Available online at www.sciencedirect.com
SCIENCE DIRECT
Applied Catalysis A: General 303 (2006) 72–80
APPLIED CATALYSIS A: GENERAL
www.elsevier.com/locate/apcata

Evaluation of catalyst library optimization algorithms: Comparison of the Holographic Research Strategy and the Genetic Algorithm in virtual catalytic experiments

András Tompos [a], József L. Margitfalvi [a,*], Ernő Tfirst [a], Lajos Végvári [b]

[a] *Institute of Surface Chemistry and Catalysis, Chemical Research Center, Hungarian Academy of Sciences, 1525 Budapest, PO Box 17, Hungary*

[b] *Meditor General Innovation Bureau, 2623 Kismaros, Zrínyi u. 16, Hungary*

Received 12 July 2005; received in revised form 19 January 2006; accepted 23 January 2006

Abstract

In this study two catalyst library optimization methods, the Holographic Research Strategy (HRS) and the Genetic Algorithm (GA) were compared based on their ability to find the optimum compositions in a given multi-dimensional experimental space. Results obtained in three different case studies were used to investigate both the rate and the certainty of the optimum search. In these case studies the activity–composition relationships were established using Artificial Neural Networks (ANNs) trained with catalytic data published earlier. The above relationships were used in ''virtual optimization experiments'' using both HRS and GA for catalyst library optimization. Upon using the stochastic GA its exceedingly divers mode of sampling often resulted in poor catalytic materials in the next catalyst generation. This fact resulted in a decreased rate of convergence to the optimum. In contrast, in HRS, which is a deterministic optimization algorithm, a moderate level of diversity in the catalyst library can easily be achieved. In this way an acceptable rate in optimum search can be accomplished. The visualization ability of HRS allows the illustration of all virtually tested compositions in a two-dimensional form regardless the optimization algorithm used. Upon using HRS a structured arrangement of experimental points in the virtual holograms was observed. However, when GA was applied for virtual optimization ''starry sky''-like arrangement of compositions in the virtual holograms was obtained. Therefore based on virtual holograms, upon using HRS the relationship between the composition of catalytic materials and their performance can be qualitatively revealed, while no similar correlation can be obtained using GA.

Keywords: Catalyst library design; Combinatorial catalysis; Holographic Research Strategy; Genetic Algorithm; Visualization

1. Introduction

The number of catalytic tests that could be made in parallel way strongly depends on the type of reactions and the experimental conditions used. In gas phase reactions several hundreds catalysts can simultaneously be tested. However, in high-pressure liquid phase catalytic reactions only 8–16 parallel experiments can be performed. In the former case huge catalyst libraries can be designed, while in the latter one the rational approach does not allow to test libraries containing more than 250–300 catalysts. Consequently, the informatic platform used to design catalyst libraries for high-pressure liquid phase reactions should have fast optimization and information mining tools.

It has to be mentioned that in most of the cases in catalyst library design the final goal is to develop multi-component catalysts. The key parameters in the design of heterogeneous catalyst libraries are: (i) the number of variable elements, (ii) the maximum number of components in one composition, (iii) the amount or ratios of components, and (iv) the total amount of elements. This approach results in large experimental spaces, which contain catalysts different from each other only in composition.

In the field of combinatorial catalysis different approaches are used for library design. Industrial companies, like Symix, Avantium, hte GmbH, are using their own proprietary methods. In academic research the Genetic Algorithm (GA) is widely applied [1–3]. Recently the combination of GA with Artificial

* Corresponding author. Tel.: +36 1 3257747; fax: +36 1 3257554.
E-mail address: joemarg@chemres.hu (J.L. Margitfalvi).

0926-860X/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.apcata.2006.01.028

Neural Networks (ANNs) has been reported [4–8]. In these works ANNs have been used for the establishment of composition–activity relationships in different catalyst libraries.

Recently we described a new approach, the Holographic Research Strategy (HRS) and its combination with ANNs [9–11] for catalyst library optimization. It was shown previously that HRS is a deterministic approach, i.e. the route of optimization is unequivocally determined by the applied setting parameters [9]. These settings are: (i) the total number of components and their concentration levels, (ii) the number and arrangement of experimental points of the initial catalyst library in the experimental hologram, (iii) the number of the best hits around which the new catalyst generation is created, and (iv) the size and the form of the experimental regions used to design the next catalyst generation [9,11]. Presumably all of these parameters can affect in some extent the rate and the certainty of optimum search. In this respect the importance of the size of the experimental region has already been discussed in our previous study [9]. A straightforward correlation was found, i.e. the smaller is the experimental window the faster is the rate of the optimization. However, there is a limit in reducing the size of the experimental region. The smaller the experimental region, the higher is the chance to miss the global optimum. If the size of the experimental region is not properly chosen HRS can get stuck in a local optimum.

