PARAMETER IDENTIFICATION OF AXIAL DISPERSION MODEL

D. Bártová, B. Jakeš, J. Kukal

Institute of Chemical Technology, Department of Computing and Control Engineering, Technická 5, Prague 6

Abstract

The paper is devoted to the parameter identification of the axial dispersion model in the closed–open canal. It is oriented to the identification precision for the two variants of the flow model with various variance levels. The identification method is based on the transfer function application for the solution of given model PDE. When the impulse function g(*t*) is obtained, the characteristics of this function are obtained analytically as functions of the model parameters. The impulse function was used for the generation of artificial response with the additive Gaussian noise of various dispersions. Then the parameter identification was computed from experimental data. The theory enables to estimate the model parameters via moment method. Then the non-linear regression by the least square method was used. The point parameter estimate was obtained in the sense of the maximum likelihood method. The Hessian matrix was also calculated and then used for the standard deviation determination.

1 Introduction

Hydrodynamic models of flow structure inside apparatus, which corresponds with our ideas about the inner hydrodynamic situation, are sometimes used for the description of dynamical behavior of chemical–technological systems. The variety of models consists of perfectly mixed model, plug (piston) flow model, cascade of ideal mixers without or with recycling, axial and radial dispersion models, compartment models and their parallel and serial combinations, which are sometimes modified via bypassing, dead space (zone), recirculation etc. All the flow models in typical applications are parametric and described via linear operator.

It is useful to notify the task of the mathematical operator synthesis consists of two parts for previous models in this context. It is necessary to select, discriminate and verify the flow structure first. The identification of operator parameters is the second step.

The characteristic behavior of operators is described via basic flow dynamic equation in time, frequency or operator domain as the transfer function. The weight (impulse response) function g(*t*) is used as general characteristics of the hydrodynamic situation (flow structure), which can be statistically interpreted as a probability distribution function of the residence time of individual flow elements (particles) in the apparatus. The probability theory enables to characterize the function g(*t*) by statistical moments, mode, median etc.

This paper is devoted to the second step of the operator determination – the methodology of parameter identification of flow structures. The methodology will be demonstrated on one type of axial dispersion model with two types of open–close boundary conditions.

2 Axial dispersion model

Let $c = \mathrm{c}(x, t)$ be concentration profile in given tube as a function of space coordinate and time. Let D, v, L be effective diffusivity, flow velocity and tube length. The axial dispersion model (for $K = 1$, $c_0 = 0$) is described via partial differential equation

$$\mathrm{c}_t(x,t) = D\mathrm{c}_{xx}(x,t) - v\mathrm{c}_x(x,t), \quad t > 0, x > 0 \tag{1}$$

with initial condition $\mathrm{c}(x, 0) = c_0 = 0$ for $x \geq 0$ and two boundary conditions. They depend on the model choice:

- Enforced input concentration or Danckwerts condition in the left closed part
- Open condition the right open part.

The first model has enforced input concentration u(t) in the tube input as c(0, t) = u(t) for $t > 0$.
The right boundary is in infinity, so the boundary condition is c(+∞, t) = 0 for $t > 0$.
The output signal is represented as the concentration in the tube output y(t) = c(L, t).

Denoting mean residence time and Peclet criterion as $T = L/v$, $Pe = v\,L/D$ and applying Laplace transform, we obtain the transfer function

$$\mathrm{F}(L,p) = \frac{\mathrm{Y}(L,p)}{\mathrm{U}(p)} = \exp\left(\frac{Pe}{2}(1-q)\right) \tag{2}$$

where

$$q = \sqrt{1 + \frac{4Tp}{Pe}} \tag{3}$$

in the sense of main value of complex square root. In the case of Dirac impulse response we have U(p) = K = 1 and the inverse Laplace transform comes to the resulting formula

$$\mathrm{y}(t) = \frac{1}{T}\mathrm{g}\left(\frac{t}{T}, Pe\right) \tag{4}$$

where

$$\mathrm{g}(\theta, Pe) = \frac{1}{2}\sqrt{\frac{Pe}{\pi\theta^3}}\exp\left(-\frac{Pe}{4\theta}(1-\theta)^2\right) \tag{5}$$

for $\theta = t/T > 0$. Anyway g(θ,Pe) = 0.

