IL NUOVO CIMENTO Vol. 30 C, N. 2 Marzo-Aprile 2007
DOI 10.1393/ncc/i2007-10003-y

# Estimation of emission rate from experimental data

D. R. Roberti(1)(*), D. Anfossi(2)(**), H. F. de Campos Velho(3)(***) and G. A. Degrazia(1)(**)

(1) *Department of Physics, Federal University of Santa Maria (UFSM) Santa Maria (RS), Brazil*
(2) *ISAC, CNR - Torino, Italy*
(3) *LAC, INPE - São José dos Campos (SP), Brazil*

(ricevuto l' 1 Marzo 2007; approvato il 30 Aprile 2007; pubblicato online il 16 Luglio 2007)

**Summary.** — The estimation of the source pollutant strength is a relevant issue for atmospheric environment. This characterizes an inverse problem in the atmospheric pollution dispersion studies. In the inverse analysis, a time-dependent pollutant source is considered, where the location of such source term is assumed known. The inverse problem is formulated as a non-linear optimization approach, whose objective function is given by the least-square difference between the measured and simulated by the mathematical model, pollutant concentration, associated with a regularization operator. The forward problem is addressed by a Lagrangian model, and a quasi-Newton method is employed for minimizing the objective function. The second-order Tikhonov regularization is applied and the regularization parameter is computed by using the L-curve scheme. The inverse-problem methodology is verified with data from the tracer Copenhagen experiment.

PACS 47.27.tb – Turbulent diffusion.

## 1. – Introduction

The pollutant dispersion in the atmosphere is a topic with high interest nowadays. Models for air monitoring are not only important for describing the pollutant impact over urban or rural areas, or for urban planning consideration, but other issues are also relevant, such as the pollutant source strength estimation, $CO_2$ diurnal cycle, and total ozone in the atmosphere. The last three issues are examples of inverse problems in

(*) E-mail: droberti@smail.ufsm.br
(**) E-mail: anfossi@isac.cnr.it
(***) E-mail: haroldo@lac.inpe.br
(**) E-mail: degrazia@ccne.ufsm.br

atmospheric pollution modelling. Inverse problems are usually solved by an *implicit technique*, where the inverse problem is formulated as a constrained non-linear optimization problem: the forward problem is iteratively solved for successive approximations of the unknown parameters. The associated forward problem is the solution of the dispersion model. As pointed out in [1], the major obstacle to a good inversion result is the accuracy of the dispersion model and the representativeness of the measurements.

The purpose of this paper is to employ an inverse-problem methodology for estimating the emission rate from a measured dataset in Copenhagen experiment using a Lagrangian particle model, named LAMBDA [2, 3].

## 2. – Forward model

The Lagrangian particle model LAMBDA was developed to study the transport process and pollutants diffusion, starting from the Brownian random walk modelling. This model is based on a three-dimensional form of the Langevin equation for the random velocity [4]. LAMBDA was tested against many data sets, and in particular with the Copenhagen experiment, giving satisfactory results [2, 3].

LAMBDA uses a large number of fictitious particles to simulate the atmospheric diffusion. Each particle can be marked for a mass, and the spatial distribution of particles in the computational domain allows calculating the three-dimensional concentration field, $C^{\mathrm{Mod}}(\vec{\mathbf{r}}) = C(x, y, z, t)$, through the calculation of how many of them lie in a cell or imaginary volume centred in $x$, $y$, $z$, as follows:

$$C(x, y, z, t) = m_p \frac{N_v}{V_c}\,, \tag{1}$$

where $m_p$ is the mass of each particle, $N_v$ is the particle number in the cell and $V_c$ is the cell volume. The mass of each particle is determined from

$$m_p = \frac{Q(t)\Delta t}{N_p}\,, \tag{2}$$

where $Q(t)$ is the emission rate, $\Delta t$ is the time step and $N_p$ is the number of particle emitted per time step.

