Inverse Problems, Design and Optimization Symposium
Rio de Janeiro, Brazil, 2004

Generalized Discrepancy Principle

Élcio H. Shiguemori, Haroldo F. de Campos Velho
José Demisio S. da Silva

Laboratory for Computing and Applied Mathematics – LAC
National Institute for Space Research – INPE
São José dos Campos, SP, Brazil
[elcio, haroldo, demisio]@lac.inpe.br

ABSTRACT

The Morosov's discrepancy principle has been used as a general criterion to compute the regularization parameter in inverse problems. The Morosov's principle is established when the discrepancy of the corresponding regularized solution is just equal to the measurement error. Considering the measurement error as a random variable, the goal of this paper is to present a generalized discrepancy principle for distributions in which the second moment is not defined. The generalized discrepancy principle is applied to several distributions: uniform, Gaussian, Cauchy, t-Student, Tsallis. The estimation of the initial condition in a heat transport problem is used as a test-problem.

INTRODUCTION

The Morosov's discrepancy principle [1-3] has been used as a general criterion to compute the regularization parameter in inverse problems. Essentially, the criterion is to calculate the root of the equation:

$$\left\| K\left(x^{a(d)}\right) - Y_d \right\|_Y = d \qquad (1)$$

being $K(x)$ the forward model, Y_d is the measured quantity, d is the error of the measurement, and $\|\cdot\|_Y$ is a norm in the Y space. A heuristic procedure can be established to compute the root of equation (1) when the measured error is normally distributed (the probability density function is Gaussian), with zero mean and s^2 variance (see [2], page 238). In this work a generalized discrepancy principle is presented, considering distributions in which the second moment is not defined. The generalized discrepancy principle is applied to several distributions: uniform, Gaussian, Cauchy, t-Student, Tsallis. The estimation of the initial condition in a heat transport problem is used as a test-problem.

FORWARD TEST PROBLEM

The direct (forward) problem consists of a transient heat conduction problem in a slab with adiabatic boundary condition and initially at a temperature denoted by *f(x)*. The mathematical formulation of this problem is given by the following heat equation

$$\begin{aligned}
&\frac{\partial T(x,t)}{\partial t} = \frac{\partial^2 T(x,t)}{\partial x^2} && (x,t)\in \Lambda\times\Re^+ \\
&\frac{\partial T(x,t)}{\partial x} = 0 && (x,t)\in \partial\Lambda\times\Re^+ \qquad (2)\\
&T(x,0) = f(x) && (x,t)\in \Lambda\times\{0\}
\end{aligned}$$

where *T(x,t)* (temperature), *f(x)* (initial condition), x (spatial variable) and t (time variable) are dimensionless quantities and $\Lambda=[0,1]$. The set of partial differential equations is solved by using a central finite difference approximation for space variable $O(Dx^2)$, and explit Euler method for numerical time integration $O(Dt)$ [3].

The forward problem solution, for a given initial condition $f(x)$, is explicitly obtained using separation of variables, for $(x,,t) \in \Omega \times R^+$:

$$T(x,t) = \sum_{m=0}^{+\infty} e^{-b_m^2 t} \frac{1}{N(b_m)} X(b_m,x) \int_0^1 X(b_m,x') f(x') dx' \qquad (3)$$

where $X(b_m,x) = \cos(b_m x)$ are the *eigenfunctions* associated to the problem, $b_m = mp$ are the *eigenvalues*, and $N(b_m)= \int_\Omega X(b_m,x') f(x') dx'$

represents the integral normalization (or the *norm*) [4]. The inverse problem consists in estimating the initial temperature profile $f(x)$ for a given transient temperature distribution $T(x,t)$ at time t.

This problem has been extensively used for testing different methodologies in inverse problems [5–9], and it is badly conditioned problem [5].

INVERSE ANALYSIS

In general, inverse problems belong to the class of ill-posed problems, where existence, uniqueness and stability of their solutions cannot be ensured. Following the Tikhonov's approach [10], a regularized solution is obtained by choosing the function f^* that minimizes the following functional

$$J_a\left[\tilde{T}, f\right] = \left\|\tilde{T} - T(f)\right\|_2^2 + a\Omega[f] \qquad (3)$$

where $\tilde{T} = \tilde{T}(x,t)$ is the experimental data ($t=t$), $T(f)$ is the temperature computed from the forward model at time t, $\Omega[f]$ denotes the regularization term given, a is the regularization parameter, and $\|\cdot\|_2$ is the 2-norm.

