



MAX-PLANCK-GESellschaft

Originally published as:

“2.2 Computer-Aided Design of Solids”

M. Holena, M. Baerns

In: Handbook of Heterogeneous Catalysis, Vol. **1**, (2008) 66-81

ISBN-Nummer: 978-3-527-31241-2

Reprint from

Handbook of Heterogeneous Catalysis

8 Volumes

Edited by
Gerhard Ertl, Helmut Knözinger, Ferdi Schüth,
and Jens Weitkamp

Second, Completely Revised and Enlarged Edition



© WILEY-VCH Verlag GmbH & Co. KGaA

to the application of *in-situ* spectroscopies to surface and bulk processes, and to advanced computational methods based on ab-initio quantum chemistry techniques. The results of such investigations with regards to catalysis allow the identification not only of interactions of gaseous or liquid reactants with the solid catalytic surface, but also of changes in the electronic and structural properties of the solid catalytic material. Moreover, elementary catalytic reaction steps can be frequently discovered. The information derived from these processes provides essential indications for designing a catalyst based on certain materials of catalytic impact. Such a procedure often does not lead to an optimum catalyst, however, and this is applicable to both qualitative and quantitative compositions, as well as to the method of preparation. Against this background, empirical methods – which make use of fundamental knowledge – are still required in the quest for an optimal catalyst.

In this chapter we present a brief introductory overview of the results of the theoretical methods mentioned above, but no further mention is made of the various spectroscopic techniques described elsewhere in this handbook [1]. The main emphasis of this chapter is placed on a combination of fundamental and empirical knowledge in preparing new catalysts, as well as on the design of catalytic experiments and the assessment of experimental data. A careful evaluation of the catalytic results by suitable data mining procedures may add to our fundamental understanding of catalysis [2, 3]. Combinatorial methodologies were introduced into the development of solid catalysts during the late 1990, at a time when they were already well known from the viewpoints of organic synthesis and enzymatic catalysis. One major advantage of the combinatorial techniques is their ability to cover of a wide range of variables during the development process; these variables include mostly the composition of the catalytic material, as well as other properties of the solid material, and finally also the reaction conditions to which the catalytic material is exposed (see Section 2.2.3). An extensive overview of related subjects, notably with respect to organic chemistry, has been provided elsewhere [4]. So-called expert systems, which have been described earlier [4–7], have not proven to be especially successful in heterogeneous catalysis; hence, they are not considered within this chapter.

2.2

Computer-Aided Design of Solid Catalysts

*Martin Holena and Manfred Baerns**

2.2.1

Introduction

Our understanding of heterogeneous catalysis has increased significantly during the past 10 years, due mainly

2.2.2

Theoretical Methods

In an earlier review, the use of ab-initio molecular-orbital calculation had already been discussed for the elucidation of catalytic processes; in this way, the method offers a potential for predicting the catalytic

* Corresponding author.

properties of solid materials and hence, in the widest sense, for designing catalysts [8]. Likewise, it was pointed out later [9] that an understanding of the catalytic and adsorptive properties of solid materials at atomic and electronic levels is essential in the design of novel catalysts. In this way, computer simulations can significantly contribute to a rational interpretation of experimental results, and they can also suggest modifications of new catalytic materials. In 2004, Motoki and Shiga [10] developed a reaction simulator, which is an intermolecular interaction analyzer based on the theories of paired interacting orbitals and localized frontier orbitals. For specific situations, activities and selectivities, as well as molecular weight in olefin polymerization, can be predicted. By using density functional theory (DFT) calculations, various catalytic performances became predictable; for example, the difference in dehydrogenation activity for cyclohexane by Pt and Ni [11], as well as the activation of C–H, C–C and C–I bonds by Pd and *cis*-Pd(CO)₂I₂ [12].

A first-principles method was illustrated by Linic et al. [13] for the identification of bimetallic alloy catalysts. The elementary reaction steps for ethylene epoxidation were derived from surface science experiments, and DFT calculations used for the catalytic chemistry. This molecular-level mechanistic information was used as an input in computational screening of potential bimetallic alloy catalysts that might offer greater selectivity to ethylene oxide than the traditional monometallic silver catalyst. In this way, a formulation of a novel Cu/Ag alloy leading to improved selectivity was found which was later verified experimentally. The main point of discussion with regards to this example was the confirmation of the great value of first-principles studies in catalyst design.

The above-described methods are often referred to as *combinatorial computational chemistry approach*, a term introduced by Kubo et al. [14]. Originally, the term had been applied mainly to the synthesis of organic compounds, but was then extended to include drugs and enzyme catalysis; today, it is successfully applied also to inorganic chemistry and hence, to the preparation of solid catalysts. Examples of this approach include the design of metal catalysts for methanol synthesis [15] and deNO_x chemistry [16].

The selected studies mentioned above indicate that significant progress has been made in computational chemistry, and these advances are progressively coming to bear on catalyst design. In the past, computational chemistry usually served the purpose of elucidating the physico-chemical properties of materials, which then further contributed to their improved design by computation. Unfortunately, first-principle quantum-chemical calculations, which represent the most powerful

approach, suffered from certain disadvantages, the main problem being that as the quality of the computational results depends largely on the assumed cluster size of the catalytic material, significant amounts of computer time are needed for large clusters. A variety of simplifications have been introduced, such as the neglect of long-range electrostatic interactions. More recently, Selvam et al. [17] dealt with the implications related to materials development and introduced a new approach, termed the *tight-binding quantum chemical molecular dynamics method*, whereby the computing time was reduced by a factor of 5000. This method can obviously be applied not only to single-site but also to nano-scale catalysis. In their report, the authors described comprehensively the design of materials and an understanding of its performance for a variety of applications, of which only those related to catalysis are mentioned. These include chemical reaction dynamics over organometallic catalysts, electronic states of supported metal catalysts, interface characteristics of precious metals-on-zirconia catalysts, chemical reaction dynamics on a Ziegler–Natta catalyst, adsorption and electron transfer dynamics on a metal surface, and chemical reaction dynamics of materials-synthesis processes. The effective simulation of a number of physical and chemical systems at reaction temperatures is of special importance, as such simulations cannot yet be studied by static first-principles calculations and classical molecular dynamics simulations.

In similar manner to the studies described above, a concept of micro-kinetic analysis of heterogeneous catalysis was pioneered by Stoltze and Norskov [18, 19], as well as by Dumesic et al. [20]. For the kinetics of the elementary reaction steps the kinetic parameters were derived by computational studies. Hereby, the catalytic performance, especially of catalytic metal sites, could be predicted, and this then served as a basis for selecting catalytic elements for the reaction under consideration.

In spite of all the recent progress that has been made in computational chemistry coupled with combinatorial approaches, there is an ongoing demand for experimental studies related to catalyst design. By introducing the concept of combinatorics to computational chemistry, a theoretical high-throughput screening of catalysts has become feasible, as long as sufficient knowledge can be made available for carrying out the computational chemistry calculations. Unfortunately, at present this is still not the case in many situations, and consequently a variety of empirical combinatorial approaches have been established. However, these require the rational design of experiments using all available empirical

and fundamental knowledge, a suitable analysis of the catalytic data obtained from these experiments, and their subsequent use in the design of new experiments to develop improvements in catalytic performance. Moreover, the data analysis comprises extensive data mining in order to extract all available empirical and fundamental knowledge. No mention is made here of the experimental details of any high-throughput experimentation based on the design of experiments, but the interested reader is referred to extensive reference material available on the subject [21, 22] (see also Chapter 9.3).

2.2.3

Rational Methods for the Design of Catalytic Experiments

Many properties and conditions influence the ability of a material to serve as the catalyst for a particular reaction. First, the composition of the material is important, it being defined by the present chemical elements or compounds, in addition to their respective masses or molar fractions within the material. In addition, the material's structure and texture, both of which are mainly determined by the preparation method used, play important roles. The composition and structure of the bulk and the surface are interconnected by further properties of the material, such as acidity, basicity, redox potential, electronic conductivity; all such properties are normally referred to as *descriptors* (cf. [23, 24]). The catalytic performance of a material depends not only on the material itself, but also on the conditions to which it is exposed in the reaction, at a particular temperature, at partial pressures of the reactants, total pressure, and the space velocity of the feed. Taken together, the descriptors of the material and the reaction conditions are termed the *input variables* in the sequence.

In a catalytic experiment, a certain number of catalytic materials are tested, and these are chosen from a very comprehensive set of potential materials. The number of materials to be examined is affected by the following aspects:

- The *ranges* of the individual input variables.
- Whether a particular input variable can have *only particular prespecified values*, or whether it can assume *any value* within the corresponding range, restricted only through the finite discernibility due to experimental error.
- The *experimental error* of a particular input variable. As there always exists some experimental error for a continuous variable, the number of discernible values within its range is always finite, but may be quite high. For example, a range of fractions of 0 to 20% and an

experimental error of 0.1% will entail 201 discernible values.