## 3. – Inverse model

In order to set up the inverse analysis, it is assumed that the concentration obtained with the mathematic model is given by $C^{\mathrm{Mod}}(\vec{\mathbf{r}}, \mathbf{Q})$, where $\mathbf{Q} = Q(t)$ is the emission rate and $C^{\mathrm{Exp}}(\vec{\mathbf{r}})$ are data from experimental data of concentration. The solution of the inverse problem is a function $\mathbf{Q}$ that minimizes the following objective function:

$$J(\lambda, \mathbf{Q}) = \left\| C^{\mathrm{Exp}}(\vec{\mathbf{r}}) - C^{\mathrm{Mod}}(\vec{\mathbf{r}}, \mathbf{Q}) \right\|_2^2 + \lambda \Omega(\mathbf{Q})\,. \tag{3}$$

The least-square difference between experimental data and calculated values is represented by the first term in eq. (3); $\Omega(\mathbf{Q})$ is a regularization operator, a *a priori* piece of information, that mathematically is a restriction about the function.

The regularization operator can be expressed by Tikhonov scheme [5]:

$$\Omega(\mathbf{Q}) = \sum_{m,j=0}^{p} \kappa_{m,j} \left\| \mathbf{Q}^{(m)} \right\|_2^2; \tag{4}$$

here $\mathbf{Q}^{(m)}$ denotes the $m$-th difference. In general the parameter $\kappa_{m,j}$ is chosen as $\kappa_{m,j} = \delta_{mj}$ (Kronecker delta) and the regularization is named Tikhonov-$j$ regularization operator, where $j$ denotes the order of the regularization. Here, the second-order regularization (Tikhonov-2) is employed.

**3**˙1. *Optimization algorithm.* – The optimization problem is iteratively solved by the quasi-Newtonian optimiser routine E04UCF, from the NAG Fortran Library. This algorithm is designed to minimize an arbitrary smooth function subject to constraints (simple bound, linear or non-linear constraints), using a sequential programming method.

This routine has been successfully used in several previous works: in geophysics, hydrological optics, heat transfer, and meteorology [6-12].

## 4. – Simulating the Copenhagen experiments with forward problem

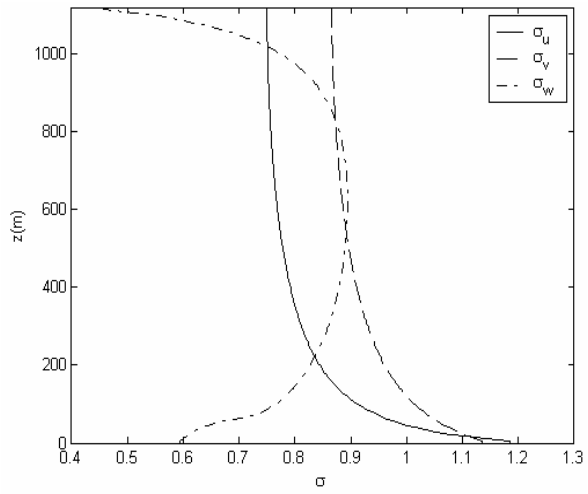
The choice of this experiment derives from the amount and quality of the measurements. For each experiment, meteorological information, like wind velocity standard deviations at the emission height, mixing height and Obukhov length, were available for each hour; and wind speed and wind direction were available at each 10 min, and at three levels along a tower. The ground level concentration was measured using an array of samplers.

Copenhagen tracer dispersion experiments were carried out in the northern part of Copenhagen [13, 14]. In this experiment, the pollutant ($SF_6$) was released without buoyancy from a tower at a height of 115 m, and collected at ground level in three crosswind arcs located at 2–6 km from the releasing point. The site was mainly residential with a roughness length of 0.6 m. The Copenhagen region lies on an island. Tracer concentrations were measured and an average was computed at each 20 min. We considered the exercise performed on 19 October 1978, and the concentrations evaluated in the second period (12:33h-13:53h). The Copenhagen tracer dispersion experiment used in this analysis was selected due to the amount and quality of the concentration and meteorological data set.