Contrary to HRS, GA is a stochastic, multivariate approach. In case of GA even in one given experimental space using two different qualitative settings (such as, e.g. two different rates of qualitative and/or quantitative mutations) numerous optimum search can be performed [1–3]. In this study the rate and certainty of convergence to the optimum of HRS will be compared to that of the GA. In this work three case studies were investigated. In these case studies the following approach was used: previously published experimental data obtained in different catalysts library design were analyzed by ANNs creating a definite mathematical expression describing the activity–composition relationship for the given reaction. The obtained exact mathematical relationships between the composition and the activity have been applied for ''virtual optimization'' of catalyst libraries using both HRS and GA optimization algorithms.

ANNs were trained with results obtained in the following three catalytic reactions: (i) methane oxidation [10], (ii) propane oxidation [2] and (iii) methane oxidative coupling [5]. Further details will be given in Section 2.3.

2. Methods

2.1. Characteristic procedures in HRS

The construction of experimental holograms is described in detail in our previous study [9]. In the two-dimensional representation of a multi-dimensional experimental space the discrete concentration levels of components are represented by lines. The level of each component increases gradually till it reaches its maximum then it decreases gradually again. This mode of representation leads to wavelike arrangement of levels (see Figs. 1–3). It is very important to emphasize that the levels of components are arranged in such a way that each component has different periodicity. Accordingly, moving along any axis from one experimental point to the next one only one level of one variable is changed.

The initial catalyst library has been created using elements of symmetry of experimental spaces according to our previous routine [9–11]. The initial experimental points have been fixed as small clusters: (i) in the center, (ii) along the symmetry axes and (iii) in the corners of experimental spaces resulting in 48 different catalyst compositions. The forthcoming generations have been created by a rectangular-shaped experimental region 4×4 in size around the best three hits of the preceding generations [9–11]. Prior to the formation of the experimental regions a variable position change has to be done according to our previous routines [9–11]. This step is considered as the main driving force in HRS [9]. As a result of variable position changes the arrangement of compositions in the experimental hologram is altered, eventually after variable position changes each hit has a new neighborhood.

In this work catalyst libraries are virtually designed and tested for three reactions, such as methane oxidation, propane oxidation and methane oxidative coupling. The corresponding objective functions are the conversion for methane and propane, and the yield of C_2 hydrocarbons, respectively. Artificial Neural Networks have been used for virtual catalytic tests. ANNs provide the quantitative relationship between composition and catalytic performance. ANNs describing the objective functions in the above experimental spaces were previously trained with data represented in Refs. [10,2,5], respectively, as discussed in detail in Section 2.3. The concentration levels of different components in the three above mentioned catalyst libraries are given in Tables 1–3.

2.2. Characteristic procedures in Genetic Algorithm

Genetic Algorithm is a stochastic approach for catalyst library optimization. The initial catalyst library is created by a random procedure, which is followed by an iterative process during which the catalysts of the next generation are determined by using catalytic results of the previous one.

In this report the method described in studies of Baerns et al. [1] was used with small modification adapted to the discrete concentration levels applied in our approach. Correspondingly, the following operators have been involved in the process leading for the creation of the forthcoming generations:

- crossover,
- qualitative mutation, and
- quantitative mutation.

The probabilities of these operators are determined by the following formulas as described by Baerns et al. [1]:

$$W_{\text{cross}} = A\frac{P_{\text{best}} - BP_{\text{mean}}}{P_{\text{best}}} \quad (1)$$

$$W_{\text{quant}} = A\frac{BP_{\text{mean}}}{P_{\text{best}}} \quad (2)$$

$$W_{\text{mut}} = 1 - W_{\text{cross}} - W_{\text{quant}} \quad (3)$$

where the so called control parameters A and B are equal to 0.5 in this study. Similar values were used in earlier studies [1]. P_{mean} and P_{best} are the mean and the best value of catalytic performance.

The action of different operators mentioned above is discussed in detail in Ref. [1]. Correspondingly, during a crossover one randomly selected component of two different catalytic materials is exchanged, while during qualitative mutation yes/no inversion of one randomly selected component of a catalyst has been performed. However, as far as in our approach the concentration of the components have only discrete levels, the process of quantitative mutation has been altered in comparison to that described in Ref. [1]. As a consequence, the concentration level of component i, L_i in a quantitative mutation is controlled by a random integer number, t (0; 1):

$$L_i^{\text{new}} = L_i^{\text{preceding}} + (-1)^t \quad (4)$$

This is an adaptation of GA using discrete levels for components. In order to analyze the effect of the above modification on the rate and certainty of optimum search it seems to be inevitable to implement also the original version of Genetic Algorithm using continuos domain for component loadings [1]. In the latter case the new concentration of a component, that come through quantitative mutation can be calculated as follows:

$$x_i^{\text{new}} = x_i^{\text{preceding}} + (-1)^t \frac{x_i^{\text{preceding}}}{2} \quad (5)$$

where x_i represents concentration value of the selected component. Using this version of GA the compositions are not forced to stay within the experimental spaces defined by sets of levels given in Tables 1–3. The effect of the higher ''freedom'', resulted from the use of continuos domain, on the certainty of the convergence to the optimum in comparison to GA using discrete levels will be discussed in Section 3.

