Optimum tuning parameters for Encapsulated Evolution Strategies: Results for a nonlinear regression problem

C. Kracht H. Geyer P. Ulbig S. Schulz

University of Dortmund, Department of Chemical Engineering, Institute for Thermodynamics, 44221 Dortmund (Germany)

July 20, 1998

Abstract

The prediction of certain thermodynamic properties of pure substances and mixtures with calculation methods is a frequent task during the process design in chemical engineering. Group contribution models divide the molecules into functional groups and if the model parameters for theses groups are known, predictions of compounds that comprise these groups are possible. The model parameters have to be fitted to experimental data, which leads to a multi-parameter multimodal optimization problem. In this paper the optimization of the tuning parameters of Evolution Strategies and different methods of parameter fitting regarding the number of parameters are presented.

1 Introduction

Group contribution models are used for the prediction of certain thermodynamic properties, such as activity coefficients, excess enthalpies or heats of vaporization to assist in design and simulation of chemical processes. These thermodynamic properties originate in physical interactions between molecules. Group contribution models split the molecules

*This work is a product of the Collaborative Research Center 531, "Computational Intelligence", at the University of Dortmund financially supported by the Deutsche Forschungsgemeinschaft

into functional groups and the physical interactions between the functional groups can be determined, if the model parameters for these functional group interactions exist.

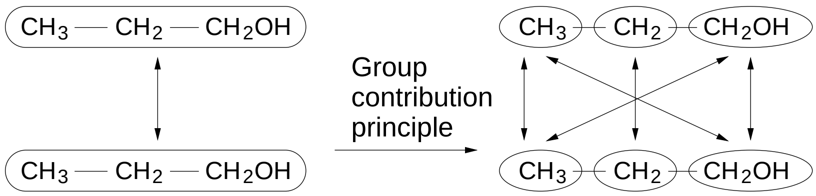

Figure 1: Group contribution concept

In many modern methods, such as mod. UNIFAC [2, 15], the model parameters appear in sums of exponential terms, especially if temperature dependencies are described. This usually leads to a nonlinear regression problem with a multimodal objective function. The topology of the objective function can be visualized by varying two of the model parameters while the remaining parameters are kept constant (Fig. 2).

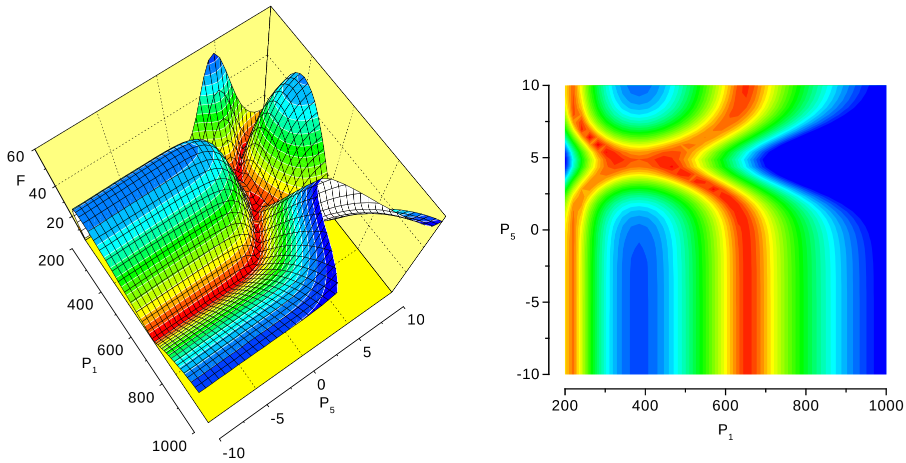

Figure 2: Adaptive surface of a mod. UNIFAC group contribution model by varying 2 parameters

A good prediction of thermodynamic quantities depends on well fitted interaction parameters, i. e. six parameters per one functional group combination. The parameters of the group contribution model EBGCM [8, 13] are fitted to a database of 2315 activity coefficients at infinite dilution γ^{∞} and 1240 excess enthalpy data sets h^{E} with 18746 single data points (i. e. the mole fraction of the considered binary mixture and its corresponding excess enthalpy value) at different temperatures.

The objective function is not the model equation in which the model parameters occur but an error criterion which establishes the relation between the calculated value and the respective experimental data point.

