DEMOKRITOS

NATIONAL CENTRE FOR SCIENTIFIC RESEARCH

DEVELOPMENT AND FIRST TESTS OF A DATA ASSIMILATION ALGORITHM IN A LAGRANGIAN PUFF ATMOSPERIC DISPERSION MODEL

V. Tsiouri[1,3], I. Kovalets[2], S. Andronopoulos[1], J.G. Bartzis[3]

[1]Environmental Reasearch Laboratory, Institute of Nuclear Technology and Radiation Protection, NCSR 'Demokritos', Greece
E-mail: vtsiouri@ipta.demokritos.gr

[2]Department of Environmental Modelling, Institute of Mathematical Machines & System Problems, National Academy of Sciences of Ukraine

[3]Department of Engineering and Management of Energy Resources, University of Western Macedonia, Greece

11th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
Cambridge, United Kingdom, July 2nd - 5th, 2007

“DATA ASSIMILATION”

is one of the most challenging problems in atmospheric dispersion modelling

MODEL

- **Forecast capabilities**
- **High spatial / temporal resolution**

- **Uncertainties in input data**
- **Assumptions / Approximations**

OBSERVATIONS

- **Representation of reality**

- **Incomplete spatial and temporal coverage**
- **No forecast**

DATA ASSIMILATION

- combines both in an optimal way, to enhance their advantages and reduce their disadvantages
- solves the inverse problem of determining the model state from the observations

DATA ASSIMILATION TECHNIQUES

SEQUENTIAL Approach (Linear Filters)

Compute the analysis
by
solving the optimal estimation equations

A filter analyzes the system state each time that observation data become available

VARIATIONAL Approach (Variational Methods)

Compute the analysis
by
minimizing a cost function

Variational methods are based on minimization of a cost function within a time interval

Optimal Estimator Equations

e.g. Kalman Filter

Time Update equations
("predict")

Measurement Update equations
("Correct")

Cost function

Sum of the squared deviations of the analysis values from the observations, weighted by the accuracy of the observations

+

Sum of the squared deviations of the forecast fields from the analyzed fields, weighted by the accuracy of the forecast

This term is added to make sure that the analysis does not drift too far away from observations and forecasts that are usually known to be reliable.

Problem to be solved:

Atmospheric Dispersion forecast with Assimilation of the available concentration measurements Data

Solution method:

- Development of a Data Assimilation algorithm based on *Variational* approach and its implementation in a *Lagrangian puff* dispersion model
- The algorithm evaluates the “true” *source emission rate* of pollutant

THE ATMOSPHERIC DISPERSION MODEL

The puff model used is a simplified (Gaussian) version of the DIPCOT model

A parcel of pollutant is characterized by

↓

Gaussian distribution of concentration inside the parcel

The concentration at a given location is obtained by

↓

Summing the impacts of all puffs

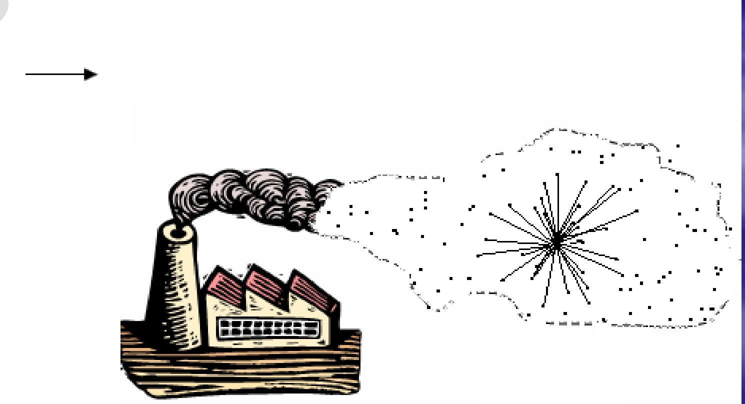
The model-calculated concentration at a point (*X*,*Y*,*Z*) :

$$C^M(X,Y,Z) = \frac{1}{(2\pi)^{3/2}} \sum_{P=1}^{L} \frac{M_p}{\sigma_{xp}\sigma_{yp}\sigma_{zp}} \exp\left[-\frac{1}{2}\frac{(X_p - X)^2}{\sigma_{xp}^2}\right] \exp\left[-\frac{1}{2}\frac{(Y_p - Y)^2}{\sigma_{yp}^2}\right]$$

$$\left\{\exp\left[-\frac{1}{2}\frac{(Z_p - Z)^2}{\sigma_{zp}^2}\right] + \exp\left[-\frac{1}{2}\frac{(Z_p + Z - 2Z_g)^2}{\sigma_{zp}^2}\right]\right\}$$

$$M_p = q_p \times \Delta t$$

- M_p is the "load" of puff p
- q_p is the source emission rate during the release of the puff p
- Δt is the release duration of puff p



Why Variational Data Assimilation?

For the case of the "puff" model Variational assimilation is more suitable, because the state vector of the model (puff positions) is nonlinearly related to the measurements (concentrations).

In that case one has to use the so-called "Extended Kalman Filter", which is valid only if the true values of concentrations are close enough to the first guess model predictions This is not the case in real practice: error in source functions and hence in first guess model predictions can be large.

VARIATIONAL METHOD

- A set of unknown model-parameters ("control variables") is selected
- The control variables are calculated such as to minimize the functional

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(\sigma_{nk}^{O}\right)^{-2} \left(C_{nk}^{O} - C_{nk}^{M}\right)^{2}$$

Where: N observation times and K measurement locations, and

σ_{nk}^{O}	RMS error of the observation at time *n* and location *k*
C_{nk}^{O}	Measured concentration at time *n* and location *k*
C_{nk}^{M}	Model-calculated concentration at time *n* and location *k*

Selected Control Variables

The selected control variables are the values of the source emission rate during the release of each model puff $\mathbf{q} = \left(q_1, q_2, \ldots, q_L\right)^T$

Rationale: in the event of accidental gas releases, there is a high uncertainty regarding the release rate which is important for forecasting the gas concentration

The model-calculated concentrations are expressed as function of the control variables vector: $\mathbf{C}^M = \mathbf{G} \cdot \mathbf{q}$

THE MODEL EQUATIONS

The cost functional becomes:

$$J = \left(\mathbf{C}^O - \mathbf{G}\cdot\mathbf{q}\right)^T \mathbf{O}^{-1}\left(\mathbf{C}^O - \mathbf{G}\cdot\mathbf{q}\right) + \left(\mathbf{q}-\mathbf{q}_B\right)^T \mathbf{B}^{-1}\left(\mathbf{q}-\mathbf{q}_B\right)$$