- *Constraints* on input variables, such as the constraint that the fractions of all components should sum up to 100%, or constraints on the number of non-zero fractions. The latter actually express constraints on the number of components in a catalytic material, for example, two to three active components plus one dopant.

2.2.3.1 The Statistical Design of Experiments

The task of finding, for a given set of potential catalytic materials, a subset of representatives which convey the required information about the whole set has, for almost a century, been addressed by methods of *statistical design of experiments* (DOE). A very simple situation called *factorial design* occurs if the required information represents the impact of any combination of possible values of some n independent factors. If for $i = 1, \dots, n$ the i -th factor can assume f_i different values, the factorial design needs $f_1 \cdot \dots \cdot f_n$ representatives. As an example, let a catalytic material have a fixed fraction of support chosen from two possibilities, active components chosen from three possibilities, among which one is always present, whereas each of the remaining two either is absent or its ratio to the first one can assume one of three given values, and can have one dopant which either is absent or assumes a given value. Then $n = 4$ independent factors exist, $f_{\text{support}} = 2$, $f_{2\text{nd active component}} = f_{3\text{rd active component}} = 4$, $f_{\text{dopant}} = 2$, and the factorial design needs:

$$f_{\text{support}} \cdot f_{2\text{nd active component}} \cdot f_{3\text{rd active component}} \\ \times f_{\text{dopant}} = 64 \text{ representatives.}$$

On the other hand, assuming that there is no interaction between different factors and the only information required is about the impact of any single factor in isolation, one needs only $n + 1$ representatives. In the above example, $n + 1 = 5$.

In both of the above methods, the number of representatives can vary from experiment to experiment, depending on the number of independent factors, and in factorial design also on the number of their possible values. *Computer-aided statistical DOE methods* have been developed only since the 1970s, and these are used to solve more ambitious and computationally much more demanding tasks: the main target here is to choose, for any given set of possible materials, a *representative subset of particular size such that the amount of required information about the whole set is maximal among all subsets of that size*.

Among the computer-aided DOE methods available, the *D-optimal design* is the most frequently used. This

relies on a model assuming a dependent variable, such as yield, conversion or selectivity, which depends linearly on input variables, and interactions between them. The information maximized in the D-optimal design is an information measure called the *Fisher information* for that model; more precisely, it is the determinant of the matrix of that information measure (actually, the term "D-optimal" relates to the fact that the determinant is maximized). Moreover, from statistics it is known that maximizing this information is equivalent to minimizing the volume of the confidence ellipsoid for parameter estimates in the underlying linear model. To illustrate these principles again by an example, a catalytic material is considered that consists of a support material and four active components. For this material, it is assumed that the yield y_i of the reaction catalyzed with the i -th catalytic material from the representative subset depends on the fraction $x_{i,0}$ of the support and the fractions $x_{i,1}$, $x_{i,2}$, $x_{i,3}$ of the active components 1, 2, and 3 in this material. The fraction $x_{i,4}$ of the active component 4 is obtained from the condition $x_{i,0} + x_{i,1} + x_{i,2} + x_{i,3} + x_{i,4} = 1$. Moreover, the yield y_i is also assumed to depend on interactions between any two of the fractions $x_{i,1}$, $x_{i,2}$, $x_{i,3}$, and on interactions between $x_{i,0}$ and any pair of those three active-component fractions. Rewritten in a formula, these assumptions read:

$$y = \beta_0 x_{i,0} + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \beta_3 x_{i,3} + \beta_{01} x_{i,0} x_{i,1} \\ + \beta_{02} x_{i,0} x_{i,2} + \beta_{03} x_{i,0} x_{i,3} + \beta_{12} x_{i,1} x_{i,2} \\ + \beta_{13} x_{i,1} x_{i,3} + \beta_{23} x_{i,2} x_{i,3} + \beta_{012} x_{i,0} x_{i,1} x_{i,2} \\ + \beta_{013} x_{i,0} x_{i,1} x_{i,3} + \beta_{023} x_{i,0} x_{i,2} x_{i,3}$$

where $\beta_0, \beta_1, \dots, \beta_{023}$ are the unknown model parameters. The information, which is maximized in the D-optimal design, is then the determinant $\det(R'R)$, with R denoting the 13-column matrix with rows

$$(x_{i,0}, x_{i,1}, x_{i,2}, x_{i,3}, x_{i,0}x_{i,1}, \dots, x_{i,2}x_{i,3}, x_{i,0}x_{i,1}x_{i,2}, \\ x_{i,0}x_{i,1}x_{i,3}, x_{i,0}x_{i,2}x_{i,3})$$

and R' representing the transpose of R . As was recalled above, maximizing that information is equivalent to minimizing the volume of the confidence ellipsoid for the estimates of $\beta_0, \beta_1, \dots, \beta_{023}$.

For a comprehensive explanation of various statistical DOE methods, competent monographs are available, both of a general form [25–27] and those addressing specifically chemical applications [28, 29]. Examples of such applications in catalytic research can be found elsewhere [30–34].

2.2.3.2 Optimization Methods for Empirical Objective Functions

Different statistical methods for the design of experiments use different criteria to choose experimentally tested materials from among those considered possible for testing. Nonetheless, the respective criteria are always applied uniformly to the entire set of possible candidate materials. This is to be quite impractical if the most interesting catalytic materials are not distributed uniformly in the parameter space of possible materials, but rather form only one or several small clusters. This often occurs when interest is expressed only in catalytic materials of sufficiently high performance (in terms of yield, conversion, selectivity, etc.). Indeed, high values of such performance measures are typically achieved only in small areas of contiguous compositions (Fig. 1).

In such situations, it is more relevant to use a method designed specifically to seek, for any given objective function, locations at which the function values are maximized. Since each such method equivalently seeks locations in which the negative of the objective function takes its minimal value, the more general terms *optima/extremes of the objective function* are used rather than *maxima* and *minima*. The process of searching such optima by means of a corresponding algorithm is termed *function optimization*, and the involved computer-aided methods *optimization methods*. Before presenting an overview of these types of methods, two introductory points should be clarified:

- Whether a location is an optimum of the objective function in the space of input variables depends on

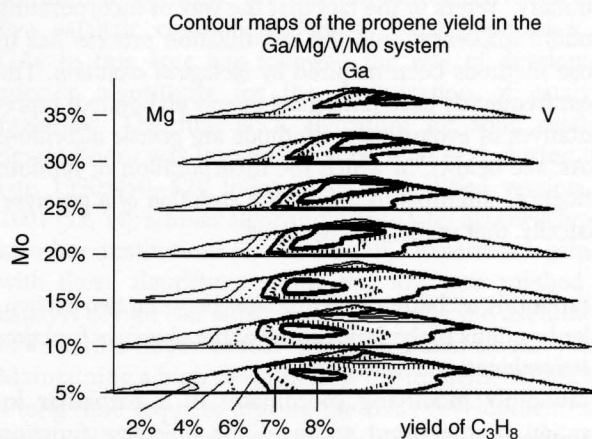


Fig. 1 Example of dependency of yield on catalyst composition. This example shows the dependency of propene yield on the fractions of Mg, V, Ga, and Mo in the oxidative dehydrogenation of propane. The approximation was obtained by means of artificial neural networks on data analyzed in Ref. [35].

References see page 79

other locations with which it is compared. On occasion, the function value in a location is not surpassed by values in locations within a certain neighborhood, yet outside that neighborhood there are locations with a higher function value. For example, varying the fraction of individual elements in the catalytic material within a certain small range does not lead to a material with a better performance, whereas varying some of those fractions outside that range can lead to such a performance. The material is then said to have a *locally optimal performance* or, in mathematical terms, to be a *local optimum* of the function describing the dependency of its performance on the composition. On the other hand, if even varying any of the fractions within the whole range of their admissible values does not lead to a better performance, the considered catalytic material is said to correspond to a *global optimum* of that function.

- In most other applications (including, e.g., reaction kinetics), values of the objective function may be obtained analytically – that is, either as the result of setting the function input into a mathematical expression, or as the solution of an equation described with a mathematical expression (e.g., of a differential equation). In contrast, values of functions describing the dependence of catalyst performance on its composition are obtained *empirically*, through experimental measurements.

2.2.3.2.1 Evolutionary Methods Evolutionary optimization methods are stochastic methods, which means that the available information about the objective function is complemented with random influences. The term “evolutionary” refers to the fact that the way of incorporating random influences into the optimization process has in those methods been inspired by *biological evolution*. The most frequently used (and most deeply elaborated) representatives of evolutionary methods are *genetic algorithms* (GAs; see below), in which the incorporation of random influences attempts to mimic the *evolution of a genotype*. Basically, that method comprises:

- randomly exchanging coordinates between two particular locations in the input space of the objective function (*recombination, crossover*);
- randomly modifying coordinates of a particular location in the input space of the objective function (*mutation*); and
- *selecting* the locations for crossover and mutation (parent locations) according to a probability distribution, either uniform or skewed towards locations at which the objective function takes high values (the latter being a probabilistic expression of the survival-of-the-fittest principle).