The generalization of previous model can have the form

$$\mathrm{y}(t) = K\,\mathrm{g}\left(\frac{t - T_d}{T}, Pe\right) + c_0 \tag{6}$$

where T_d, K, c_0 are dead time, gain factor and steady state concentration. The vector of model parameters can be denoted as $\boldsymbol{a} = (K, Pe, T, T_d, c_0)$ with default values $T_d = c_0 = 0$. Thus the first model of axial dispersion has three, four or five free parameters.

The second model differs only in the left boundary condition, which is designed according to Danckwerts as

$$\mathrm{c}(0,t) = \mathrm{u}(t) + \frac{D}{v}\mathrm{c}_x(0,t) \tag{7}$$

for $t > 0$.

The adequate transfer function of the second model is

$$\mathrm{F}(p) = \frac{\mathrm{Y}(p)}{\mathrm{U}(p)} = \frac{2}{1+q}\exp\left(\frac{Pe}{2}(1-q)\right) \tag{8}$$

After the inverse Laplace transform we have alternative impulse response

$$\mathrm{g}(\theta, Pe) = \sqrt{\frac{Pe\,\theta}{\pi}}\exp\left(\frac{-Pe}{4\theta}(1-\theta)^2\right) - \frac{Pe}{2}\exp(Pe)\mathrm{erfc}\left(\frac{\sqrt{Pe}}{2\theta}(1+\theta)\right) \tag{9}$$

as a kernel of the second model with three, four or five free parameters.

3 Moments and parameter estimation

Let y(t) be impulse response of previous models for $c_0 = 0$. We define absolute moments as

$$m_k = \frac{M_k}{M_0} \tag{10}$$

where

$$M_k = \int t^k\,\mathrm{y}(t)\mathrm{d}t \tag{11}$$

is an integral over infinite domain. Central moments are then defined as

$$\mu_k = \int (t - m_1)^k \, y(t) \mathrm{d}t / M_0 \tag{12}$$

The moments m_1, μ_2 represents mean value and dispersion of impulse response in physical units. Well known dimensionless skew and kurtosis are defined as

$$\begin{aligned} skew &= \frac{\mu_3}{\mu_2^{3/2}} \\ kurt &= \frac{\mu_4}{\mu_2^2} - 3 \end{aligned} \tag{13}$$

The moments can be estimated from experimental data using numerical integration over finite time domain. The trapezoidal rule was used to evaluate all the previous integrals.

The second way of moment evaluation is based on the model analysis. The first model has the moment characteristics:

$$\begin{aligned} m_1 &= T + T_d \\ \mu_2 &= \frac{2T^2}{Pe} \\ skew &= 3\sqrt{\frac{2}{Pe}} \\ kurt &= \frac{30}{Pe} \end{aligned} \tag{14}$$

The second model has a little bit complex moment characteristics:

$$\begin{aligned} m_1 &= \left(1 + Pe^{-1}\right)T + T_d \\ \mu_2 &= \frac{2T^2}{Pe}\left(1 + \frac{3}{2Pe}\right) \\ skew &= 4\frac{3Pe + 5}{(2Pe + 3)^{3/2}} \\ kurt &= 30\frac{4Pe + 7}{(2Pe + 3)^2} \end{aligned} \tag{15}$$

The parameter estimation via moment method is an inverse task which is based on the enforced model. It consists of several steps which are here discussed on the example of the first model:

- Estimation of c_0 from the last measured concentration or setting $c_0 = 0$.
- Substraction of c_0 from measured concentrations y.
- Evaluation of m_1, μ_2, *skew* from measured concentrations by trapezoidal method.
- Estimation of $Pe = 2m_1^2/\mu_2$, $T = m_1$ in the case when $T_\mathrm{d} = 0$.
- Estimation of $Pe = 2(3/skew)^2$, $T = 3\mu_2^{1/2}/skew$, $T_\mathrm{d} = m_1 - T$ when T_d was unknown.
- Estimation of dimensionless maximum time $\theta_{\max} = (1 + 9/Pe^2)^{1/2} - 3/Pe$ which is the mode (maximum concentration) coordinate.
- Gain estimation as $K = y_{\max}/\, \mathrm{g}(\theta_{\max}, Pe)$ from maximum response concentration $y_{\max}$.

This estimate can be used as initial point for nonlinear regression task.