A scheme to determine the regularization parameter a is the Morozov's discrepancy principle: assuming that a bound d (or the 'statistics') of the measurement error is known, i.e., $\left\|T_{\text{exact}} - \tilde{T}\right\|_2 \leq d$.

The Morosov's Discrepancy Principle

For establishing a scheme to compute the regularization parameter, it is necessary to define the quantites: the residue $R(f_a)$ and the error $E(f_a)$ are defined by

$$R(f_a) = \left\|\tilde{T} - T(f_a)\right\|_2^2 \qquad (4)$$

.

$$E(f_a) = \left\|f_a - f_{\text{exact}}\right\|_2^2 \qquad (5)$$

The Morosov's standard discrepancy principle is an *a-posteriori* parameter choice rule. It demands that a suitable regularized solution can be obtained under the condition:

$$R(f_*) \cong N_x s^2 \qquad (6)$$

corresponding the *optimum* value for a - the regularization parameter, and assuming that s^2 is the variance associate to a Gaussian distribution. The last hypothesis can be justified by *central limit theorem* [11], and considering that the components of the random vector are uncorrelated (*white Gaussian noise*). The condition (6) is a particular case of the Morosov's discrepancy principle.

Optimization Algorithm

The optimization problem is iteratively solved by the quasi-newtonian optimizer routine from the NAG Fortran Library [12], with variable metrics. This algorithm is designed to minimize an arbitrary smooth function subject to constraints (simple bound, linear or nonlinear constraints), using a sequential programming method.

This routine has been successfully used in several previous works: in geophysics, hydrologic optics, and meteorology.

ESTIMATING INITIAL CONDITION

Numerical experiments are carried out using two test functions, the triangular function

$$f(x) = \begin{cases} 2x & x \in [0, 0.5] \\ 2(1-x) & x \in (0.5, 1] \end{cases} \qquad (6)$$

and semi-triangular function

$$f(x) = \begin{cases} 0.55 & 0 \leq x \leq 0.2 \\ 8/3\,x & 0.2 < x < 0.5 \\ -28/5\,x + 23/5 & 0.5 < x < 0.75 \\ 2/9 & 0 < x \leq 1 \end{cases} . \qquad (7)$$

In order to simulate the experimental data (measured temperatures at a time τ>0), which intrinsically should contain errors, a random perturbation is added to the exact solution of the direct problem, such that

$$\tilde{T} = T_{\text{exact}} + s\,n \qquad (8)$$

where s is the standard deviation of the errors and n is a random variable taken from a statistical distribution, with zero mean and unitary variance.

All tests were carried out using 5% of noise (s=0.05).

It is important to observe that the spatial grid consists of 101 points (N_x=101), and the time-integration is performed up to t=0.01. If we effectively want to apply some kind of regularization, which means a>0 in Eq. (5), then the discrepancy principle - an *a-posteriori* parameter choice rule - implies that a suitable regularized solution can be obtained. Since the spatial resolution is N_x=101, the *optimum* a is reached for $R(f_*) \sim N_x s^2 = 0.2525$ (according to the condition (6)).

The parameter vector was always subjected to the following simple bounds: $1.2 \geq f_k \geq -0.2$ for the triangular test function, and $1.2 \geq f_k \geq 0$ for the semi-triangular test function, with $k = 1, 2, \ldots, N_x$.

Probability Density Functions (PDF)

For generating the random variable in Eq. (8), several distributions have been considered.

- *Uniform distribution*:

$$r(x) = u(x; c, d) = \begin{cases} 1/(c-d) & \text{for } c \leq x \leq d \\ 0 & \text{otherwise} \end{cases} \qquad (9)$$

- *Normal (Gaussian) distribution*:

$$r(x) = \frac{1}{s\sqrt{2p}} e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2} \qquad (10)$$

where m and s are mean and standard deviation, respectively.

- *Cauchy's distribution*:

$$r(x) = \frac{s}{p\left((x-t)^2 + s^2\right)} \qquad (11)$$

where t is the location parameter and s is the scale parameter. The case where t=0 and s=1 is called the *standard Cauchy distribution*, and the PDF reduces to

$$r(x) = \frac{1}{p(1+x^2)} \qquad (12)$$

- *Student's t distribution*:

$$r(x) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\sqrt{pa}\,\Gamma\left(\frac{a}{2}\right)\left(1+\frac{x^2}{a}\right)^{\frac{a+1}{2}}} \qquad (13)$$

where a is a parameter, and Γ is the Gamma function.