Detailed treatment of various types of genetic algorithms, as well as of other evolutionary optimization methods, can be found in specialized monographs [36–45]. In this chapter, only those four features of GAs will be highlighted which are particularly important for the development of catalytic materials.

Feature 1 The meaning of the individual coordinates of locations in the input space of the objective function is strongly problem-dependent. In catalyst design, the coordinates typically convey some of the following meanings:

- (i) The *qualitative composition* of the catalytic material; that is, of which active components does it consist, whether it contains dopants (and which ones), whether it is supported, and what is its support.
- (ii) The *quantitative composition* of the catalytic material; that is, the fractions of the various components mentioned in (i).
- (iii) The *preparation of the catalytic material*, its individual steps and their quantitative characterizations, such as temperatures or times they need.
- (iv) The *reaction conditions* of the catalyzed reaction.

There is an intimate connection between qualitative and quantitative composition of catalytic materials. The presence of a particular component in the catalytic material is equivalent to the fraction of that component being non-zero. There are consequences for the algorithms accomplishing the operations of recombination and mutation. Namely, the algorithms must guarantee that this equivalence cannot be invalidated through the respective operation. For example, if the presence of a particular component in one of the parent materials, and the absence of that component in the other parent material, are exchanged during recombination, then the fraction of that component has also to be changed at the same time. Similarly, if a mutation eliminates a certain component from the catalytic material, the fraction of that component must be set simultaneously to zero. In this context, it is useful to differentiate between *quantitative mutation* (which modifies only the quantitative composition of the catalytic material, without affecting its qualitative composition) and *qualitative mutation* (which modifies its qualitative composition) (Fig. 2).

Feature 2 Crossover and mutation operations can be applied to many individuals simultaneously; thus, the GA can *follow many optimization paths in parallel*. Moreover, optimization proceeds between subsequent iterations for different paths independently. Because of the biological inspiration of GAs, individual iterations of a GA are

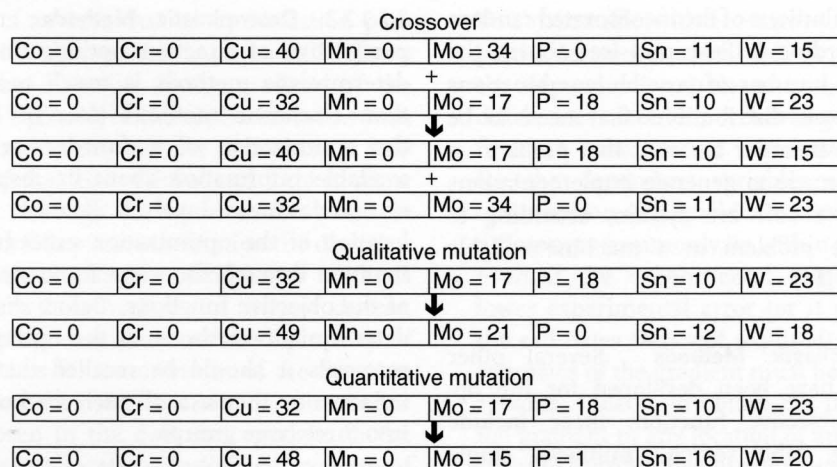


Fig. 2 Illustration of operations used in genetic algorithms; the values in the examples are mass fractions of active elements in the catalytic material expressed in mol%. (Reproduced from Ref. [46].)

referred to as *generations*, and all locations in which the value of the objective function is considered in a particular generation (e.g., all catalytic materials of which the performance has been measured in that generation) are denoted as *population*. The fact that GAs follow many optimization paths in parallel is actually the main reason for their attractiveness in high-throughput catalyst development, because a straightforward correspondence can be established between those optimization paths and channels of the high-throughput reactor in which the materials are experimentally tested.

Feature 3 As GAs do not use derivatives, they are not attracted to local optima. On the contrary, the random variables incorporated into recombination, mutation, and selection enable the optimization paths to *leave the attraction area of the nearest local optimum*, and to continue *searching for a global one*.

Feature 4 Due to the biological inspiration of the GAs, a particular distribution of the incorporated random variables cannot be justified mathematically, but its choice is an heuristic task. The most important heuristic parameters describing such a distribution are the *overall probability of any modification* of an individual, the *ratio between the conditional probabilities* of crossover and qualitative or quantitative mutation, conditioned on any modification, and the *distribution of the intensity of quantitative mutation*. The *population size* is also sometimes a matter of heuristic choice, although in the practice of catalyst development it is usually determined by the number of channels in the testing reactor.

As a consequence of the problem-dependency of GA inputs, for solving optimization problems in specific

application areas such as in the development of catalytic materials, it is quite difficult to use general GA software, such as Matlab's Genetic Algorithm and Direct Search Toolbox [47]. Indeed, such general GA software optimizes only functions within input spaces of low-level data types, examples of which are vectors of real numbers and bit-strings. However, encoding the qualitative and quantitative composition of catalytic materials, their preparation and reaction conditions with low-level data types is tedious and error-prone. Moreover, it requires a great deal of mathematical erudition.

For the reasons highlighted, it is not surprising that—apart from early attempts to use general GA software to optimize the distribution of active sites of two catalytic components [48, 49]—the application of GAs in this area has followed the way of developing specific algorithms for the optimization of catalytic materials. The first such algorithm for solid catalyst development was implemented by Wolf et al. during the late 1990s [50–52], followed by a modified version in 2001 [53, 54]; similar algorithms were later also developed at other institutions [2, 32, 55–59]. Experience gained with those algorithms shows that any new method of catalyst design can substantially decrease the usefulness of an earlier implemented specific GA after several years. Maintaining a high usefulness of the implementation for a long time requires (during the GA's development) the anticipation of all catalytic materials and their preparation methods for which the implemented algorithm might need to be used in the future. Nevertheless, there is no guarantee that all of these factors can be anticipated during development. In addition, the more possibilities the implementation of a GA attempts to cover, the broader the

class of possible distributions of the incorporated random variables. Thus, according to Feature 4 (see above), the greater becomes the number of possible combinations of parameters of those distributions that need to be heuristically chosen. In order to avoid that problem, a recent proposal was made to generate implementations of specific GAs by a software system, according to a description of the problem in a machine-readable description language [46].

2.2.3.2.2 Other Stochastic Methods Several other stochastic methods have been developed for the optimization of an objective function; these include simulated annealing, tabu search, multilevel single linkage, topographical optimization, stochastic hill climbing, stochastic tunneling, stochastic branch and bound [60, 61]. Like GAs, these methods typically only compare function values at different locations; hence, they tend to find global optima rather than local ones. On the other hand, as function values alone contain only little information, the search for the optimum progresses only very slowly.

Simulated annealing has sometimes been used in heterogeneous catalysis, although less frequently than GAs [49, 62–65]. Basically, the way in which the simulated annealing method searches for the global minimum of an objective function mimics the way in which a crystalline substance reaches the ground state with minimal energy in a process of heating and then slowly cooling. The method works by iteratively proposing changes and either accepting or rejecting each of them. There are various criteria for the acceptance or rejection of proposed changes, the most often encountered being a criterion called *Metropolis*. This accepts the change unconditionally if the objective function decreases; otherwise it accepts it only with an *exponentially distributed probability* that depends on a driving-force parameter being called *temperature*, which is equal to the mean of that exponential distribution. Thus, if the temperature is very large, nearly all changes are accepted and the method simply moves the system through various states, irrespective of the values of the objective function. Consequently, if the method starts at a higher temperature, which is then gradually decreased, it brings the system to a state with a minimum of the objective function, but still allows it to escape that state if the minimum is not global (hence, the temperature is not zero). In particular, when searching for catalyst materials with high performance, the simulated annealing method can leave a locally optimal catalyst, and continues with lower-performance materials, in order to ultimately identify a globally optimal catalyst.

2.2.3.2.3 Deterministic Methods From the overall perspective of function optimization, the plethora of deterministic methods is much more frequently used than stochastic methods [66, 67]. In those methods, the optimization of a function relies solely on the available information about its response surface, with no randomness involved. Hence, the same starting location of the optimization procedure always leads to iteration through the same locations in the input space of the objective functions. Before discussing the role of deterministic methods in the optimization of catalytic materials it should be recalled that, according to the information they use, all such methods could be divided into three large groups.

Group 1 Methods These include methods that use only information about function values. As mentioned above in the context of stochastic optimization, such methods tend to find a global optimum rather than a local one (in contrast to Groups 2 and 3, below), and they are very slow. Their most frequently encountered example is the *simplex method*. This produces a sequence of simplexes; that is, of polyhedra with $(n + 1)$ vertices, where n refers to the dimension of the input space of the objective function. The vertices of the first simplex of the sequence are typically chosen at random, and any further simplex is obtained from the previous one through replacing the vertex in which the value of the objective function was the worst (i.e., the lowest value in the case of maximization, and the highest value in the case of minimization). Other methods of this group, used in the design of catalytic materials, include the *holographic research strategy* and the *sequential weight increasing factor technique*.