Artificial Neural Networks (ANNs) have been used to obtain the composition–performance (conversion or yield) relationship and in such a way ANNs have been applied to perform virtual catalytic experiments using GA. Upon using GA, both applying continuos domain and discrete levels, 50 optimizations have been performed in each experimental space. For the analysis of the performance of GA the best optimization resulted in the fastest convergence to best attainable catalytic performance was selected. In each optimization the population size of one catalyst generation was 30 and 15 consecutive generations have been created. The best five hits in the preceding generation become automatically members of the next one. In the process of creation of the new generation the remaining 25 catalysts have been determined one by one according to the following procedure:

1. First a decision was made which operator to use in the creation of a new catalyst. This decision was based on the probabilities calculated by Eqs. (1)–(3).
2. Catalyst or catalysts considered as a parent/parents are selected from the whole preceding generation. In this process each member of the preceding generation, i.e. 30 catalysts, take part:

- fitness-proportional selection for quantitative mutation,
- random selection for qualitative mutation, and
- crossover between one catalytic material selected in a fitness-proportional manner and the other catalytic material chosen randomly.

2.3. Artificial Neural Networks

ANNs are usually created for the evaluation of the objective functions describing quantitatively the activity–composition relationships in a given multi-dimensional experimental space. The formation of ANNs requires previously measured experimental data. Accordingly, in the present study results obtained in earlier studies in the optimization of three different catalysts libraries [2,5,10] were used. It has to be emphasized that in the above three studies different optimization algorithms have been used.

For methane oxidation 167 different catalyst compositions optimized by HRS [10] have been explored. In low temperature oxidation of propane 150 different catalyst compositions obtained in three generations using GA [2] were examined. In oxidative coupling of methane 76 catalysts optimized by applying a hybrid algorithm based on alternating use of GA and ANNs [5] were investigated.

For proper formation of ANNs and for assessment of their predictive ability the available data of each catalyst library have been divided into three well-distinguishable sets: (i) training, (ii) validation and (iii) testing. It has to be emphasized that in each set all sorts of catalysts from poor to good catalytic activity are involved with the same rate. In each experimental space investigated in this study the ratio of training, validation and testing sets was set to approximately 70: 15:15, respectively.

The process of training and validation has been described in detail elsewhere [7]. The networks are trained with resilient back-propagation algorithm [7]. Training is stopped if the validation error increases for more then two consecutive epochs, an epoch is defined as a pass through the entire set of training and validation patterns. This protocol is invoked to prevent over-memorization during the training phase [7].

Nineteen different network architectures proposed by Cundari et al. [7] were investigated to achieve acceptable model accuracy. Every neural network architecture has been trained 1000 times (each training has been initialized with different, random node-to-node weights) [7]. According to the average mean square errors (MSE) (for both training and validation patterns) the resulted 19,000 networks were ranked. The best 100 networks, i.e. networks with the smallest MSEs have been involved into optimal linear combination [12],

during which so called OLC-network has been created. Eventually, in the three catalyst libraries the resulting three OLC-networks have been applied in this study for virtual catalytic tests in combination either with HRS or with GA in order to assess differences between the peculiarities of these two optimization strategies.

It has to be mentioned that in case of methane oxidative coupling only 76 data were available. Therefore optimal linear combination cannot be initiated from the best 100 networks, as the number of component networks cannot exceed the number of data. In this case only the best 50 networks have been characterized and used.

3. Results and discussion

3.1. Characterization of trained ANNs

The best 100 networks formed in three different experimental spaces are characterized in Table 4. It is remarkable that the first four-seven network architectures having the less number of connection weights seldom appeared among the best 100 nets. Furthermore, the structure itself seems to play a decisive role in the determination of model accuracy of a network. In all of the three investigated case studies among the best 100 nets, the occurrence of networks with at least two hidden layers in the architecture is significantly larger than that of the networks containing only one hidden layer.

In order to get the OLC-network linear combination of the best 100 networks formed in the training phase have been made. As emerges from Tables 5–7, upon using the OLC-network the mean square error on the test set is lower than that obtained by using any of the component networks. This finding indicates on a fairly good correlation between measured and predicted catalytic results and it supports the use of OLC-approach, since OLC-networks in all of these cases have superior performance to any component networks. The results show also that in the best 100 networks involved in the linear combination only 9, 6 and 4 networks have combination weights different from zero in the three investigated case studies, respectively.

3.2. Comparison of HRS and GA

As described in Section 2 50 optimum searches with GA, using both discrete levels and continuos domains for concentration of components, have been initiated in each experimental space. The result of the best runs, i.e. which provided the fastest search for the best catalytic performance, are shown and compared with results of holographic optimization. These results are given in Tables 8–10. As emerges from these data, HRS performs better than any kind of GA when the rate of optimum search is compared. In each case less number of generations and less number of catalysts were needed to find the global optimum. It has to be mentioned that the global optima found are valid only for the sets of discrete levels predefined in Tables 1–3. Further improvement in the catalytic activity can only be attained by establishing new concentration levels for the components around the present optimum. It allows refining the resolution as well, i.e. reducing the increments between the discrete values of levels as it was applied in our pervious studies [9,10]. Similar approach is also accepted in the classical factorial DOE method.