Setting derivative with respect to $\mathbf{q}$ equal to zero:

$$\left(\mathbf{I} + \left(\mathbf{O}^{-1}\mathbf{G}\mathbf{B}\right)^T \mathbf{G}\right)\cdot\mathbf{q} = \left(\mathbf{O}^{-1}\mathbf{G}\mathbf{B}\right)^T \mathbf{C}^O + \mathbf{q}_B$$

Background and observational errors uncorrelated and constant

$$\left(r + \mathbf{G}^T\mathbf{G}\right)\cdot\mathbf{q} = \mathbf{G}^T\mathbf{C}_O + r\mathbf{q}_B$$

Where: $r = \sigma_O^2 / \sigma_B^2$

Observational errors proportional to observed concentrations

$$\left(r_{\mathrm{mod}} + \mathbf{G}^T_{\mathrm{mod}}\mathbf{G}\right)\cdot\mathbf{q} = \mathbf{G}^T_{\mathrm{mod}}\mathbf{C}^O + r_{\mathrm{mod}}\mathbf{q}_B$$

Where: $r_{\mathrm{mod}} = a^2 / \sigma_B^2$

The above equations are solved for $\mathbf{q}$

RESULTS

"Identical twin" experiments were used to evaluate the performance of the data assimilation methodology for the puff model.

"Twin" experiment is a method for testing a DA algorithm, when real observations are not available.
"Twin" experiments include two parts:

'True' run and 'Simulation' run.

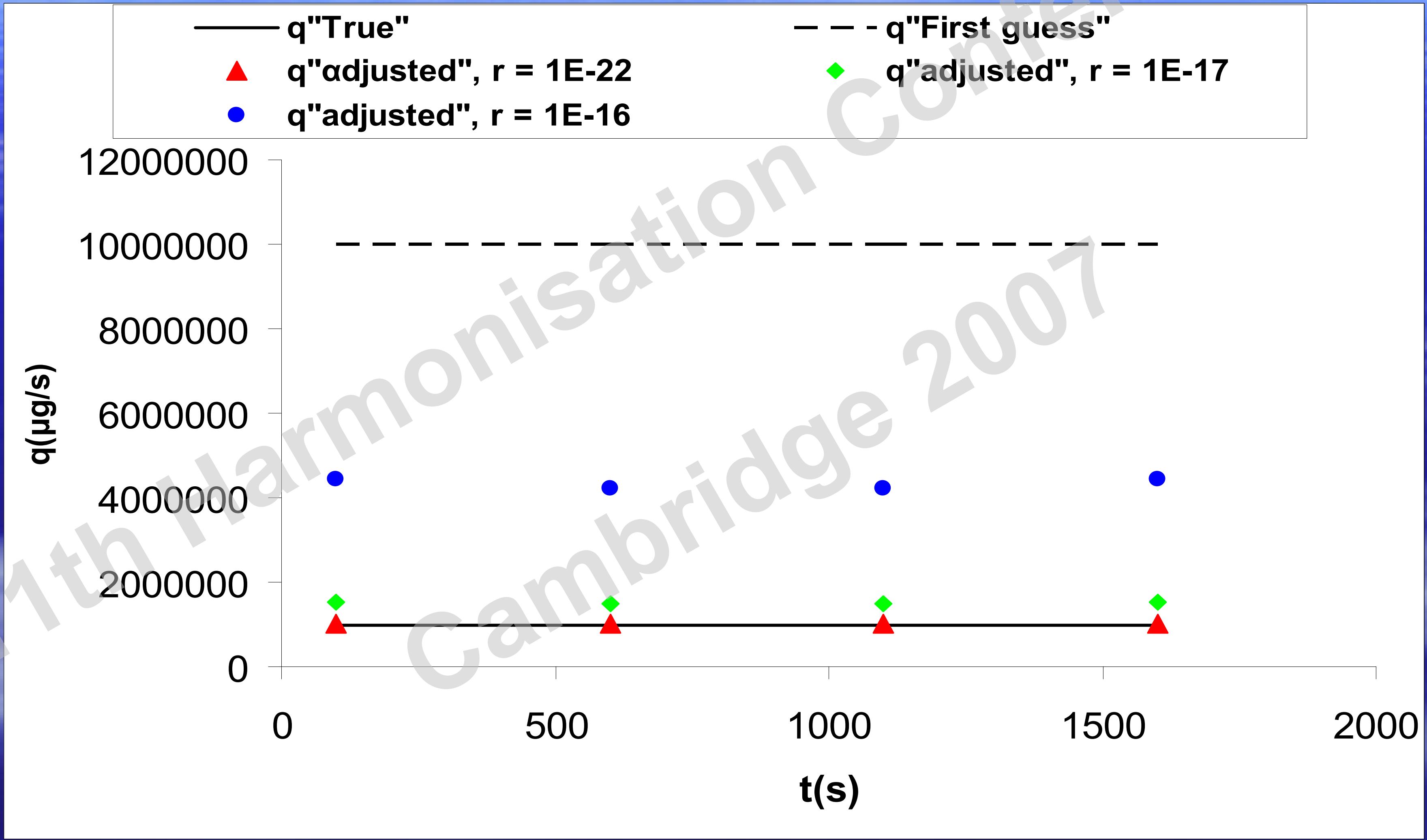
The dispersion model generates "concentration observations" using a "true" source term function.
Then the model is run again using an "assumed" source term and assimilating the observations with the aim to evaluate the true source function.