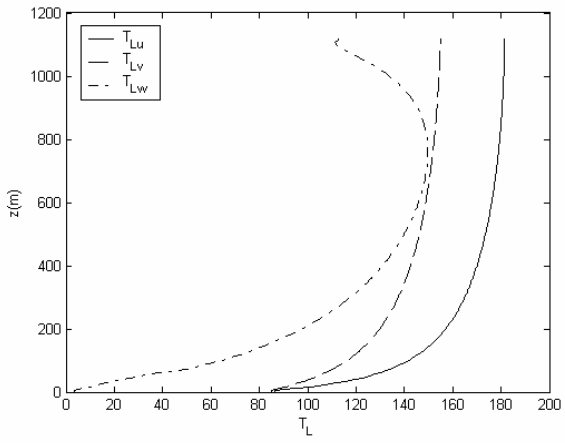
Meteorological data, available each 10 min (see table I), were used to create the input for the simulation in the forward model (LAMBDA). The profiles of wind standard deviations ($\sigma_i$, being $i = u, v, w$), and the Lagrangian decorrelation time scales ($\tau_{L_i}$) were calculated according to the turbulence parameterisation scheme suggested by Degrazia *et al.* [15] (figs. 1a and b). The measured wind velocity variances $\sigma_i^2$ ($i = v, w$) were used to fit the analytical model for the wind variance. It is worth mentioning that the ratios of these variances, $(\sigma_v^2)_{\mathrm{Model}}/(\sigma_v^2)_{\mathrm{Exp}}|_{z=10\ \mathrm{m}}$ and $(\sigma_w^2)_{\mathrm{Model}}/(\sigma_w^2)_{\mathrm{Exp}}|_{z=10\ \mathrm{m}}$, were found to be very close to one (1.08, 1.05, respectively), and this gives a further verification of the good quality of the parameterisations used. The longitudinal wind standard deviation was set equal, at each height, to the crosswind one. It was verified that the observed vertical profile of the wind speed $U(z)$ was in a very good agreement with the classical

TABLE I. – *Meteorological data.*

| | $U$ (m/s) | | | $\theta$ (°) | | |
|---|---|---|---|---|---|---|
| hour | 10 m | 120 m | 200 m | 10 m | 120 m | 200 m |
| 12:05 | 2.6 | 5.7 | 5.7 | 290 | 310 | 310 |
| 12:15 | 2.6 | 5.1 | 5.7 | 300 | 310 | 310 |
| 12:25 | 2.1 | 4.6 | 5.1 | 280 | 310 | 320 |
| 12:35 | 2.1 | 4.6 | 5.1 | 280 | 310 | 320 |
| 12:45 | 2.6 | 5.1 | 5.7 | 290 | 310 | 310 |




(a)




(b)

Fig. 1. – (a) Wind standard deviations ($\sigma_i$); (b) Lagrangian decorrelation time scales ($\tau_{L_i}$) from Copenhagen experiment (12:33h-12:53h, 19/10/1978).

logarithmic profile, namely

$$U(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right), \tag{5}$$

where $u_*$ is the local friction velocity, $\kappa = 0.4$ is the von Karman constant, and $z_0$ the roughness length. Equation (5) is used to compute $u_*$ at the three levels. Wind speed for all levels in the domain was also computed from eq. (5). The wind direction was linearly interpolated from experimental values. The mixing height $z_i$, Obukhov length $L$, and wind standard deviations were only given as values averaged at each 1 hour and, therefore, were kept constant during the simulation: $z_i = 1120\,\mathrm{m}$, $L = -71\,\mathrm{m}$, $\sigma_v = 0.85\,\mathrm{ms}^{-1}$ and $\sigma_w = 0.68\,\mathrm{ms}^{-1}$. In this experiment the emission rate was constant and equal to $3.2\,\mathrm{g/s}$.

The horizontal LAMBDA domain was determined according to maximum sampler distance and the vertical domain was set equal to $z_i$. The time step was maintained constant ($\Delta t = 1\,\mathrm{s}$). One hundred particles were released at each time step during 2880 time steps. The Gram-Charlier probability density function (PDF) truncated to the third order was chosen to represent the PDF of the vertical wind velocity [16, 17].

Since the meteorological information is available at every 10 min, in the simulation all the meteorological parameters, except $z_i$ and $L$, were allowed to vary each 10 min. The simulation starts 28 min before the second sampling period, in order to better mime the experimental conditions (thus allowing the particles previously emitted to "fill" the computation domain as it occurred with the tracer in the experiment). To better estimate the ground level concentration (g.l.c.) in this non-stationary case, g.l.c. is computed 10 times (each 2 min) during each 20 min period, and then the average is computed.