In case of GA exceedingly large diversity of materials is required in the process of optimization, which is due to the stochastic nature of this optimization algorithm. It can be mentioned that moderate level of compositional diversity is inevitable even in HRS. The large diversity of materials in GA necessarily leads to testing numerous catalytically poor materials. Probably, this fact is responsible for the relatively slow rate of approximation of the optimum when GA is used.

The easiest way to compare the diversity of compositions in optimized catalyst libraries is the visualization of experimental data, i.e. the representation of catalytic results as a function of composition in comprehensible figures. In our previous studies it has already been shown that beyond its excellent optimization feature HRS has also strong visualization ability [9,10]. This ability of HRS was also applied in this study. Upon using the ANNs, which provide the composition–activity relationship, the whole experimental space can be visualized by a two-dimensional hologram [10]. It should be emphasized that ANNs can be developed for any catalyst libraries regardless the optimization algorithm used provided that the number of

Table 1
Concentration levels of the components in methane oxidation

Levels	Concentration of components						
	Molar ratio to Ce				w/w% to support		
	Co	Zr	Cr	La	Pt	Pd	Au
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.5	0.5	0.1	0.1	0.1	0.1	0.1
3	1.0	1.0	0.2	0.2	0.2	0.2	0.2
4	2.0	2.0	0.4	0.4	1.5	1.5	1.0
5					1.9	3.0	1.5

Table 2
Concentration levels of the components in propane oxidation

Levels	Concentration of components (w/w% to support)							
	Pt	Pd	Rh	Ru	Au	Cu	Ag	Mn
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
2	0.1	0.1	0.1	1.0	0.1	0.1	0.1	0.4
3	0.2	0.2	0.2	2.0	0.2	0.2	0.2	0.6
4	0.3	0.3	0.3	3.0	0.3	0.3	0.3	0.8

Table 3
Concentration levels of the components in methane oxidative coupling

Levels	Concentration of components (mmol/g SiO_2)					
	Na	S	W	P	Zr	Mn
1	0.00	0.00	0.00	0.00	0.00	0.00
2	1.50	0.40	0.08	0.25	0.70	0.13
3	3.00	0.80	0.16	0.50	1.40	0.26
4	4.50	1.20	0.24	0.75	2.10	0.39
5	6.00	1.60	0.32	1.00	2.80	0.52
6	7.50	2.00	0.40	1.25	3.50	0.65

Table 4
Characterization of the best 100 networks obtained in the three case studies

Structure of ANNs[a]			Occurrence of a structure amongst the best 100 networks		
1st hidden layer	2nd hidden layer	3rd hidden layer	Methane oxidation	Propane oxidation	Methane oxidative coupling[b]
5			0	0	0
5	3	2	0	0	0
5	5		0	0	0
10			0	0	0
15			0	2	0
20			1	1	0
25			2	0	0
10	15		5	14	0
14	10		8	8	5
15	10		5	14	6
30			2	3	0
35			6	2	1
15	15		13	8	7
20	10		15	17	7
40			3	1	1
45			5	0	0
20	10	10	17	20	17
50			3	1	0
20	20		15	9	6

[a] Structure of ANNs shows the number of nodes (neurons) in the hidden layers.
[b] Only the best 50 networks are characterized.

Table 5
Predictive ability of the OLC-network obtained in methane optimization

		Component networks									OLC-net
		1st	2nd	5th	13th	25th	34th	40th	46th	49th	
Structure[a]	1st hidden layer	20	20	14	45	20	20	10	20	15	
	2nd hidden layer	10	10	10			20	15	10	10	
	3rd hidden layer		10						10		
MSE × 10^3	Training + validation[b]	0.6	0.7	0.8	0.9	1.0	1.0	1.0	1.1	1.1	0.4
	Test[b]	22.2	16.1	14.5	17.9	11.7	20.9	12.3	13.7	11.8	9.1
Combination weights		0.14	0.29	0.09	0.06	0.14	0.07	0.04	0.07	0.09	

[a] Structure shows the number of nodes (neurons) in the hidden layers.
[b] One hundred and forty-four catalysts have been involved in training and validation patterns, whereas the test set consists of 24 catalysts. In each pattern all sorts of catalysts from poor to good catalytic activity are involved with the same rate.

Table 6
Predictive ability of the OLC-network obtained in propane oxidation

		Component networks						OLC-network
		1st	5th	12th	17th	21st	37th	
Structure[a]	1st hidden layer	10	15	14	10	10	20	
	2nd hidden layer	15		10	15	15	10	
	3rd hidden layer						10	
MSE × 10^3	Training + validation[b]	1.1	1.6	1.7	1.8	1.9	1.9	0.6
	Test[b]	66.7	59.6	30.8	35.6	47.3	62.4	30.0
Combination weights		0.28	0.08	0.29	0.16	0.09	0.11	

[a] Structure shows the number of nodes (neurons) in the hidden layers.
[b] One hundred and twenty-nine catalysts have been involved in training and validation patterns, whereas the test set consists of 21 catalysts. In each pattern all sorts of catalysts from poor to good catalytic activity are involved with the same rate.