For the optimization to activity coefficients at infinite dilution γ^{∞} the mean relative error was used:

$$\mathrm{MRE} = \frac{1}{n} \sum_{i}^{n} \left| \frac{\gamma^{\infty}_{i_{calc}} - \gamma^{\infty}_{i_{exp}}}{\gamma^{\infty}_{i_{exp}}} \right| \qquad (1)$$

The fitting to excess enthalpy data h^{E} was carried out with the mean relative range related error ($\mathrm{MRE_R}$, Eq. 2)). The range is defined as the difference between the largest and the smallest h^{E} value of the isotherm considered. By this the deviation between experimental and calculated data points near the zero line (at the edges of the concentration range, for instance) is not weighted too strong.

$$\mathrm{MRE_R} = \frac{1}{Nn} \sum_{j}^{N} \sum_{i}^{n} \left| \frac{h^{\mathrm{E}}_{i,j_{calc}} - h^{\mathrm{E}}_{i,j_{exp}}}{h^{\mathrm{E}}_{max_{exp}}(R_j)} \right| \qquad (2)$$

Additionally the interaction parameters of all three models were fitted simultaneously to both h^{E} and γ^{∞} data. This means that for one parameter set the deviation between the calculated and experimental h^{E} data as well as between the calculated and experimental γ^{∞} data had to be determined. These individual deviations were added. The sum then served as the objective function for the simultaneous optimization. The error criteria were the same as for the separate fittings to h^{E} data and γ^{∞} data, the mean range related relative error ($\mathrm{MRE_R}$) and the mean relative error (MRE), respectively.

2 Selected algorithms

In the first steps of parameter optimization of different group contribution models only local search algorithms like Simplex-Nelder-Mead [9], Gauss-Newton, Levenberg etc. were used. Because of the multimodal character of the non-linear regression problems which are to be treated here they are not suitable due to their deterministic way of action. This lead to the development of a Genetic Algorithm [3], that delivered substantial results which could be interpreted.

As, however, the optimization problem which is to be solved has a real valued character it was obvious to implement a similar but purely real valued acting algorithm. According to the theory of Rechenberg [10] it was tried to optimize parameters of group contribution models with the help of encapsulated Evolution Strategies using only a one dimensional step length (n_σ=1) and the 1/5 success rule. However, several tests with different tuning parameters delivered unacceptable results only. The development and implementation of multimembered non-encapsulated Evolution Strategies could be made immediately after contacting SCHWEFEL, Chair of Systems Analysis, Department of Computer Science, University of Dortmund. These new developments use multi-dimensional and, if necessary, correlated control of step length ($n_\sigma = n$, $n_\alpha = (n^2 - n)/2$). The first result was that conventional multimembered (μ, λ)- and $(\mu + \lambda)$-Evolution Strategies could not cope with non-linear regression problems without further ado. Especially (μ, λ)-Evolution Strategies showed worse results than $(\mu + \lambda)$-Evolution Strategies although they were thought to be more suitable for the self-adaption of the strategic variables [1].

Finally only the combination of the theories of RECHENBERG and SCHWEFEL lead to satisfying results. Here multimembered (μ, λ)- and $(\mu + \lambda)$-Evolution Strategies with multi-dimensional (correlated) step length control in the encapsulated version were used and delivered better results than the Genetic Algorithm [4, 5].

When using an encapsulated ES, the parent individuals on the second (higher) level are created by duplicating the parent individuals on the first (lower) level. Here a sequential isolated optimum seeking process is carried out in several planes (mostly two):

$$\left[r_1^{\vec{x}} r_1^{\vec{\sigma}} r_1^{\vec{\alpha}} \; \mu_1[s_1]\lambda_1 \; (r_2^{\vec{x}} r_2^{\vec{\sigma}} r_2^{\vec{\alpha}} \; \mu_2[s_2]\lambda_2)^{\gamma_2} \right]^{\gamma_1} - \mathrm{ES} \tag{3}$$

The 3-number letter code used in the strategy notation marks the used recombination mechanism for each plane in the following order: objective variables, standard deviations, and rotation angles, if necessary. The recombination operators can be chosen as $r_i^{\vec{x}}, r_i^{\vec{\sigma}}, r_i^{\vec{\alpha}} \in \{-, d, D, i, I, g, G\}$ [1]. A 2-number letter code indicates a renunciation of a correlated step length control. The selection mechanism s_i can be chosen as plus [+] or as comma [,] for each plane independently.

Every offspring λ_1 is the founder of a new sub-population for γ_2 new generations during the main iteration steps γ_1. This sub-population can act totally independently of the population on the first level. The offspring λ_1 coming from the first level is duplicated μ_2 times for every parent in the second plane. The offspring then works as a starting value for the iteration on level two. This encapsulated iteration always starts using newly initialized strategic variables (standard deviations $\vec{\sigma}$ and rotation angles $\vec{\alpha}$). After γ_2 generations usually the best individual which was found is then returned to the bottom level. The returned individuals when using an encapsulated ES are not assigned with the original set of variables which they got during the recombination-mutation-procedure in the course of the $(\mu_1[s_1]\lambda_1)$-ES on the bottom level, but the returned individuals are exposed to the selection procedure on the bottom level in the usual way. The symbol N marks a standardization of the definition areas of the parameters which were to be fitted. The parent individuals (objective variables) are always chosen by random during their initialization. When using an encapsulated ES, the parent individuals on the second (higher) level are created by duplicating the parent individuals on the first (lower) level.

3 Optimization procedure

The fitting procedure was done one maingroup interaction after the other. Most of the binary mixtures contain more than two maingroups though, therefore the fitting procedure of the parameters of the maingroup interaction of interest refers to already fitted parameters.