The model performance is shown in table II and figs. 2a and b. Table II presents the comparison between observed ($C^{\mathrm{Exp}}$) and predicted ($C^{\mathrm{Mod}}$) values of ground-level concentration in the samplers. Figure 2a shows the scatter diagram between observed and predicted concentrations. Table III shows the result of a standard statistical analysis made with the data presented in table II. The statistical indices are the following:

| | |
|---|---|
| Normalized Mean Square Error, | $\mathrm{NMSE} = \overline{(C^{\mathrm{Exp}} - C^{\mathrm{Mod}})^2}/\overline{C^{\mathrm{Exp}} C^{\mathrm{Mod}}}$; |
| Fractional Bias, | $\mathrm{FB} = (\overline{C^{\mathrm{Exp}}} - \overline{C^{\mathrm{Mod}}})/(0.5(\overline{C^{\mathrm{Exp}}} + \overline{C^{\mathrm{Mod}}}))$; |
| Fractional Standard Deviation, | $\mathrm{FS} = 2(\sigma^{\mathrm{Exp}} - \sigma^{\mathrm{Mod}})/(\sigma^{\mathrm{Exp}} + \sigma^{\mathrm{Mod}})$; |
| Correlation Coefficient, | $\mathrm{COR} = \overline{(C^{\mathrm{Exp}} - \overline{C^{\mathrm{Exp}}})(C^{\mathrm{Mod}} - \overline{C^{\mathrm{Mod}}})}/\sigma^{\mathrm{Exp}}\sigma^{\mathrm{Mod}}$; |
| Factor 2, | $\mathrm{FA2} = 0.5 \leq C^{\mathrm{Exp}}/C^{\mathrm{Mod}} \leq 2$, |

where $\sigma$ are concentration standard deviations.

From fig. 2b, the observed and computed g.l.c. at each sampler of the three arcs is shown, suggesting that the model was able to capture the general behaviour of the diffusion experiment. However, the simulated plume is wider than the observed one. Thus the maxima at each arc are lower than the observed ones, whereas at the further samplers the model computes more concentration than measured.

## 5. – Results for the emission rate estimation

In the estimation of the emission rates it was assumed that the pollutant emission rate is changing with time, *i.e.* $Q = Q(t)$. Actually, in the Copenhagen experiment, the emission rate is constant and equal to $3.2\,\mathrm{gs}^{-1}$. The pollutant source term is estimated by

TABLE II. – *Ground-level concentration* ($C^{\mathrm{Exp}}$) *measured during the Copenhagen experiment (12:33h-12:53h, 19/10/1978) and predicted concentrations* ($C^{\mathrm{Mod}}$) *by LAMBDA model using turbulence parameterisation given by Degrazia et al.* [15].

| Number | $X$ (m) | $Y$ (m) | $C^{\mathrm{Exp}}$ ($\mu$g/s) | $C^{\mathrm{Mod}}$ ($\mu$g/s) |
|---|---|---|---|---|
| 1 | 1398 | −1312 | 0 | 0.721 |
| 2 | 1404 | −1214 | 0 | 1.350 |
| 3 | 1492 | −1131 | 0 | 2.271 |
| 4 | 1516 | −1044 | 0.614 | 3.266 |
| 5 | 1582 | −964 | 1.816 | 4.478 |
| 6 | 1592 | −884 | 5.455 | 5.105 |
| 7 | 1602 | −798 | 7.016 | 5.727 |
| 8 | 1703 | −767 | 6.770 | 5.582 |
| 9 | 1766 | −681 | 5.472 | 5.077 |
| 10 | 1800 | −593 | 3.806 | 4.067 |
| 11 | 1877 | −485 | 1.114 | 2.842 |
| 12 | 1921 | −405 | 0.919 | 1.969 |
| 13 | 2067 | −371 | 0.077 | 1.902 |
| 14 | 2061 | −284 | 0 | 1.001 |
| 15 | 2055 | −180 | 0 | 0.463 |
| 16 | 2818 | −2134 | 0 | 1.400 |
| 17 | 2920 | −1987 | 0.107 | 1.579 |
| 18 | 3002 | −1830 | 0.840 | 2.635 |
| 19 | 3075 | −1704 | 1.478 | 2.691 |
| 20 | 3204 | −1629 | 3.133 | 2.580 |
| 21 | 3380 | −1367 | 2.563 | 2.249 |
| 22 | 3448 | −1231 | 2.225 | 2.033 |
| 23 | 3518 | −1093 | 0.538 | 1.282 |
| 24 | 3558 | −919 | 0 | 1.050 |
| 25 | 3729 | −787 | 0 | 0.361 |
| 26 | 3837 | −550 | 0 | 0.121 |
| 27 | 4027 | −3616 | 0 | 0.306 |
| 28 | 4283 | −3447 | 0.018 | 0.788 |
| 29 | 4390 | −3277 | 0.021 | 1.087 |
| 30 | 4459 | −3010 | 0.085 | 1.117 |