Table 7
Predictive ability of the OLC-network obtained in methane oxidative coupling

		Component networks				OLC-network
		7th	27th	28th	35th	
Structure[a]	1st hidden layer	35	20	15	20	
	2nd hidden layer		10	10	10	
	3rd hidden layer		10			
MSE × 10^3	Training + validation[b]	1.3	1.5	1.5	1.5	1.0
	Test[b]	36.9	21.4	18.4	18.7	17.6
Combination weights		0.20	0.23	0.20	0.36	

[a] Structure shows the number of nodes (neurons) in the hidden layers.
[b] Sixty-four catalysts have been involved in training and validation patterns, whereas the test set consists of 12 catalysts. In each pattern all sorts of catalysts from poor to good catalytic activity are involved with the same rate.

Table 8
Summary of optimization using HRS and GA in methane oxidation

OM[a]	Gen.[b]	Num.[c]	Composition of the best catalysts								α^d
			Molar ratio					w/w% to support			
			Ce	Co	Zr	Cr	La	Pt	Pd	Au	
HRS	6	213	1.00	0.00	0.00	0.00	0.00	1.90	3.00	0.00	98.1
GA_d[e]	13	390	1.00	0.00	0.00	0.00	0.00	1.90	3.00	0.00	98.1
GA_c[e]	13	390	0.96	0.00	0.00	0.00	0.04	1.85	2.97	0.0	98.0

Reaction temperature: 350 °C; catalyst loading: 100 mg; feed: 1% CH_4, 10% O_2 and 89% N_2; flow rate: 10 ml/min. Conditions are cited from Ref. [10].
[a] Optimization method.
[b] Generations investigated.
[c] Accumulated number of virtual catalytic test.
[d] Conversion of methane.
[e] GA_d and GA_c represent GA using discrete levels and continuos domain, respectively.

Table 9
Summary of optimization using HRS and in propane oxidation

OM[a]	Gen.[b]	Num.[c]	Composition of the best catalysts (w/w% to support)								α^d
			Pt	Pd	Rh	Ru	Au	Cu	Ag	Mn	
HRS	4	163	0.00	0.30	0.00	2.00	0.00	0.00	0.00	0.80	94.8
GA_d[e]	11	330	0.00	0.30	0.00	2.00	0.00	0.00	0.00	0.80	94.8
GA_c[e]	13	390	0.13	0.25	0.00	2.24	0.00	0.01	0.00	0.78	94.6

Reaction temperature: 150 °C; catalyst loading: 200 mg; feed: 0.1% propane, 20% O_2, He balance; flow rate: 6 ml/min. Conditions are cited from Ref. [2].
[a] Optimization method.
[b] Generations investigated.
[c] Accumulated number of virtual catalytic test.
[d] Conversion of propane.
[e] GA_d and GA_c represent GA using discrete levels and continuos domain, respectively.

Table 10
Summary of optimization using HRS and GA in oxidative coupling of methane

OM[a]	Gen.[b]	Num.[c]	Composition of the best catalysts (mmol/g SiO_2)						Y^d
			Na	S	W	P	Zr	Mn	
HRS	6	240	1.50	0.40	0.40	0.00	0.00	0.65	37.4
GA_d[e]	12	360	1.50	0.40	0.40	0.00	0.00	0.65	37.4
GA_c[e]	14	420	1.87	0.32	0.39	0.00	0.00	0.59	35.9

Reaction temperature: 796 °C; catalyst loading: 300 mg; feed: 75% CH_4, 25% O_2; flow rate: 40 ml/min. Conditions are cited from Ref. [5].
[a] Optimization method.
[b] Generations investigated.
[c] Accumulated number of virtual catalytic test.
[d] Yield of C_2 hydrocarbons.
[e] GA_d and GA_c represent GA using discrete levels and continuos domain, respectively.

experimental points exceeds a critical number. Irrespectively of the source of catalytic data used for training, the obtained ANNs, similarly to previously mentioned example in Ref. [10], can be used for the visualization of experimental spaces in a two-dimensional hologram. In such a way, for example, results of virtual catalytic tests obtained by different optimization algorithms can also be visualized in experimental holograms.

In the present study based on the composition–activity relationship established by ANNs the composition of catalytic materials was optimized virtually, using both HRS and GA. Virtual catalytic results as a function of composition have been arranged in two-dimensional holograms. Accordingly, catalyst libraries for methane oxidation, propane oxidation and methane coupling obtained in ''virtual catalyst library optimization'' using HRS and GA on discrete concentration levels are shown in Figs. 1–3.

In Figs. 1–3A and B the arrangements of virtually tested compositions are shown for HRS and GA, respectively. It has to be mentioned that in our previous studies the performance of tested catalysts was substituted by different colors or shades in the experimental holograms [9,10]. Contrary to that in this study only the location of the tested compositions is given. Consequently, only the arrangement of the tested experimental points will be analyzed and discussed, since in the present case the catalytic performance has no relevance. Therefore, all tested compositions are highlighted uniformly as small black points.

Similarly to our previous results well-defined periodic arrangements of data can be observed in catalyst libraries obtained in HRS, whereas in case of GA ''unstructured'' large diversity of the catalyst libraries is evidenced. As emerges from Figs. 1–3 HRS avoids the points that are far away from the optimum. Similar results were obtained in our earlier studies [9–11]. In addition, the arrangements of tested catalysts reveal qualitatively the composition–activity relationships.