An optimization procedure is done in the following way: the maingroup interaction, for which parameters are to be fitted is marked. If any already fitted parameters of other maingroup interactions are to be taken into account, then these interactions are also marked. For example: the optimization of the interaction between H_2O and CH_nOH without referring to other interaction parameters is carried out with experimental data of water/methanol mixtures only. If the interactions $CH_n - CH_nOH$ and $CH_n - H_2O$ are marked in a way that their (previously fitted) parameters apply, the optimization of $H_2O - CH_nOH$ is carried out with experimental data of water/n-alkanol mixtures (including water/methanol systems).

3.1 4+2 optimization

During the first calculations it became apparent that repeated optimizations for one group combination lead to different parameter sets with almost the same results. Another problem was the following: if a group combination which parameters have to be optimized refers to already fitted parameters of other group combinations (which remained constant), EBGCM lead to only unsatisfying results, i. e. incompatibilities between different parameter sets.

These problems lead to the idea to simplify the optimization procedure for the algorithms. Since the complexity of the optimization space raises with increasing number of parameters, the optimization was splitted in two steps. Due to the form of the temperature dependency of EBGCM with respect to a reference temperature it is possible to determine only four out of six parameters (h_{kj}, h_{jk}, a_{jk1}, and a_{kj1}) in the first step if only experimental data at reference temperature apply. After this the remaining two parameters (a_{jk2} and a_{kj2}) can be fitted to data not equal to the reference temperature (the previously fitted four parameters remain constant). Finally all six parameters are transferred to a deterministically acting algorithm, like the Simplex method of Nelder and Mead [9].

The used algorithms were the following:

4-parameter:	$[\texttt{dGG 3+6(dGG 6+20)}^{200}]^{10..20}$ - ES *followed by* Simplex-Nelder-Mead with 500..2000 iteration steps
2-parameter:	$[\texttt{dGG 3+6(dGG 6+20)}^{200}]^{10..20}$ - ES *followed by* Simplex-Nelder-Mead with 100..200 iteration steps
6-parameter:	Simplex-Nelder-Mead with 500..2000 iteration steps

3.2 2+2+2 optimization

To further minimize the dimension of the optimization space it was searched for a possibility to split the 4-parameter optimization (h_{kj}, h_{jk}, a_{jk1}, and a_{kj1}). The fractions h_{kj}/a_{jk1} and h_{jk}/a_{kj1} serve as weighting factors of the sums in the model equations. Additionally the h_{kj} and h_{jk} do not appear at any other place. So the parameter fractions were replaced by the new parameters h'_{kj} and h'_{jk}.

In the first step of the splitting of the 4-parameter optimization only the two parameters a_{jk1} and a_{kj1} can be fitted to data at reference temperature with the parameters h'_{kj} and h'_{jk} equal to 1. In the second step the just fitted a_{jk1} and a_{kj1} remain constant and the h'_{kj} and h'_{jk} are optimized. Then all four parameters are transferred to a deterministically acting algorithm. The final steps are the same as for the 4+2 optimization: the fitting of the parameters which describe the temperature dependency (a_{jk2} and a_{kj2}) and the transfer of all six parameters to a deterministically acting algorithm.

The algorithms used for this optimization were the same as for the 4+2 optimization. During the fitting process it became apparent that the optimization spaces for the 2-parameter problems exhibited only minor multimodality or were even unimodal (see section 4.2). So the parameters of some interactions were fitted with a Genetic Algorithm (that was less time consuming than an Evolution Strategy) followed by the Simplex-Nelder-Mead method, or just with the Simplex-Nelder-Mead method omitting the Genetic Algorithm.

3.3 Optimization of all six parameters at a time

After calculating the model parameters with the splitted optimizations, an attempt for the optimization of all six parameters per maingroup interaction at a time an encapsulated Evolution Strategy (ES) with optimized tuning parameters was made. Therefore a test system with three excess enthalpy data sets was created and different strategies were tested.

Table 1 shows an extract of all on the basis of the 6-parametric test system tested Evolution Strategy. Besides the notation of the analyzed ES the arithmetic mean of all 50 $\mathrm{MRE_R}$-results in per cent are shown. The start step length was either defined absolutely (e. g.: 25) for every parameter the same or was defined in per cent (e. g.: 20%) of each total definition area (absolute values) divided by $\sqrt{n}$ (n = number of objective variables), in order to make the standard deviations independent from n and from different definition areas of the parameters. The rotation angles - if used - were initialized by random between $[-\pi, +\pi]$ as suggested in [1, 12]. Besides that the number of the needed function calls per run is listed as well as the best and the worst determined error of the 50 runs and finally the found out standard deviation for all runs is shown to characterize the reliability and ability of reproduction of every used ES. The symbol N marks a standardization of the definition areas of the parameters which were to be fitted. The parent individuals (objective variables) are always chosen by random during their initialization.

