STATISTICAL SOLUTION OF INVERSE PROBLEMS

Helcio R. B. Orlande

Department of Mechanical Engineering, DEM/PEM
Escola Politécnica/COPPE
Federal University of Rio de Janeiro, UFRJ
Cid. Universitária, Cx. Postal: 68503
Rio de Janeiro, RJ, 21941-972
Brazil
helcio@mecanica.coppe.ufrj.br

INVERSE PROBLEM

MAXIMUM LIKELIHOOD OBJECTIVE FUNCTION

$$\boxed{S_{ML}(\mathbf{P}) = [\mathbf{Y} - \mathbf{T}(\mathbf{P})]^T \mathbf{W} [\mathbf{Y} - \mathbf{T}(\mathbf{P})]}$$

where $\mathbf{P}$ = vector of unknown parameters
$\mathbf{Y}$ = vector of measured temperatures
$\mathbf{T}(\mathbf{P})$ = vector of estimated temperatures
$\mathbf{W}$ = Inverse of the covariance matrix of the measurements

INVERSE PROBLEM

Hypotheses:

- The errors are additive, with zero mean and normally distributed.
- The statistical parameters describing the errors are known.
- There are no errors in the independent variables.
- There is no prior information about **P**.

INVERSE PROBLEM

THE LEVENBERG-MARQUARDT METHOD

$$\boxed{\mathbf{P}^{k+1} = \mathbf{P}^{k} + [\mathbf{J}^{T}\mathbf{W}\mathbf{J} + \lambda^{k}\mathbf{\Omega}^{k}]^{-1}\mathbf{J}^{T}\mathbf{W}[\mathbf{Y} - \mathbf{T}(\mathbf{P}^{k})]}$$

where λ^k is the *damping parameter* and $\mathbf{\Omega}^k$ is a *diagonal matrix.*

- The Levenberg-Marquardt Method is related to ***Tikhonov's regularization*** approach.
- Compromise between steepest-descent method and Gauss' method.
- Simple, powerful and straightforward iterative procedure.
- Capable of treating complex physical situations.
- Easy to program.
- Stable and converges fast.

INVERSE PROBLEM

Remark: With the statistical hypotheses described above, the minimization of the least-squares norm yields *maximum likelihood* estimates, that is, the values estimated for the unknown parameters **P** are those most likely to produce the measured data **Y**.

Remark: Although very popular and useful in many situations, the minimization of the least-squares norm is a non-Bayesian estimator. A Bayesian estimator is basically concerned with the analysis of the *posterior probability density*, which is the conditional probability of the parameters **P** given the measurements **Y.**

<u>INVERSE PROBLEM</u>

The **<u>statistical inversion approach</u>** is based on the following principles:

1. All variables included in the model are modeled as random variables.
2. The randomness describes our degree of information concerning their realizations.
3. The degree of information concerning these values is coded in the probability distributions.
4. The solution of the inverse problem is the posterior probability distribution.

• Jari P. Kaipio and Erkki Somersalo, *Computational and Statistical Methods for Inverse Problems*, Springer, 2004.

• S. Tan, C. Fox, G. Nicholls, *Inverse Problems*, Course Notes for Physics 707, University of Auckland

INVERSE PROBLEM

BAYES' FORMULA

$$\boxed{\pi_{posterior}(\mathbf{P}) = \pi(\mathbf{P}|\mathbf{Y}) = \frac{\pi_{prior}(\mathbf{P})\pi(\mathbf{Y}|\mathbf{P})}{\pi(\mathbf{Y})}}$$

Where: $\pi_{\text{posterior}}(\mathbf{P})$ = posterior probability density (conditional probability of the parameters **P** given the measurements **Y**)

$\pi_{\text{prior}}(\mathbf{P})$ = prior density (information about the parameters prior to the measurements)

$\pi(\mathbf{Y}|\mathbf{P})$ = likelihood function (expresses the likelihood of different measurement outcomes **Y** with **P** given)

$\pi(\mathbf{Y})$ = probability density of the measurements (normalizing constant)

$$\boxed{posterior \propto prior \times likelihood}$$

INVERSE PROBLEM

Hypotheses:

- The errors are additive, with zero mean and normally distributed.
- The statistical parameters describing the errors are known.
- There are no errors in the independent variables.
- **P** is a random vector with known mean μ and known covariance matrix **V**.
- **P** is distributed normally and is independent of **Y**.

MAXIMUM A POSTERIORI

Likelihood

$$\boxed{\pi(\mathbf{Y}|\mathbf{P}) = (2\pi)^{-I/2}\left|\mathbf{W}^{-1}\right|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{Y}-\mathbf{T})^T\mathbf{W}(\mathbf{Y}-\mathbf{T})\right]}$$

where I = number of observations
$\mathbf{W}$ = inverse of the covariance matrix of the measurements

For uncorrelated measurements:

$$\mathbf{W} = \begin{bmatrix} 1/\sigma_1^2 & & & 0 \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ 0 & & & 1/\sigma_I^2 \end{bmatrix}$$

MAXIMUM A POSTERIORI

Normal Prior

$$\boxed{\pi(\mathbf{P}) = (2\pi)^{-N/2}\left|\mathbf{V}\right|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{P}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{P}-\boldsymbol{\mu})\right]}$$

where N = number of parameters
$\boldsymbol{\mu}$ = known mean for **P**
V = known covariance matrix for **P**

Bayes' Formula: $$\boxed{\pi_{posterior}(\mathbf{P}) = \pi(\mathbf{P}\left|\mathbf{Y}\right.) \propto \pi_{prior}(\mathbf{P})\pi(\mathbf{Y}\left|\mathbf{P}\right.)}$$

MAXIMUM A POSTERIORI

$$\ln\left[\pi(\mathbf{P} \mid \mathbf{Y})\right] = -\frac{1}{2}\left[(I+N)\ln 2\pi + \ln|\mathbf{W}^{-1}| + \ln|\mathbf{V}| + S_{MAP}\right]$$

Maximum a Posteriori Objective Function

$$S_{MAP}(\mathbf{P}) = \left[\mathbf{Y} - \mathbf{T}(\mathbf{P})\right]^T \mathbf{W}\left[\mathbf{Y} - \mathbf{T}(\mathbf{P})\right] + (\mu - \mathbf{P})^T \mathbf{V}^{-1}(\mu - \mathbf{P})$$

MAXIMUM A POSTERIORI

For the minimization of $S_{MAP}(\mathbf{P})$:

$$\frac{\partial S_{MAP}(\mathbf{P})}{\partial P_1} = \frac{\partial S_{MAP}(\mathbf{P})}{\partial P_2} = \cdots = \frac{\partial S_{MAP}(\mathbf{P})}{\partial P_N} = 0$$

$$\boxed{-2\,\mathbf{J}^T\,\mathbf{W}[\mathbf{Y} - \mathbf{T}(\mathbf{P})] - 2\mathbf{V}^{-1}[\mu - \mathbf{P}] = 0}$$

where $\mathbf{J}$ is the *Sensitivity Matrix*.

MAXIMUM A POSTERIORI

$$-2\,\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P})]-2\mathbf{V}^{-1}[\mu-\mathbf{P}]=0$$

Linear Problems: $\mathbf{J}$ does not depend on $\mathbf{P}$ ⇨ $\mathbf{T}(\mathbf{P})=\mathbf{J}\,\mathbf{P}$

$$\boxed{\mathbf{P}=[\mathbf{J}^T\mathbf{W}\,\mathbf{J}+\mathbf{V}^{-1}]^{-1}[\mathbf{J}^T\mathbf{W}\,\mathbf{Y}+\mathbf{V}^{-1}\mu]}$$

Nonlinear Problems: $\mathbf{J}\equiv\mathbf{J}(\mathbf{P})$ ⇨ $\mathbf{T}(\mathbf{P})=\mathbf{T}(\mathbf{P}^k)+\mathbf{J}^k(\mathbf{P}-\mathbf{P}^k)$

$$\boxed{\mathbf{P}^{k+1}=\mathbf{P}^k+[\mathbf{J}^T\mathbf{W}\mathbf{J}+\mathbf{V}^{-1}]^{-1}\{\mathbf{J}^T\mathbf{W}[\mathbf{Y}-\mathbf{T}(\mathbf{P}^k)]+\mathbf{V}^{-1}(\mu-\mathbf{P}^k)\}}$$

SEQUENTIAL PARAMETER ESTIMATION TECHNIQUE

- Utilizes the measurements in a sequential manner in order to estimate the parameters.
- Avoids matrix inversions.
- Permits the identification of improper mathematical models.
- Possible to identify if a sufficient number of transient measurements and if a sufficiently long experimental time have been used in the experiment.

COMPUTATIONAL ALGORITHM FOR THE NONLINEAR CASE

Step 1. Initialize the iterative procedure by setting the iteration index k to 0 and making $\mathbf{P}^0 = \mu$.

Step 2. Compute the estimate for the vector of unknown parameters sequentially, for i=0,…,(I-1), by using

$$\mathbf{A} = \mathbf{V}_i \mathbf{J}_{i+1}^T$$

$$\Delta = \mathbf{J}_{i+1} \mathbf{A} + W_{i+1}^{-1}$$

$$\mathbf{K} = \mathbf{A}\Delta^{-1}$$

$$E_{i+1} = Y_{i+1} - T_{i+1}(\mathbf{P}^k)$$

$$\mathbf{P}_{i+1}^{k+1} = \mathbf{P}_i^{k+1} + \mathbf{K}[E_{i+1} - \mathbf{J}_{i+1}(\mathbf{P}_i^{k+1} - \mathbf{P}^k)]$$

$$\mathbf{V}_{i+1} = \mathbf{V}_i - \mathbf{K}\,\mathbf{J}_{i+1}\mathbf{V}_i$$

COMPUTATIONAL ALGORITHM FOR THE NONLINEAR CASE

Step 3. Check convergence of the values estimated sequentially with all I measurements

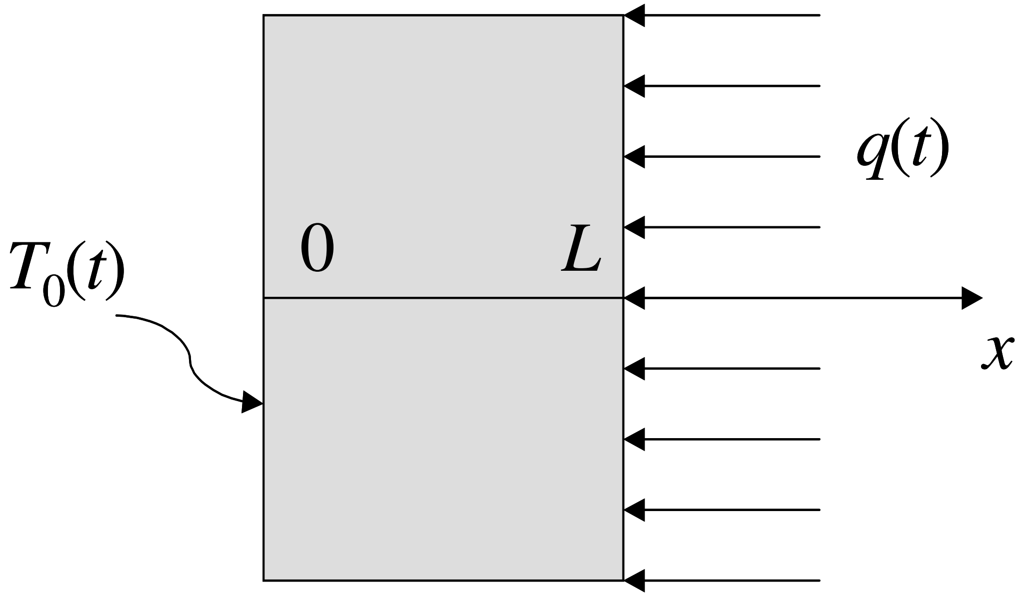
$$\left\| \mathbf{P}_I^{k+1} - \mathbf{P}_I^{k} \right\| < \varepsilon$$

If the convergence criterion is not satisfied, increment k, make

$$\mathbf{P}^k = \mathbf{P}_I^k$$

and return to step 2.