For example, in methane oxidation (see Fig. 1A) the preferential role of Pd is verified as along the *Y*-axis the periodic appearance of experimental points follows the waves of Pd levels. Additionally, catalysts perform better at low levels of Zr and La as the majority of tested materials are located at the smallest Zr and La levels.

Pronounced periodicity in low temperature oxidation of propane can also be seen in Fig. 2A, where the ''lines'' along the *Y*-axis represent the highest level of manganese. The presence of Rh and Ag has a negative effect on the propane conversion, as there are no experimental points at the highest Rh and Ag levels.

In oxidative coupling of methane (see Fig. 3A) sulfur has a negative effect on the C_2 yield as there are no experimental points in areas containing sulfur. In contrast to the effect of sulfur the presence of Mn has an advantageous effect on C_2 yield, as experimental points follows the highest Mn loading.

As emerges form Figs. 1B–3B in GA, contrary to HRS, the compositions are arranged in stochastic manner in all case studies. Consequently, the use of GA cannot provide any exact conclusion about the activity–composition relationship in the given experimental space. The large diversity of virtual experimental points leads to the preparation and test of large variety of catalytic materials. Apparently, our results show that in GA numerous catalytically poorly active materials are unnecessary tested resulting in slow rate in the way to find the optimum catalyst composition.

It has to be mentioned that in the field of catalyst library design and optimization GA has been used for more than 5 years to search for the composition of the best performing catalysts using continuos domain for concentration of components. However, in our study discrete levels have been created for all components. In this respect it has to be emphasized that after optimization by GA the holographic method of visualization of the catalyst libraries requires the use of discrete concentration levels. Moreover, in this study our main goal was to apply equivalent conditions for both optimization algorithms in order to compare both the rate and certainty of optimum search. In addition, we should like to demonstrate the advantage of the use of discrete concentration

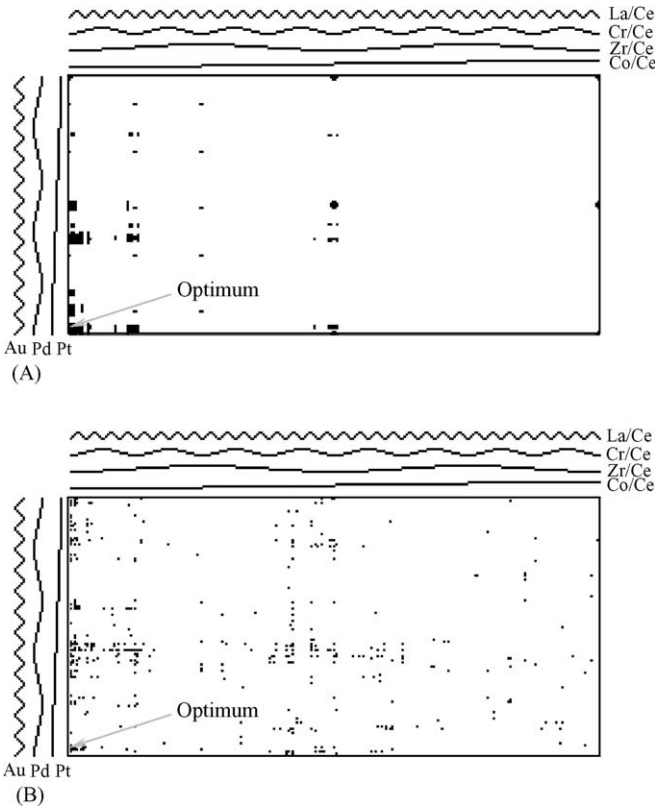

Fig. 1. The arrangement of virtual conversion data in the experimental holograms after optimization with HRS and GA (A and B) in the experimental space of methane oxidation.

Table 11
Comparison of GA using continuos domain to GA using discrete levels

Experimental space	Average of optimal values[a]		Interval of confidence[b]	
	GA_continuos	GA_discrete	GA_continuos	GA_discrete
Methane oxidation	96.0	96.4	±4.2	±2.7
Propane oxidation	92.3	93.7	±5.0	±3.2
OCM	31.0	32.2	±4.6	±3.9

[a] Average of optimal values in 50 virtual optimization.
[b] At 5% level of significance.

levels over the use of continuous domain. Therefore upon using GA in each experimental space 50 optimizations have been initiated using continuos domain and discrete levels. As emerges from data given in Table 11 the average of the optimum values obtained upon using discrete levels exceeds that of obtained by using continuous domain. In this respect, it has to be mentioned that taking into account the intervals of confidence at 5% level of significance (see Table 11) the above differences are not so significant. Nevertheless, in each experimental space the intervals of confidence are smaller when discrete levels were used in GA. It means that in the given case study the convergence to the optimum value is more probable when discrete values are used than when continuous domain is applied. We consider this new finding as a very important information.