FIRST TESTS

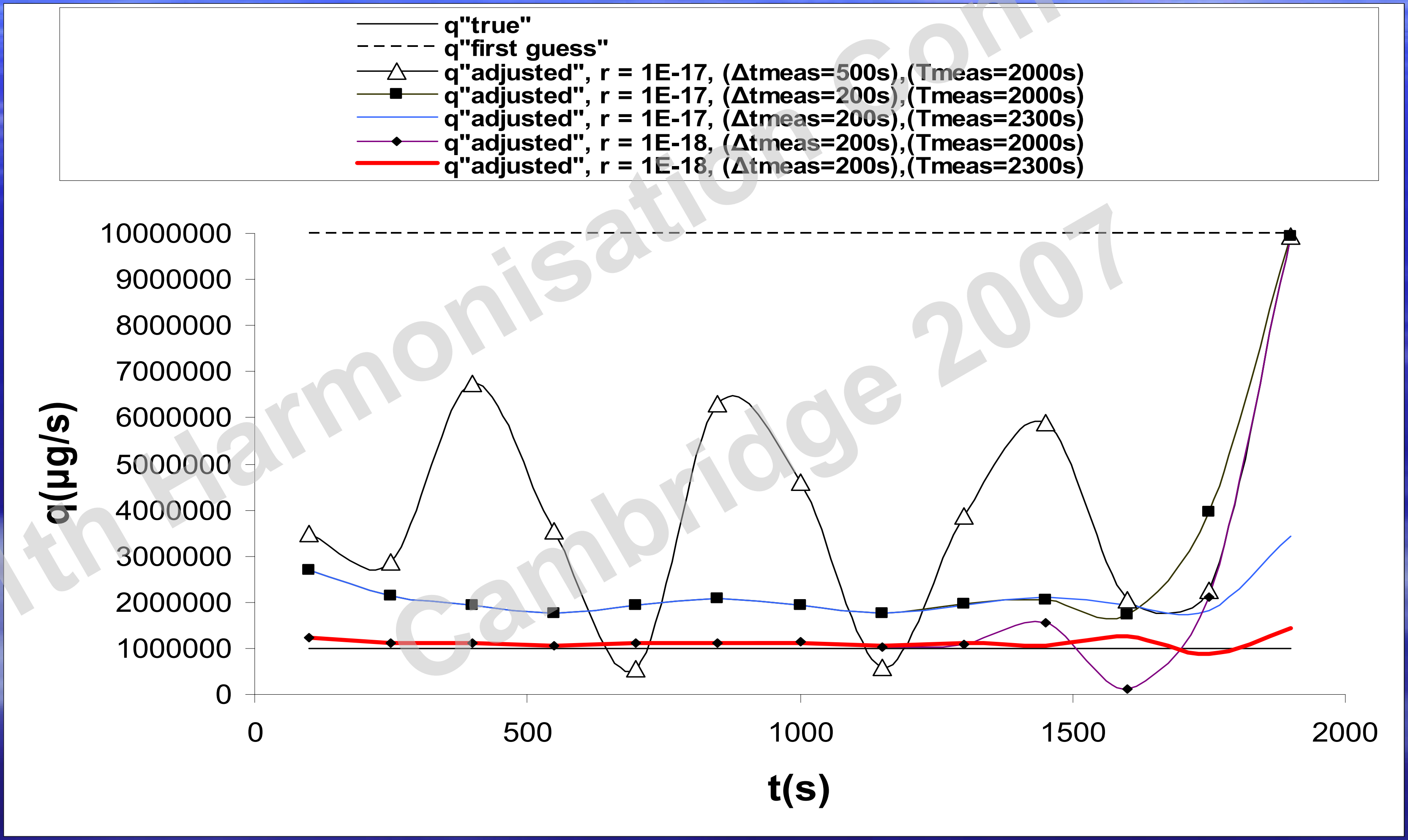
- 1-dimensional dispersion
- Constant in time source rate
- Constant wind speed
- **Constant root mean square error of the observations**
- 1 observation point

The initially assumed source term function, differs by a factor of 10 from the "true" one

Effect of different *r* values on the estimation of "true" source function



The influence of the total time duration of the measurements and of the measurement time interval on the analysed source term (Tsource= 2000 s)



CONCLUSIONS:

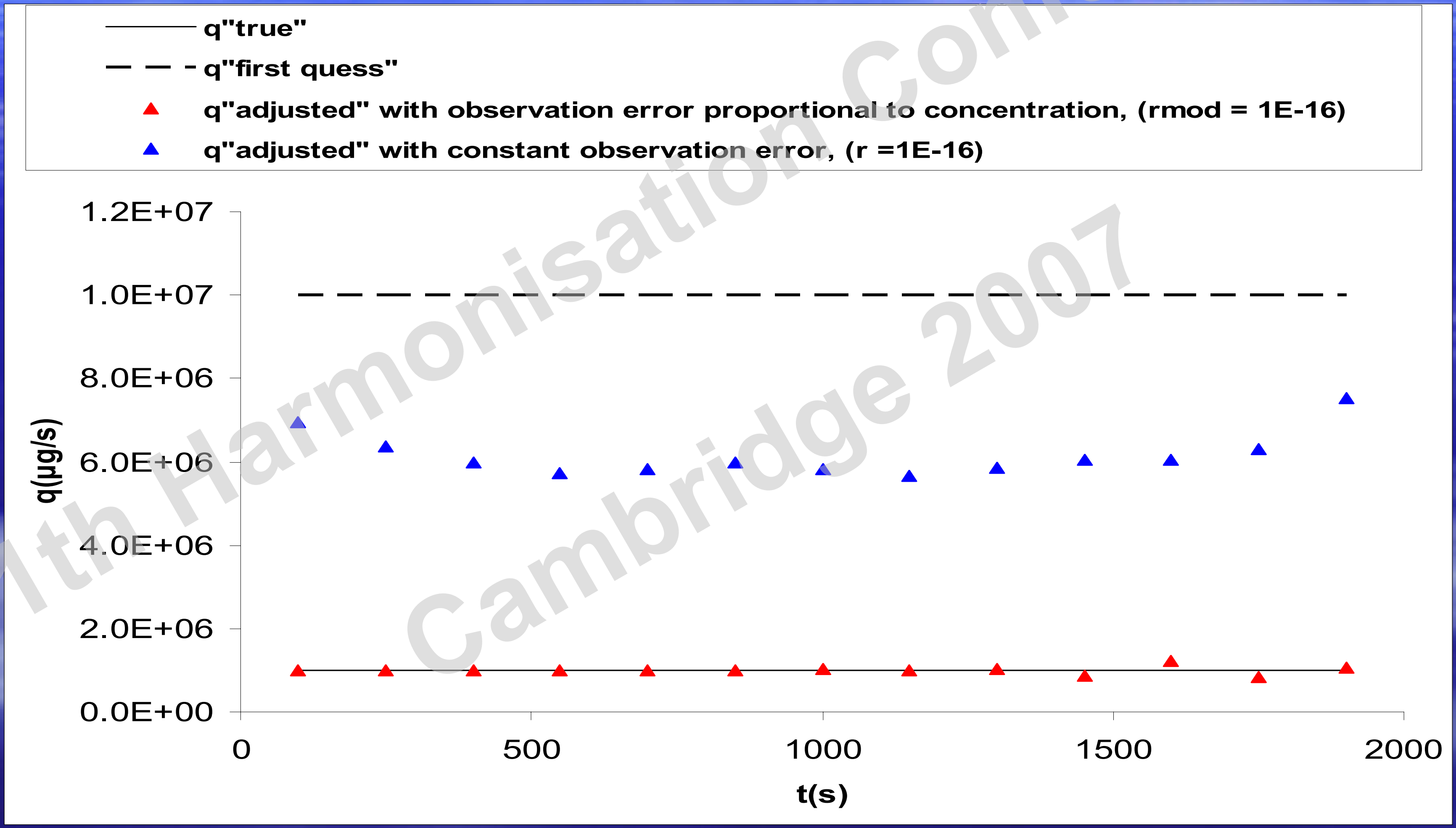
- The adjustment of source function strongly depends on the *r* value: it improves as *r* decreases. Small *r* means small error of measurements, and in this case the measurements are given higher weights.
- The performance of the method improves as the time frequency of observations increases: the "load" of more model-puffs is adjusted with more frequent observations.
- If the total measurements time does not cover the source release duration, not all the model puffs can be adjusted.