EXAMPLE 1



EXAMPLE 1

$$C(T)\frac{\partial T(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[k(T)\frac{\partial T}{\partial x}\right] \qquad \text{in } \; 0 < x < L,\, t > 0$$

$$T = T_0(t) \qquad \text{at } \; x = 0, \quad t > 0$$

$$k(T)\frac{\partial T}{\partial x} = q(t) \qquad \text{at } \; x = L, \quad t > 0$$

$$T = T_{ini} \qquad \text{for } t = 0, \text{ in } \; 0 < x < L$$

EXAMPLE 1

$$C(T)\frac{\partial T(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[k(T)\frac{\partial T}{\partial x}\right] \quad \text{in} \quad 0 < x < L,\ t > 0$$

$$T = T_0(t) \quad \text{at} \quad x = 0, \quad t > 0$$

$$k(T)\frac{\partial T}{\partial x} = q(t) \quad \text{at} \quad x = L, \quad t > 0$$

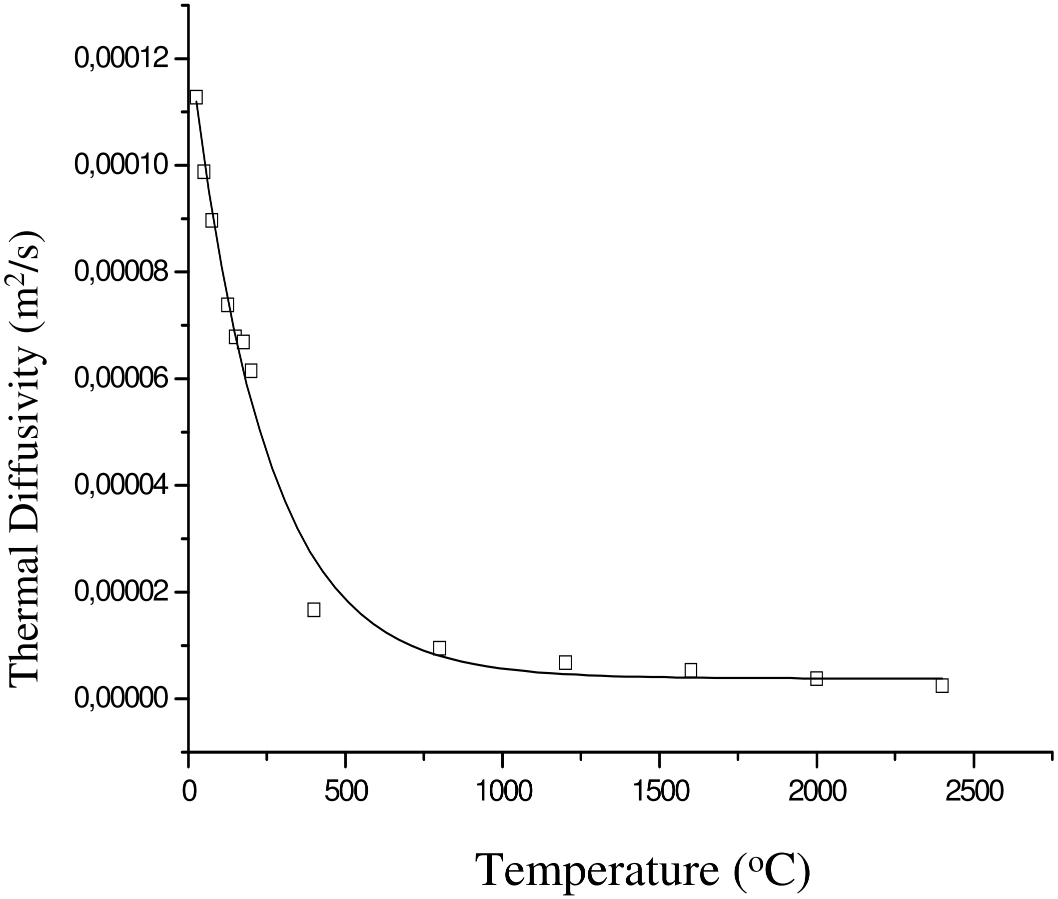
$$T = T_{ini} \quad \text{for } t = 0, \text{ in} \quad 0 < x < L$$

EXAMPLE 1

- Heat Flux: $q(t)$
- Thermal Conductivity: $k(T) = B_1 + B_2 e^{-T/B_3}$
- Volumetric Heat Capacity: $C(T) = A_1 + A_2 e^{-T/A_3}$

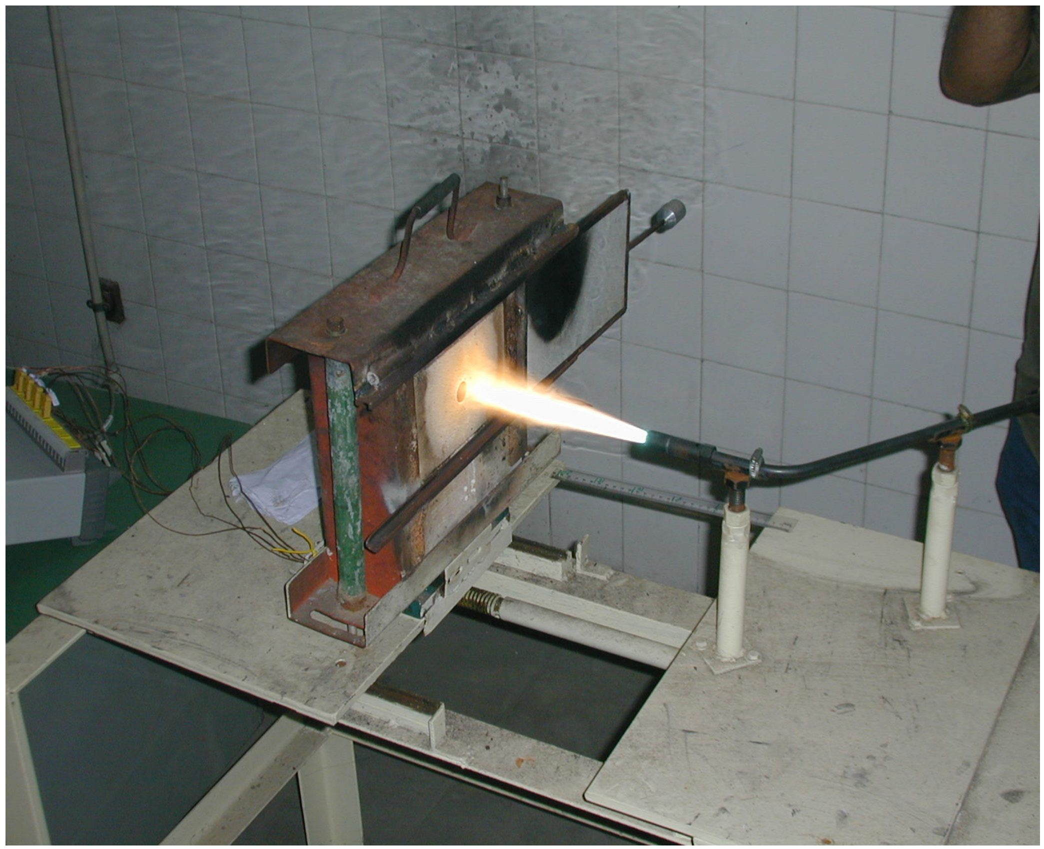
$$\boxed{\mathbf{P} = [q_1, q_2, \ldots, q_I, A_1, A_2, A_3, B_1, B_2, B_3]}$$

EXAMPLE 1



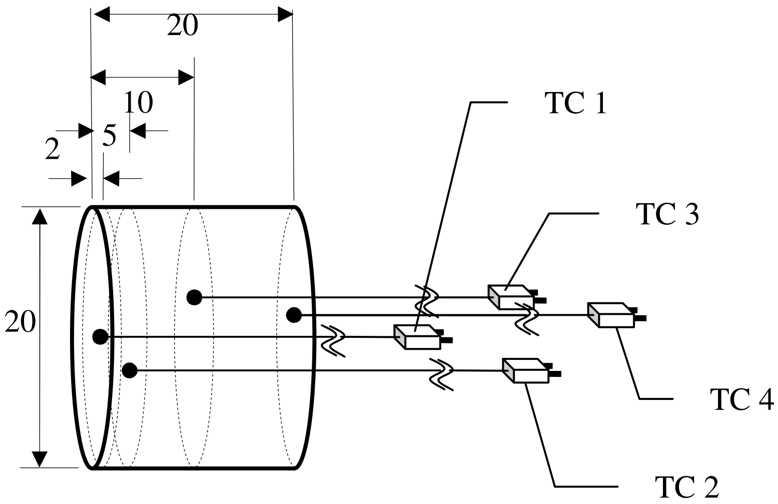
Thermal Diffusivity (m²/s)
0,00012
0,00010
0,00008
0,00006
0,00004
0,00002
0,00000
0
500
1000
1500
2000
2500
Temperature (°C)