- *Tsallis's distribution*:

A non-extensive form of entropy has been proposed by Tsallis [15]:

$$S_q(\boldsymbol{p}) = \frac{k}{q-1}\left(1 - \sum_{i=1}^{N_p} p_i^q\right) \qquad (14)$$

where p_i is a probability, and q is a free parameter – it is called the *non-extensivity parameter*, and the parameter q has a central role in Tsallis' thermostatiscs. In thermodynamics the parameter k is known as the Boltzmann's constant. Tsallis' entropy reduces to the the usual Boltzmann-Gibbs-Shanon formula

$$S(p) = -k \sum_{i=1}^{N_p} p_i \ln p_i \qquad (15)$$

in the limit $q \rightarrow 1$.

As for extensive form of entropy, the equiprobability condition produces the maximum for the non-extensive entropy function, and this condition leads to special distributions:

<u>$q > 0$</u>:

$$r(x) = \frac{1}{s}\left[\frac{q-1}{p(3-q)}\right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{3-q}{2(q-1)}\right)} \frac{1}{\left(1+\frac{q-1}{3-q}\frac{x^2}{s^2}\right)^{\frac{1}{q-1}}}$$

(16)

<u>$q = 0$</u>:

$$r(x) = \frac{1}{s}\left[\frac{1}{2p}\right]^{\frac{1}{2}} e^{-(x/s)^2/2} \qquad (17)$$

<u>$q < 0$</u>:

$$r(x) = \frac{1}{s}\left[\frac{1-q}{p(3-q)}\right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{5-3q}{2(1-q)}\right)}{\Gamma\left(\frac{2-q}{(1-q)}\right)} \left[1 - \frac{(1-q)}{(3-q)}\frac{x^2}{s^2}\right]^{\frac{1}{(1-q)}}$$

(18)

if $x < s\,[(3-q)/(1-q)]^{1/2}$, $r(x)$=0 otherwise. These PDFs are shown in Figure 1.

For q<5/3, the standard central limit theorem applies, implying that if p_i is written as a sum of M random independent variables, in the limit case $M\to\infty$, the probability density function for p_i in the distribution space is the *normal* (Gaussian) distribution [6]. However, for 5/3<q<3 the Levy-Gnedenko's central limit theorem applies, resulting for $M\to\infty$ the Levy distribution as the probability density function for the random variable p_i . The index in such Levy distribution is g=(3-q)/(q-1) [16].

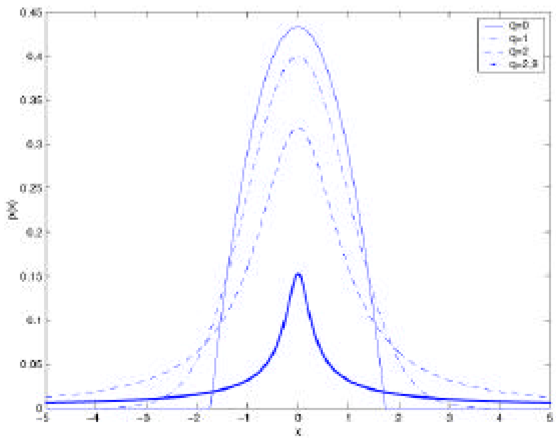


Figure 1. Tsallis' Distribution with q=0, q=1, q=2 and q=2.9.

Generalized Discrepancy Principle

Should be noted that Cauchy's distribution (11), Studend's *t* distribution, and Tsallis's distribution (with q > 5/3), them second statistical moments diverge ($s^2 \to \infty$).

Firstly, all distributions are normalized (when this is possible) for s^2=1. For distributions with divergent variance, a modified PDF is considered, in a such way that a parameter ***d*** is chosen for satisfying the relation

$$\int_{-d}^{d} x^2\, r(x)\,dx = 1\,. \qquad (19)$$

The PDF and this *new* domain [-*d* , *d*] is now used for generating the random variable n in Eq. (8), and the discrepancy principle can be applied for any distribution type.

Numerical Results

The numerical experiments were carried out for many realizations (10). For each experiment, new different random numbers were generated. For all cases the deviation was assumed as s=0.05, as mentioned before, and the discrepancy principle (6) was used to compute the regularization parameter.

Tables 1 and 2 show computed values for regularization parameter using the generalized discrepancy principle. Inverse solutions are depicted in figures below.