Table 1: Optimization results of different Evolution Strategies [4]

ES-notation	$\mathrm{MRE_R}$/% average	step-width	func. calls	$\mathrm{MRE_R}$/% best	$\mathrm{MRE_R}$/% worst	$\sigma_{(50)}$ /%
$\texttt{(dI 120,800)}^{380}$	10.61	25	304000	10.30	10.82	0.09
$\texttt{(dI 120+800)}^{380}$	10.16	25	304000	5.53	10.70	1.18
$\texttt{15*(dI 15,100)}^{200}$	7.75	25	300000	0.25	10.13	1.95
$\texttt{15*(dI 15+100)}^{200}$	5.85	25	300000	0.19	10.02	2.57
$\texttt{[GG 4,8(GG 7,19)}^{200}\texttt{]}^{10}$	5.05	20%/4%	304080	0.22	9.61	2.06
$\texttt{[GG 4+8(GG 7,19)}^{200}\texttt{]}^{10}$	4.39	20%/4%	304080	0.11	9.04	2.02
$\texttt{[GG 4,8(GG 7+19)}^{200}\texttt{]}^{10}$	2.91	20%/4%	304080	0.03	9.05	2.37
$\texttt{[GG 4+8(GG 7+19)}^{200}\texttt{]}^{10}$	1.19	20%/4%	304080	0.02	8.25	1.99
$\texttt{[GG 4+8(GG 7+19)}^{200}\texttt{]}^{10}$	0.81	10%/4%	304080	0.00	5.95	1.59
$\texttt{[GG 4+8(GG 7+19)}^{200}\texttt{]}^{10}$	0.47	5%/4%	304080	0.00	5.38	1.25
$\texttt{[GG 4+8(GG 7+19)}^{200}\texttt{]}^{10}$	0.77	20%/2%	304080	0.00	5.53	1.48
$\texttt{[GG 4+8(GG 7+19)}^{200}\texttt{]}^{10}$	1.33	20%/1%	304080	0.00	5.53	1.92
$\texttt{[-- 4+8(-- 7+19)}^{200}\texttt{]}^{10}$	3.61	20%/4%	304080	0.00	8.92	2.55
$\texttt{[dI 4+8(dI 7+19)}^{200}\texttt{]}^{10}$	2.31	20%/4%	304080	0.01	5.55	2.28
$\texttt{[dG 4+8(dG 7+19)}^{200}\texttt{]}^{10}$	1.92	20%/4%	304080	0.02	8.69	2.38
$\texttt{[GG 2+8(GG 7+19)}^{200}\texttt{]}^{10}$	1.16	20%/4%	304080	0.00	7.83	1.77
$\texttt{[GG 4+8(GG 2+19)}^{200}\texttt{]}^{10}$	1.86	20%/4%	304080	0.04	6.14	2.09
$\texttt{[GG 4+8(GG 4+19)}^{200}\texttt{]}^{10}$	1.50	20%/4%	304080	0.01	5.79	1.95

Table continued

ES-notation		MRE$_R$/% average	step-width	function calls	MRE$_R$/% best	worst	$\sigma_{(50)}$ /%
$[GG\ 4+8(GG\ 15+19)^{200}]^{10}$		2.29	20%/4%	304080	0.02	7.63	2.40
$[GGG\ 4+8(GGG\ 7+19)^{200}]^{10}$	N	5.15	20%/4%	304080	0.24	9.73	2.13
$[GGG\ 4+8(GGG\ 7+19)^{200}]^{10}$	N	4.65	20%/4%	304080	0.11	9.24	2.41
$[GG\ 4+8(GGG\ 7+19)^{200}]^{10}$	N	1.77	20%/4%	304080	0.00	5.53	2.30
$[GGG\ 4+8(GGG\ 7+19)^{200}]^{10}$		11.27	20%/4%	304080	2.58	26.65	6.09
$[GG\ 3+8(GG\ 7+19)^{100}]^{40}$		0.02	5%/2%	608320	0.00	0.20	0.03

The obtained results in table 1 clearly show that conventional multimembered non-encapsulated ES even at high population values only deliver very unsatisfying results. The results show that [+]-strategies can obviously act better within the solution space of the given optimization problem than [,]-strategies. This in all probability has the reason that the adaptive surface at the n-dimensional space of variables is characterized by narrow long and bent valleys [3, 13]. This seems to lead to a more difficult self-adaption of the strategic variables when using [,]-ES than [+]-ES. Evidently ES show difficulties following narrow ridge-like search paths (especially at rising number of dimensions). Considering that it can be understood why repeated non-encapsulated ES - started from new randomly generated starting points by keeping the fittest during the optimum seeking process (e. g. $15*(GG\ 15+100)^{200}$) - lead to better optimization results: by the repetition of the search for the optimum an isolation in the solution space occurs which has the consequence that the probability of finding the global or a very good local optimum is increased. The difficulties of the given optimization problem can already be seen when looking at the results of repeatedly starting ES,- the adaptive surface is so complicated that it is crucial to find convenient starting points which is guaranteed by a continuously repeated start of the algorithm.