EXAMPLE 1

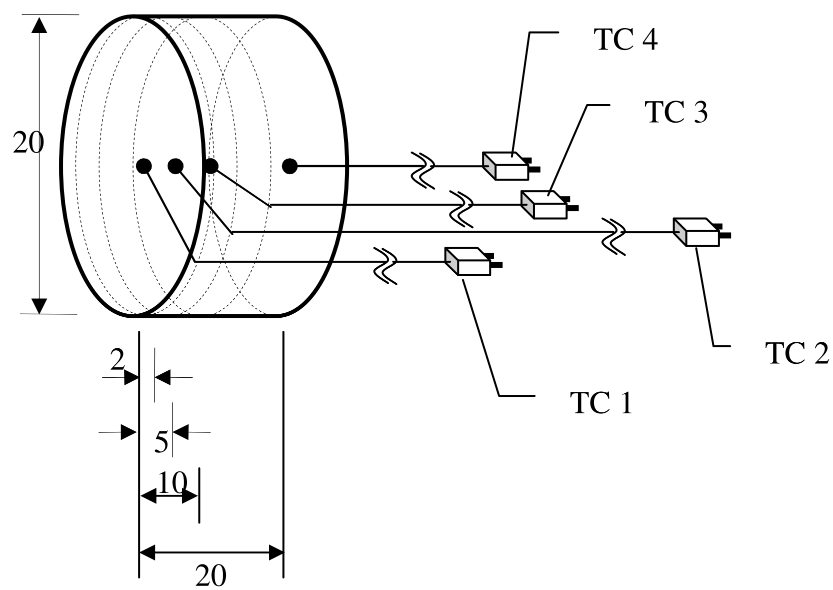


EXAMPLE 1

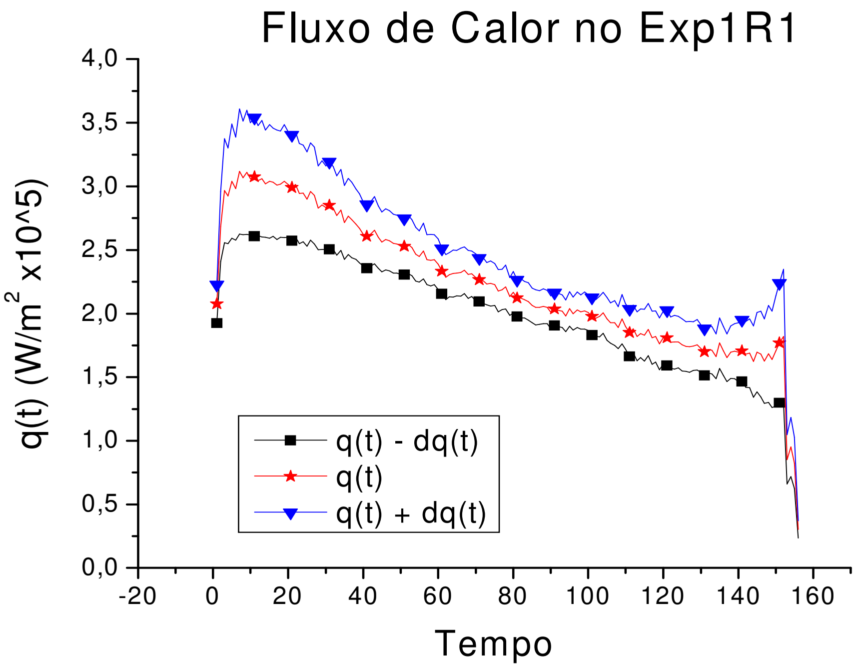
SAMPLE 1

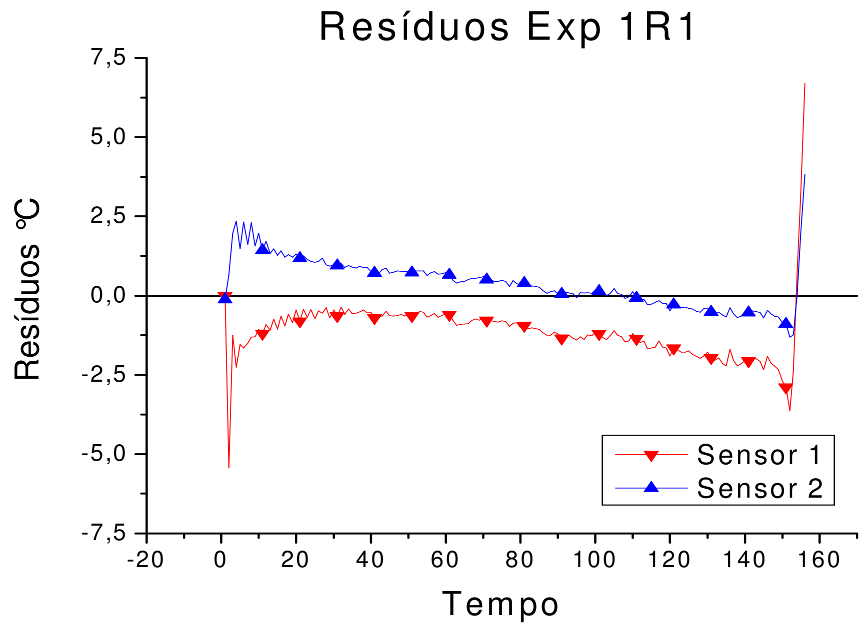


SAMPLE 2

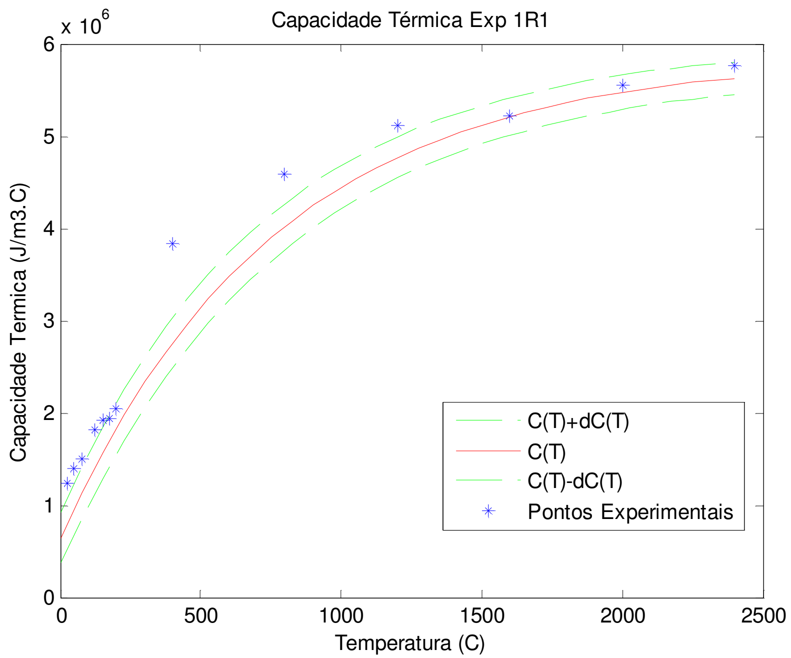


EXAMPLE 1

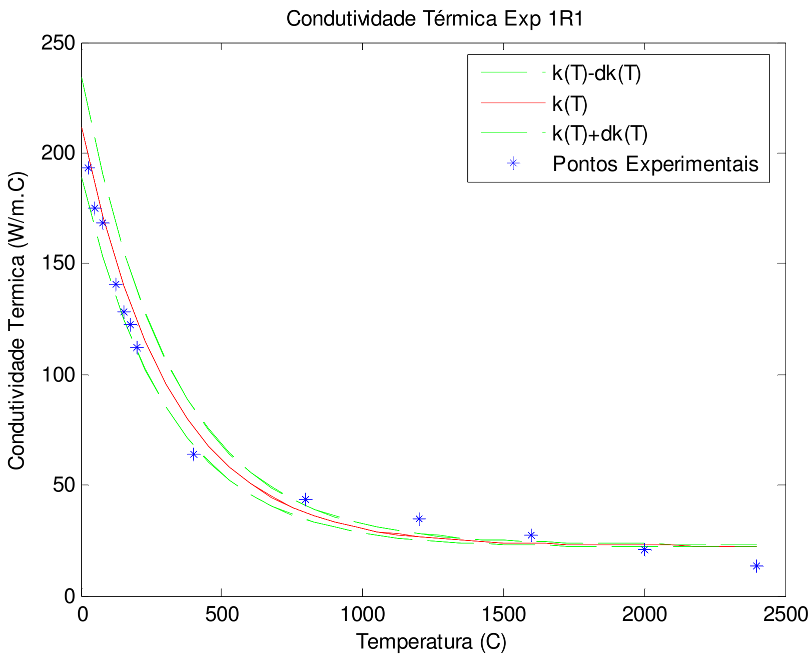

Fluxo de Calor no Exp1R1
q(t) (W/m2 x10^5)
Tempo
q(t) - dq(t)
q(t)
q(t) + dq(t)
4,0
3,5
3,0
2,5
2,0
1,5
1,0
0,5
0,0
-20
0
20
40
60
80
100
120
140
160


Resíduos Exp 1R1
Resíduos °C
Tempo
Sensor 1
Sensor 2
7,5
5,0
2,5
0,0
-2,5
-5,0
-7,5
-20
0
20
40
60
80
100
120
140
160

EXAMPLE 1

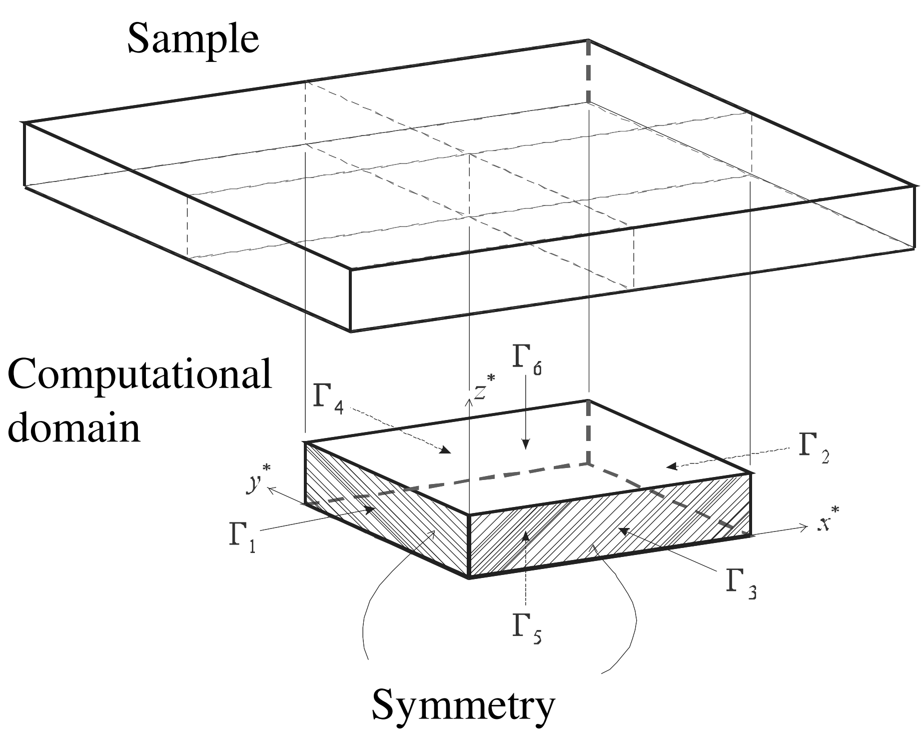


Capacidade Térmica Exp 1R1
x 10^6
6
5
4
3
2
1
0
0
500
1000
1500
2000
2500
Capacidade Termica (J/m3.C)
Temperatura (C)
C(T)+dC(T)
C(T)
C(T)-dC(T)
Pontos Experimentais

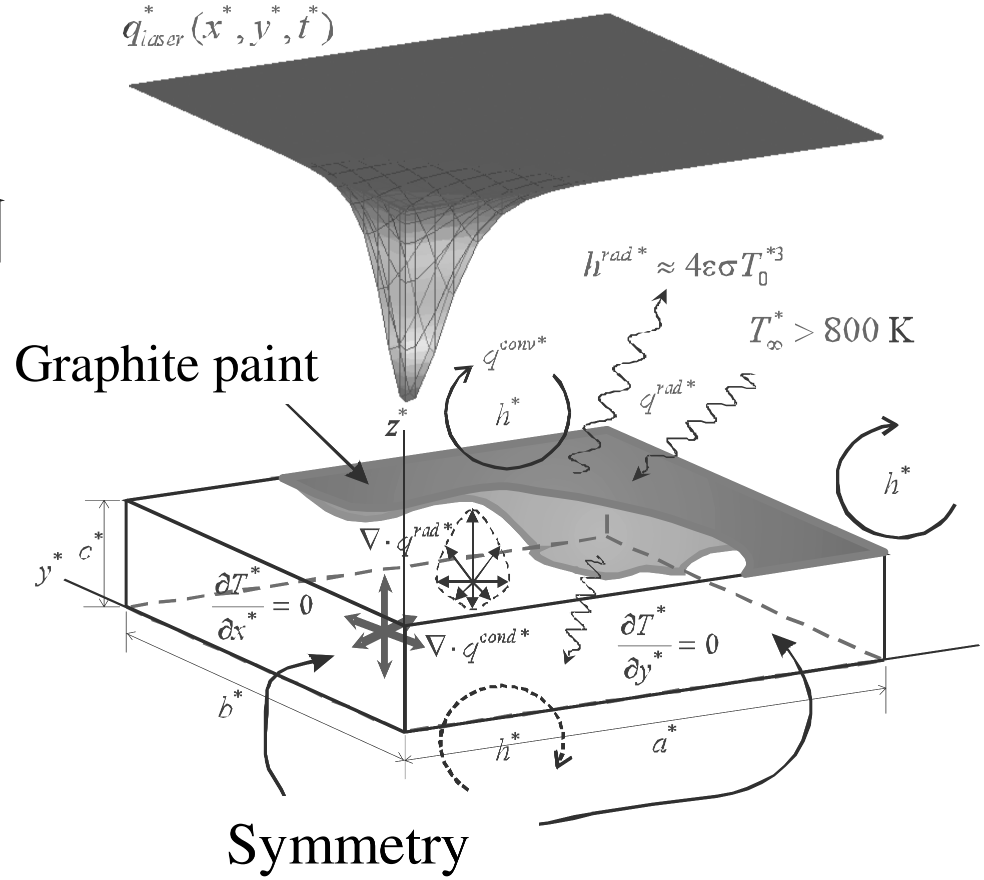


Condutividade Térmica Exp 1R1
250
200
150
100
50
0
0
500
1000
1500
2000
2500
Condutividade Termica (W/m.C)
Temperatura (C)
k(T)-dk(T)
k(T)
k(T)+dk(T)
Pontos Experimentais

EXAMPLE 2



Sample
Computational domain
Γ_1
Γ_2
Γ_3
Γ_4
Γ_5
Γ_6
x^*
y^*
z^*
Symmetry



$q^*_{laser}(x^*, y^*, t^*)$
$h^{rad*} \approx 4\varepsilon\sigma T_0^{*3}$
$T^*_\infty > 800$ K
Graphite paint
q^{conv*}
h^*
q^{rad*}
z^*
h^*
$\nabla \cdot q^{rad*}$
y^*
c^*
$\frac{\partial T^*}{\partial x^*} = 0$
$\nabla \cdot q^{cond*}$
$\frac{\partial T^*}{\partial y^*} = 0$
b^*
h^*
a^*
Symmetry

Equation of Radiative Transfer

$$\xi \frac{\partial I^l}{\partial x} + \eta \frac{\partial I^l}{\partial y} + \mu \frac{\partial I^l}{\partial z} = -\left(\kappa_a + \sigma_s\right) I^l + S^l \quad \text{in } 0 < x < a,\, 0 < y < b,\, 0 < z < c$$

where

$$S^l = \kappa_a n_r^2 I_b\left(T\right) + \frac{\sigma_s}{4\pi} \int_{\Omega' = 4\pi} I^{l'} p\left(\vec{s}\,' \rightarrow \vec{s}\right) d\Omega'$$