Group 2 Methods These include methods which use, in addition to function values, information about first partial derivatives; that is, about the gradient. The gradient of a function has the property that its direction coincides with the direction of the fastest increase of function values, and is opposite to the direction of their fastest decrease. Thus, if the optimization path follows from some location in the input space the direction of the objective function gradient, then the function value increases with the highest possible speed along that path (at least in the immediate neighborhood of that location). Similarly, if the path follows the direction opposite to its gradient, then function values decrease with the highest possible speed. However, following the direction of the gradient (or opposite to the gradient) of the objective function does not generally allow the optimization path to reach its global maximum (or minimum), but only a local one. More precisely, for every maximum (global as well

as local) there exists an area that an optimization path will never leave if it starts within that area and follows in each location the direction of the gradient of the objective function. The area is called the *attraction area* of the considered maximum, and the maximum is said to be an *attractor* of that area. Global and local minima are also attractors, although the optimization paths in their attraction areas follow in each location the direction opposite to the gradient. Simple examples of methods of this group are various variants of *steepest descent*, which directly employs the above recalled property of gradients and searches for a minimum of the function in such a way, that the next location in a sequence of iterations is chosen in the direction opposite to the gradient in the current location. Individual variants of this method differ in their particular way of choosing the next location in that direction. More sophisticated representatives of this group include several *methods of conjugate gradients*.

Group 3 Methods These include methods which use, in addition to function values, information about partial derivatives up to the second order. Like gradient-based methods, they search for a maximum or minimum of the objective function only within its attraction area. However, second-order derivatives allow the construction of a quadratic approximation of the function, whereas gradient approximates the function only linearly. Close to a maximum or minimum, a quadratic approximation of a function is more accurate than a linear one, and therefore methods of this group can localize searched optima faster, once they reach their proximity. On the other hand, linear approximations are frequently more accurate far from any optima. Therefore, methods of this group are most often used as combined methods, which switch between the behavior of second-order methods close to an optimum and the behavior of a gradient-based method far from the optimum. The most frequently encountered method of that type is the *Levenberg–Marquardt method*, whereas using only second-order derivatives leads in the most simple case to the *Gauss–Newton method*.

For the optimization of empirical functions in the development of solid catalytic materials, it is, unfortunately, impossible to employ methods using partial derivatives – that is, methods from Groups 2 and 3. The reasons for this impossibility are connected with the *character of catalytic experiments*, and can be summarized as follows:

- To obtain sufficiently accurate *numerical estimates* of the partial derivatives, small differences between values of the dependent variable must be recorded; typically, differences should not be larger than 0.1% of the function value in the respective location. However, for the dependent variables occurring in catalyst optimization (e.g., yield, conversion, selectivity), such differences commonly lie within the *experimental error*.
- Even if the experimental setting would lead to a lower experimental error (or if sufficient accuracy of the estimates were not required), obtaining numerical estimates of the gradient must be given up for practical reasons. Indeed, to obtain the numerical estimate of the gradient in any location of an n -dimensional input space of the objective function would require empirical evaluation of the function in $(n + 1)$ locations located *very close to each other*. For example, imagine that the objective function describes the dependence of yield on the composition of catalytic material on choosing its components from a pool of 15 elements. Then, to estimate its gradient numerically in one single location would require testing 16 materials with nearly the same composition. That is clearly not affordable for reasons of cost and time.

Therefore, in heterogeneous catalysis only deterministic optimization methods of Group 1 have been occasionally used [64, 68–70]. On the other hand, in the context of high-throughput development of solid catalysts, it was pointed out that their best known representative – the simplex method – allows an easy adaptation to produce any prescribed number of sequences of simplexes in parallel [64]. Consequently, the method can propose in each step a *prescribed number of new catalytic materials in parallel*, which provides a straightforward correspondence with their subsequent testing in a multichannel reactor, as in the case of evolutionary algorithms.

2.2.3.3 Concluding Remarks

The choice of computer-aided methods for the design of catalytic experiments depends on whether the space of potential catalytic materials should be investigated in some systematic way, or whether the search for catalysts with the best performance is its only objective. The former case is a task for statistical DOE methods, whereas optimization methods are appropriate for the latter. Within the computer-aided branch of statistical DOE methods, the D-optimal design is most universally employed. An optimization method has to require solely function values to be applicable to the optimization of catalytic materials. Among such methods, GAs are most

frequently used, especially those developed specifically for catalytic optimization.

One problem with GAs is the crucial role of their heuristic parameters, a correct tuning of which requires a great deal of preliminary knowledge. In part, that knowledge can be obtained by means of data analysis and data mining methods, which will be surveyed in the following section. Therefore, GAs are frequently *combined and interleaved* with such methods, especially with neural networks [54, 56–58, 68, 71–73].

2.2.4

Data Analysis and Data Mining

The term *data analysis* is traditionally used for statistically assessing the extent to which data support particular relationships between the various objects, properties and phenomena that they describe. Until the early 1970s, researchers hypothesized such relationships, and computers were employed only for their assessment. However, the steadily increasing amount of experimental data, which became available following the advent of database technology, required an automated, data-driven search for such relationships within a new branch of statistics, referred to as *exploratory statistics*. On the other hand, the increasing significance of computers also led to a rapid development of *artificial intelligence* and *machine learning*, which attempt to mimic the way in which human beings themselves hypothesize relationships from observations. The most important types of machine learning methods include rule-based methods, decision trees, inductive logic programming, support vector machines, and artificial neural networks (ANNs). Methods based on ANNs have been frequently used also in heterogeneous catalysis (see Section 2.2.4.2 below). Although methods of exploratory statistics and machine-learning rely on different paradigms, they deal with the same data and employ the same information technologies, especially databases, the world-wide web, and various object-oriented technologies (Fig. 3). Together, they form a new interdisciplinary area known as *data mining*. In the context of heterogeneous catalysis, the distinction between traditional data analysis and data mining should be taken into account when choosing between the two conceptually different approaches to the design of libraries of solid catalysts – the optimization approach and the discovery approach [74].

2.2.4.1 Statistical Methods: An Overview

The simplest use of statistical methods is to provide summary parameters characterizing important statistical properties of input variables and of various measures

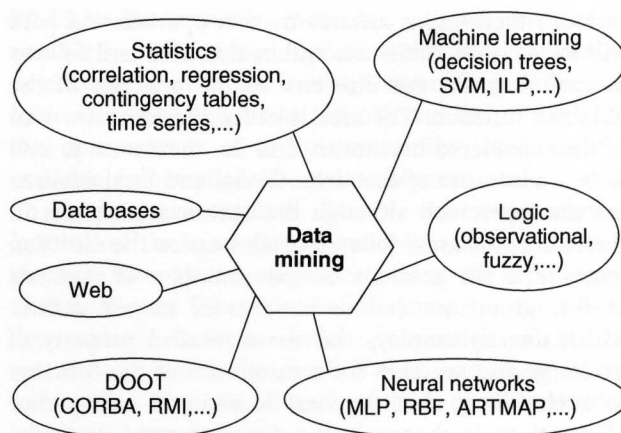


Fig. 3 Main data mining approaches and supporting technologies. ILP, inductive logic programming; MLP, multilayer perceptron; RBF, radial base functions network; ARTMAP, adaptive resonance theory mapping network; DOOT, distributed object-oriented technologies; CORBA, common object request broker; RMI, remote method invocation.

of catalyst performance (yield, conversion, etc.), or relationships between them. Such summary parameters are usually called *descriptive statistics*; their common representatives are *mean, median* and other *quantiles, variance, standard deviation, skewness, kurtosis, covariance* and *correlation*.

