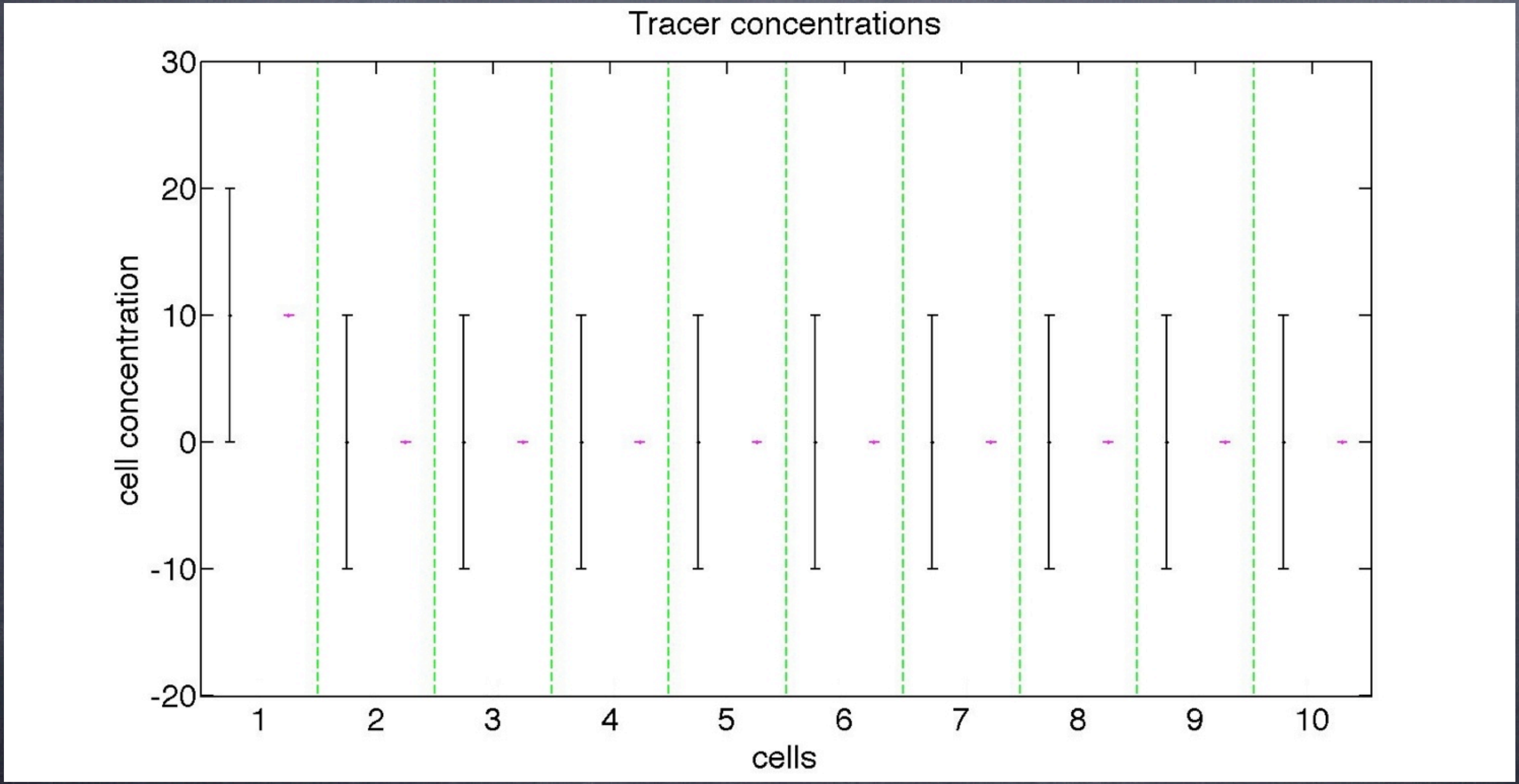


Cost function value

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T B^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}_i - H\mathbf{x}_i)^T R^{-1}(\mathbf{y}_i - H\mathbf{x}_i)$$