Boundary Conditions

$$I(\xi,\eta,\mu) = I(-\xi,\eta,\mu) \qquad \text{at } \Gamma_1 : \begin{cases} x = 0 \\ 0 < y < b \\ 0 < z < c \end{cases}$$

$$I(-\xi,\eta,\mu) = \varepsilon n_r^2 I_b + \frac{1-\varepsilon}{\pi} \int_{\xi'>0} I(\xi',\eta',\mu') \;\; \xi' \, d\Omega' \qquad \text{at } \Gamma_2 : \begin{cases} x = a \\ 0 < y < b \\ 0 < z < c \end{cases}$$

$$I(\xi,\eta,\mu) = I(\xi,-\eta,\mu) \qquad \text{at } \Gamma_3 : \begin{cases} 0 < x < a \\ y = 0 \\ 0 < z < c \end{cases}$$

$$I(\xi,-\eta,\mu) = \varepsilon n_r^2 I_b + \frac{1-\varepsilon}{\pi} \int_{\eta'>0} I(\xi',\eta',\mu') \;\; \eta' \;\; d\Omega' \qquad \text{at } \Gamma_4 : \begin{cases} 0 < x < a \\ y = b \\ 0 < z < c \end{cases}$$

$$I(\xi,\eta,\mu) = \varepsilon n_r^2 I_b + \frac{1-\varepsilon}{\pi} \int_{\mu'<0} I(\xi',\eta',\mu') \;\; \mu' \;\; d\Omega' \qquad \text{at } \Gamma_5 : \begin{cases} 0 < x < a \\ 0 < y < b \\ z = 0 \end{cases}$$

$$I(\xi,\eta,-\mu) = \varepsilon n_r^2 I_b + \frac{1-\varepsilon}{\pi} \int_{\mu'>0} I(\xi',\eta',\mu') \;\; \mu' \;\; d\Omega' \qquad \text{at } \Gamma_6 : \begin{cases} 0 < x < a \\ 0 < y < b \\ z = c \end{cases}$$

EXAMPLE 2

Energy Conservation Equation

$$C\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(k_x\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k_y\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(k_z\frac{\partial T}{\partial z}\right)-\nabla\cdot q^{rad}$$

in $0<x<a,\ 0<y<b,\ 0<z<c,$ for $t>0$

where: $$\nabla\cdot q^{rad}=\frac{\kappa_a\tau_0}{N_{\mathrm{pl}}}\left[4\pi n_r^2 I_b-\int_{\Omega=4\pi}I^l\,d\Omega\right]$$

Boundary Conditions

$$\frac{\partial T}{\partial x} = 0 \qquad \text{at } \Gamma_1 \text{ for } t > 0$$

$$k_x \frac{\partial T}{\partial x} + Bi^{rad} T = \frac{\varepsilon \tau_0}{N_{\mathrm{pl}}} \left[\int_{\xi > 0} I^l . \xi . d\Omega - n_r^2 \pi I_b \right] + Bi^{rad} T_\infty \qquad \text{at } \Gamma_2 \text{ for } t > 0$$

$$\frac{\partial T}{\partial y} = 0 \qquad \text{at } \Gamma_3 \text{ for } t > 0$$

$$k_y \frac{\partial T}{\partial y} + Bi^{rad} T = \frac{\varepsilon \tau_0}{N_{\mathrm{pl}}} \left[\int_{\eta > 0} I^l . \eta . d\Omega - n_r^2 \pi I_b \right] + Bi^{rad} T_\infty \qquad \text{at } \Gamma_4 \text{ for } t > 0$$

$$-k_z \frac{\partial T}{\partial z} + Bi^{rad} T = \frac{\varepsilon \tau_0}{N_{\mathrm{pl}}} \left[\int_{\mu < 0} I^l . \mu . d\Omega - n_r^2 \pi I_b \right] + Bi^{rad} T_\infty \qquad \text{at } \Gamma_5 \text{ for } t > 0$$

$$k_z \frac{\partial T}{\partial z} + Bi^{rad} T = \frac{\varepsilon \tau_0}{N_{\mathrm{pl}}} \left[\int_{\mu > 0} I^l . \mu . d\Omega - n_r^2 \pi I_b \right] + Bi^{rad} T_\infty + \varepsilon_{10.6\,\mu\mathrm{m}} q_{laser}(x, y, t) \qquad \text{at } \Gamma_6 \text{ for } t > 0$$

Initial Condition

$$T = 0 \qquad \text{in } 0 < x < a,\ 0 < y < b,\ 0 < z < c, \text{ for } t = 0$$

***Table 1*: Estimation techniques**

Technique	Objective Function	Method	Model for the Direct Problem	Model for the Gradient
1	Least-squares	Levenberg-Marquardt	Complete	Surrogate
2	Least-squares	Levenberg-Marquardt	Complete	Complete
3	Maximum a Posteriori	Gauss	Complete	Surrogate
4	Maximum a Posteriori	Gauss	Complete	Complete
5	Least-squares	1st step: Hybrid 2nd step: Levenberg-Marquardt	1st step: Surrogate 2nd step: Complete	1st step: Surrogate 2nd step: Surrogate
6	Least-squares	1st step: Hybrid 2nd step: Levenberg-Marquardt	1st step: Surrogate 2nd step: Complete	1st step: Surrogate 2nd step: Complete
7	1st step: Least-squares 2nd step: MAP	1st step: Hybrid 2nd step: Gauss	1st step: Surrogate 2nd step: Complete	1st step: Surrogate 2nd step: Surrogate
8	1st step: Least-squares 2nd step: MAP	1st step: Hybrid 2nd step: Gauss	1st step: Surrogate 2nd step: Complete	1st step: Surrogate 2nd step: Complete

__Table 2__: Results obtained with an initial guess close to the exact parameters

$\left(C^{*0} = 2.8\times10^{6}\ \mathrm{J/m^3.K},\ k_x^{*0} = k_y^{*0} = k_z^{*0} = 8\ \mathrm{W/m.K},\ h^{rad*0} = 800\ \mathrm{W/m^2.K}\right)$

Technique	Number of Iterations	CPU Time	Estimates				
			$C^{*}\times10^{6}$ $\mathrm{Jm^{-3}K^{-1}}$	k_x^{*} $\mathrm{Wm^{-1}K^{-1}}$	k_y^{*} $\mathrm{Wm^{-1}K^{-1}}$	k_z^{*} $\mathrm{Wm^{-1}K^{-1}}$	h^{rad*} $\mathrm{Wm^{-2}K^{-1}}$
1	16	5h35m18s	2.51 ± 0.03	4.99 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
2	16	6h14m04s	2.51 ± 0.03	5.00 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
3	13	4h33m16s	2.51 ± 0.03	5.00 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
4	6	2h36m5s	2.51 ± 0.03	5.00 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
5	50	1h17m26s	2.19	5.74	5.80	3.6	1246
	15	5h50m55s	2.51 ± 0.03	4.99 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
6	50	1h17m26s	2.19	5.74	5.80	3.6	1246
	16	6h33m02s	2.51 ± 0.03	5.00 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
7	50	1h17m26s	2.19	5.74	5.80	3.6	1246
	11	4h09m59s	2.51 ± 0.03	5.00 ± 0.07	5.01 ± 0.07	5.0 ± 0.2	1373 ± 5
8	50	1h17m26s	2.19	5.74	5.80	3.6	1246
	4	1h59m11s	2.51 + 0.03	5.00 − 0.07	5.01 + 0.07	5.0 + 0.2	1373 + 5

***Table 3**: Results obtained with an initial guess far from the exact parameters*

$\left(C^{*0} = 0.1\times10^6 \text{ J/m}^3\text{.K},\ \boldsymbol{k}_x^{*0} = \boldsymbol{k}_y^{*0} = \boldsymbol{k}_z^{*0} = 50 \text{ W/m.K},\ \boldsymbol{h}^{rad*0} = 5 \text{ W/m}^2\text{.K}\right)$

Technique	Number of Iterations	CPU Time	Estimates				
			$C^*\times10^6$ $\text{Jm}^{-3}\text{K}^{-1}$	k_x^* $\text{Wm}^{-1}\text{K}^{-1}$	k_y^* $\text{Wm}^{-1}\text{K}^{-1}$	k_z^* $\text{Wm}^{-1}\text{K}^{-1}$	h^{rad*} $\text{Wm}^{-2}\text{K}^{-1}$
1	NC	-	-	-	-	-	-
2	NC	-	-	-	-	-	-
3	NC	-	-	-	-	-	-
4	NC	-	-	-	-	-	-
5	50 21	1h19m46s 7h44m03s	2.04 2.51 ± 0.03	5.22 4.99 ± 0.07	5.26 5.01 ± 0.07	2.9 5.0 ± 0.2	1224 1373 ± 5
6	50 16	1h19m46s 6h32m44s	2.04 2.51 ± 0.03	5.22 5.00 ± 0.07	5.26 5.01 ± 0.07	2.9 5.0 ± 0.2	1224 1373 ± 5
7	50 10	1h19m46s 3h54m49s	2.04 2.51 ± 0.03	5.22 5.00 ± 0.07	5.26 5.01 ± 0.07	2.9 5.0 ± 0.2	1224 1373 ± 5
8	50 5	1h19m46s 2h22m53s	2.04 2.51 ± 0.03	5.22 5.00 ± 0.07	5.26 5.01 ± 0.07	2.9 5.0 ± 0.2	1224 1373 ± 5

SAMPLED SOLUTIONS TO INVERSE PROBLEMS

- In many cases, the Posterior Probability Distribution is analytically intractable, p. ex., if the prior probability distribution involves information which is difficult to express in analytic terms.
- Draw samples from the set $\mathbf{\Omega}$ of all possible **P**'s, each sample with probability $\pi(\mathbf{P}|\mathbf{Y})$.
- Get a set $\Theta = \{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_M\}$ of samples distributed like the posterior distribution.
- Inference on $\pi(\mathbf{P}|\mathbf{Y})$ becomes inference on $\Theta = \{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_M\}$, for example the mean of the samples in Θ give us an estimation of the average values of $\pi(\mathbf{P}|\mathbf{Y})$.
- We generally need the constant that normalizes the probability distribution to sample: **MARKOV CHAIN MONTE-CARLO METHODS**

METROPOLIS-HASTINGS ALGORITHM

- Draws a sample from a *Candidate Density* $q(\phi|\theta)$.
- If $\pi(\theta)\, q(\phi|\theta) > \pi(\phi)\, q(\theta|\phi)$ the process moves from θ to ϕ too often and from ϕ to θ too rarely.
- Introduce a probability $\alpha(\phi|\theta)$, called the *Probability of Moving* so that

$$\pi(\theta)\, q(\phi \mid \theta)\, \alpha(\phi \mid \theta) = \pi(\phi)\, q(\theta \mid \phi)$$

- At the same time, we do not want to reduce the number of moves from ϕ to θ and we make

$$\alpha(\phi \mid \theta) = \min\left[1, \frac{\pi(\phi)\, q(\theta \mid \phi)}{\pi(\theta)\, q(\phi \mid \theta)}\right]$$

METROPOLIS-HASTINGS ALGORITHM

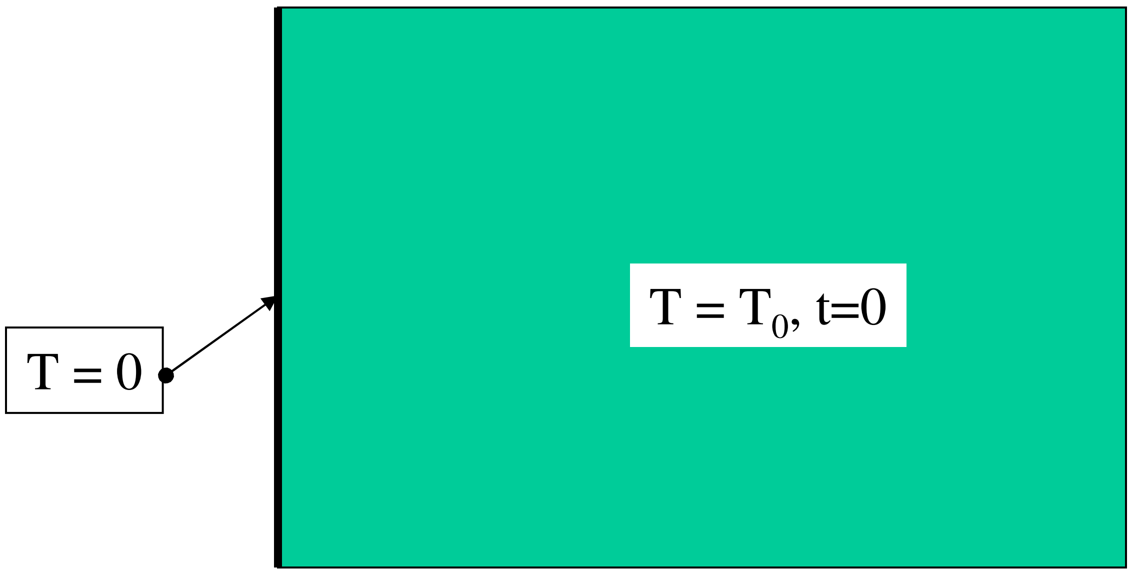
1. Sample a *Candidate Point* θ^* from a *Candidate Density* $q(\theta^* \mid \theta^{(t-1)})$.
2. Calculate:

$$\alpha = \min\left[1, \frac{\pi(\theta^*)\, q(\theta^{(t-1)} \mid \theta^*)}{\pi(\theta^{(t-1)})\, q(\theta^* \mid \theta^{(t-1)})}\right]$$

3. Generate a random value U which is uniformly distributed on (0,1).
4. If $U \leq \alpha$, define $\theta^{(t)} = \theta^*$; otherwise, define $\theta^{(t)} = \theta^{(t-1)}$.
5. Return the sequence $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(n)}\}$.