In the special case of three parameter model with $T_\mathrm{d} = c_0 = 0$ we can correct the original parameters as:

$$\begin{aligned} Pe_{\mathrm{corr}} &= \frac{1}{2}\left(Pe - 2 + \sqrt{2Pe + Pe^2}\right) \\ T_{\mathrm{corr}} &= \frac{T}{1 + Pe^{-1}} \end{aligned} \tag{16}$$

4 Nonlinear regression

Let $\boldsymbol{a}$ = (K, Pe, T, T_d, c_0) be vector of parameters. Let (t_k,y_k) be experimental response for k = 1, ... , m. The objective function for nonlinear regression is then well known sum of squares

$$\mathrm{SSQ}(\mathbf{a}) = \sum \left(K\,\mathrm{g}\left(\frac{t_k - T_d}{T}, Pe \right) + c_0 - y_k \right)^2 \tag{17}$$

Let SSQ reaches its minimum value SSQ_{opt} in point $\boldsymbol{a}_{\mathrm{opt}}$. Let $\boldsymbol{H} = \partial^2\mathrm{SSQ}(\boldsymbol{a}_{\mathrm{opt}})/\partial\boldsymbol{a}^2$ be positive definite Hessian matrix in the maximum likelihood estimate. Now we can estimate the model error $s_e = (SSQ_{\mathrm{opt}}/(m-n))^{1/2}$ and one-standard deviation of parameters as vector

$$\mathbf{s} = s_e \sqrt{\mathrm{diag}\left(2\mathbf{H}^{-1}\right)} \tag{18}$$

Here m is the number of experimental data and n is the number of estimated parameters. In our case we have n = 3, 4, 5, depending on the additional conditions $c_0 = 0$, $T_d = 0$. The last question is how to minimize the function SSQ in the neighborhood of moment estimate.

5 Optimization

Three aspects are necessary for the reliable optimization for non-convex functions:

- The suitable domain of optimization, where the solution is expected
- The sophisticated initial estimate of searched parameters
- The effective method of convex optimization

In this specific case, the moment estimates were used as the initial parameter values. The optimization domain was created by 90 % decreasing, respectively increasing of the initial nominal values. The Predictioned conjugate gradient method [2] was used for the task solution, because it is the recommended standard for the problems of this specific type. The Matlab realization of this method is well-known as `fmincon` function which attempts to find a constrained minimum of a scalar function of several variables starting at an initial estimate. One of the possible types of the function application is:

```
x = fmincon(fun,x0,A,b,Aeq,beq,lb,ub)
```

where `x` is optimum parameter vector, `fun` is the name of the objective function, `x0` is the initial estimate vector, `lb` and `ub` are the lower, respectively upper bound on the optimization domain. The remaining parameters `A, b, Aeq, beq` represent the other linear constrains and there are empty in this specific task.

6 Results

The two upper described models served as generators of initial data in given time samples for K = 100, Pe = 10 and T = 10. The remaining parameters (T_d, c_0) were set to zero and were not the optimization subjects. The additive Gaussian noise of the various one-standard deviations was put on the obtained theoretical data to form the artificial experimental data. The noise level σ was defined relatively to the maximum signal intensity (in the mode) in the range form 0.1 to 5 %.

The moment estimates were obtained from the artificial experimental data according to Eqs (14) and (15). These estimates were consequently used as the starting optimization values. The results of the optimization there are both the maximum likelihood estimates `x`, and their standard deviations – see Eq (18) – plus the model standard deviation s_e.

The identification results are presented in Table 1 for the simple axial model and Table 2 for the axial model with Danckwerts condition. The LSQ is the minimum of SSQ and SD is the individual value of parameter standard deviation in the tables. The individual SD's are the components of the vector $\mathbf{s}$ – see Eq (18).

The tables demonstrate the noise level influence to the accuracy of moment and maximum likelihood estimates.

There are some general dependences in the tables:

- Increasing of the noise level decreases the estimation accuracy
- The moment estimates are less accurate then the maximum likelihood ones
- The difference between these two estimates rapidly increases with the noise level σ
- The moment estimates are suitable as initial values for the optimization only in the given σ range

Table 1: Simple axial model

parameter	estimate	noise level σ [%]					
		0.100	0.200	0.500	1.000	2.000	5.000
K	theoretic	100.000	100.000	100.000	100.000	100.000	100.000
	moment	98.725	100.467	101.880	102.538	110.087	123.072
	LSQ	100.027	100.019	99.794	100.565	99.113	101.548
	SD	0.039	0.077	0.168	0.248	0.619	1.251
Pe	theoretic	10.000	10.000	10.000	10.000	10.000	10.000
	moment	10.126	9.472	9.076	9.008	6.083	5.110
	LSQ	9.998	9.989	10.095	9.924	9.696	9.964
	SD	0.011	0.021	0.047	0.068	0.170	0.340
T [s]	theoretic	20.000	20.000	20.000	20.000	20.000	20.000
	moment	19.972	20.087	20.046	20.343	21.162	21.272
	LSQ	20.002	20.010	9.970	20.098	20.123	20.184
	SD	0.006	0.012	0.026	0.039	0.100	0.191