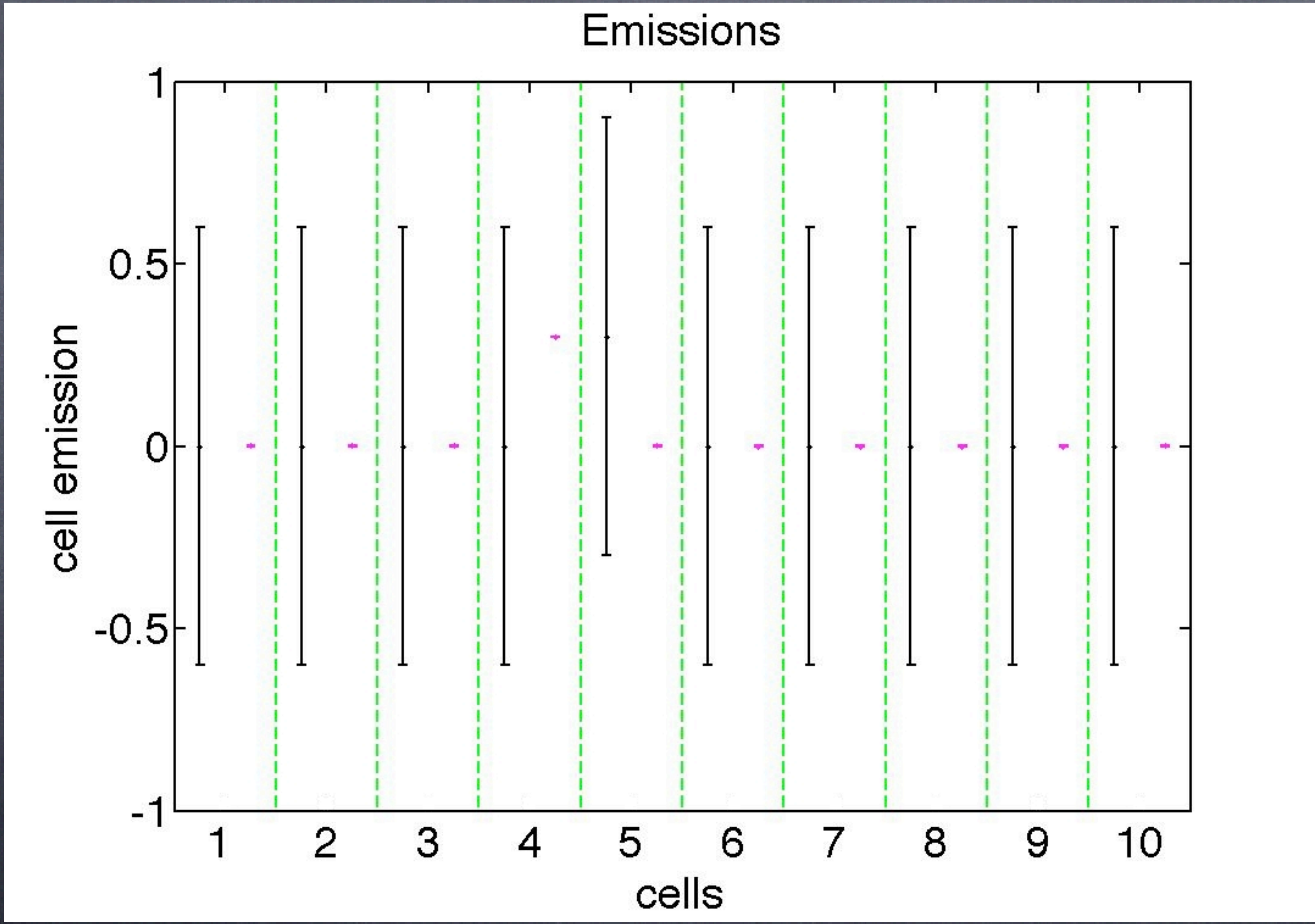


Result



Tracer concentrations
30
20
10
0
-10
-20
cell concentration
1
2
3
4
5
6
7
8
9
10
cells