Remark: Ignore $\theta^{(i)}$ until the chain has reached equilibrium.

EXAMPLE



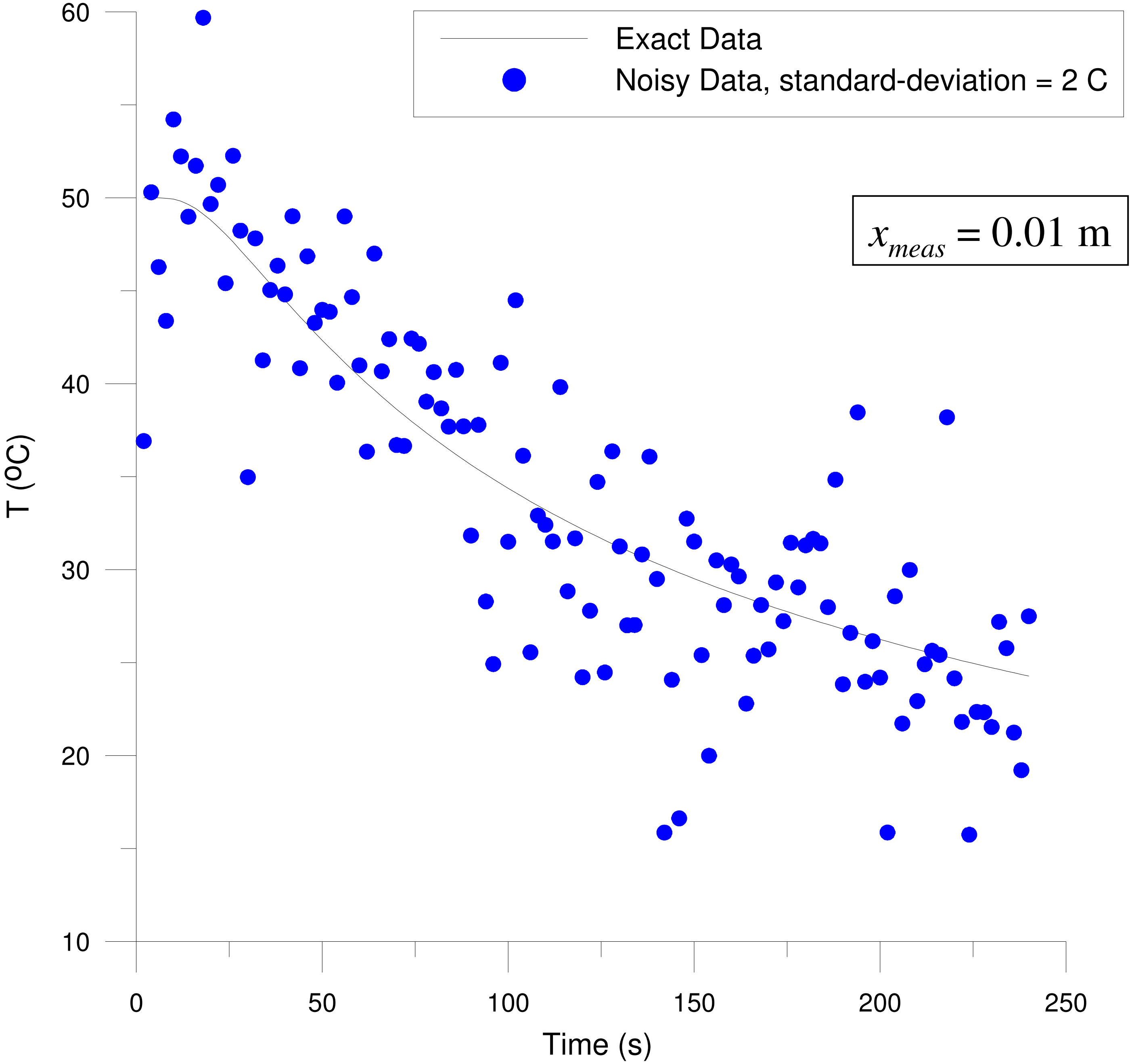
$$T(x,t) = T_0\, erf\left(\frac{x}{\sqrt{4\alpha t}}\right)$$

$T_0 = 50\ ^{o}C$

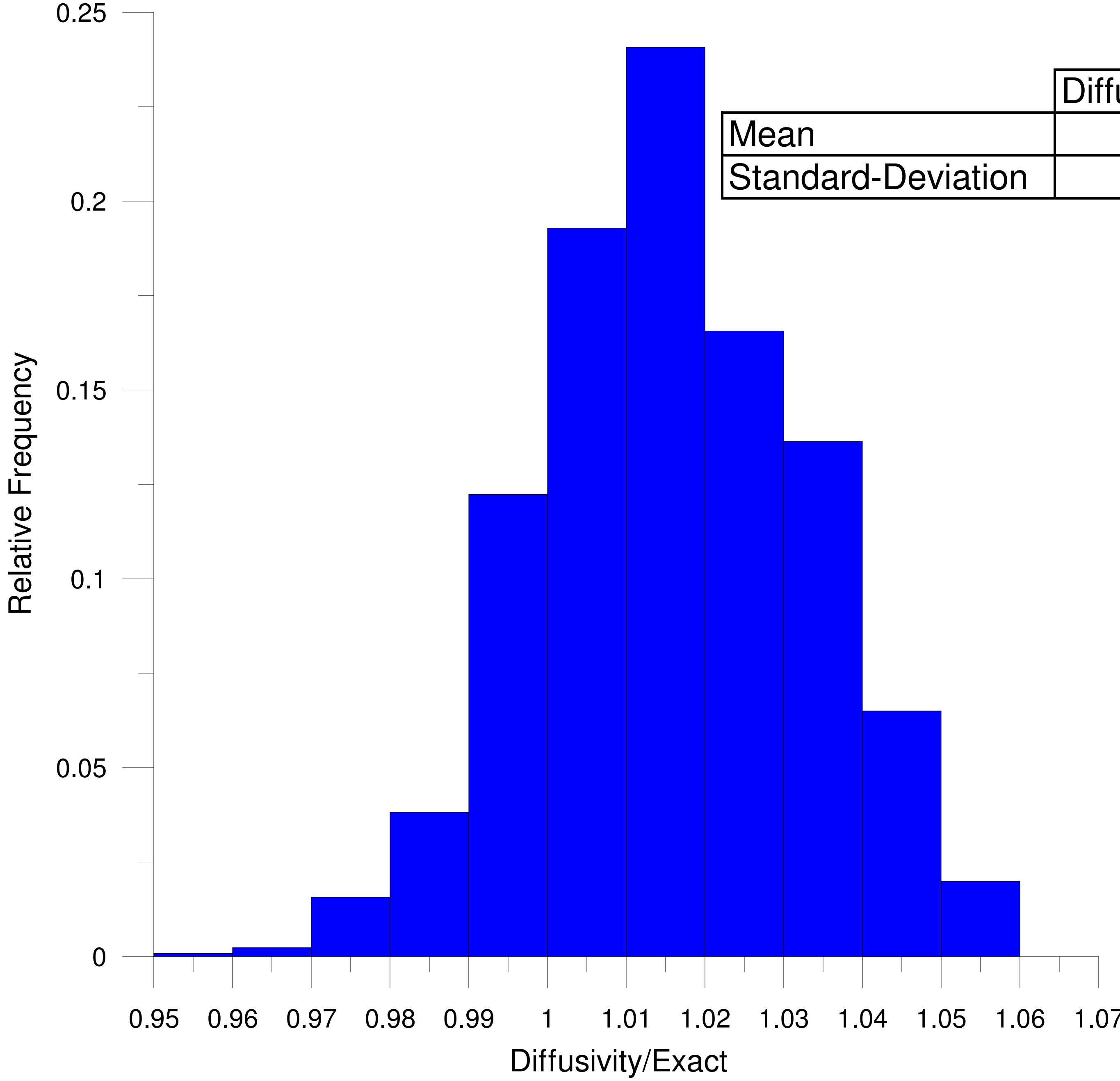
Concrete: $\alpha = 4.9 \times 10^{-7}\ m^2/s$

EXAMPLE

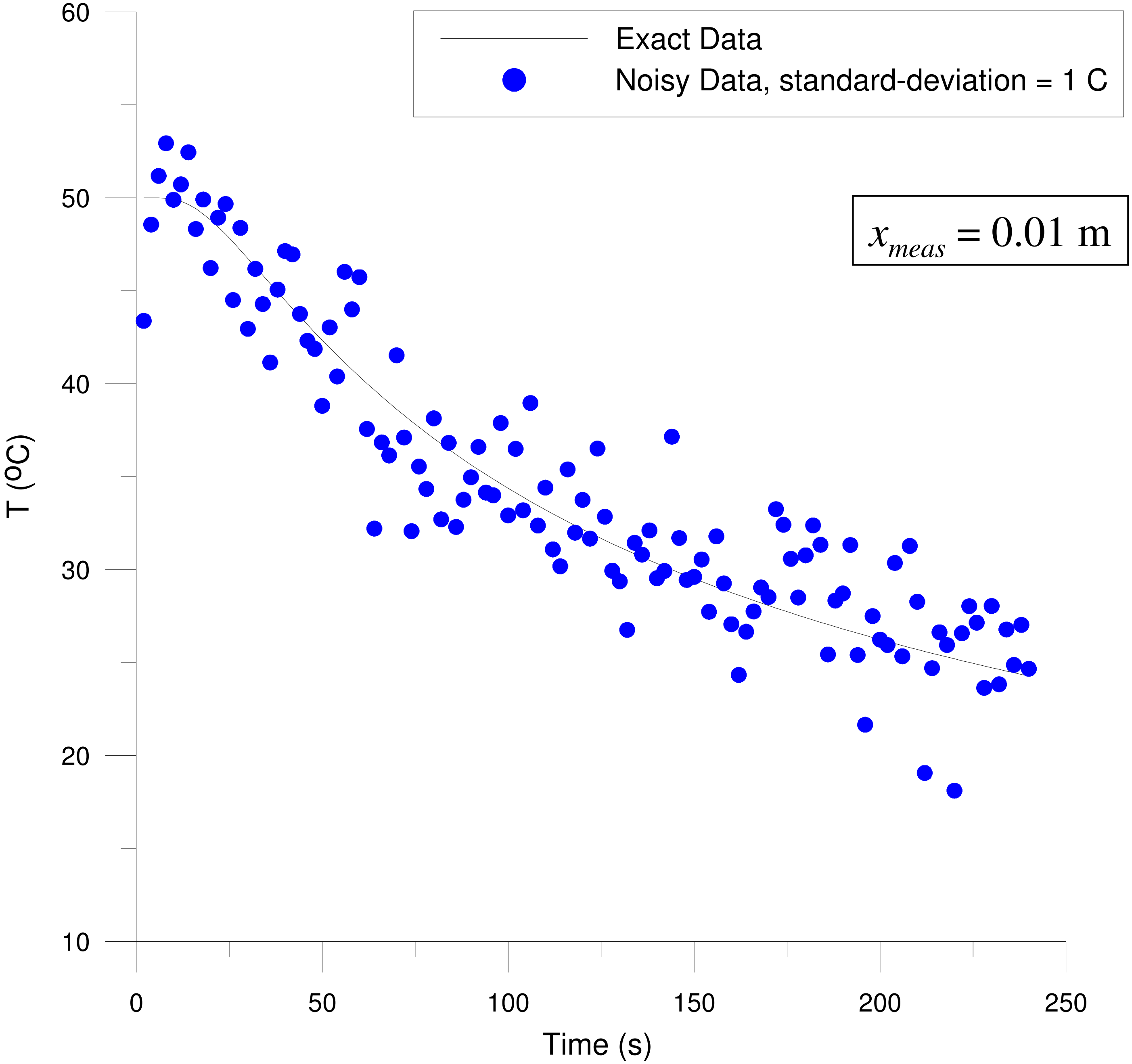
- Estimation of α.
- Prior for α: Uniform distribution (10^{-7},10^{-5}) m^2/s.
- Start the chain in the middle of the interval.