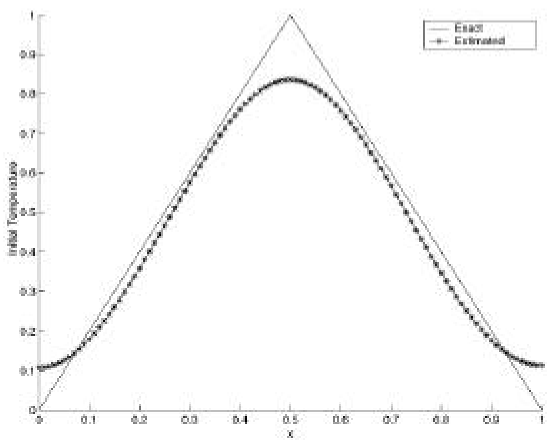
Table 1. Triangular test function

Distribution	$\bar{a}$	$\bar{E}(f_a)$	$sE(f_a)$
Uniform	722.57	0.3929	0.0396
Normal	719.42	0.3738	0.0327
Cauchy	685.69	0.3956	0.0499
Student's *t*	719.10	0.4112	0.0576
Tsallis (q=1.5)	648.49	0.3589	0.0594
Tsallis (q=0.5)	767.34	0.4268	0.0470

Table 2. Semi-triangular test function

Distribution	$\bar{a}$	$\bar{E}(f_a)$	$sE(f_a)$
Uniform	682.01	0.4698	0.0895
Normal	676.58	0.4571	0.0509
Cauchy	632.44	0.4777	0.1045
Student's *t*	677.19	0.5311	0.1197
Tsallis (q=1.5)	601.23	0.4580	0.0786
Tsallis (q=0.5)	737.94	0.5383	0.0726

- *Uniform distribution*:



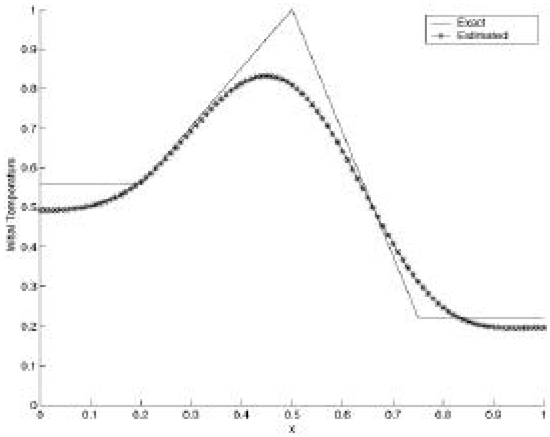


Figure 2. Initial condition estimation, with uniform distribution noise.

- *Normal (Gaussian) distribution*:

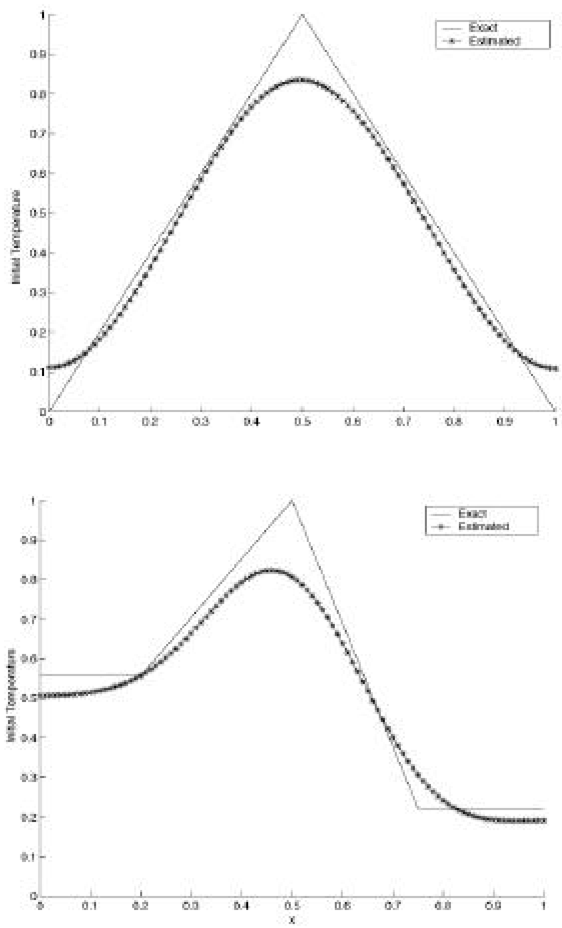


Figure 3. Initial condition estimation, with normal distribution noise.