TABLE II. – *Continued.*

| | | | | |
|---|---|---|---|---|
| 31 | 4572 | −2795 | 0.800 | 1.552 |
| 32 | 4668 | −2514 | 1.502 | 1.236 |
| 33 | 4824 | −2260 | 2.035 | 1.297 |
| 34 | 5029 | −2108 | 1.112 | 1.063 |
| 35 | 5286 | −1939 | 0.434 | 0.778 |
| 36 | 5378 | −1570 | 0.053 | 0.124 |
| 37 | 5395 | −1399 | 0 | 0.049 |
| 38 | 5375 | −1139 | 0 | 0.109 |
| 39 | 5323 | −913 | 0 | 0.047 |

minimizing the following functional:

$$F(\lambda, \mathbf{Q}) = \sum_{i=1}^{40} \left[ C_i^{\mathrm{Mod}}(\mathbf{Q}) - C_i^{\mathrm{Exp}} \right]^2 + \lambda \sum_{k=1}^{5} \left( Q_{k+1} - 2Q_k + Q_{k-1} \right)^2, \tag{6}$$

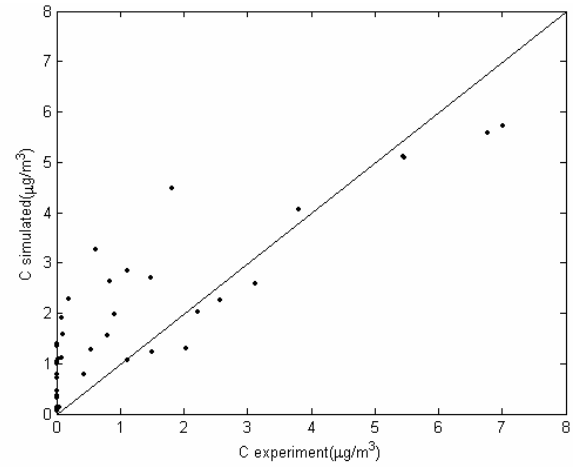
$C^{\mathrm{Mod}}$ being the concentration computed from the Lagrangian model, and $C_i^{\mathrm{Exp}}$ the concentration measured by the $i$-th sensor. The time period for our estimation was 50 minutes. The source term $Q(t)$ is assumed to be constant into the period $t \in [t, t + 10\,\mathrm{min}]$. Therefore, the unknown source term can be represented by the vector: $\mathbf{Q} = [Q_1, Q_2, Q_3, Q_4, Q_5]^{\mathrm{T}}$, where $Q_i = Q(t_0 + i\Delta t)$ with $\Delta t = 10\,\mathrm{min}$.

In order to complete the inverse analysis, the regularization parameter $\lambda$ must be determined. This parameter was found by using the procedure proposed by Hansen [18]: some value on the corner of the L-curve, relating "Regularization operator term × Square Difference" of the concentration— the second and first terms of the right-hand side of eq. (6), respectively. Figure 3 displays the L-curve for different regularization parameters, using the Tikhonov-2 regularization scheme. From the plot, it can be seen that $\lambda = 0.6$ is a good choice.

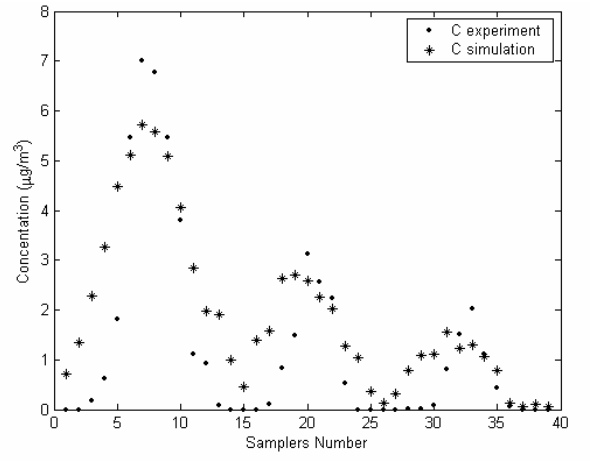
The estimation results are presented in table IV and fig. 4, using the second-order Tikhonov regularization scheme. It can be realized that the regularization is the very important issue in the inversion procedure. The estimation without regularization produced wrong estimation, indicating that the least-square scheme itself is not able to correctly address the model to find good answers. However, a bad choice of the regularization parameter (for example, $\lambda = 40$) will not produce a good inversion either. Therefore, for $\lambda \to 0$ and $\lambda \to \propto$ implies a wrong inverse procedure. The appropriate inverse regular-