The two last-mentioned descriptive statistics – covariance and correlation – allow one to summarize the relationship between a performance measure and a particular input variable. The situation becomes substantially more complicated if one is interested in the relationship between a performance measure and a whole set of input variables. Indeed, then not only parameters corresponding to individual variables, but also parameters corresponding to various levels of interactions between them are needed. In such situations, the parameters are usually combined with an assumption about the form of the dependency of the performance measure on the input variables (such as linear dependency, polynomial dependency, or dependency derived from some theoretical model), for once an assumption about its form is made, the parameters already fully determine that dependency. In statistics, this approach is known as *regression* or *response surface modeling*, and the parameters determining the dependency are called *regressors*. For example, the regression of yield γ on three component fractions x_1, x_2, x_3 uses in the case of a *linear regression* four parameters $\alpha_0, \alpha_1, \alpha_2, \alpha_3$:

$$\gamma = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$$

whereas in the case of a quadratic regression, it uses 10 parameters $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_{1,1}, \alpha_{1,2}, \dots, \alpha_{3,3}$:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_{1,1} x_1^2 + \alpha_{2,2} x_2^2 + \alpha_{3,3} x_3^2 + \alpha_{1,2} x_1 x_2 + \alpha_{1,3} x_1 x_3 + \alpha_{2,3} x_2 x_3$$

Descriptive statistics can also not be used in the case of other properties of catalytic materials that depend on a whole set of input variables. Examples of such properties are similarity between different materials, dependence on unobservable factors, or classification of materials according to their catalytic behavior in particular reactions. In order to characterize such properties, methods of *multivariate analysis* are needed. The multivariate approaches which are most relevant to the analysis of catalytic data are:

- *Principal component analysis*, which reduces data dimensionality through concentrating on those linear combinations of input variables that are most responsible for the variability of the data set (unfortunately, such combinations usually do not convey any real meaning).
- *Factor analysis*, which explains input variables as combinations of a smaller number of unobservable factors.
- Analysis of the relationship between factors influencing input variables and those influencing performance measures by means of an approach called *partial least squares*.
- *Cluster analysis*, which entails grouping catalytic materials into clusters or a hierarchy of clusters according to similarity among the values of the input variables or according to a similar catalytic performance, which can be measured with various *similarity measures*.
- *Classification* of new materials, according to values of their input variables, with respect to their usability as catalytic materials in particular reactions. An important feature of classification is that the discrimination between classes relies solely on data with known correct classification, for example, on already tested materials.

Frequently, the primary purpose of data analysis is to check the compatibility of the available data with certain assumptions about the probability distribution governing those data. To this end, methods of *statistical hypotheses testing* are needed. The most commonly tested hypotheses include:

- that the probability distribution of data belongs to a certain family of distributions, for example normal distributions or exponential distributions
- that certain parameters characterizing the distribution, such as mean or variance, have a particular value, or that their value lies within a particular range
- that probability distributions governing two data sets are identical, or that some of their parameters are identical.

One important example of testing hypotheses of the last-mentioned type is an analysis of the influence of varying the values of individual input variables on the performance of a catalytic material by means of an approach called *analysis of variance*.

For detailed information about statistical methods, the reader is referred to comprehensive monographs about statistics [75–78], or chemometrics [79, 80]. Recent applications of statistical methods to heterogeneous catalysis have been reported in [23, 24, 33, 34, 81–88].

2.2.4.2 Artificial Neural Networks

ANNs are distributed computing systems that attempt to implement a greater or smaller part of the functionality characterizing biological neural networks. Their most basic concepts are a *neuron*, the biologically inspired meaning of which is an elementary signal processing unit, and a *connection* between neurons enabling the transmission of signals between them. In addition to signal transmission between different neurons, signal transmission between neurons and the environment can also take place. Neurons, connections between them, and connections between the environment and neurons form together the *architecture* of the ANN. Those neurons that receive signals from the environment but not from other neurons are called *input neurons*, whereas those sending signals to the environment but not to other neurons are called *output neurons*. Finally, neurons that receive and send signals only from and to other neurons are called *hidden neurons* (Fig. 4).

Partitioning of the set of neurons into input, hidden, and output neurons allows a large variety of architectures; however, for almost all types of ANNs frequently encountered in practical applications, the architecture is basically the same. It is a *layered architecture*, in which all signals propagate from layer to layer in the same direction, denoted as forward direction. During that propagation, a *neuron-specific linear transformation* of signals from the preceding layer is performed in each hidden or output neuron, whereas a *network-specific nonlinear function*, called *activation function*, is subsequently applied to the results of the linear transformations in hidden and

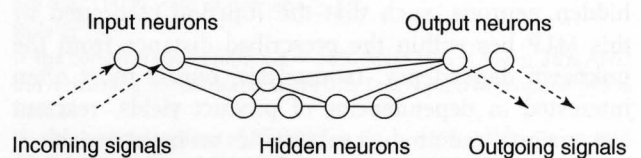


Fig. 4 Simple generic artificial neural network architecture. (Reproduced from Ref. [102].)

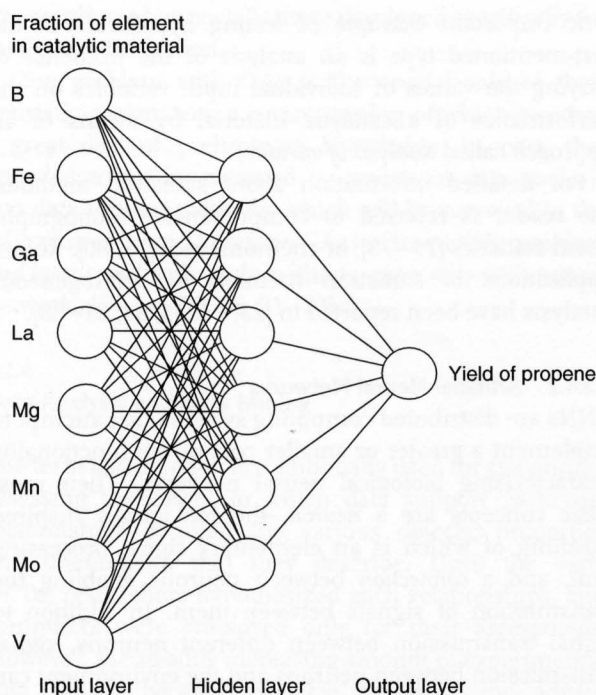


Fig. 5 Example of multilayer perceptron with one layer of hidden neurons, employed to approximate an unknown dependency of propene yield on catalyst composition. (Reproduced from Ref. [102].)

sometimes also in output neurons. Such ANNs are called feedforward neural networks, and their most prominent example is the *multilayer perceptron* (MLP; see Fig. 5), in which the activation functions are S-formed functions known as *sigmoidal functions*.

A MLP with n_I input neurons and n_O output neurons computes a function from the space of n_I -dimensional vectors to the space of n_O -dimensional vectors. The precise form of that function depends on the specific linear transformations connected with individual hidden and output neurons. It is the set of computable functions that accounts for the most useful feature of multilayer perceptrons, which is their *universal approximation property*. This property means that even to a *very general unknown dependency*, and to any prescribed arbitrarily small distance in an appropriate function space, a MLP can always be found, with as few as one layer of hidden neurons, such that the function computed by this MLP lies within the prescribed distance from the unknown dependency. In catalysis, one is most often interested in dependencies of product yields, reactant conversions, and product selectivities on input variables.

In order to obtain a neural network that computes such a function, two crucial steps are needed:

1. To choose an appropriate architecture which, in the case of a MLP, means choosing an appropriate number

of hidden layers plus an appropriate number of neurons in each of them.

2. To connect, in a process called *network training*, to each hidden and output neuron the linear transformation that fits the available data best. The quality of the overall fit of all those linear transformations may be assessed using various measures [89]. The most commonly used measure is the mean sum of squared errors; that is, the mean sum of squared distances between the output values that the network computes for a given sequence of inputs, and the output values that for those inputs have been obtained experimentally.

In reality, the choice of the number of hidden neurons cannot be separated from network training, since to assess the appropriateness of a particular architecture for the available data, a number of networks with that architecture must first be trained.

A better fit to the data used for training does not necessarily entail a better approximation of the ultimate unknown dependency. To assess the quality of the desired approximation, an independent set of data called *test data* is needed, obeying the unknown dependency but unseen by the network during training. The phenomenon of *overtraining* – that is, of a good fit to the training data accompanied by a bad fit to the test data – is the main problem faced during neural network training. Up to now, various methods have been developed that reduce that phenomenon, such as *early stopping* or *Bayesian regularization* (Fig. 6). Nevertheless, to fully exclude an influence of overtraining to the choice of the number of neurons, that choice must be based on the fit to the test data. (For details, see Ref. [89].)