It has to be emphasized that in oxidative coupling of methane (OCM) when the continuos domain has been applied in GA none of the 50 optimizations has been able to approximate the 37.4% C_2 maximum yield, which has been found by GA using discrete levels (see Table 10). It can be

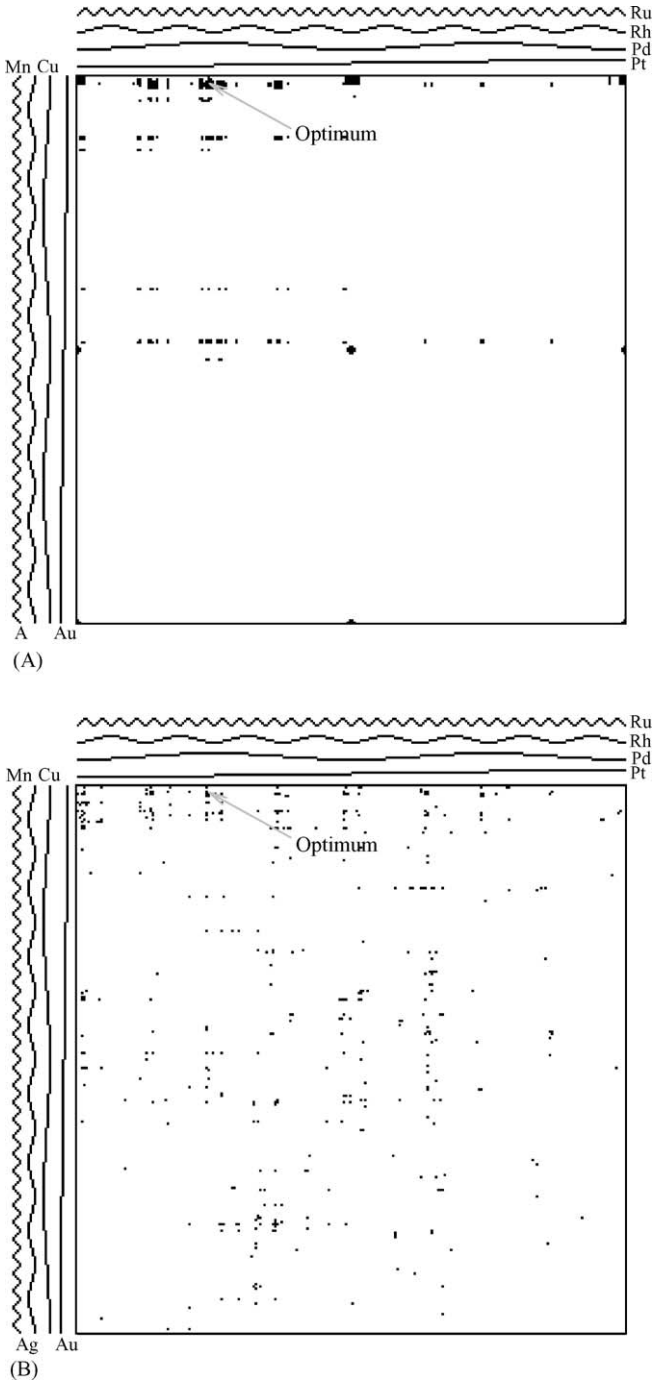

Fig. 2. The arrangement of virtual conversion data in the experimental holograms after optimization with HRS and GA (A and B) in the experimental space of propane oxidation.

