NONLINEAR PARAMETER ESTIMATION USING INTERVAL ANALYSIS

Chao-Yang Gau and Mark A. Stadtherr
Department of Chemical Engineering
University of Notre Dame
Notre Dame, IN 46556 USA

Abstract

The reliable solution of nonlinear parameter estimation problems is an important computational problem in chemical process engineering, both in on-line and off-line applications. Conventional solution methods may not be reliable since they do not guarantee convergence to the global optimum sought in the parameter estimation problem. We demonstrate here a technique, based on interval analysis, that can solve the nonlinear parameter estimation problem with complete reliability, providing a mathematical and computational guarantee that the global optimum is found. As an example, we consider the estimation of parameters in vapor-liquid equilibrium (VLE) models. Twelve VLE data sets are fit to the Wilson equation. Results indicate that several sets of published parameter values correspond to local optima only, with new globally optimal parameter values found by using the interval approach.

Keywords

Parameter estimation, Global optimization, Interval analysis, Vapor-liquid equilibrium

Introduction

Parameter estimation is a common problem in many areas of process modeling, both in on-line applications such as real time optimization and in off-line applications such as the modeling of reaction kinetics and phase equilibrium. The goal is to determine values of model parameters that provide the best fit to measured data, generally based on some type of least squares or maximum likelihood criterion. In the most general case, this requires the solution of a nonlinear and frequently nonconvex optimization problem.

It is not uncommon for the objective function in nonlinear parameter estimation problems to have multiple local optima. However, the standard methods used to solve these problems are local methods that offer no guarantee that the global optimum, and thus the best set of model parameters, has been found. Thus, there is a need for global optimization in nonlinear parameter estimation. One approach that has been suggested is the use of convex underestimating functions in connection with a branch and bound procedure (Esposito and Floudas, 1997). An alternative approach for global optimization is the use of interval analysis (e.g., Hansen, 1992). We demonstrate here the use of interval methods for determining a global optimum in nonlinear parameter estimation problems of interest in process engineering.

As an example, we consider the estimation of parameters in vapor-liquid equilibrium (VLE) models. We demonstrate that even for simple models, such as the Wilson equation, multiple local optima can occur in parameter estimation. It is also shown that for some data sets, published parameter values (Gmehling *et al.*, 1977-1990) correspond to a local but not global optimum. We then demonstrate how a simple global optimization procedure based on interval analysis can be used to reliably determine the globally optimal parameter values. The method used involves the use of an interval Newton technique combined with interval branch and bound. This method provides a mathematical and computational

guarantee of global optimality in parameter estimation. The reliability of the method is demonstrated using several VLE data sets, and the globally optimal parameters compared to published values obtained using local methods.

Background

Good introductions to the parameter estimation problem are provided by Bard (1974), Gallant (1987) and Seber (1989). Suppose that observations $y_{\mu i}$ of $i = 1,...,q$ response variables from $\mu = 1,...,p$ experiments are available, and that the responses are to be fit to a model of the form $y_{\mu i} = f_i(\mathbf{x}_\mu, \boldsymbol{\theta})$, with independent variables $\mathbf{x}_\mu = (x_{\mu 1},x_{\mu 2},...,x_{\mu m})^T$ and parameters $\boldsymbol{\theta} = (\theta_1,\theta_2,...,\theta_n)^T$. To determine optimal values of the parameters (i.e., the "best" fit), a maximum likelihood criterion is most appropriate in many circumstances. However, with some assumptions this can be simplified to the widely used relative least squares criterion, which requires minimizing the function

$$\phi(\boldsymbol{\theta}) = \sum_{i=1}^{q} \sum_{\mu=1}^{p} [(y_{\mu i} - f_i(\mathbf{x}_\mu, \boldsymbol{\theta})) / y_{\mu i}]^2.$$

This can be treated either as a constrained or, if the experimental observations are substituted directly into the objective function, unconstrained minimization problem. We will consider only the unconstrained formulation of the problem here. For minimizing ϕ, a wide variety of standard minimization techniques are available. However, in general, these are local methods that provide no assurance that a global minimum has been found. What is needed is a technique that can find the global minimum of ϕ, and do so with mathematical and computational certainty. The use of interval analysis provides such a technique.