- *Cauchy's distribution*:

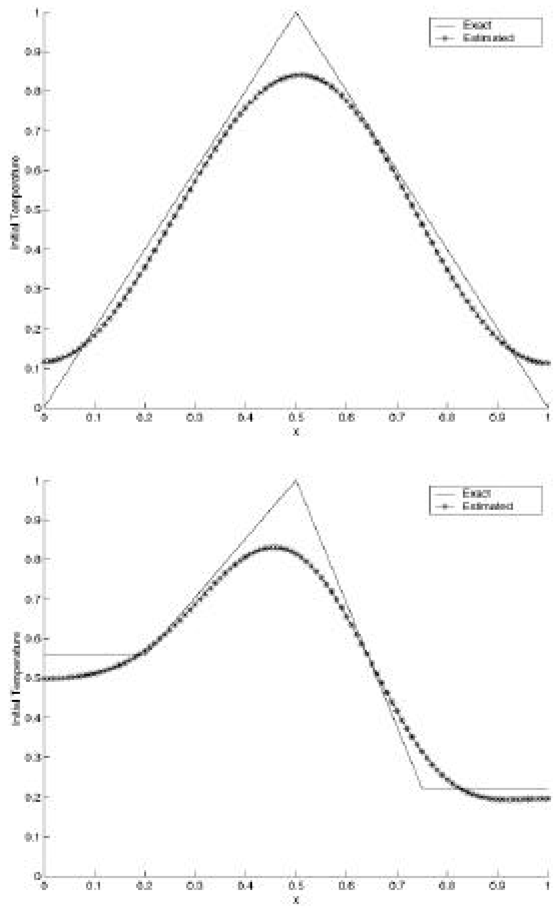
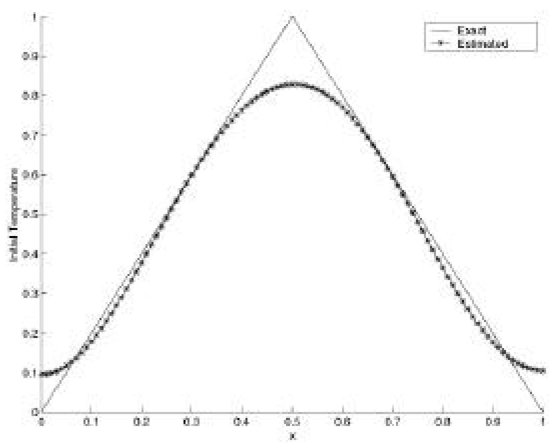


Figure 4. Initial condition estimation, with Cauchy distribution noise.

- *Tsallis's distribution*:



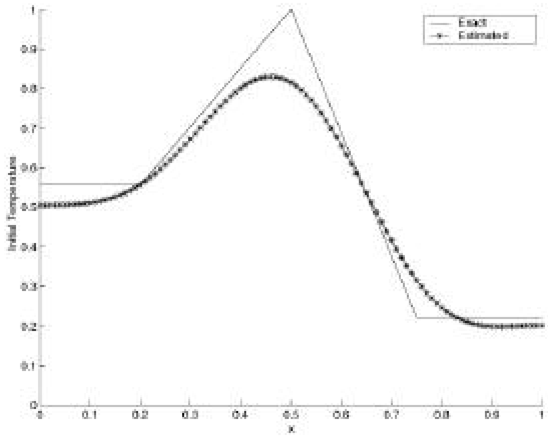


Figure 5. Initial condition estimation, with Tsallis distribution noise (q=1.5).