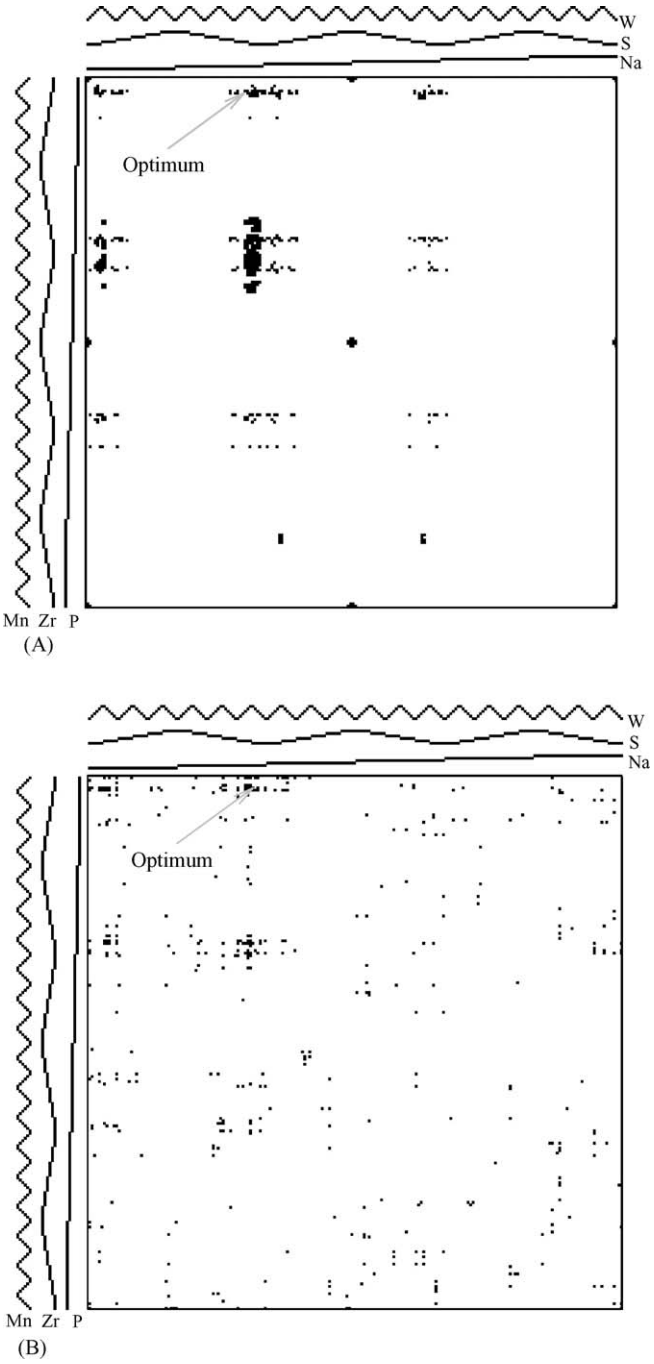

Fig. 3. The arrangement of virtual conversion data in the experimental holograms after optimization with HRS and GA (A and B) in the experimental space of methane oxidative coupling.

concluded that according to our expectation higher certainty of optimum search can be observed when, instead of the use of continuous domain, discrete concentration levels are applied in GA.

Theoretically, the maximum on continuos domain should be at least as high as using discrete levels. Since in case of OCM the optimum yield value found by GA on the continuous domain is lower than the value found on discrete domain (see Table 10), it unambiguously indicates that the algorithm did not converge to the global optimum on the continuous domain. Hence, it either converged to a local one, or did not converge at all within the available number of generations. A possible remedy to the convergence problems of the GA could be done by tuning its control parameters, in particular the heuristic parameters A and B and the proportion of elitist individuals. But, in real experimental optimization these parameters has to be set before the optimization and cannot be tuned during the optimization process. In GA we applied the parameters that proved to be the best ones according to previous studies from the related art [1,2]. Additionally, it can be mentioned that the parameters of HRS were not tuned either. We applied the previously used parameters and methods with respect to the number and size of experimental regions and the way of variable position changes [9–11].

4. Summary

Results obtained in this study support the excellent optimization features of HRS. In all cases HRS resulted in faster rate of optimization than GA. HRS decreased the number of generations and/or the number of necessary catalytic tests to obtain the optimum catalyst composition. The pronounced difference in the diversity of compositions tested by HRS and GA has been demonstrated using holographic visualization of virtual catalyst libraries obtained. The analysis of virtual holograms shows that the main drawback of GA with respect to HRS is the unstructured arrangement of tested compositions in the given experimental space represented by the virtual hologram. Upon using HRS qualitative relationships between compositions and activities can be recognized, whereas in catalyst libraries obtained by GA the experimental points are arranged more or less randomly. Therefore upon using GA no conclusions can be drawn with respect to the activity–composition relationship by visual analysis of the holograms, what represent the given experimental space. In GA the large compositional diversity of catalysts leads to testing of numerous catalytically poor materials.

Acknowledgment

Partial financial support given to A.T. (OTKA grant no. F 049742) is greatly acknowledged.

References

[1] D. Wolf, O.V. Buyevskaya, M. Baerns, Appl. Catal. A: Gen. 200 (2000) 63.
[2] U. Rodemerck, D. Wolf, O.V. Buyevskaya, P. Claus, S. Senkan, M. Baerns, Chem. Eng. J. 82 (2001) 3.
[3] D. Wolf, M. Baerns, in: J.N. Cawse (Ed.), Experimental Design for Combinatorial and High Throughput Materials Development, John Wiley & Sons Inc., Hoboken, New Jersey, 2003, p. 147.
[4] M. Holeňa, M. Baerns, in: J.N. Cawse (Ed.), Experimental Design for Combinatorial and High Throughput Materials Development, John Wiley & Sons Inc., Hoboken, New Jersey, 2003, p. 163.
[5] K. Huang, F.Q. Chen, D.W. Lü, Appl. Catal. A: Gen. 219 (2001) 61.
[6] M. Holeňa, M. Baerns, Catal. Today 81 (2003) 485.
[7] T.R. Cundari, J. Deng, Y. Zhao, Ind. Eng. Chem. Res. 40 (2001) 5475.
[8] A. Corma, J.M. Serra, E. Argente, V. Botti, S. Valero, Chem. Phys. Chem. 3 (2002) 939.
[9] L. Végvári, A. Tompos, S. Gőbölös, J. Margitfalvi, Catal. Today 81 (2003) 517.
[10] A. Tompos, J.L. Margitfalvi, E. Tfirst, L. Végvári, M.A. Jaloull, H.A. Khalfalla, M.M. Elgarni, Appl. Catal. A: Gen. 285 (2005) 65.
[11] A. Tompos, J.L. Margitfalvi, E. Tfirst, L. Végvári, Appl. Catal. A: Gen. 254 (2003) 161.
[12] S. Hashem, Neural Networks 10 (1997) 599.