Interval Analysis

A real *interval Z* is defined as the set of real numbers lying between (and including) given upper and lower bounds; i.e., $Z = [z^L,z^U] = \{z \in \Re \mid z^L \le z \le z^U\}$. A real interval vector $\mathbf{Z} = (Z_1,Z_2,...,Z_n)^T$ has n real interval components and can be interpreted geometrically as an n-dimensional rectangle. Note that in this section lower case quantities are real numbers and upper case quantities are intervals. Several good introductions to interval analysis are available (e.g., Neumaier, 1990; Hansen, 1992; Kearfott, 1996).

Of particular interest here is the interval Newton technique. Given a nonlinear equation system with a finite number of real roots in some initial interval, this technique provides the capability to find (or, more precisely, narrowly enclose) *all* the roots of the system within the given initial interval. For the unconstrained minimization of the relative least squares function $\phi(\boldsymbol{\theta})$, a common approach is to use the gradient of $\phi(\boldsymbol{\theta})$ and seek a solution of $\mathbf{g}(\boldsymbol{\theta}) \equiv \nabla\phi(\boldsymbol{\theta}) = \mathbf{0}$. The global minimum will be a root of this nonlinear equation system, but there may be many other roots as well, representing local minima and maxima and saddle points. Thus, for this approach to be reliable, the capability to find *all* the roots of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ is needed, and this is provided by the interval Newton technique. In practice, the interval Newton procedure can also be combined with an interval branch and bound technique, so that roots of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ that cannot be the global minimum need not be found.

For the system of nonlinear equations $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ with $\boldsymbol{\theta} \in \boldsymbol{\Theta}^{(0)}$, the basic iteration step in interval Newton methods is, given an interval $\boldsymbol{\Theta}^{(k)}$, to solve the linear interval equation system $G'(\boldsymbol{\Theta}^{(k)})\ (\mathbf{N}^{(k)} - \boldsymbol{\theta}^{(k)}) = -\mathbf{g}(\boldsymbol{\theta}^{(k)})$ for a new interval $\mathbf{N}^{(k)}$, where k is an iteration counter, $G'(\boldsymbol{\Theta}^{(k)})$ is an *interval extension* of the Jacobian of $\mathbf{g}(\boldsymbol{\theta})$, i.e., the Hessian of $\phi(\boldsymbol{\theta})$, over the current interval $\boldsymbol{\Theta}^{(k)}$, and $\boldsymbol{\theta}^{(k)}$ is a point in the interior of $\boldsymbol{\Theta}^{(k)}$. The interval extension of a real function over an interval is an enclosure of the range of the function over the interval, and can be computed by substituting interval quantities for the corresponding real quantities and using interval arithmetic, or in other ways. It can be shown (Moore, 1966) that any root $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^{(k)}$ of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ is also contained in the *image* $\mathbf{N}^{(k)}$, implying that if there is no intersection between $\boldsymbol{\Theta}^{(k)}$ and $\mathbf{N}^{(k)}$, then no root exists in $\boldsymbol{\Theta}^{(k)}$, and suggesting the iteration scheme $\boldsymbol{\Theta}^{(k+1)} = \boldsymbol{\Theta}^{(k)} \cap \mathbf{N}^{(k)}$. In addition to this iteration step, which can be used to tightly enclose a solution, the following property can be proven (e.g., Neumaier, 1990; Kearfott, 1996): If $\mathbf{N}^{(k)}$ is contained completely within $\boldsymbol{\Theta}^{(k)}$, then there is *one and only one root* contained within $\boldsymbol{\Theta}^{(k)}$. This property is quite powerful, as it provides a mathematical guarantee of the existence and uniqueness of a root when it is satisfied. The foregoing suggests a series of tests to determine whether a stationary point (root of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$) that might be the global minimum of $\phi(\boldsymbol{\theta})$ can be contained in $\boldsymbol{\Theta}^{(k)}$:

1. (Function Range Test) Compute an interval extension $\mathbf{G}(\boldsymbol{\Theta}^{(k)})$ containing the range of $\mathbf{g}(\boldsymbol{\theta})$ over the current interval $\boldsymbol{\Theta}^{(k)}$ and test to see whether it contains zero. Clearly, if $\mathbf{0} \notin \mathbf{G}(\boldsymbol{\Theta}^{(k)}) \supseteq \{\mathbf{g}(\boldsymbol{\theta}) \mid \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(k)}\}$, then there can be no solution of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ in $\boldsymbol{\Theta}^{(k)}$ and this interval need not be further tested since it cannot contain a stationary point of $\phi(\boldsymbol{\theta})$.
2. (Objective Range Test) Compute an interval extension $\Phi(\boldsymbol{\Theta}^{(k)})$ containing the range of $\phi(\boldsymbol{\theta})$ over the current interval $\boldsymbol{\Theta}^{(k)}$. If the lower bound of $\Phi(\boldsymbol{\Theta}^{(k)})$ is greater than a known upper bound on the global minimum of $\phi(\boldsymbol{\theta})$, then $\boldsymbol{\Theta}^{(k)}$ cannot contain the global minimum and need not be further tested (see step 3b).
3. (Interval Newton Test) Compute the image $\mathbf{N}^{(k)}$ as described above.

a. If $\mathbf{\Theta}^{(k)} \cap \mathbf{N}^{(k)} = \varnothing$, then there is no root of $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ in $\mathbf{\Theta}^{(k)}$ and it need not be further tested since it cannot contain a stationary point of $\phi(\boldsymbol{\theta})$.
b. Evaluate $\phi(\boldsymbol{\theta}^{(k)})$ and use to determine and update an upper bound on the global minimum for use in step 2.
c. If $\mathbf{N}^{(k)} \subset \mathbf{\Theta}^{(k)}$, then there is *exactly one* root of $\mathbf{g}(\boldsymbol{\theta})$ in $\mathbf{\Theta}^{(k)}$, which may correspond to the global minimum.
d. If neither of the above is true, then no further conclusion can be drawn.

In the last case, one could then repeat the root inclusion test on the next interval Newton iterate $\mathbf{\Theta}^{(k+1)}$, assuming it is sufficiently smaller than $\mathbf{\Theta}^{(k)}$, or one could bisect $\mathbf{\Theta}^{(k+1)}$ and repeat the root inclusion test on the resulting intervals. This is the basic idea of interval Newton/generalized bisection (IN/GB) methods. A more detailed description of an IN/GB algorithm has been given by Schnepper and Stadtherr (1996). Through the addition of steps 2 and 3b, it has been combined with a simple interval branch and bound scheme. Our current implementation of the IN/GB method is based on appropriately modified routines from the packages INTBIS (Kearfott and Novoa, 1990) and INTLIB (Kearfott *et al.*, 1994). The *worst-case* computational complexity of the IN/GB algorithm is exponential in the number of variables. However, process modeling problems involving over a hundred variables have been successfully solved using this approach (Schnepper and Stadtherr, 1996).

Example

As an example, we consider the estimation from binary vapor-liquid equilibrium (VLE) data of the energy parameters in the Wilson equation for liquid phase activity coefficient. Expressed in terms of the reduced excess Gibbs energy g^E for a binary system and the liquid-phase mole fractions x_1 and x_2, the Wilson equation is $g^E = -x_1\ln(x_1 + \Lambda_{12}x_2) - x_2\ln(x_2 + \Lambda_{21}x_1)$, from which expressions for the activity coefficients γ_1 and γ_2 are readily obtained. The binary parameters Λ_{12} and Λ_{21} are given by $\Lambda_{12} = (v_2/v_1)\exp(-\theta_1/RT)$ and $\Lambda_{21} = (v_1/v_2)\exp(-\theta_2/RT)$, where v_1 and v_2 are the pure component liquid molar volumes, T is the system temperature and θ_1 and θ_2 are the energy parameters that must be estimated.