Remark:

- If σ_o is constant then the performance of the method depends on the magnitude of the measured concentrations, i.e., on the distance of the measurement location(s) from the source: small measured concentrations (i.e., far from the source) would be given less weight (because the ratio σ_0/C_0 increases.

The way to overcome this undesirable behaviour is to abandon the assumption of constant σ_o and assume that σ_o is proportional to observed concentrations.

Source Rate Estimation with the RMS of the Observations Proportional to the Values of Concentration



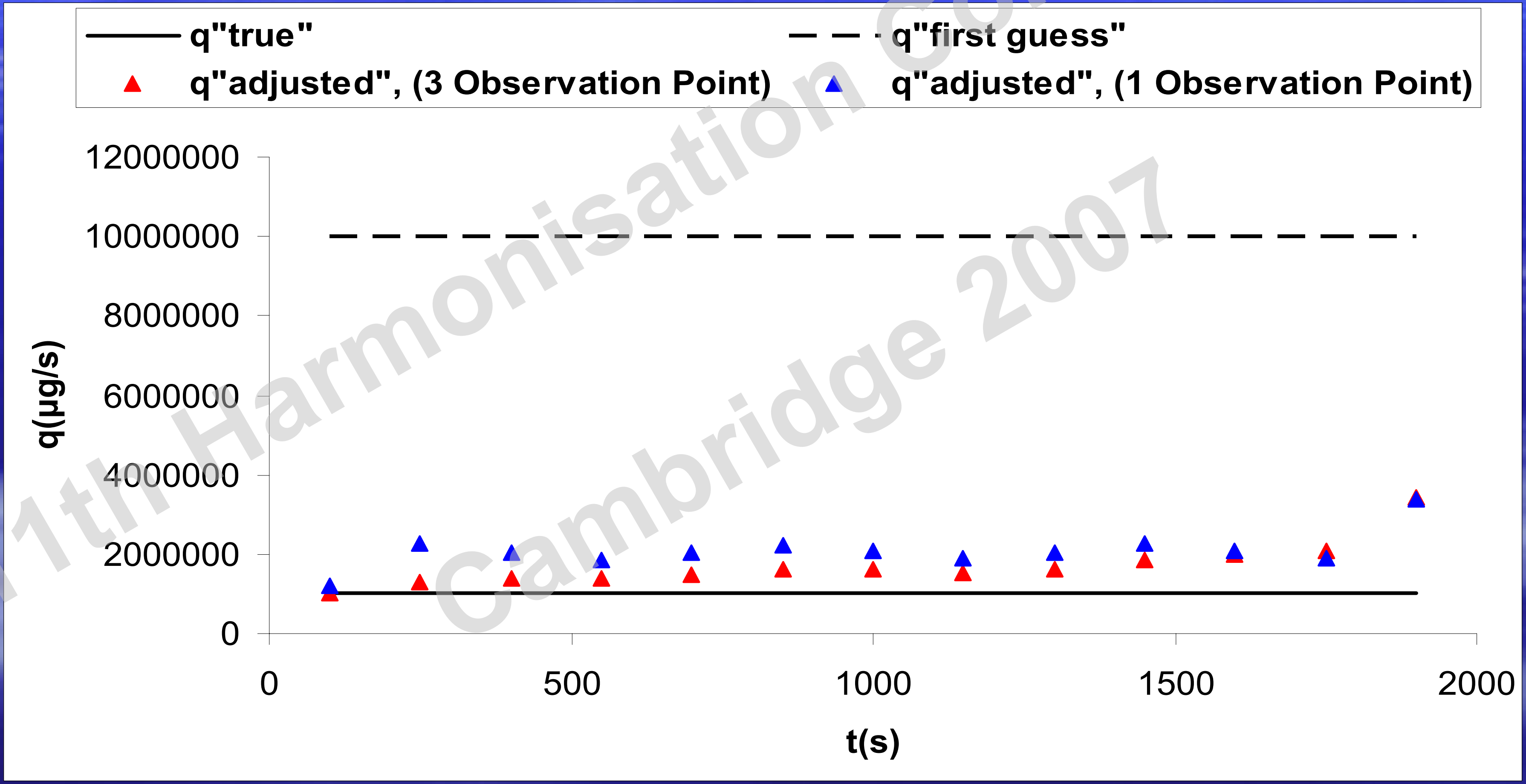
Conclusion:

- When the RMS error of observations is proportional to the observed values of concentration, the performance of the method is improved, even for larger *r* values