Table 2: Axial model with Danckwerts condition

parameter	estimate	noise level σ [%]					
		0.100	0.200	0.500	1.000	2.000	5.000
K	theoretic	100.000	100.000	100.000	100.000	100.000	100.000
	moment	82.315	82.087	81.780	87.717	94.903	110.680
	LSQ	100.056	100.001	99.676	99.640	99.440	97.541
	SD	0.034	0.055	0.155	0.363	0.681	1.426
Pe	theoretic	10.000	10.000	10.000	10.000	10.000	10.000
	moment	11.153	11.150	11.160	10.479	9.776	8.390
	LSQ	9.994	10.008	10.043	10.068	9.900	10.037
	SD	0.009	0.015	0.0417	0.098	0.183	0.400
T [s]	theoretic	20.000	20.000	20.000	20.000	20.000	20.000
	moment	28.024	28.015	28.0291	28.0888	28.337	28.451
	LSQ	19.966	20.008	20.029	20.030	20.050	19.851
	SD	0.005	0.009	0.025	0.060	0.113	0.234

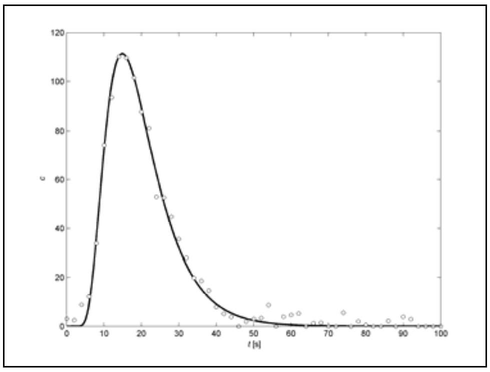

Figure 1 Simulated response of 1st model (σ = 5 %)

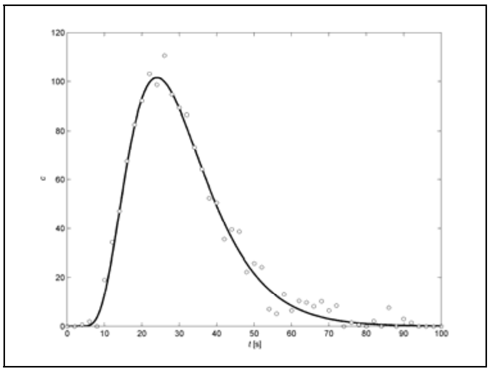

Figure 2 Simulated response of 2nd model (σ = 5 %)