Result



Fewer stations/times

- Stations in cells: 1,4,6,9
- # measurement times = 51

Fewer stations/times

- Stations in cells: 1,4,6,9
- # measurement times = 51

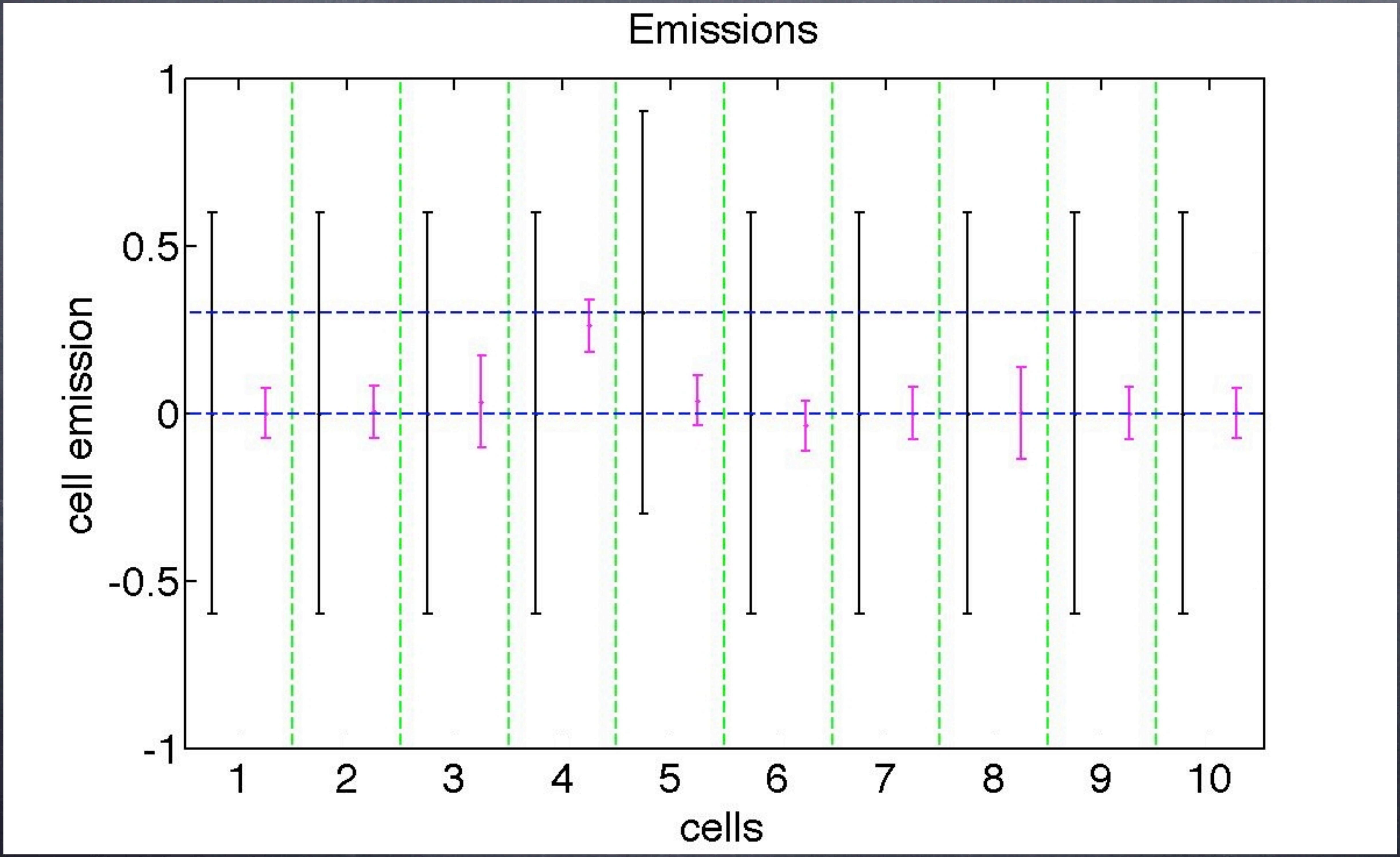
↓

The initial concentrations and emissions will be less constrained by the observations, hence, we expect that the background estimate is not completely neglected anymore

Result

Tracer concentrations
cell concentration
cells
30
20
10
0
-10
-20
1
2
3
4
5
6
7
8
9
10

Result



Problems with CONGRAD

- Not always possible to achieve preset reduction in gradient norm -> crash Especially if only a few observation stations are used.

- CONGRAD should be equivalent to BLUE analysis but is not if a small amount of observations is used.

Conclusion

- CONGRAD equivalent to BLUE analysis if many stations and observations are used

- If a few stations are used CONGRAD breaks down/does not reach tolerance

- If we use many stations but only a few measurement times, CONGRAD yields another solution than BLUE analysis.

Future research

Inversion of CO emissions using:

- Ground stations observations and satellite measurements
- TM5-model with CONGRAD minimization
- A priori estimate from inventory from literature