From VLE measurements, experimental values $\gamma_{1,\text{exp}}$ and $\gamma_{2,\text{exp}}$ of the activity coefficients can be obtained. For the parameter estimation problem, the relative least squares objective

$$\phi(\boldsymbol{\theta}) = \sum_{i=1}^{2}\sum_{\mu=1}^{p} \left[(\gamma_{\mu i,\text{exp}} - \gamma_{\mu i,\text{calc}}(\boldsymbol{\theta})) / \gamma_{\mu i,\text{exp}}\right]^2.$$

can be used, where $\gamma_{\mu i,\text{calc}}(\boldsymbol{\theta})$ is determined from the Wilson equation at the same conditions (temperature, pressure and composition) as in the measurement of $\gamma_{\mu i,\text{exp}}$. This parameter estimation problem has been solved for a large number of systems and results presented in the DECHEMA Vapor-Liquid Equilibrium Data Collection (Gmehling *et al.*, 1977-1990), along with the raw VLE data.

For the example here we consider the binary system water(1)–formic acid(2). Twelve VLE data sets, at various pressures, from the DECHEMA Collection were studied. For each data set, the DECHEMA Collection gives the raw VLE data and the results of parameter estimation for θ_1 and θ_2 based on the relative least squares objective. Since Gmehling *et al.*, (1977-1990) use a local method for minimizing $\phi(\boldsymbol{\theta})$ in this parameter estimation, it is possible that the values of θ_1 and θ_2 obtained do not correspond to a *global* minimum in $\phi(\boldsymbol{\theta})$. To investigate this, we resolved each parameter estimation problem for the global minimum using our modification of INTBIS.

Results and Discussion

The results for θ_1 and θ_2 and $\phi(\boldsymbol{\theta})$ from DECHEMA and from the interval approach (INTBIS) suggested here are summarized in Table 1, along with the number of local minima found for each problem (for purposes of determining the number of local minima, the branch and bound steps 2 and 3b were turned off). It can be seen that each problem has multiple local minima, and that in five of the twelve cases (data sets 7–11) the results presented in DECHEMA are not globally optimal. As shown schematically in Fig. 1, the parameter values given in DECHEMA are clustered in two different regions, with results from five data sets falling in one region ($\theta_2 > \theta_1$) and results from seven others falling in the second region ($\theta_1 > \theta_2$). When the global optimum is obtained in the parameter estimation problem the results are much more consistent, with results from ten of the twelve data sets clustered in one region ($\theta_2 > \theta_1$) and only two yielding substantially different results.

We now look more closely at the results for one data set, namely data set 10. For this case, INTBIS (with the branch and bound steps turned off) found five stationary points, three minima and two saddle points, in the initial interval $\Theta_1^{(0)} = \Theta_2^{(0)} = [-10000, 10000]$. These results are summarized in Table 2. The global minimum at $\boldsymbol{\theta} = (-329, 1394)^T$ (root P5) has an objective function value $\phi(\boldsymbol{\theta}) = 0.0819$ that is only about half the magnitude of the local minimum at $\boldsymbol{\theta} = (452, -664)^T$ (root P3) found by Gmehling *et al.* (1977-1990) and reported in the DECHEMA Collection. As is often the case in least squares problems of this sort, all the minima found lie in a relatively flat valley in the parameter space.

The performance of the two different parameter sets, corresponding the local minimum P3 (DECHEMA) and global minimum P5 (INTBIS), in predicting the activity

coefficients for water and formic acid with the Wilson equation is shown in Fig. 2. It is clear that when the globally optimal parameter values from INTBIS are used in the Wilson equation, it results in less deviation from the experimental values in comparison to the case in which the locally optimal parameters reported in DECHEMA are used.

It should be emphasized that the sort of difficulties observed in the water–formic acid system, namely the failure of standard local optimization techniques to find the globally optimal parameters, is not restricted to this system and model. This difficulty can be observed in other systems reported in the DECHEMA Collection. This should not be surprising, since with traditional local solution techniques, it is extremely unlikely that the global optimum will always be found.