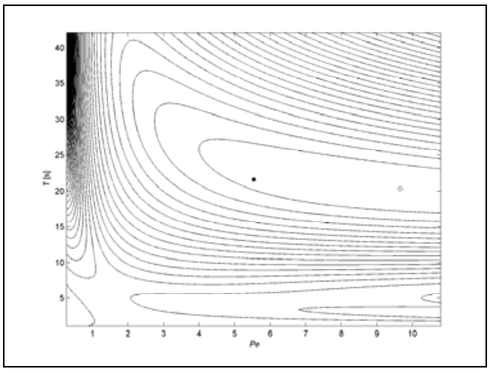

Figure 3 Geometry of SSQ function for 1st model and optimum K

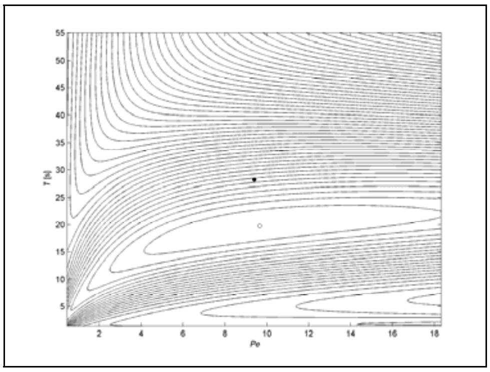

Figure 4 Geometry of SSQ function for 2nd model and optimum K

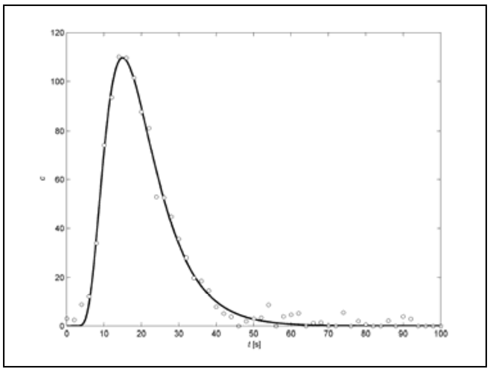

Figure 5 Identification result of 1st model

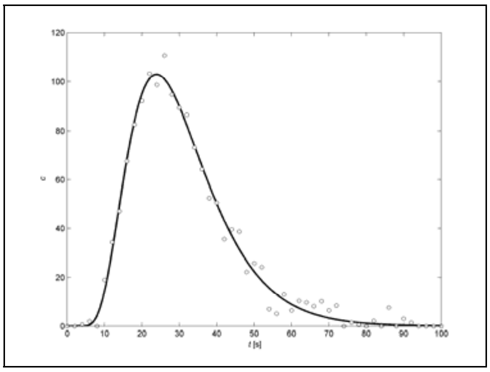

Figure 6 Identification result of 2nd model

The most illustrative situation occurs for the 5 % noise level. The initial theoretical impulse responses for the both models are depicted in the Fig 1 and Fig 2 as lines. The artificial experimental data are presented in the same figures as circles. These two figures also demonstrate that the Danckwerts condition decreases the kurtosis.

The geometry of SSQ function in the optimum neighborhood is depicted in the Fig 3 and Fig 4. The full circles represent the moment estimates (the initial values for optimization), the empty circles are the optimum points and the lines represent the contours of SSQ function in 2-D cut.

The final results of maximum likelihood identification are depicted in the Fig 5 and Fig 6 where the lines correspond to the results of identification and the circles have the same meaning as in Fig 1 and Fig 2.

Both the identification methodology and its Matlab realization are suitable for the parameter estimation of axial dispersion models in the case of small noise levels. The procedure will be used both for the more complicated models identification and the real systems investigation in the future.

7 Acknowledgement

This work has been supported by the Ministry of Education of the Czech Republic (program No. MSM 6046137306)

References

[1] Čermáková, J., Scargiali, F., Siyakatshana, N., Kudrna, V., Brucato, A., Machoň, V. *Axial dispersion model for solid flow in liquid suspension in system of two mixers in total recycle*. Chemical Engineering Journal, Vol. 117, Issue 2, April 2006, pp. 101-107

[2] Coleman, T.F., Verma, A. *A predictioned conjugate gradient approach to linear equality constrained minimization*. Computational Optimization And Applications, 20(1), 2001, pp. 61-72

[3] Himmelblau, D., Bischoff, K.B. *Process Analysis and Simulation* : *Deterministic Systems*. John Wiley and Sons, Inc., New York, 1968

[4] Kudrna, V., Jahoda, M., Siyakatshana, N., Čermáková, J., Machoň, V. *General solution of the dispersion model for a one-dimensional stirred flow system using Danckwerts' boundary conditions*. Chemical Engineering Science, Vol. 59, Issue 14, July 2004, pp. 3013-3020

[5] Levenspiel, O., Smith, W.K. *Notes on the diffusion-type model for the longitudinal mixing of fluids in flow*. Chemical Engineering Science, 1957, Vol. 6, No. 4-5, pp. 227-233

[6] Martin, A.D. *Interpretation of residence time distribution data*. Chemical Engineering Science, 2000, Vol. 55, No. 23 (December), 5907-5917

[7] Sorensen, D.C. *Minimization of a Large Scale Quadratic Function Subject to an Ellipsoidal Constraint*. Department of Computational and Applied Mathematics, Rice University, Technical Report TR94-27, 1994.

[8] Wen C.Y., Fan L.T. *Models for Flow Systems and Chemical Reactors*. Marcel Dekker, Inc., New York, 1975

Darina Bártová
Institute of Chemical Technology, Department of Computing and Control Engineering, Technická 5, 166 28 Prague 6, tel.: +420 220 444 170, e-mail: dbartova@vscht.cz

Bohumil Jakeš
Institute of Chemical Technology, Department of Computing and Control Engineering, Technická 5, 166 28 Prague 6, tel.: +420 220 444 170, e-mail: jakesb@vscht.cz

Jaromír Kukal
Institute of Chemical Technology, Department of Computing and Control Engineering, Technická 5, 166 28 Prague 6, tel.: +420 220 444 212, e-mail: kukalj@vscht.cz