The architecture of a trained neural network and the mapping computed by the network inherently represent the knowledge contained in the data used to train the network, i.e., deriving and knowledge about relationships between the input and output variables, for example, about relationships between component fractions and yield. Such a representation is not very human-comprehensible, however, as it is far from the symbolic and modular manner in which people represent knowledge by themselves. Therefore, various *knowledge extraction* methods have been developed with the aim of transforming the network architecture and the computed mapping into a representation that is more acceptable to humans, namely into representation by means of *logical rules*. For dependencies of output variables on input variables, logical rules are of the general form:

- IF the input variables fulfill an input condition C_{input}

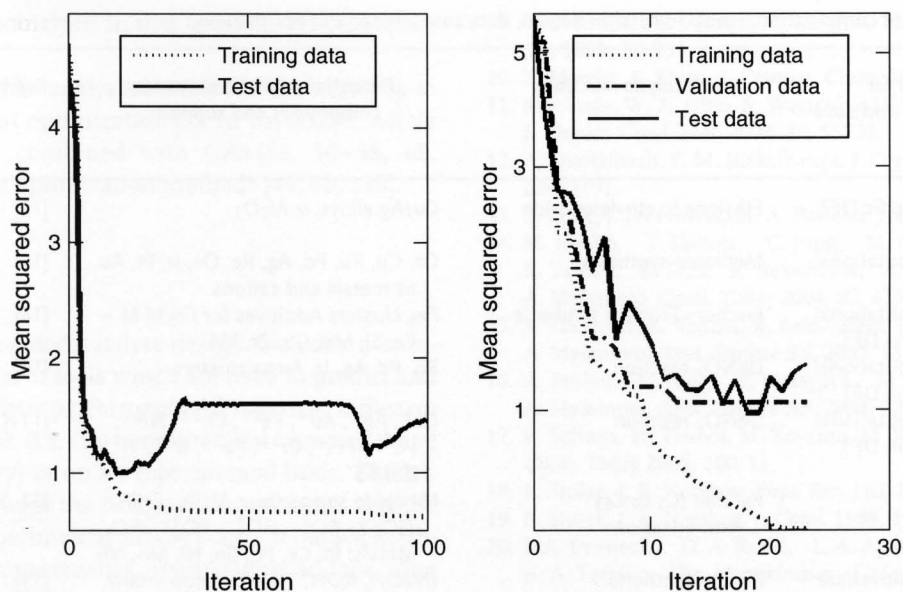


Fig. 6 An example of overtraining encountered when training a MLP with the data analyzed in Ref. [35], using the Levenberg–Marquardt training method (left), and an example of overtraining reduction using the early-stopping method (right). (Reproduced from Ref. [102].)

- THEN the output variables are likely to fulfill an output condition *Coutput*.

For examples of such rules, see Fig. 7.

Detailed information about ANNs is available in specialized monographs [90–96]. Applications of multilayer perceptrons to catalysis on solid materials have been reported in [23, 24, 57, 87, 97–107], along with applications of other types of neural network [23, 56, 58, 71, 86, 87, 108].

2.2.4.3 Concluding Remarks

In Section 2.2.4 we dealt with the relationships between various objects, properties and phenomena described by catalytic data, which primarily are relationships between input variables and catalyst performance in particular reactions. The relationships to be investigated may either be hypothesized by the researcher, or searched for based on the available data. In the former case, statistical data analysis methods are used, whereas in the latter, data mining methods are appropriate, notably methods of exploratory statistics or machine learning. Among the latter, ANNs have been frequently used in catalytic research. As the area of machine learning is still undergoing rapid development, other relevant methods may become comparably important in the next few years, including *decision trees* [23, 24, 86, 87] and *support vector machines* [109].

Both, data analysis and data mining are performed in a stepwise manner, and the result of any one particular

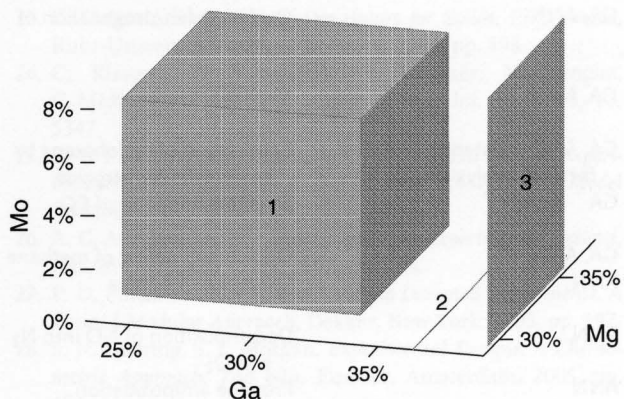


Fig. 7 Visualization of the three-dimensional projection to the dimensions corresponding to Ga, Mg, and Mo of the antecedents (left-hand sides) of three logical rules extracted from an artificial neural network trained with data on oxidative dehydrogenation of propane. (Reproduced from Ref. [102].)

Rule 1:

IF the concentrations fulfill $24\% \leq \text{Ga} \leq 33\%$ AND $31\% \leq \text{Mg} \leq 39\%$ AND $0\% \leq \text{Mo} \leq 7\%$ AND the remaining concentrations are 0%, THEN yield of propene $\geq 8\%$

Rule 2:

IF the concentrations fulfill $\text{Ga} \approx 36\%$ AND $28\% \leq \text{Mg} \leq 38\%$ AND the remaining concentrations are 0%, THEN yield of propene $\geq 8\%$

Rule 3:

IF the concentrations fulfill $0\% \leq \text{Fe} \leq 12\%$ AND $\text{Ga} \approx 38\%$ AND $29\% \leq \text{Mg} \leq 36\%$ AND $0\% < \text{Mo} < 9\%$ AND the remaining concentrations are 0%, THEN yield of propene $\geq 8\%$.

Tab. 1 Application of computational tools in catalyst design, data assessment and optimization

Computational tools for experimental design and data assessment	Catalytic reactions	Potential catalytic elements, compounds and supports	References	Year
First principles, periodic DFT calculations	Ethylene to ethylene oxide	Cu/Ag alloys, α -Al ₂ O ₃	[13]	2004
Combinatorial computational DFT chemistry	Methanol synthesis	Co, Cu, Ru, Pd, Ag, Re, Os, Ir, Pt, Au as metals and cations	[15]	2002
Combinatorial computational chemistry based on DFT	Fischer–Tropsch synthesis	Fe ₆ clusters Additives for Fe ₅ M M = Fe, Si, Mn, Ge, Zr, Mo	[14]	2004
Combinatorial computational chemistry based on DFT	DeNO _x reaction	Rh, Pd, Ag, Ir, Au as clusters	[16]	2004
Combinatorial computational chemistry based on DFT	DeNO _x reaction	Cu ⁺ , Ag ⁺ , Au ⁺ , Fe ²⁺ , Co ²⁺ , Ni ²⁺ , Pd ²⁺ , Pt ²⁺ , Cr ³⁺ , Fe ³⁺ , Ir ³⁺ , Ti ³⁺ ZSM-5	[111]	2000
GA, factorial design	Paraffin (C ₅ to C ₆) isomerization	Me oxide supports: γ -Al ₂ O ₃ , ZrO ₂ , TiO ₂ (+ SO ₄ ²⁻ , BO ₃ ³⁻ , PO ₄ ³⁻ , WO _x . Metals: Pt, Ce, Pd, Sn, Ni, Mn, Nb	[32, 55]	2003
Factorial design, multivariate analysis	Zeolite synthesis	(NaOH, KOH, 1,4-dibromopentane, 1-methylpiperidine, Sylobloc 47 (Grace), aluminium-tri-iso-propoxide	[33]	2003
GA	Low-temperature oxidation of low-concentration propane in air	Pt, Pd, Rh, Ru, Au, Cu, Ag, Mn on TiO ₂	[52]	2001
GA, ANN	Oxidative dehydrogenation of propane	Redox MeO of V, Mo, Mn, Fe, Ga; non-reducible MeO of La; acidic & basis MeO of B, Mg.	[54]	2004
GA, ANN	Methanol synthesis	Oxides of Cu, Zn, Al, Sc, B, Zr (at different preparation conditions)	[56, 73]	2004
GA, ANN, clustering and factorial analysis	Epoxidation of cyclohexene by <i>tert</i> -butylhydroperoxide	Synthesis conditions of Ti-silicate-based catalysts	[58, 86]	2005
GA	Selective oxidation of CO	Au, Cu, Pt and oxides of Mo, Nb, V supported on TiO ₂ , ZrO ₂ , CeO ₂	[59]	2005
GA, ANN	Oxidative coupling of methane	Oxides of Na, W, P, S, Zr, Mn	[68, 100]	2003, 2001
ANN	Oxidation of butane	MgO, Al ₂ O ₃ , SiO ₂ , TiO ₂ , ZnO, ZrO ₂ , SnO ₂ , Bi ₂ O ₃	[97]	1995
ANN	Decomposition of NO into N ₂ and O ₂	Cu _x /ZSM-5x: variable & different conditions	[98]	1995
ANN	Propane ammoxidation	Oxides of V, Sb, W, Sn (P, K, Cr, Mo) on Al ₂ O ₃ /SiO ₂	[99]	1997
ANN	CO ₂ hydrogenation to hydrocarbons	Ag, Al, Au, Ce, Co, Cr, Cu, Fe, Ga, La, Mn, Mo, Ni, Pd, Pt, Ru, Si, Ti, W, Zn, Zr; supports ZrO ₂ , Al ₂ O ₃ , ZnO, TiO ₂	[101]	2001
GA, ANN	Oxidative dehydrogenation of ethane	Oxides of Ga, Cu, Mn, Mo, W, Sn, Cr, Co, Zr, Ca, La, Au	[53, 72]	2003, 2002
Holographic research strategy and ANN	Total oxidation of methane	Oxides of Co, Zr, Cr, La, Cu, Pt, Pd, Au; Support CeO ₂ , La ₂ O ₃	[110]	2005
Factorial design, ANN	Selective CO oxidation in excess H ₂	Co _x /SrCO ₃ x: variable	[107]	2005

GA, genetic algorithms; ANN, artificial neural networks.

step might then be used in later steps. This is especially apparent in the following two situations:

- The objective of an initial stage of data analysis or data mining is to determine which input variables play the

most important role in the investigated relationships, or which of them play in that relationship similar roles. Such an initial stage is called *feature selection*, and this leads ultimately to the *dimensionality reduction* of the space of input variables, allowing subsequently a

more thorough analysis in that less-dimensional space [23, 24, 86].