In determining the global minimum with the interval approach, initial parameter intervals of $\Theta_1^{(0)} = \Theta_2^{(0)} =$ [−10000, 10000] were used for each data set, which should be wide enough to enclose any physically feasible solution. The ability to provide a wide initial interval, as opposed to an initial point guess, means that the method is essentially initialization independent. For each data set, the computation time needed to perform the global optimization was from roughly 10 to 50 seconds on a Sun Ultra 2/1300 workstation. The difference in times is due to the differing number of data points in each data set, and the differing number of stationary points found. It should be emphasized that at this point, no significant efforts have been made to optimize the efficiency of the code. The use of techniques for tightening the evaluation of interval function extensions, as suggested by Tessier (1997) and Hua *et al.* (1998) can potentially provide an order of magnitude improvement in computational efficiency.

While, in comparison to traditional local methods, additional computation time will typically be required to implement the interval approach, this may be well compensated by the guaranteed reliability of the results. Continuing advances in computing hardware (both in single processor performance and multiprocessing) and software (e.g., compiler support for interval arithmetic) will make this approach even more attractive.

Concluding Remarks

We have described here a new method for reliably solving nonlinear parameter estimation problems. The method is based on interval analysis, in particular an interval Newton/generalized bisection algorithm. The approach provides a mathematical and computational guarantee that the global optimum in the parameter estimation problem is found. We applied the technique here to several data sets in which the Wilson activity coefficient model was used. However, the technique is model independent and can be applied in connection with any thermodynamic model for vapor-liquid equilibrium. The approach presented is general purpose and can also be used in connection with other objective functions, such as maximum likelihood, and other types of VLE measurements. It can also be applied to a wide variety of other nonlinear parameter estimation problems in chemical process engineering.

Acknowledgements

This work has been supported in part by the donors of The Petroleum Research Fund, administered by the ACS, under Grant 30421-AC9, by the National Science Foundation Grants DMI96-96110 and EEC97-00537-CRCD, and by a grant from Sun Microsystems, Inc.

References

Bard, Y. (1974). *Nonlinear Parameter Estimation*. Academic Press, New York

Esposito, W., and C. A. Floudas (1997). Parameter estimation of nonlinear algebraic models via global optimization. Presented a*t* AIChE Annual Meeting, Los Angeles, CA, November.

Gallant, A. R. (1987). *Nonlinear Statistical Models.* Wiley, New York.

Gmehling, J., U. Onken and W. Arlt (1977-1990). *Vapor-Liquid Equilibrium Data Collection*. Chemistry Data Series, Vol. I, Parts 1-8, DECHEMA, Frankfurt/Main, Germany.

Hansen, E. (1992). *Global Optimization Using Interval Analysis*. Marcel Dekker, New York.

Hua, J. Z., J. F. Brennecke and M. A. Stadtherr (1998). Enhanced interval analysis for phase stability: Cubic equation of state models. *Ind. Eng. Chem Res.*, in press.

Kearfott, R. B. (1996). *Rigorous Global Search: Continuous Problems*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Kearfott, R. B., and M. Novoa III (1990). Algorithm 681: INTBIS, a portable interval Newton/bisection package. *ACM Trans. Math. Soft.*, **16**, 152–157.

Kearfott, R. B., M. Dawande, K.-S. Du and C.-Y. Hu (1994). Algorithm 737: INTLIB, a Portable FORTRAN 77 interval standard function library. *ACM Trans. Math. Soft.*, **20**, 447–459.

Moore, R. E. (1966). *Interval Analysis*. Prentice-Hall, Englewood Cliff, NJ.

Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge, England.

Seber, G. A. F. (1989). *Nonlinear Regression.* Wiley, New York.

Schnepper, C. A., and M. A. Stadtherr (1996). Robust process simulation using interval methods. *Comput. Chem. Eng.*, **20**, 187–199.

Tessier, S. R. (1997). *Enhanced Interval Analysis for Phase Stability: Excess Gibbs Energy Models*. M. S. Thesis, University of Notre Dame, Notre Dame, IN.