The last paragraph of table 1 finally shows the result of an ES based on the determined optimum tuning parameters concluded from all results shown by way of extract in table 1. The optimum tuning parameters can only be transferred partly to the parameter optimization of different group contribution models or optimization problems of different types of non-linear regression. The optimum tuning parameters which were found are to be understood as indicatory values which can be used for different optimization problems of similar type.

Detailed information about the determined results of the different strategy types can be obtained in [4, 5].

4 Results and Discussion

4.1 Optimization to h^{E} data

The 4+2 optimization was carried out with h^{E} data, as well as the optimization of all six parameters at once. The overall average mean relative range related error ($\mathrm{MRE_R}$) of all 31 maingroup interactions is 31.11% for the 4+2 optimization, whereas the optimization of all six parameters exhibits a deviation of nearly half the size (16.17%). The results for the single interactions are given in table 2.

Table 2: Comparison between the 4+2 and 6-parameter optimization

Maingroup interaction		data	$\mathrm{MRE_R}$/% 4+2	$\mathrm{MRE_R}$/% all 6
CH_n	cCH_n	1141	30.22	29.22
CH_n	$CH_n =$	246	30.15	29.16
CH_n	$CH_n \equiv$	312	58.91	15.30
CH_n	aCH_n	1774	25.20	21.34
CH_n	H_2O	1064	18.88	15.40
CH_n	CH_nOH	4008	14.44	14.43
CH_n	gCH_nOH	167	7.20	23.98
CH_n	$CH_m(O)CH_n$	1272	19.55	24.32
CH_n	CH_nCHO	159	5.56	5.56
CH_n	$CH_m(CO)CH_n$	441	9.55	9.33
cCH_n	$CH_n =$	148	11.75	11.49
cCH_n	aCH_n	826	17.01	17.06
cCH_n	H_2O	29	2.67	6.38
cCH_n	CH_nOH	1978	22.56	15.22
cCH_n	$CH_m(O)CH_n$	364	18.10	22.71
cCH_n	CH_nCHO	18	6.96	11.80
cCH_n	$CH_m(CO)CH_n$	477	6.58	6.18
$CH_n =$	aCH_n	12	50.67	0.67
$CH_n =$	$CH_m(O)CH_n$	139	37.33	8.48
aCH_n	H_2O	57	2.28	2.17
aCH_n	CH_nOH	723	37.76	17.32
aCH_n	$aCOH$	57	2.28	2.17
aCH_n	$CH_m(O)CH_n$	300	274.29	45.42
aCH_n	CH_nCHO	112	67.45	25.00
aCH_n	$CH_n(CO)CH_n$	335	87.00	55.61
H_2O	CH_nOH	174	19.51	11.45
H_2O	gCH_nOH	116	27.13	11.33

Table continued

Maingroup interaction		data	$MRE_R/\%$	
			4+2	all 6
H_2O	$aCOH$	57	2.28	2.17
H_2O	$CH_m(CO)CH_n$	406	4.11	5.20
CH_nOH	gCH_nOH	167	7.20	23.98
CH_nOH	$CH_m(CO)CH_n$	1667	39.95	11.47
Overall average		18746	31.11	16.17

Examination of the single maingroup interactions shows that there are some interactions that have smaller deviations for the 4+2 optimization. For the interaction $cCH_n - CH_nCHO$ both optimization methods lead to the same parameter set but different deviations. The fitting procedure of this interaction refer to parameter sets of two other interactions $CH_n - cCH_n$ and $CH_n - CH_nCHO$. The parameter set (and the deviation) of $CH_n - CH_nCHO$ are the same for both fitting procedures. The optimization of $CH_n - cCH_n$ lead to different parameter sets and a smaller deviation for the 6-parameter optimization (29.22%, 4+2: 30.22%). Although the 6-parameter optimization found a better parameter set, it is obviously not the best parameter set for the fitting of the interaction $cCH_n - CH_nCHO$. The same can be noticed for the optimiztion of the interactions $cCH_n - CH_m(O)CH_n$ and $CH_n - gCH_nOH$.