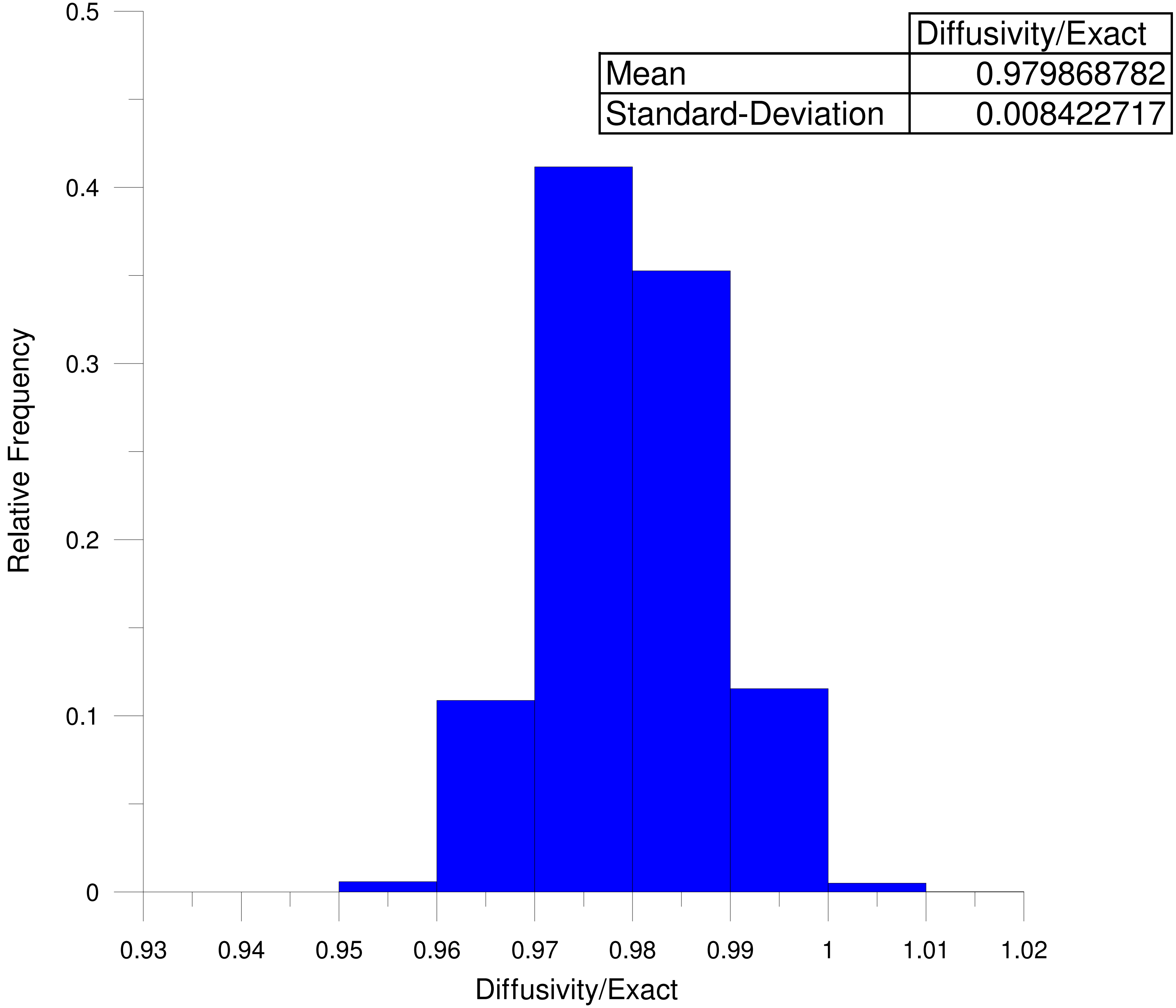


Exact Data
Noisy Data, standard-deviation = 2 C
$x_{meas} = 0.01$ m
T (oC)
Time (s)
60
50
40
30
20
10
0
50
100
150
200
250



	Diffusivity/Exact
Mean	1.016224796
Standard-Deviation	0.017660818

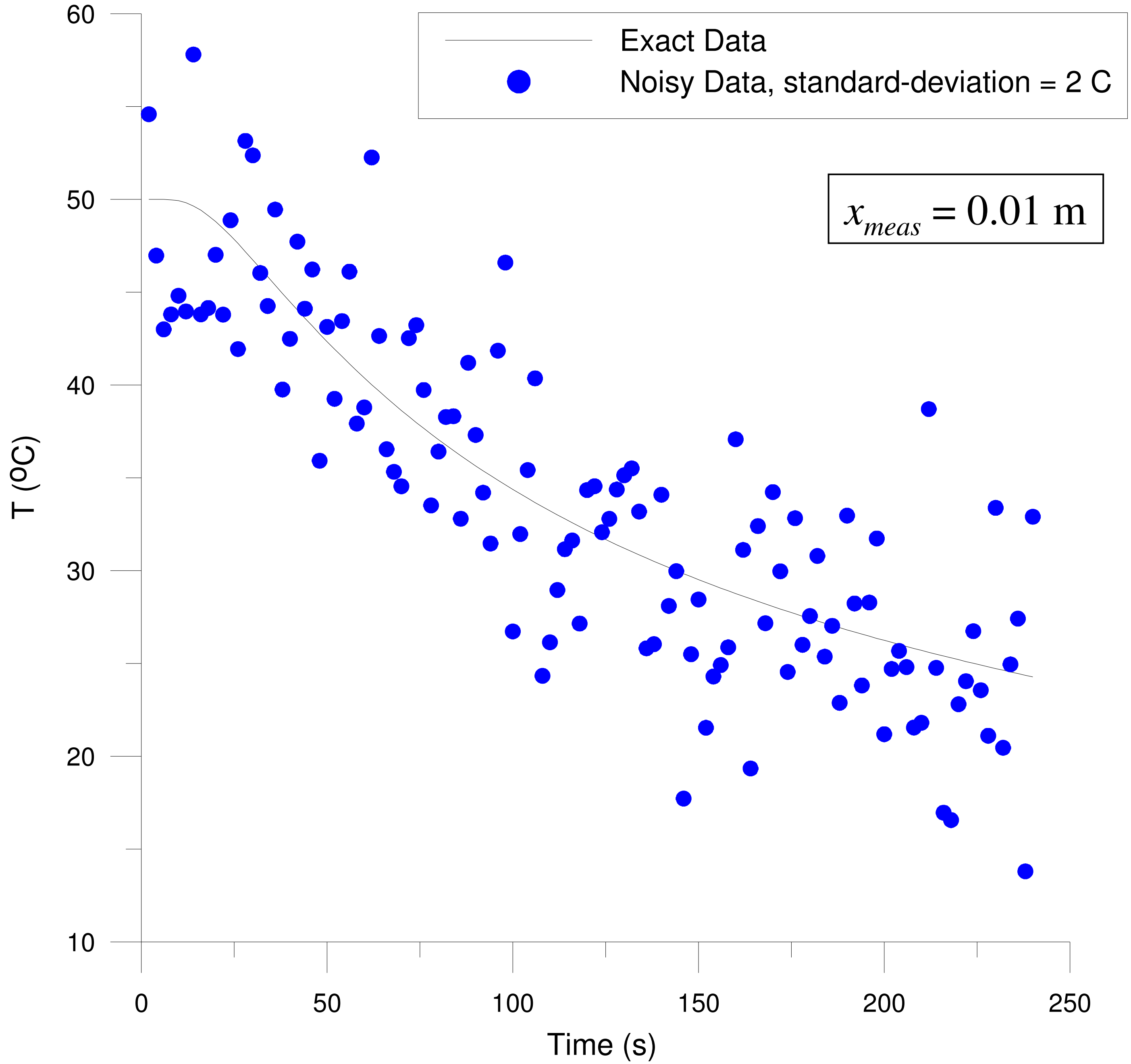

Exact Data
Noisy Data, standard-deviation = 1 C
$x_{meas} = 0.01$ m
T (ºC)
Time (s)
60
50
40
30
20
10
0
50
100
150
200
250

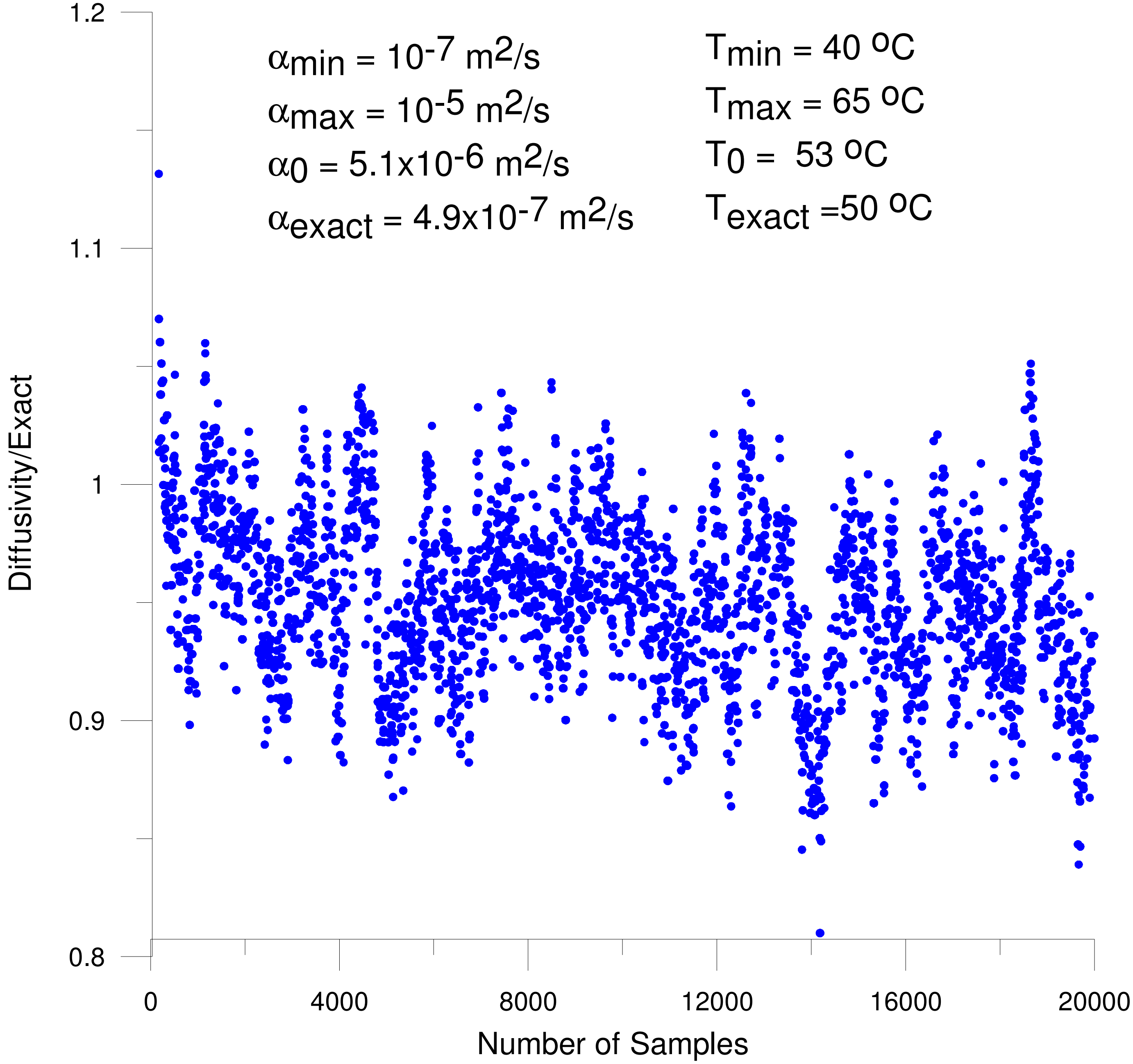


Diffusivity/Exact
Mean 0.979868782
Standard-Deviation 0.008422717
Relative Frequency
0
0.1
0.2
0.3
0.4
0.5
0.93
0.94
0.95
0.96
0.97
0.98
0.99
1
1.01
1.02
Diffusivity/Exact