TABLE III. – *Statistical indices of the LAMBDA model performance for the Copenhagen experiment simulated.*

| NMSE | COR | FA2 | FB | FS |
|---|---|---|---|---|
| 0.51 | 0.87 | 0.36 | −0.38 | 0.19 |




(a)




(b)

Fig. 2. – (a) Scatter diagram between observed and predicted samplers ground-level concentrations; (b) experimental and predicted samplers ground-level concentrations (the samplers at the three arcs are sequentially plotted) from Copenhagen experiment (12:33h-12:53h, 19/10/1978).

TABLE IV. – *Emission rate estimation.*

| Time (h:min) | $Q^{\text{true}}$ ($\text{gs}^{-1}$) | $Q^{\text{estimated}}$ ($\text{gs}^{-1}$) | | |
|---|---|---|---|---|
| | | $\lambda = 0$ | $\lambda = 0.6$ | $\lambda = 40$ |
| 12:05 | 3.20 | 3.57 | 3.185 | 0.655 |
| 12:15 | 3.20 | 1.79 | 3.185 | 0.595 |
| 12:25 | 3.20 | 4.82 | 3.187 | 0.502 |
| 12:35 | 3.20 | 2.74 | 3.188 | 0.607 |
| 12:45 | 3.20 | 0.25 | 3.188 | 0.704 |

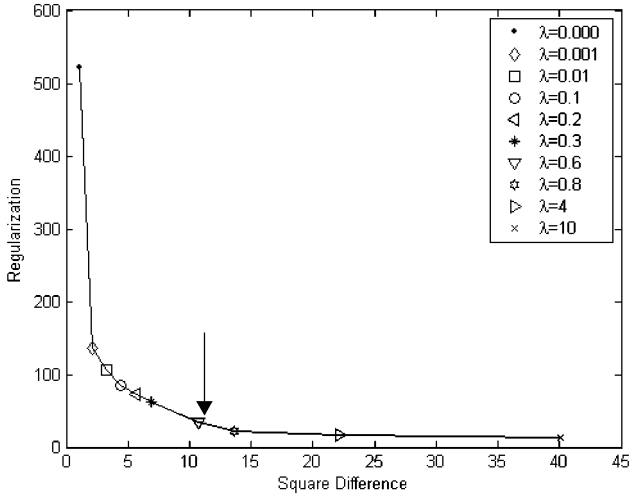



Fig. 3. – L-curve for reconstruction of emission rate.

ized solution is only obtained by using a correct value for the regularization parameter, indicating a good balance between the data adhesion (the square difference term in expression (6)) and the smoothness (regularization) of the estimated function. Table IV shows that the inversion with regularization parameter $\lambda = 0.6$ presents good results.

## 6. – Conclusions

An ineluctable feature with dealing with experimental data is the noise associated of such data. Inverse problems belong to the class of ill-posed problems, where they