Table 1. Summary of parameter estimation results for the water(1) and formic acid(2) system, showing parameters θ_1 and θ_2 and objective function $\phi(\boldsymbol{\theta})$. Values shown in bold are globally optimal parameters that differ from those given in DECHEMA.

Data Set	*P* (mm Hg)	DECHEMA			INTBIS			No. of Minima
		θ_1	θ_2	$\phi(\boldsymbol{\theta})$	θ_1	θ_2	$\phi(\boldsymbol{\theta})$	
1	760	-195	759	0.0342	-195	759	0.0342	2
2	760	-278	1038	0.0106	-278	1038	0.0106	2
3	760	-310	1181	0.0151	-308	1167	0.0151	2
4	760	-282	985	0.353	-282	984	0.353	2
5	760	-366	1513	0.0257	-365	1509	0.0257	3
6	760	1067	-1122	0.0708	1065	-1120	0.0708	2
7	200	892	-985	0.141	**-331**	**1250**	0.0914	2
8	200	370	-608	0.0459	**-340**	**1404**	0.0342	3
9	100	539	-718	0.165	**-285**	**996**	0.111	2
10	100	450	-663	0.151	**-329**	**1394**	0.0819	3
11	70	558	-762	0.0399	**-330**	**1519**	0.0372	3
12	25	812	-1058	0.0502	807	-1055	0.0502	2

Table 2. Details for roots (stationary points) found using INTBIS for data set 10. Point P3 is the local minimum presented in DECHEMA, while point P5 is the global minimum.

Root	Position (θ_1, θ_2)	Eigenvalues of Hessian	$\phi(\boldsymbol{\theta})$	Status
P1	(1958, -1251)	7.55E-5, 2.58E-7	0.164	minimum
P2	(1165, -1083)	6.83E-5, -1.44E-7	0.178	saddle
P3	(452, -664)	6.97E-5, 9.42E-8	0.151	minimum
P4	(-37.8, 38.5)	9.08E-5, -3.54E-7	0.19	saddle
P5	(-329, 1394)	1.23E-4, 1.47E-7	0.0819	global minimum

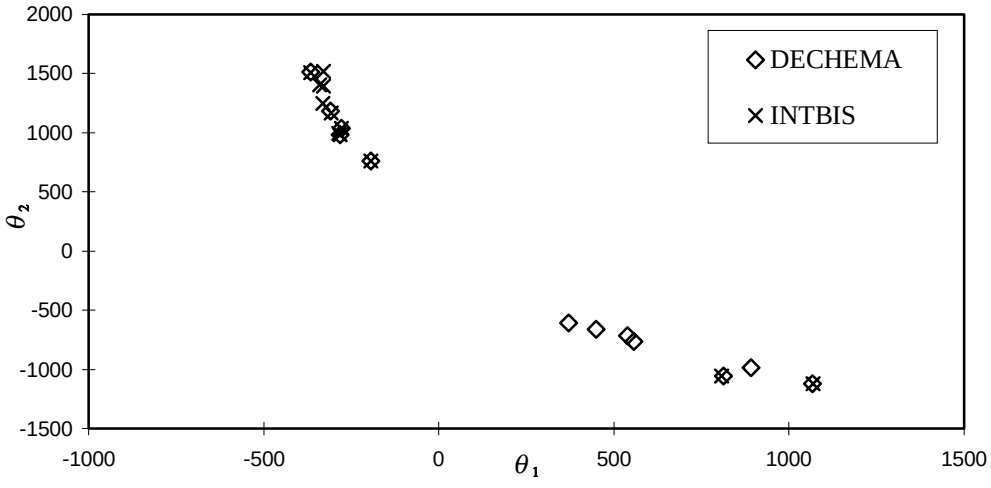

Figure 1. Comparison of estimation results from DECHEMA and INTBIS.