- The results of data analysis are used in the next step to improve catalyst optimization [2]. In particular, ANNs are frequently combined with GAs [54, 56–58, 68, 71–73] or other optimization methods [49, 68, 110].

2.2.5

Conclusions

Today, computer-aided catalyst design includes a comprehensive arsenal of tools which are used to predict and optimize the catalytic performance of materials, either on a theoretical basis (i.e., by combinatorial computational catalytic chemistry) or on an experimental basis. The latter approach includes the design of experiments and the assessment of experimental data by computer-aided methods originating from statistics, optimization, and machine learning. The fundamentals of the various procedures have been applied in many cases; selected examples, which have been reported in the literature are detailed in Table 1.

In the present authors' opinion, knowledge within the broad field of computer-aided catalyst design has progressed significantly during the past five to 10 years, and this trend will surely continue as it offers a high potential for further development. As a joint effort, combinatorial computational catalytic chemistry and high-throughput experimentation will undoubtedly foster catalyst development in the years to come.

It is also believed that the results of high-throughput experimentation based on combinatorial approaches, together with sophisticated data analysis, will in the long run contribute to a better understanding of catalysis and supplement our fundamental knowledge gained by other means.

References

1. G. Ertl, H. Knözinger, F. Schüth, J. Weitkamp (Eds.), *Handbook of Heterogeneous Catalysis*, 2nd edn, Wiley-VCH, Weinheim, 2007.
2. J. M. Caruthers, J. A. Lauterbach, K. T. Thomson, V. Venkatasubramanian, C. M. Snively, A. Bhan, S. Katare, G. Oskarsdottir, *J. Catal.* **2003**, 216, 98.
3. O. V. Buyevskaya, A. Brückner, E. V. Kondratenko, D. Wolf, M. Baerns, *Catal. Today* **2001**, 67, 369.
4. J. Gasteiger (Ed.), *Handbook of Chemoinformatics*, Vol. 3, Wiley-VCH, Weinheim, 2004, pp. 490.
5. T. Hattori, Y. Murakami, *Appl. Catal.* **1989**, 48, 107.
6. E. Körting, M. Baerns, in *Proceedings of the Symposium of the Materials Research Society*, Vol. 454, Boston, 1997, p. 187.
7. E. Körting, M. Baerns, in *Handbook of Heterogeneous Catalysis*, G. Ertl, H. Knözinger, J. Weitkamp (Eds.), Vol. 1, Wiley-VCH, Weinheim, 1997, p. 419.
8. H. Kobayashi, *Catal. Today* **1991**, 10, 167.
9. S. C. Ammal, S. Takami, M. Kubo, A. Miyamoto, *Bull. Mater. Sci.* **1999**, 22, 851.
10. T. Motoki, A. Shiga, *J. Comput. Chem.* **2004**, 25, 106.
11. M. Tsuda, W. A. Dino, S. Watanabe, H. Nakanishi, H. Kasai, *J. Physics. Cond. Mat.* **2004**, 16, S5721.
12. A. Diefenbach, F. M. Bickelhaupt, *J. Organomet. Chem.* **2005**, 690, 2191.
13. S. Linic, J. Jankowiak, M. A. Barteau, *J. Catal.* **2004**, 224, 489.
14. M. Kuobo, T. Kubota, C. Jung, M. Ando, S. Sakahura, K. Yajima, K. Seki, R. Belosludov, A. Endou, S. Takami, A. Myiamoto, *Catal. Today* **2004**, 89, 479.
15. S. Sakahura, K. Yajima, R. Belosludov, S. Takami, M. Kumoi, A. Myiamoto, *Appl. Surface Sci.* **2002**, 189, 253.
16. A. Endou, C. Jung, T. Kusagaya, M. Kuobo, P. Selvam, A. Myiamoto, *Appl. Surface Sci.* **2004**, 223, 159.
17. P. Selvam, H. Tsuboi, M. Koyama, M. Kuobo, A. Myiamoto, *Catal. Today* **2005**, 100, 11.
18. P. Stolze, J. K. Norskov, *Phys. Rev. Lett.* **1985**, 55, 2502.
19. P. Stolze, J. K. Norskov, *J. Catal.* **1988**, 110, 1.
20. J. A. Dumesic, D. A. Rudd, L. A. Aparicio, J. E. Bekoske, A. A. Trevino, *The Microkinetics of Heterogeneous Catalysis*, American Chemical Society, Washington, 1993, pp. 316.
21. U. Rodemerck, M. Baerns, in *Basic Principles in Applied Catalysis*, M. Baerns (Ed.), Springer, Berlin, 2003, p. 259.
22. A. Hagemeyer, P. Strasser, A. F. Volpe (Eds.), *High-Throughput Screening in Chemical Catalysis*, Wiley-WCH, Weinheim, 2004, pp. 319.
23. C. Klanner, *Evaluation of Descriptors for Solids*, PhD thesis, Ruhr-University Bochum, Bochum, 2004, pp. 198.
24. C. Klanner, D. Farrusseng, L. Baumes, M. Lengliz, C. Mirodatos, F. Schüth, *Angew. Chem. Int. Ed.* **2004**, 43, 5347.
25. G. E. P. Box, W. G. Hunter, J. S. Hunter, *Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building*, Wiley, New York, 1978, pp. 653.
26. A. C. Atkinson, A. N. Donev, *Optimum Experimental Designs*, Oxford University Press, 1992, pp. 328.
27. P. D. Fukenbusch, *Practical Guide to Designed Experiments. A Unified Modular Approach*, Dekker, New York, 2005, pp. 197.
28. S. N. Deming, S. L. Morgan, *Experimental Designs: A Chemometric Approach*, 2nd edn, Elsevier, Amsterdam, 2005, pp. 454.
29. R. Carlson, *Design and Optimisation in Organic Synthesis*, 2nd edn, Elsevier, Amsterdam, 2005, pp. 596.
30. M. Nele, A. Vidal, D. L. Bhering, J. C. Pinto, V. M. M. Salim, *Appl. Catal. A: General* **1999**, 178, 177.
31. R. Ramos, M. Menendez, J. Santamaria, *Catal. Today* **2000**, 56, 239.
32. A. Corma, J. M. Serra, A. Chica, *Catal. Today* **2003**, 81, 495.
33. M. Tagliabue, L. C. Carluccio, D. Ghisletti, C. Perego, *Catal. Today* **2003**, 81, 405.
34. R. J. Hendershot, W. B. Rogers, C. M. Snively, B. A. Ogannaik, J. Lauterbach, *Catal. Today* **2004**, 98, 375.
35. J. N. Cawse, M. Baerns, M. Holena, *J. Chem. Inf. Computer Sci.* **2004**, 44, 143.
36. D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, 1989, pp. 412.
37. J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, 1992, pp. 840.
38. J. R. Koza, *Genetic Programming II: Automatic Discovery of Reusable Programs*, MIT Press, Cambridge, 1994, pp. 350.