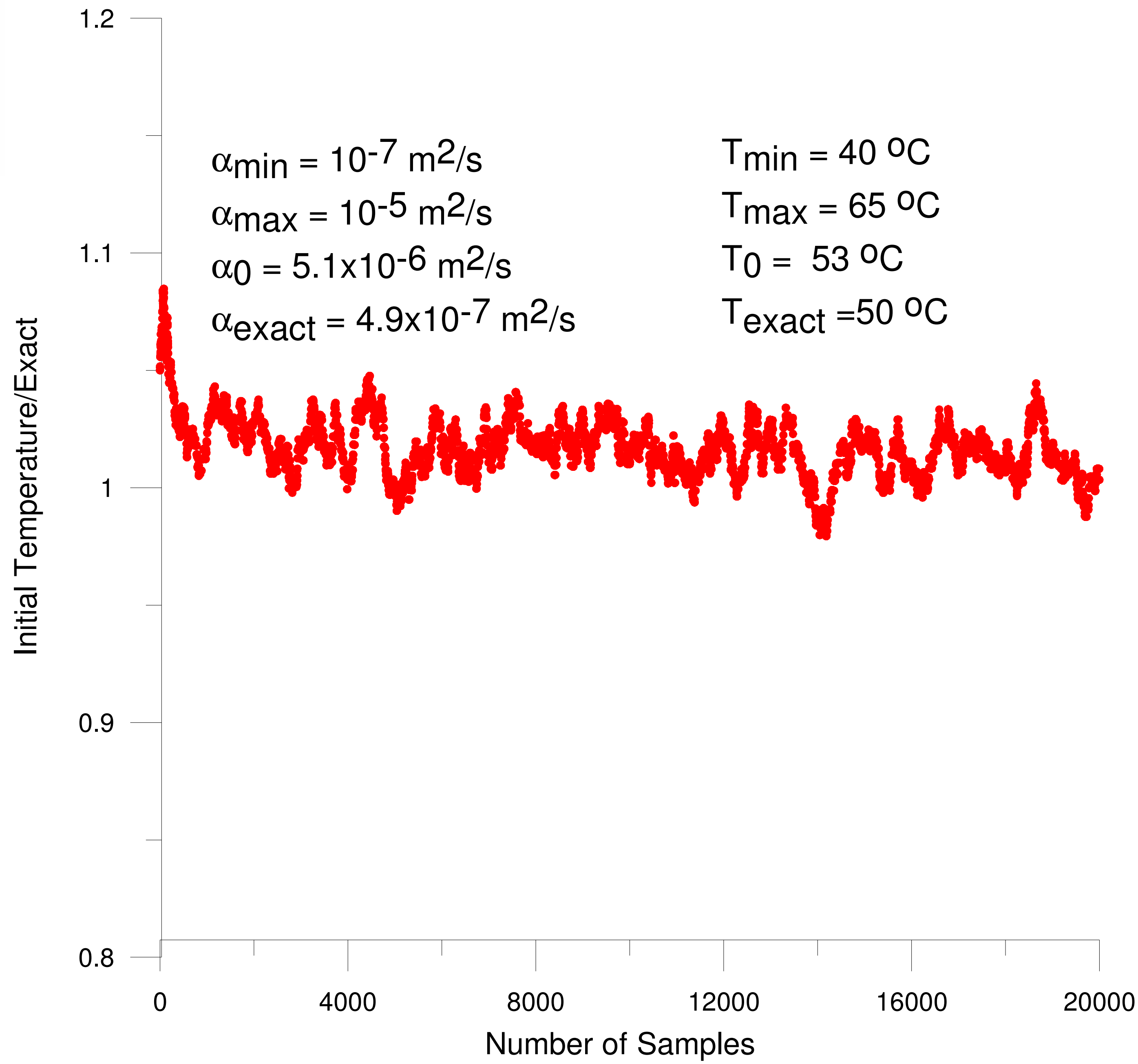
EXAMPLE

- Simultaneous estimation of T_0 and α.
- Prior for T_0 : Uniform distribution (40, 65) ^oC
- Prior for α: Uniform distribution $(10^{-7}, 10^{-5})$ m^2/s
- Start the chain in the middle of the intervals

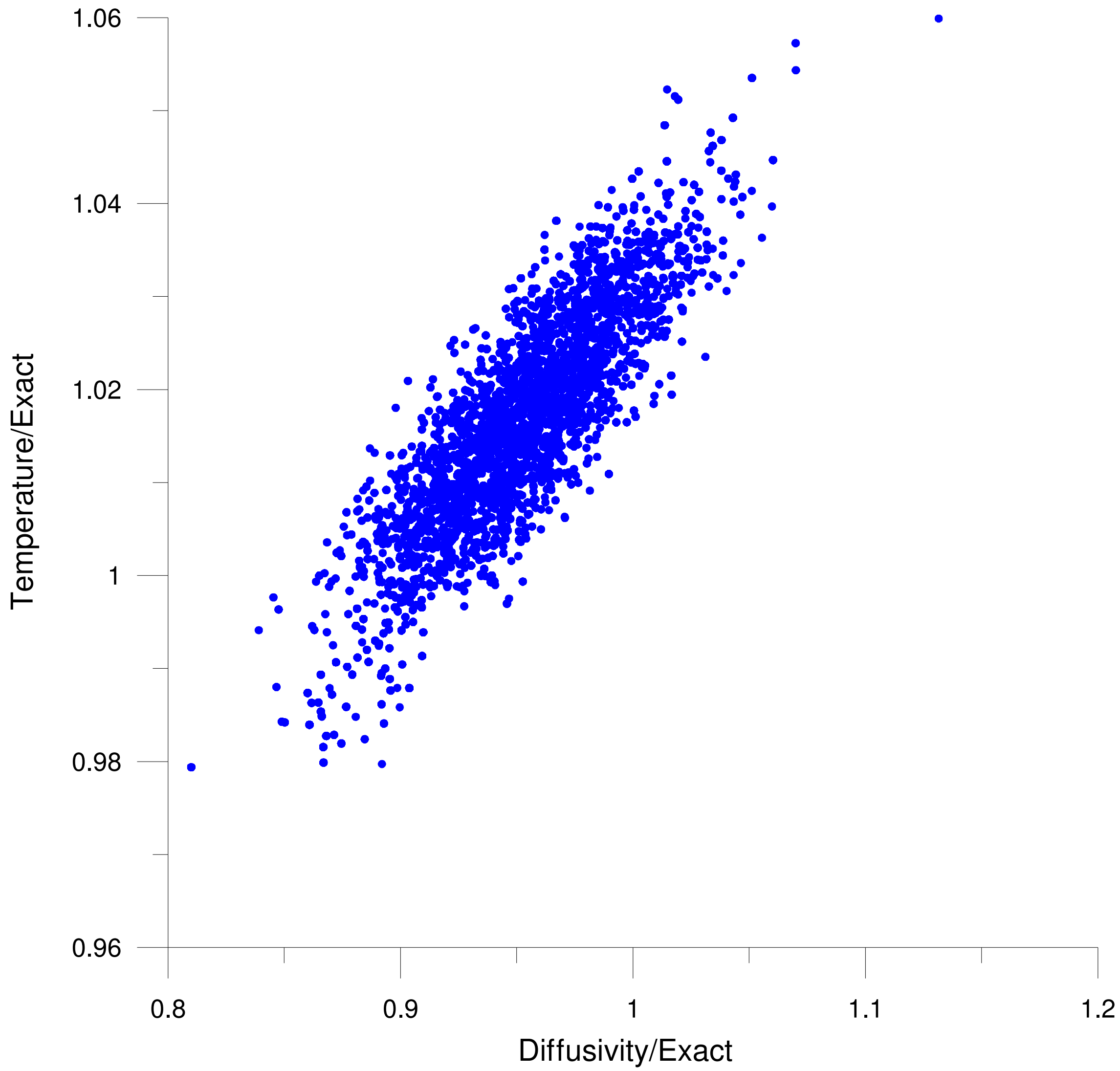

Exact Data
Noisy Data, standard-deviation = 2 C
$x_{meas} = 0.01$ m
T (^oC)
Time (s)
60
50
40
30
20
10
0
50
100
150
200
250



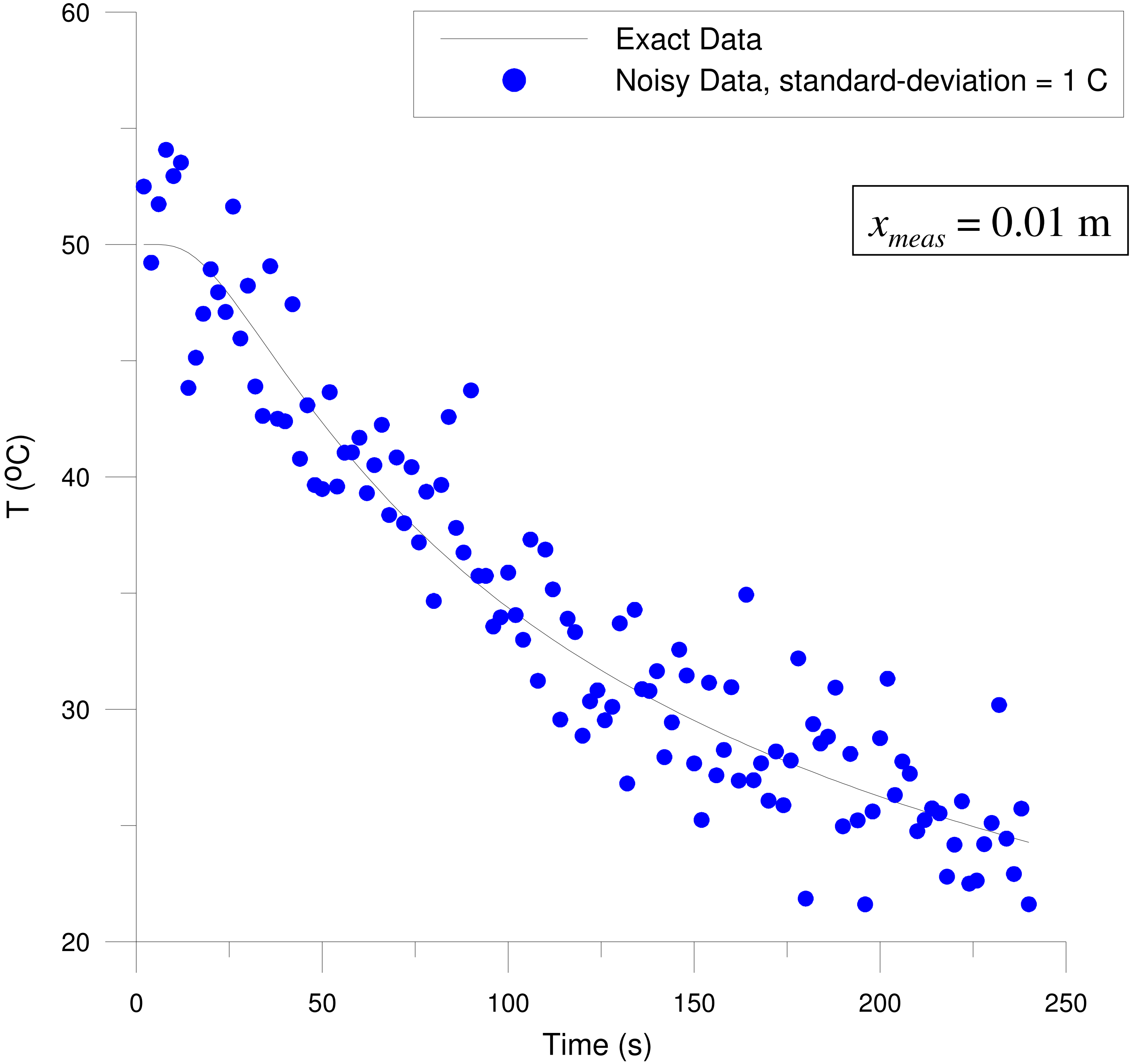
$\alpha_{min} = 10^{-7}\ m^2/s$
$\alpha_{max} = 10^{-5}\ m^2/s$
$\alpha_0 = 5.1x10^{-6}\ m^2/s$
$\alpha_{exact} = 4.9x10^{-7}\ m^2/s$
$T_{min} = 40\ ^oC$
$T_{max} = 65\ ^oC$
$T_0 = 53\ ^oC$
$T_{exact} = 50\ ^oC$
1.2
1.1
1
0.9
0.8
Diffusivity/Exact
0
4000
8000
12000
16000
20000
Number of Samples