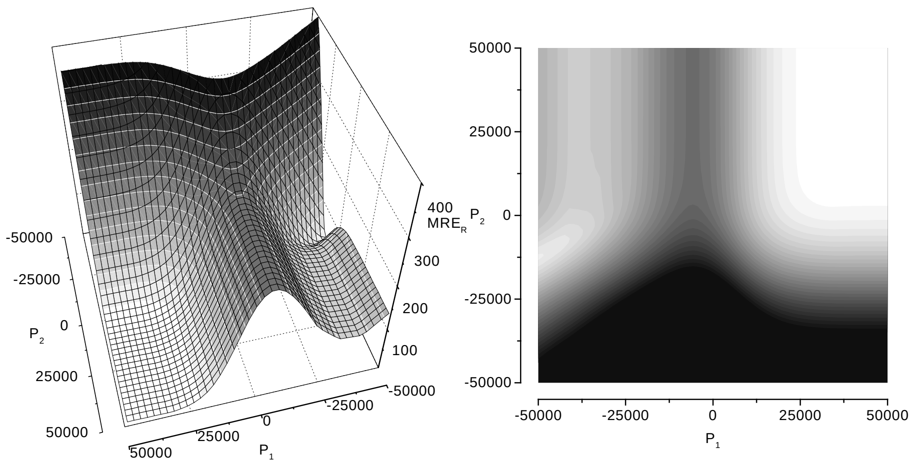

Figure 3: 4+2: adaptive surface for $aCH_n - H_2O$; a_{jk1} and a_{kj1} variied

A proof for the multimodality of the optimization space is the optimization of $CH_n - CH_nOH$. Both optimization methods lead to almost the same deviation (all 6: 14.43%, 4+2: 14.44%) with different parameter values. The same characteristics show the interactions $aCH_n - H_2O$, $aCH_n - aCOH$ and $H_2O - aCOH$, which had to be fitted simultaneously, because only experimental data of phenol/water mixtures existed. The deviations are almost the same (all 6: 2.17%, 4+2: 2.28%) with totally different parameter values. Figure 3 shows the adaptive surface of the variied parameters a_{jk1} and a_{kj1} of the interaction $aCH_n - H_2O$ while the other parameters (which were optimized with the 4+2 method) remain constant. Figure 3 show the same surface but with the parameters obtained with the 6-parameter optimization. Both plots show similar multimodal optimization spaces. Besides a plateau (at $P_1 > 25000$ and $P_2 > 0$) both plots exhibit a valley at $P_1 \approx 25000$ which begins to bend at $P_2 \approx 0$. Additionally the valley lowers at this point (and thus forms a local minimum). It is interesting to note that, despite the different parameter sets both surfaces are similar.

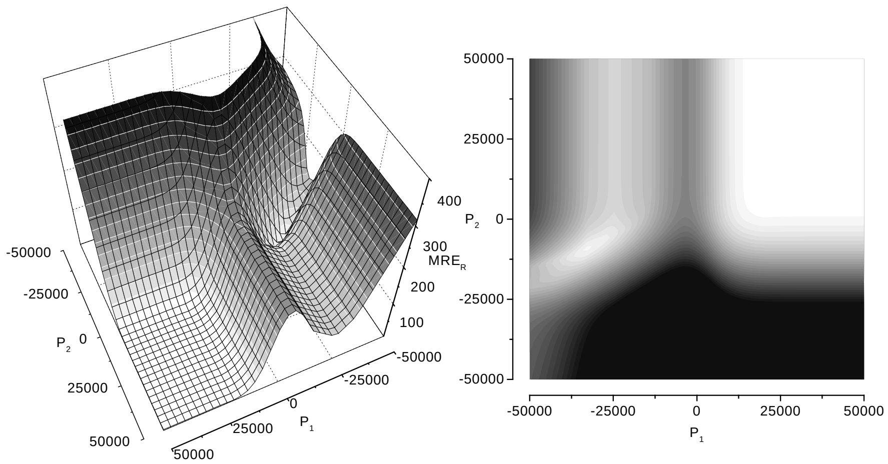

Figure 4: all 6: adaptive surface for $aCH_n - H_2O$; a_{jk1} and a_{kj1} variied

4.2 Simultaneous optimization to h^{E} and γ^{∞} data

During the fitting processes to either h^{E} or γ^{∞} data some interactions showed incompatibilities with previously fitted parameter sets, which had to be used in the optimization. This suggests that the objective function could become more definite by introducing additional constraints. A chance to achieve such a more definite objective function is the determination of the different model parameters by simultaneously fitting them to several thermodynamic properties such as γ^{∞} and h^{E}.

Initially, the simultaneous fitting to both quantities was carried out with the 2+2+2 optimization. The greatest advantage for using the 2+2+2 optimization is the reduced multimodality of the optimization space. Figures 5 to 7 show the adaptive surfaces of the optimization of $CH_n - aCH_n$. These are the real surfaces, where the algorithm has to find the global optimum. In some cases, the optimization spaces for both quantities are even unimodal. As expected, the objective function seems to get more definite if the optimization is carried out with different thermodynamic quantities.