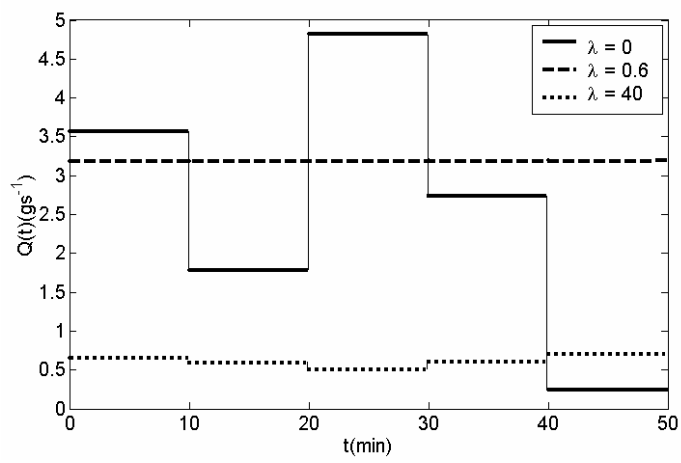



Fig. 4. – Pollutant source $Q(t)$ estimation: effect of the regularization parameter. Without regularization ($\lambda = 0$): oscillatory behavior; strong regularization ($\lambda = 40$), only regularization operator is focused in the minimization process; appropriate value ($\lambda = 0.6$), where $Q(t) \approx 3.2\,\mathrm{gs}^{-1}$.

are unstable in the presence of noise (small variation in the input data implies a wide variation in the output data). The regularized inverse solution is a strategy to give a good answer [5]. The methodology used for estimating the emission rate was effective to produce correct reconstructions of the emission rate. In this paper, these results were obtained using the second-order Tikhonov regularization.

In this paper it was shown that the inverse procedure yielded good results, even dealing with data from the real field (the Copenhagen tracer experiment), where some level of noise in the data is expected. However, one important issue is the determination of the regularization parameter. The L-curve scheme [18] permitted to find out an appropriated value for regularization operator.

* * *

The authors would like to thank to the Brazilians financial support agencies. The first author acknowledges CNPq-Brasil (Proc. 20.0046/04-7) for the sandwich doctorate fellowship at ISAC-CNR—Turin—Italy.

REFERENCES

[1] Seibert P., in *Inverse Methods in Global Biogeochemical Cycles*, edited by Kasibhatla P. *et al.* (American Geophysical Union, Washington) 2000, pp. 147-154.
[2] Ferrero E., Anfossi D., Brusasca G. and Tinarelli G., *Int. J. Environ. Pollut.*, **5** (1995) 360.
[3] Carvalho J. C., Degrazia G. A., Anfossi D., Campos C. R. J de, Roberti D. R. and Kerr A., *Meteorol. Z.*, **11** (2002) 89.
[4] Thomson D. J., *J. Fluid Mech.*, **180** (1987) 529.
[5] Tikhonov A. N. and Arsenin V. I., *Solutions of Ill-posed Problems* (John Wiley & Sons) 1977.
[6] Ramos F. M. and Giovannini A., *Int. J. Heat Mass Transfer*, **38** (1995) 101.
[7] Ramos F. M. and Campos Velho H. F., in *2nd International Conference on Inverse Problems on Engineering, Le Croisic, France, 1996*, pp. 199-206.
[8] Stephany S., Ramos F. M., Campos Velho H. F. de and Mobley C. D., *Comput. Modelling Simul. Eng.*, **3** (1998) 161.
[9] Muniz W. B., Campos Velho H. F. de and Ramos F. M., *J. Comput. Appl. Math.*, **110** (1999) 153.
[10] Campos Velho H. F. de, Moraes M. R. de, Ramos F. M., Degrazia G. A. and Anfossi D., *Nuovo Cimento C*, **23** (2000) 65.
[11] Muniz W. B., Ramos F. M. and Campos Velho H. F., *Comput. Math. Appl.*, **40** (2000) 1071.
[12] Roberti D. R., Campos Velho H. F. and Degrazia G. A., *Inverse Probl. Sci. Eng.*, **12** (2004) 329.
[13] Gryning S. E. and Lyck E., *J. Clim. Appl. Meteorol.*, **23** (1984) 651.
[14] Gryning S. E. and Lyck E., *The Copenhagen tracer experiments: reporting of measurements*. Available on line at: `http://www.risoe.dk/rispubl/VEA/ris-r-1054.htm`.
[15] Degrazia G. A., Anfossi D., Carvalho J. C., Mangia C., Tirabassi T. and Campos Velho H. F., *Atmos. Environ.*, **34** (2000) 3575.
[16] Ferrero E. and Anfossi D., in *Air Pollution Modelling and its Applications XII*, edited by Gryning S. E. and Chaumerliac N. (Plenum Press, New York) 1998, pp. 673-680.
[17] Ferrero E. and Anfossi D., *Int. J. Environ. Pollut.*, **9** (1998) 384.
[18] Hansen P. C., *SIAM Rev.*, **34** (1992) 561.