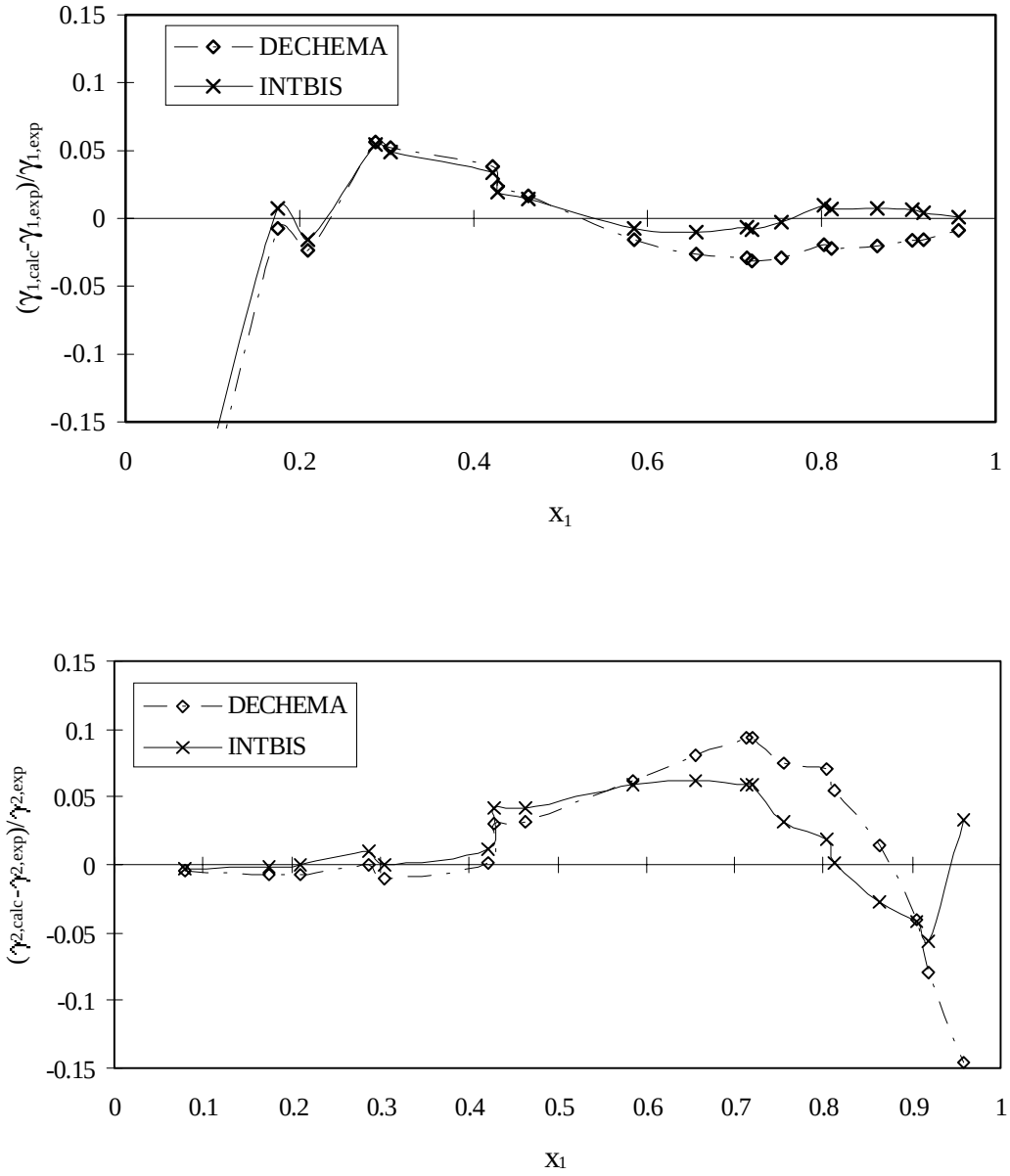

Figure 2. The relative error in data set 10 between calculated and experimental activity coefficients for water (top) and formic acid (bottom) resulting from the locally optimal DECHEMA values and the globally optimal values found using INTBIS. For water (top), the relative error for a data point at $x_1 = 0.0802$ is off scale (roughly at –0.22) for both cases.