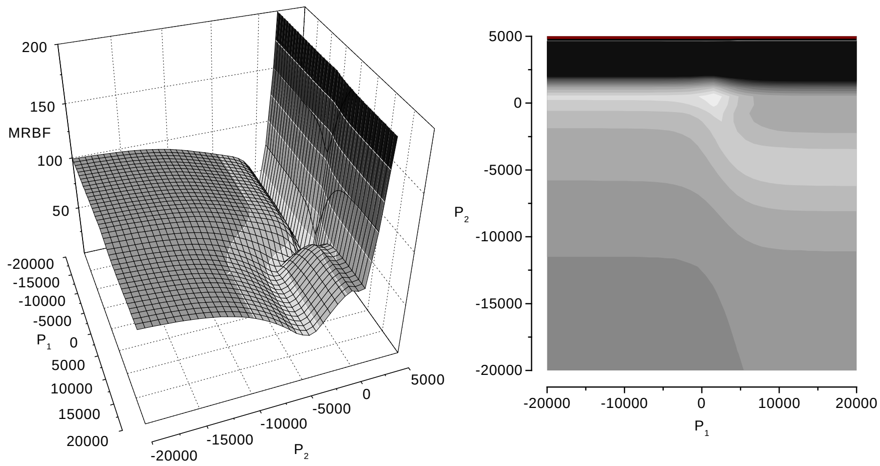

Figure 5: adaptive surface for γ^{∞}, variied parameters: a_{jk1} and a_{kj1}

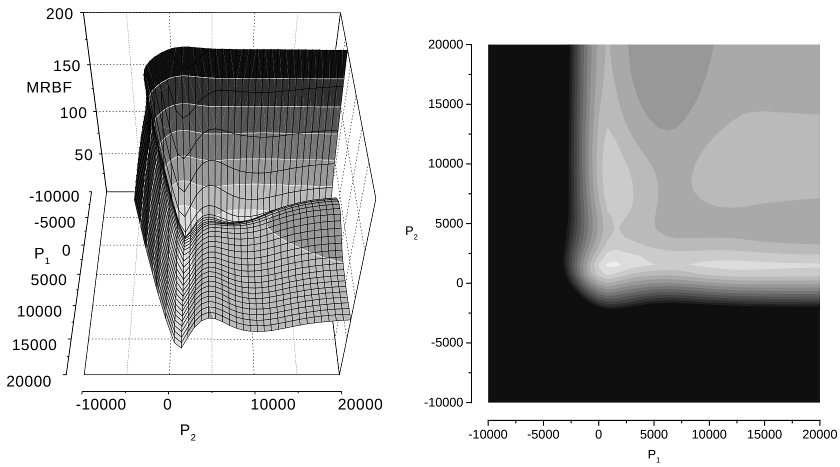

Figure 6: adaptive surface for h^{E}, variied parameters: a_{jk1} and a_{kj1}

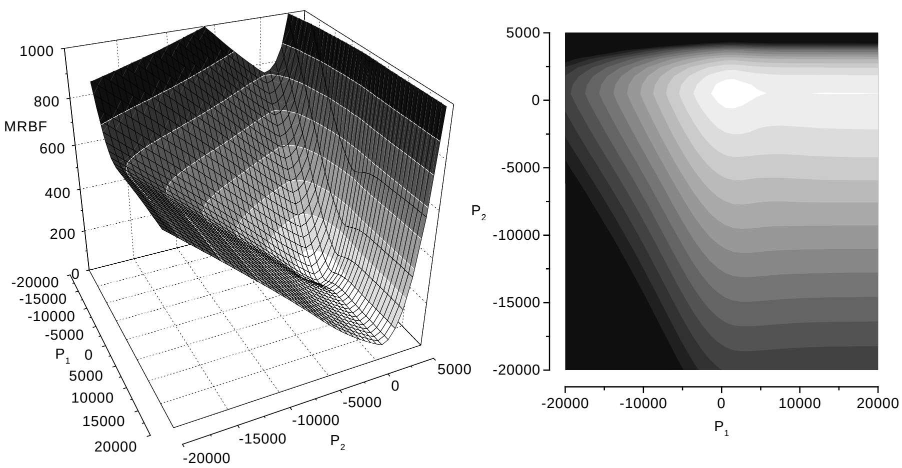

Figure 7: adaptive surface for $h^{\mathrm{E}}+\gamma^{\infty}$, variied parameters: a_{jk1} and a_{kj1}

With the newly developed Evolution Strategy a 6-parameter optimization was carried out. A comparison to the 2+2+2 optimization shows that interactions that can be fitted without other incoming parameter sets exhibits nearly the same deviations but totally different parameter sets, except for $CH_n - CH_nOH$ (see table 3). These parameter sets differ only slightly. The parameters of the interactions $CH_n - H_2O$ and $H_2O - CH_nOH$ were fitted altogether, i. e. a 12 parameter optimization. The 2+2+2 optimization lead to a much better result than the 6-parameter optimization (which, however, was a 12-parameter optimization in this case). The increased variable dimension seemed to be a problem for the unsplitted optimization procedure.

Table 3: Comparison between the 2+2+2 and 6-parameter optimization

Maingroup interaction		data h^{E}	data γ^{∞}	MRE$_{\mathrm{R}}$ (%) 2+2+2	MRE$_{\mathrm{R}}$ (%) all 6
CH_n	cCH_n	1141	82	17.88	17.47
CH_n	$CH_n=$	246	218	18.81	18.53
CH_n	aCH_n	1774	447	20.36	20.20
CH_n	H_2O	1064	12	15.92	18.78
CH_n	CH_nOH	4008	424	24.52	24.40
CH_n	$CH_m(O)CH_n$	1272	44	11.15	11.07
CH_n	CH_nCHO	159	41	6.15	4.87
CH_n	$CH_m(CO)CH_n$	441	226	11.33	11.08
cCH_n	$CH_n=$	148	11	94.35	12.83
cCH_n	aCH_n	826	26	34.61	16.94
cCH_n	CH_nOH	1929	114	33.24	26.49
cCH_n	$CH_m(O)CH_n$	364	13	—	10.44
cCH_n	$CH_m(CO)CH_n$	477	27	35.19	12.96
$CH_n=$	aCH_n	12	53	37.73	40.50
aCH_n	CH_nOH	723	167	16.76	17.59
aCH_n	$CH_m(CO)CH_n$	335	67	46.36	34.15
H_2O	CH_nOH	174	12	15.92	18.78
H_2O	$CH_m(CO)CH_n$	406	4	60.41	30.66
CH_nOH	$CH_m(CO)CH_n$	1667	147	32.01	25.24
Overall average		17166	2315	29.59	20.14