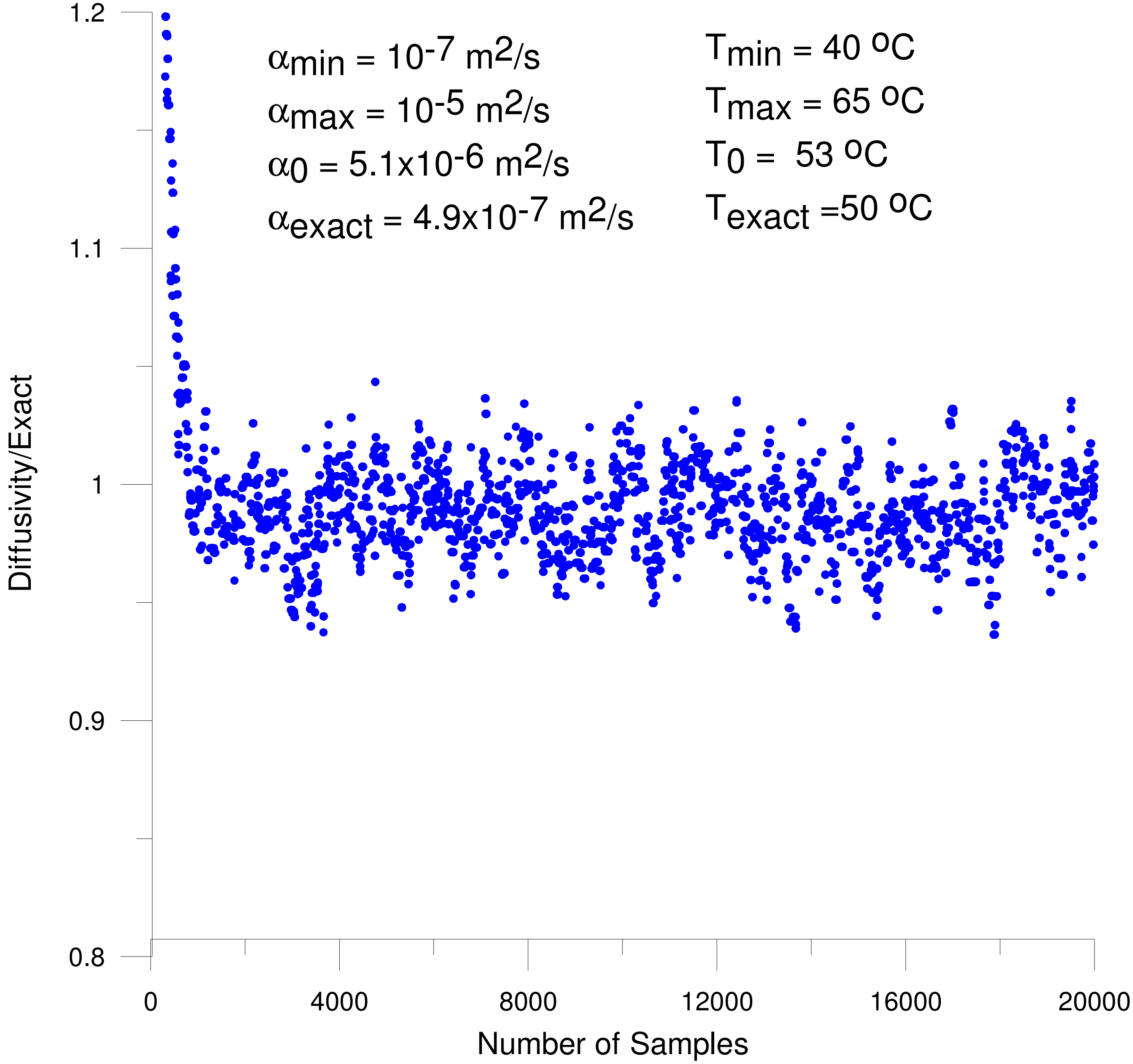
$\alpha_{min} = 10^{-7}\ m^2/s$
$\alpha_{max} = 10^{-5}\ m^2/s$
$\alpha_0 = 5.1x10^{-6}\ m^2/s$
$\alpha_{exact} = 4.9x10^{-7}\ m^2/s$
$T_{min} = 40\ ^oC$
$T_{max} = 65\ ^oC$
$T_0 = 53\ ^oC$
$T_{exact} = 50\ ^oC$
1.2
1.1
1
0.9
0.8
Initial Temperature/Exact
0
4000
8000
12000
16000
20000
Number of Samples



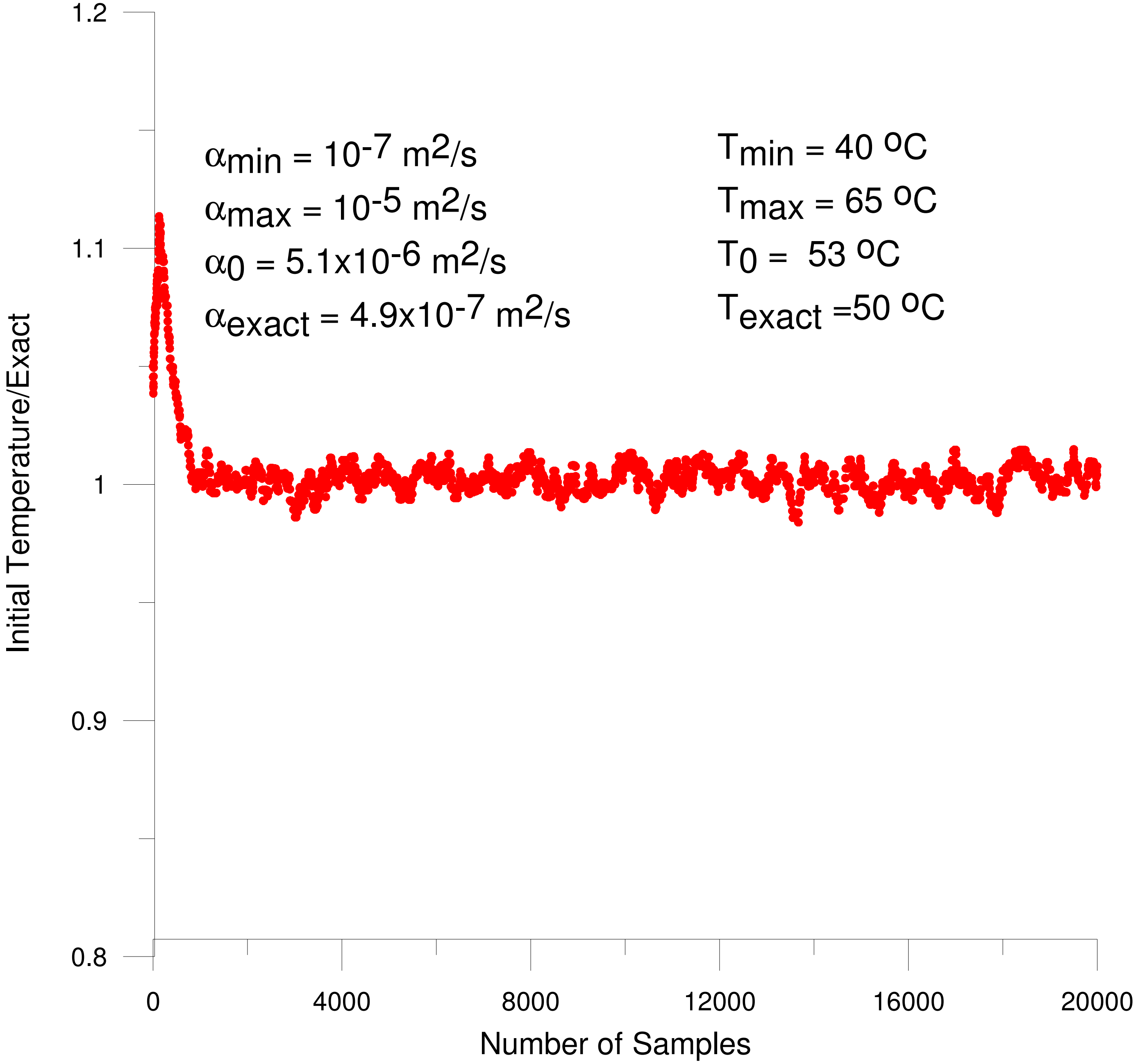
	Diffusivity/Exact	Temperature/Exact
Mean	0.950617158	1.016337295
Standard-Deviation	0.034386526	0.011192818



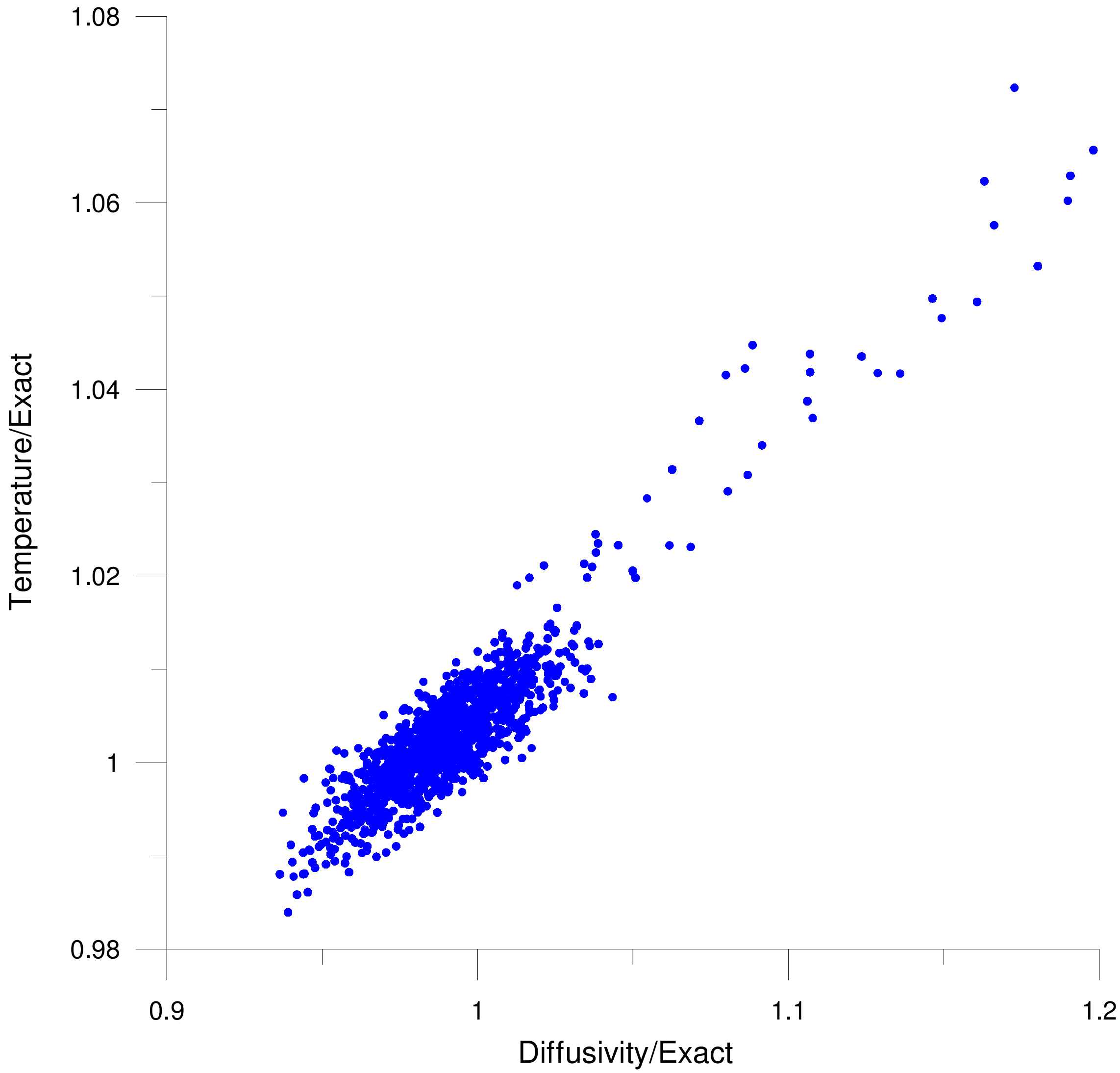
Exact Data
Noisy Data, standard-deviation = 1 C
$x_{meas} = 0.01$ m
T (°C)
Time (s)
60
50
40
30
20
0
50
100
150
200
250



$\alpha_{min} = 10^{-7}\ m^2/s$
$\alpha_{max} = 10^{-5}\ m^2/s$
$\alpha_0 = 5.1 \times 10^{-6}\ m^2/s$
$\alpha_{exact} = 4.9 \times 10^{-7}\ m^2/s$
$T_{min} = 40\ ^oC$
$T_{max} = 65\ ^oC$
$T_0 = 53\ ^oC$
$T_{exact} = 50\ ^oC$
Diffusivity/Exact
1.2
1.1
1
0.9
0.8
0
4000
8000
12000
16000
20000
Number of Samples



α_{min} = 10^{-7} m^2/s
α_{max} = 10^{-5} m^2/s
α_0 = 5.1×10^{-6} m^2/s
α_{exact} = 4.9×10^{-7} m^2/s
T_{min} = 40 ^oC
T_{max} = 65 ^oC
T_0 = 53 ^oC
T_{exact} =50 ^oC
1.2
1.1
1
0.9
0.8
0
4000
8000
12000
16000
20000
Initial Temperature/Exact
Number of Samples



	Diffusivity/Exact	Temperature/Exact
Mean	0.991879086	1.003147487
Standard-Deviation	0.030681073	0.009687663