More Observation Points

Xm(1)=5 km, Xm(2)=10 km, Xm(3)=20 km

$r_{mod} = 10^{-13}$

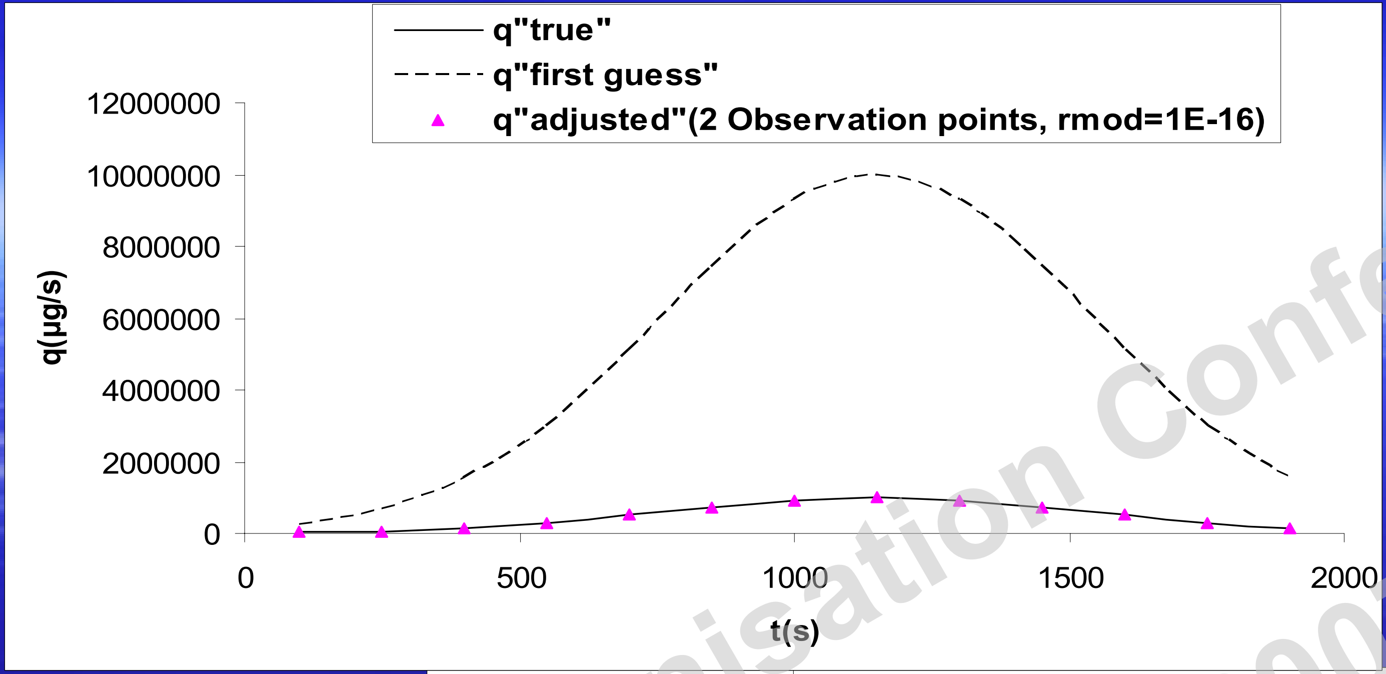


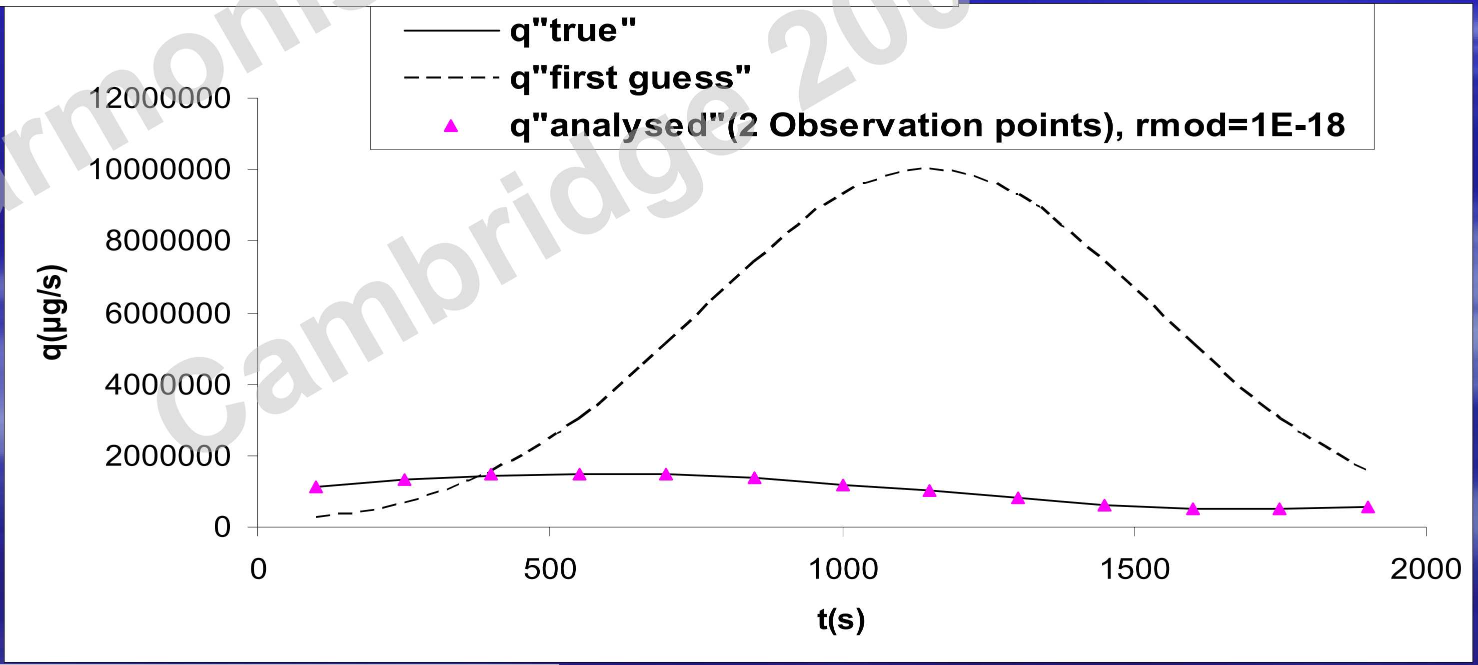
Conclusion:

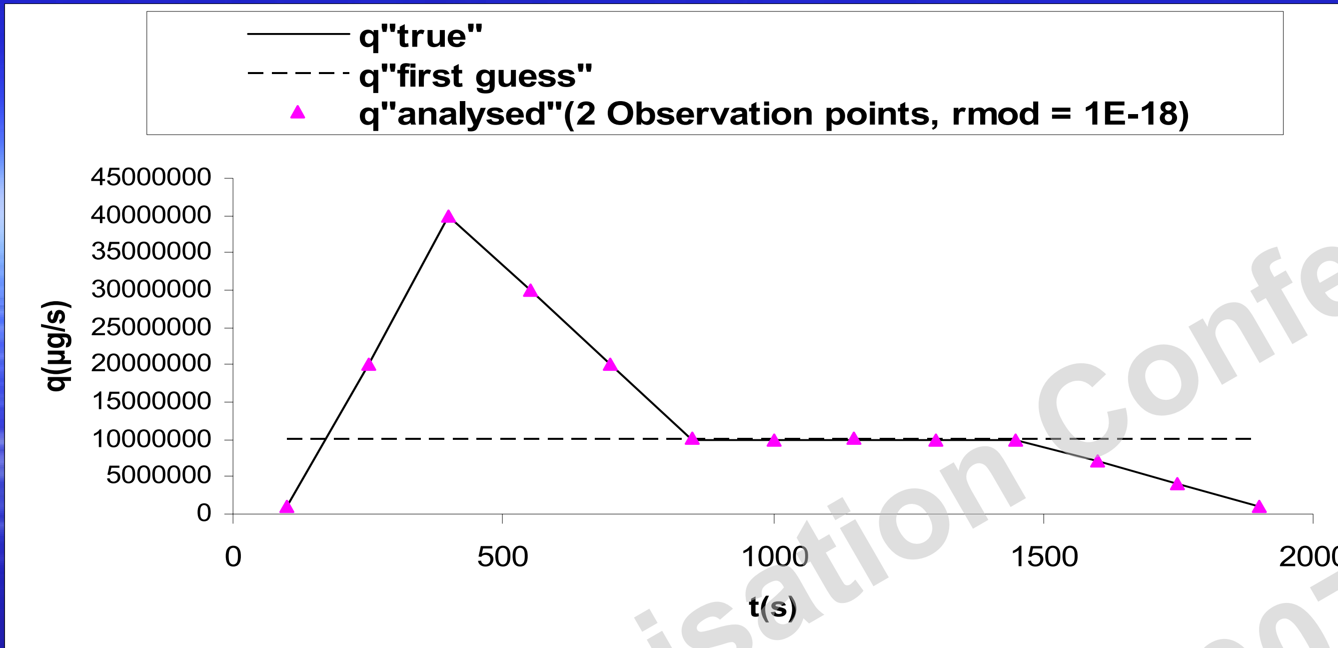
- When there are more observation points, the adjusted source rate is improved.

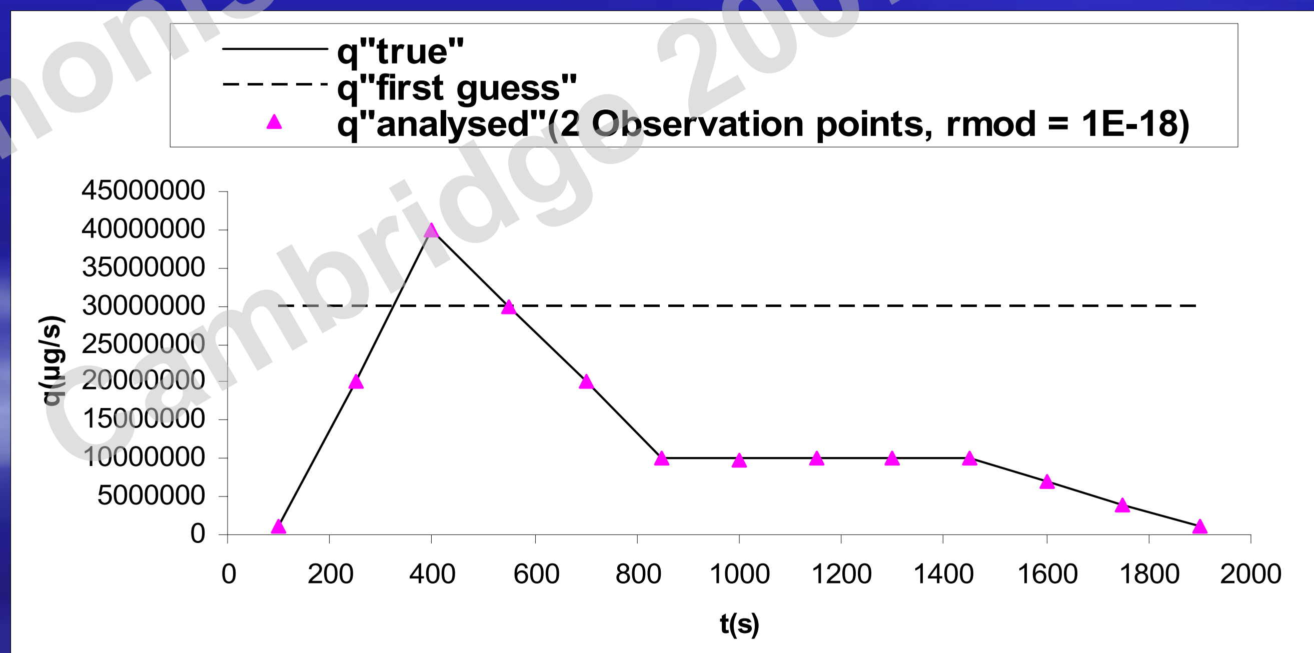
1-dimensional cases

- *Estimation of variable in time emission rate:*
  - *Gaussian-shaped function*
  - *Shifted in time*
  - *Linear variation*


q"true"
q"first guess"
q"adjusted"(2 Observation points, rmod=1E-16)
12000000
10000000
8000000
6000000
4000000
2000000
0
q(µg/s)
0
500
1000
1500
2000
t(s)


q"true"
q"first guess"
q"analysed"(2 Observation points), rmod=1E-18
12000000
10000000
8000000
6000000
4000000
2000000
0
q(µg/s)
0
500
1000
1500
2000
t(s)


q"true"
q"first guess"
q"analysed"(2 Observation points, rmod = 1E-18)
q(µg/s)
45000000
40000000
35000000
30000000
25000000
20000000
15000000
10000000
5000000
0
0
500
1000
1500
2000
t(s)