Interactions that refer to already fitted parameter sets gave only unsatisfying results with the 2+2+2 method, especially for the interactions $cCH_n - CH_n =$ and $H_2O - CH_m(CO)CH_n$. The previously fitted parameter sets which apply for the optimization seemed to be incompatible. The application of the 2+2+2 optimization compensate the advantage of the more definite objective function. This problem did not occur with the 6-parameter optimization.

5 Conclusion

The optimization of the model parameters of the group contribution model EBGCM was carried out in different ways with several Evolutionary Algorithms.

The splitting of the 6-parameter optimization into several steps in order to minimize the variable dimension lead to contrary results: on one hand the optimization results are of equal quality than for the 6-parameter optimization. Furthermore, for the simultaneous optimization of more than one maingroup interaction (i. e. at least 12 parameters altogether) results are even better. On the other hand though, the splitted optimization caused incompatibilities between different parameter sets.

This lead to the result that the 2+2+2 optimization may useful in the reduction of variable dimension if several maingroup interactions are to be fitted simultaneously to avoid incompatibilities between the parameter sets. If the maingroup interactions are to be optimized one after another, the 6-parameter optimization applies. Another advantage of the 6-parameter optimization is the less time consuming optimization process.

In the future further examinations of multiparameter optimization compared to optimizations of only one maingroup interaction will be made. The application of the 2+2+2 optimization in order to reduce the variable dimension with special regard to the appearence of incompatibilities between different parameter sets will be further tested.

References

[1] T. Bäck, *Evolutionary Algorithms in Theory and Practice*, Oxford Univ. Press, New York, 1996.

[2] A. Fredenslund, B. L. Larsen and P. Rasmussen, A modified UNIFAC group contribution model for prediction of phase equilibria and heats of mixing, *Ind. Eng. Chem. Res.* **26** (1987) 274-2286.

[3] T. Friese, P. Ulbig and S. Schulz, Use of Evolutionary Algorithms for the Calculation of Group Contribution Parameters in order to Predict Thermodynamic Properties. Part 1: Genetic Algorithms, *Computers and Chem. Eng.*, in press.

[4] H. Geyer, P. Ulbig and S. Schulz, Use of Evolutionary Algorithms for the Calculation of Group Contribution Parameters in order to

Predict Thermodynamic Properties. Part 2: Evolution Strategies, *submitted for publication.*

[5] H. Geyer, P. Ulbig and S. Schulz, Encapsulated Evolution Strategies for the Determination of Group Contribution Model Parameters in order to predict Thermodynamic Properties, *Fifth International Conference on Parallel Problem Solving from Nature*, Amsterdam, 27-30 September 1998 (accepted).

[6] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, 1989.

[7] U. Hoffmann and H. Hofmann, *Einführung in die Optimierung*, Verlag Chemie, Weinheim, 1971.

[8] C. Kracht, T. Friese, P. Ulbig and S. Schulz, Development of an Enthalpy Based Group Contribution g^{E} Model (EBGCM), *submitted for publication.*

[9] J. A. Nelder and R. Mead, A simplex method for function minimization, *Computer Journal* **7** (1965) 308-313.

[10] I. Rechenberg, *Evolutionsstrategie '94, Werkstatt Bionik und Evolutionstechnik*, Band 1, Friedrich Frommann, Stuttgart, 1994.

[11] H.-P. Schwefel, *Numerical Optimization of Computer Models*, Wiley, Chichester, 1981.

[12] H.-P. Schwefel, *Evolution and Optimum Seeking*, Wiley, New York, 1995.

[13] P. Ulbig, *Gruppenbeitragsmodelle UNIVAP & EBGCM* (PhD Thesis), Hannemann, Gelsenkirchen, 1996.

[14] P. Ulbig, T. Friese, H. Geyer, C. Kracht and S. Schulz, Prediction of thermodynamic properties for chemical engineering with the aid of Computational Intelligence, *Progress in Connectionist-Based Information Systems - Proceedings of the 1997 International Conference on Neural Information Processing and Intelligent Information Systems, Vol. 2*, Springer, New York, 1997.

[15] U. Weidlich and J. Gmehling, A modified UNIFAC Model, *Ind. Eng. Chem. Res.* **26** (1987) 1372-1381.