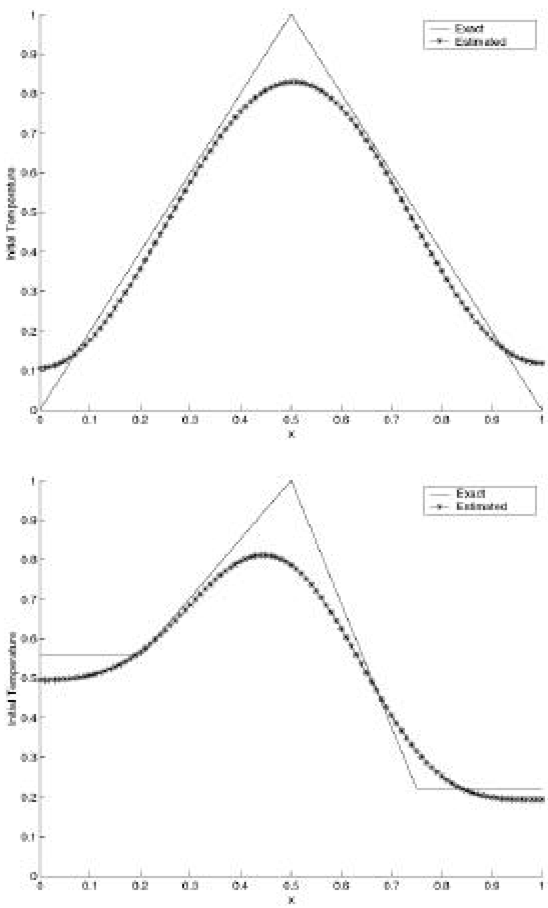


Figure 6. Initial condition estimation, with Tsallis distribution noise (q=0.5).

CONCLUSION

It was shown that the discrepancy principle can be adapted for situations where the error random variable can follow other statistical distributions instead of Gaussian distribution. Even those distributions that do not have a defined variance, the PDF of the random number generator can be modified, becoming possible to apply the discrepancy principle. For the worked examples, the regularization parameter can be determined as exposed, producing good inverse solutions.

ACKNOWLEDGEMENTS

The authors thanks to the FAPESP and CNPq, Brazilian agencies for research financial support, and an anonymous referee, whose comments improved this version of the paper.

REFERENCES

1. V. A. Morosov, On the Solution of Functional Equations by the Method of Regularization, *Soviet Math. Dokl.*, **7**, pp.414 (1966).
2. V. A. Morosov, *Methods for Solving Incorrectly Posed Problems*, Springer Verlag (1984).
3. V. A. Morozov, *Regularization Methods for Ill-Posed Problems*, CRC Press (1992).
4. J.D. Hoffmann, *Numerical Methods for Engineers and Scientists*, McGraw-Hill, 1993.
5. M. N. Özisik, *Heat Conduction*, Wiley Interscience (1980).
6. W.B. Muniz, F.M. Ramos, H.F. de Campos Velho, Entropy- and Tikhonov-based Regularization Techniques Applied to the Backwards Heat Equation, *Comput. Math. Appl.*, **40**(8/9), pp.1071 (2000).
7. W.B. Muniz, F.M. Ramos and H.F. de Campos Velho, Entropy- and Tikhonov-based regularization techniques applied to the backwards heat equation, *Comp. Math. Appl.*, **40**, pp.1071 (2000).
8. E.H. Shiguemori, H.F. de Campos Velho, J.D.S. da Silva (2002): Estimation of Initial Condition in Heat Conduction by Neural Network, *4th International Conference on Inverse Problems in Engineering: Theory and Practice* (ICIPE-2002), May 26-31, Angra dos Reis (RJ), Brazil - Proc. in CD-Rom: paper code 092 – Proc. Book: *Inverse Problems in Engineering: Theory and Pratice* (Editor Helcio R. B. Orlande), Vol. II, Part Heat and Mass Transfer, pp. 251 - [=] e-papers, Rio de Janeiro, Brazil (2002).
9. L.D. Chiwiacowsky, H.F. de Campos Velho, Different Approaches for the Solution of a

Backward Heat Conduction Problem, *Inverse Problems in Engineering*, **11**(3), pp.471 (2003).

10. L.D. Chiwiacowsky, H.F. de Campos Velho, A.J. Preto, S. Stephany, *Identifying Initial Condition in Heat Conduction Transfer by a Genetic Algorithm: A Parallel Approach*, *Iberian Latin-american Congress on Computational Methods in Engineering* (CILAMCE-2003), 29-31 October, Ouro Preto (MG), Brazil - Proc. in CD-Rom: paper code cil261-37, Abstract Book pp. 180 (2003).

11. A.N. Tikhonov and V.I. Arsenin, *Solutions of Ill-posed Problems*, John Wiley & Sons, 1977.

12. A. Papoulis, *Probability and Statistics*, Pearson Higher Education (1989).

13. E04UCF routine, NAG Fortran Library Mark 17, Oxford, UK (1995).

14. C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, *J. Statistical Physics*, **52**, 479 (1988).

15. C. Tsallis, Nonextensive statistics: theoretical, experimental and computational evidences and connections, *Braz. J. Phys.*, **29**, 1 (1999).