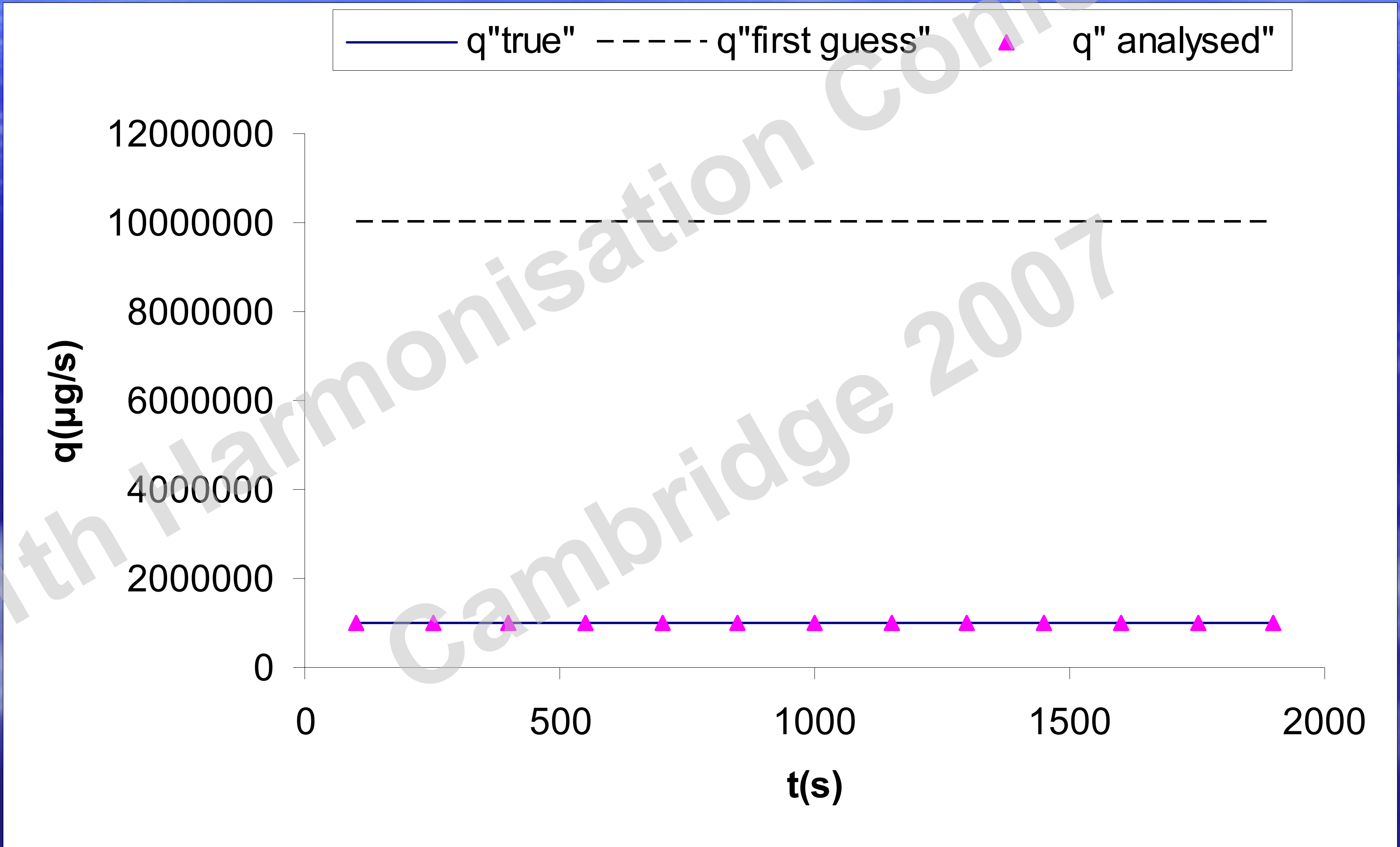

q"true"
q"first guess"
q"analysed"(2 Observation points, rmod = 1E-18)
q(µg/s)
45000000
40000000
35000000
30000000
25000000
20000000
15000000
10000000
5000000
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000
t(s)

Estimation of source emission rate for the case of Variable Wind Speed

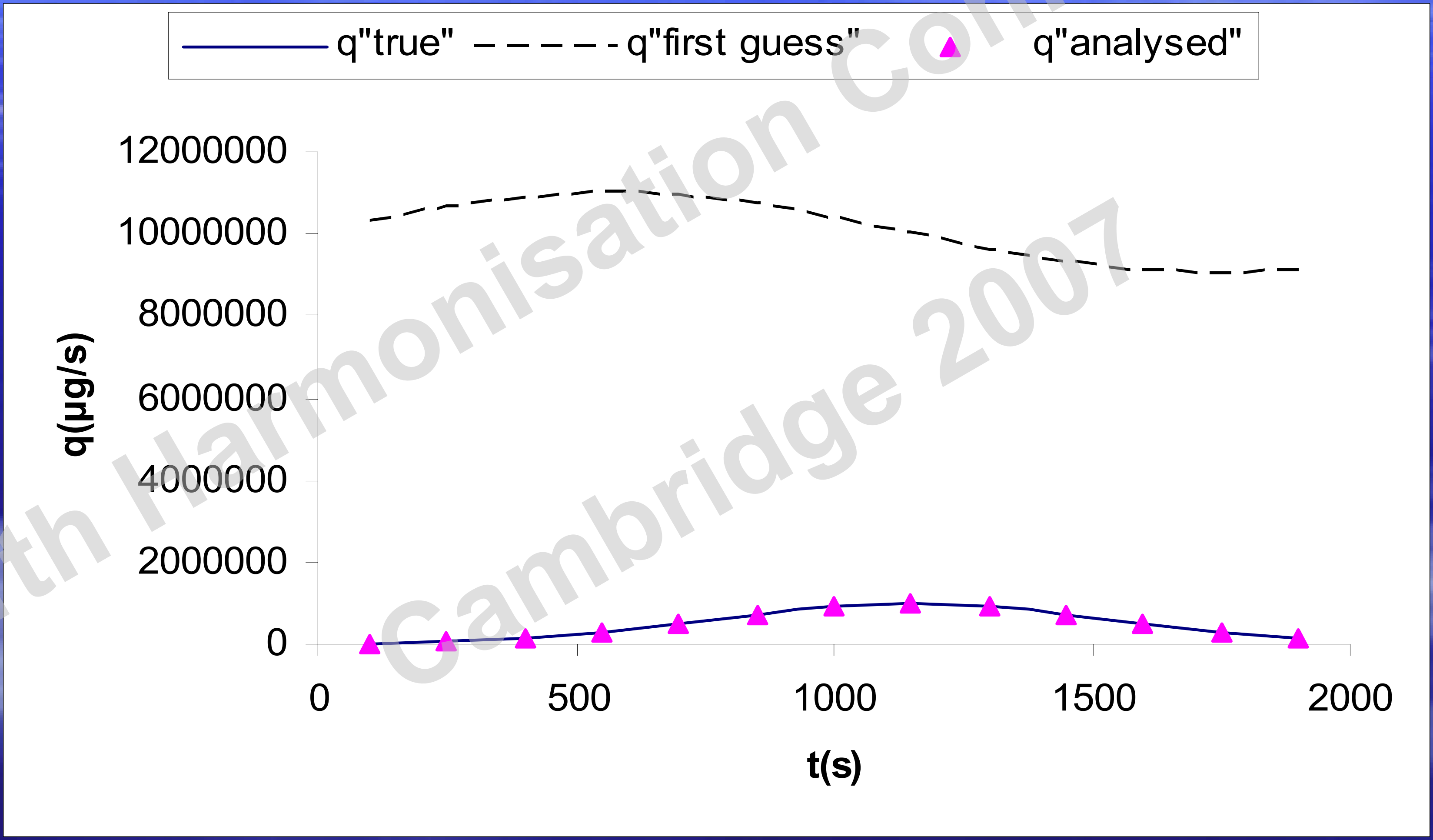
- The wind velocity varies spatially as a sinusoidal function