39. J. R. Koza, F. H. Bennett, D. Andre, M. A. Keane, *Genetic Programming III: Darwinian Invention and Problem Solving*, Academic Press, Orlando, 1999, pp. 1154.
40. J. R. Koza, M. A. Keane, M. J. Streeter, W. Mydlowec, J. Yu, G. Lanza, *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*, Kluwer, Dordrecht, 2003, pp. 590.
41. T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press, New York, 1996, pp. 314.
42. M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, 1996, pp. 224.
43. D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 2nd edn, IEEE Press, New York, 1999, pp. 270.
44. M. L. Wong, K. S. Leung, *Data Mining Using Grammar Based Genetic Programming and Applications*, Kluwer, Dordrecht, 2000, pp. 213.
45. A. A. Freitas, *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, Springer, Berlin, 2002, pp. 264.
46. M. Holena, in *High-Throughput Screening in Chemical Catalysis*, A. Hagemeyer, P. Strasser, A. F. Volpe (Eds.), Wiley-VCH, Weinheim, 2004, p. 153.
47. *Genetic Algorithm and Direct Search Toolbox*. The MathWorks, Inc., Natick, 2004, pp. 268.
48. A. S. McLeod, M. E. Johnston, L. F. Gladden, *J. Catal.* **1997**, *167*, 279.
49. A. S. McLeod, L. F. Gladden, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 981.
50. D. Wolf, O. V. Buyevskaya, M. Baerns, *Appl. Catal. A: General* **2000**, *200*, 63.
51. O. V. Buyevskaya, D. Wolf, M. Baerns, *Catal. Today* **2000**, *62*, 91.
52. U. Rodemerck, D. Wolf, O. V. Buyevskaya, P. Claus, S. Senkan, M. Baerns, *Chem. Eng. J.* **2001**, *82*, 3.
53. G. Grubert, E. Kondratenko, S. Kolf, M. Baerns, P. van Geem, R. Parton, *Catal. Today* **2003**, *81*, 337.
54. U. Rodemerck, M. Baerns, M. Holena, D. Wolf, *Appl. Surface Sci.* **2004**, *223*, 168.
55. J. M. Serra, A. Chica, A. Corma, *Appl. Catal. A: General* **2003**, *239*, 35.
56. Y. Watanabe, T. Umegaki, M. Hashimoto, K. Omata, M. Yamada, *Catal. Today* **2004**, *89*, 455.
57. A. Corma, J. M. Serra, *Catal. Today* **2005**, *107–108*, 3.
58. A. Corma, J. M. Serra, P. Serna, S. Valero, E. Argente, V. Botti, *J. Catal.* **2005**, *229*, 513.
59. R. M. Pereira, F. Clerc, D. Farrusseng, J. C. Waal, T. Maschmeyer, *QSAR Combinatorial Sci.* **2005**, *24*, 45.
60. R. H. J. M. Otten, L. P. P. Gineken, *The Annealing Algorithm*, Springer, New York, 1989, pp. 224.
61. Z. B. Zabinsky, *Stochastic Adaptive Search for Global Optimization*, Kluwer, Boston, 2003, pp. 224.
62. B. Li, P. Sun, Q. Jin, J. Wang, D. Ding, *J. Mol. Catal. A: Chemical* **1999**, *148*, 189.
63. A. Eftaxias, J. Font, A. Fortuny, J. Giralt, A. Fabregat, F. Stüber, *Appl. Catal. B: Environmental* **2001**, *33*, 175.
64. A. Holzwarth, P. Denton, H. Zanthoff, C. Mirodatos, *Catal. Today* **2001**, *67*, 309.
65. S. Senkan, *Angew. Chem. Int. Ed.* **2001**, *40*, 312.
66. C. A. Floudas, *Deterministic Global Optimization. Theory, Methods and Applications*, Kluwer, Dordrecht, 2000, pp. 739.
67. J. A. Snyman, *Practical Mathematical Optimization. An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, Springer, New York, 2005, pp. 257.
68. K. Huang, X. L. Zhan, F. Q. Chen, D. W. Lü, *Chem. Eng. Sci.* **2003**, *58*, 81.
69. L. Végvári, A. Tompos, S. Göbölös, J. F. Margitfalvi, *Catal. Today* **2003**, *81*, 517.
70. B. Lin, S. Chavali, K. Camarda, D. C. Miller, *Comput. Chem. Eng.* **2005**, *29*, 337.
71. T. Cundari, J. Deng, Y. Zhao, *Ind. Eng. Chem. Res.* **2001**, *40*, 5475.
72. A. Corma, J. M. Serra, E. Argente, V. Botti, S. Valero, *Chemphyschem* **2002**, *3*, 939.
73. K. Omata, M. Hashimoto, Y. Watanabe, T. Umegaki, S. Wagatsuma, G. Ishiguro, M. Yamada, *Appl. Catal. A: General* **2004**, *262*, 207.
74. C. Klanner, D. Farrusseng, L. Baumes, C. Mirodatos, F. Schüth, *QSAR Combinatorial Sci.* **2003**, *22*, 729.
75. A. F. Siegel, C. J. Morgan, *Statistics and Data Analysis: An Introduction*, 2nd edn, Wiley, New York, 1996, pp. 635.
76. M. Berthold, D. J. Hand, *Intelligent Data Analysis. An Introduction*, 2nd edn, Springer, Berlin, 2002, pp. 460.
77. M. Kantardzic, *Data Mining. Concepts, Models, Methods, and Algorithms*, Wiley, Chichester, 2003, pp. 345.
78. J. K. Taylor, C. Cihon, *Statistical Techniques for Data Analysis*, 2nd edn, Chapman & Hall, Boca Raton, 2004, pp. 273.
79. P. C. Meier, R. E. Zund, *Statistical Methods in Analytical Chemistry*, Wiley, New York, 1993, pp. 345.
80. M. Caria, *Measurement Analysis. An Introduction to the Statistical Analysis of Laboratory Data in Physics, Chemistry and the Life Sciences*, Imperial College Press, London, 2000, pp. 229.
81. D. Wallenstein, B. Kanz, R. H. Harding, *Appl. Catal. A: General* **1999**, *178*, 117.
82. J. T. Richardson, D. Remue, J. K. Hung, *Appl. Catal. A: General* **2003**, *250*, 319.
83. O. M. Wilkin, P. M. Maitlis, A. Haynes, M. L. Turner, *Catal. Today* **2003**, *30*, 309.
84. A. M. Amat, A. Arques, S. H. Bossmann, A. M. Braun, S. Göb, M. A. Miranda, E. Oliveros, *Chemosphere* **2004**, *57*, 1123.
85. D. G. Cantrell, L. J. Gillie, A. F. Lee, K. Wilson, *Appl. Catal. A: General* **2005**, *287*, 183.
86. A. Corma, J. M. Serra, P. Serna, M. Moliner, *J. Catal.* **2005**, *232*, 335.
87. D. Farrusseng, C. Klanner, L. Baumes, M. Lengliz, C. Mirodatos, F. Schüth, *QSAR Combinatorial Sci.* **2005**, *24*, 78.
88. R. J. Hendershot, R. Vijay, B. J. Feist, C. M. Snively, J. Lauterbach, *Measurement Sci. Technol.* **2005**, *16*, 302.
89. M. Holena, M. Baerns, in *Experimental Design for Combinatorial and High Throughput Materials Development*, J. N. Cawse (Ed.), Wiley, New York, 2003, p. 163.
90. J. Zupan, J. Gasteiger, *Neural Networks for Chemists*, Wiley-VCH, Weinheim, 1993, pp. 305.
91. C. M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, 1996, pp. 482.
92. M. T. Hagan, H. Demuth, M. Beale, *Neural Network Design*, PWS Publishing, Boston, 1996, pp. 736.
93. B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, 1996, pp. 403.
94. K. Mehrotra, C. K. Mohan, *Elements of Artificial Neural Networks*, MIT Press, Cambridge, 1997, pp. 344.
95. S. Haykin, *Neural Networks. A Comprehensive Foundation*, 2nd edn, IEEE, New York, 1999, pp. 600.

96. J. Zupan, J. Gasteiger, *Neural Networks in Chemistry and Drug Design: An Introduction*, Wiley-VCH, Weinheim, 1999, pp. 402.
97. T. Hattori, S. Kito, *Catal. Today* **1995**, *23*, 347.
98. M. Sasaki, H. Hamada, Y. Kintaichi, T. Ito, *Appl. Catal. A: General* **1995**, *132*, 261.
99. Z. Y. Hou, Q. Dai, X. Q. Wu, G. T. Chen, *Appl. Catal. A: General* **1997**, *161*, 183.
100. K. Huang, C. Feng-Qiu, D. W. Lü, *Appl. Catal. A: General* **2001**, *219*, 61.
101. Y. Liu, D. Liu, T. Cao, S. Han, G. Xu, *Computers Chem. Eng.* **2001**, *25*, 1711.
102. M. Holena, M. Baerns, *Catal. Today* **2003**, *81*, 485.
103. J. M. Serra, A. Corma, E. Argente, S. Valero, V. Botti, *Appl. Catal. A: General* **2003**, *254*, 133.
104. J. M. Serra, A. Corma, A. Chica, E. Argente, V. Botti, *Catal. Today* **2003**, *81*, 393.
105. S. Kito, A. Satsuma, T. Ishikura, M. Niwa, Y. Murakami, T. Hattori, *Catal. Today* **2004**, *97*, 41.
106. M. Moliner, J. M. Serra, A. Corma, E. Argente, S. Valero, V. Boti, *Microporous Mesoporous Mater.* **2005**, *78*, 73.
107. K. Omata, Y. Kobayashi, M. Yamada, *Catal. Commun.* **2005**, *6*, 563.
108. G. Zahedi, A. Jahanmiri, M. H. Rahimpor, *Int. J. Chem. Reactor Eng.* **2005**, *3*, A8.
109. M. W. B. Trotter, B. F. Buxton, S. B. Holden, *Measurement Control* **2001**, *34*, 235.
110. A. Tompos, J. F. Margittfalvi, E. Tfirst, L. Végvári, M. A. Jaloull, H. A. Khalfalla, M. M. Elgarni, *Appl. Catal. A: General* **2005**, *285*, 65.
111. K. Yajima, Y. Ueda, H. Tsuruya, T. Kanougi, Y. Oumi, S. S. C. Ammal, S. Takami, M. Kubo, A. Miyamoto, *Appl. Catal. A: General* **2000**, *194*, 183.