$$v = v_0 + A\sin(kX)$$

- Constant source function
- 3 Observation points , $r_{mod}=10^{-18}$
- Variable wind speed

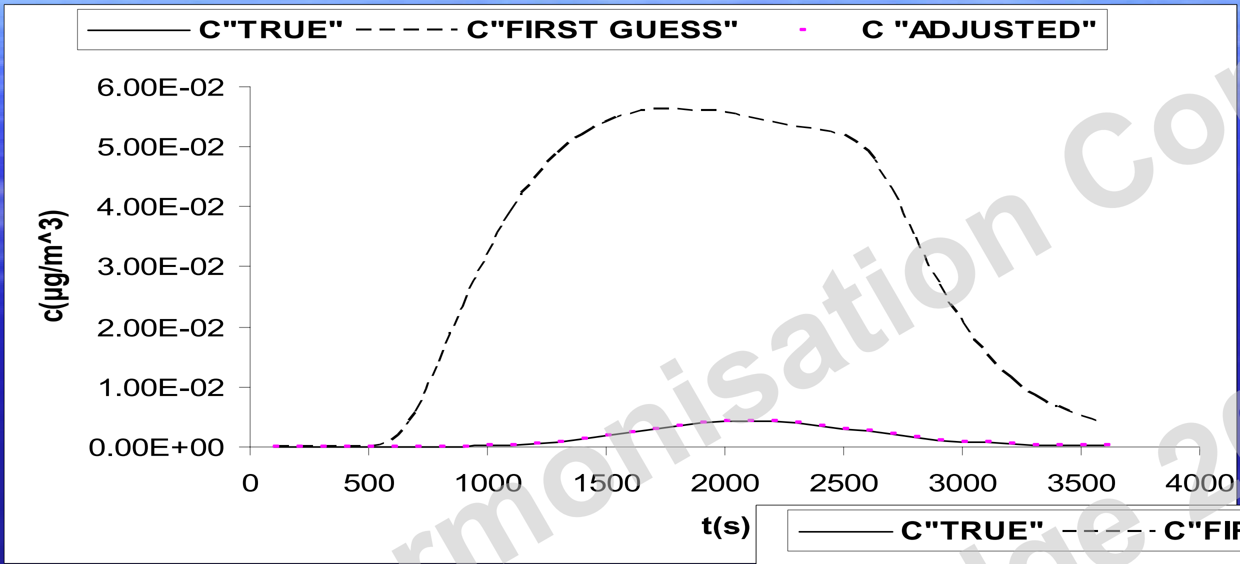


- Variable source function
- 3 Observation points, $r_{mod}=10^{-18}$
- Variable wind speed

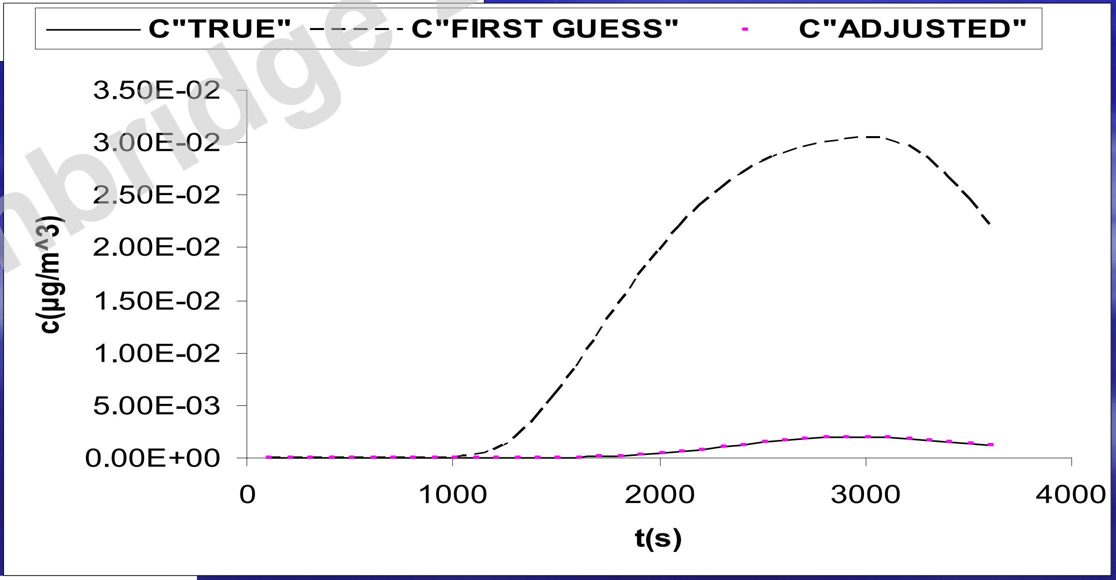


Results for the model-calculated concentration after Data Assimilation

Xm = 20 km



Xm = 10 km



- **Variable source function**
- **3 Observation points**
- **Variable wind speed**
- **$r_{mod} = 10^{-18}$**

CONCLUSIONS

- The developed algorithm of variational data assimilation allows adjustment of source function in puff/particle model for non-stationary (wind and source) conditions

- The results showed that the assumption of error of observations proportional to the concentrations is more preferable than the constant relative error

- Finally tests with variable wind speed showed that the formulation of the methodology is more general and can be extended to take into account variable meteorological conditions.

Future Work

Future work involves:

- Application of the method to more realistic situations (2-dimensional, 3-dimensional) and the required optimizations
- Extension of the DA method to evaluate the source height through adjoint equations