Modelling
Quantitative Composition Activity Relationships (QCARs)
for Heterogeneous Catalysts by Kriging
and a Multilevel B-Spline Approach

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

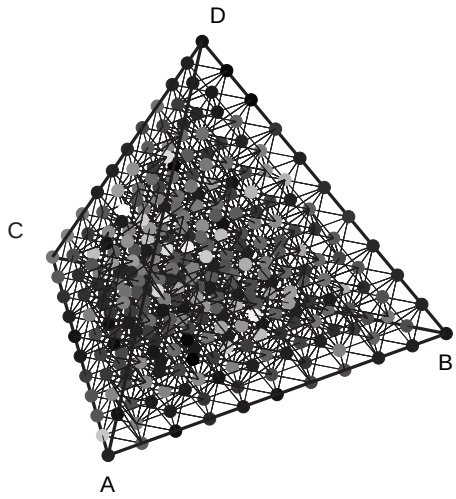
von
Simone Christine Sieg

Saarbrücken
2007

Tag des Kolloquiums:	22.11.2007
Dekan:	Prof. Dr. U. Müller
Berichterstatter:	Prof. Dr. W. F. Maier
	Prof. Dr. M. Springborg
Prüfungsvorsitzender:	Prof. Dr. G. Wenz
Akademischer Beirat:	Dr. R. Haberkorn

Simone C. Sieg

Modelling Quantitative Composition Activity Relationships (QCARs) for Heterogeneous Catalysts by Kriging and a Multilevel B-Spline Approach



Die Wissenschaft ist der Verstand der Welt, die Kunst ihre Seele.

Maksim Gorkij

Acknowledgements

First of all, I want to express my gratitude to Prof. Dr. W. F. Maier for giving me the opportunity to join his team and to start my work in a field being completely new to me. I am very thankful for all his support, his open mind and his encouraging ideas throughout the whole thesis.
Second, I thank Prof. Dr. M. Springborg for his time and interest in this thesis.

Especially, I would like to thank Prof. Dr. K. Rajan at Iowa State University, USA, for his kind invitations to join his work group and to visit him in Troy and Ames. I am very grateful to my dear colleague and friend Dr. Changwon Suh at Iowa State University who helped me to deepen my understanding of Principal Component Analysis and high-dimensional visualization techniques. Our collaboration has always been a great pleasure.

I would like to thank Prof. Dr. F. Hamprecht at IWR in Heidelberg for his kind advice in the application of Kriging and the discussions we had.

Moreover, I want to thank all my colleagues and friends in Saarbrücken for the friendly and positive atmosphere during my PhD-time and all the wonderful moments we shared together. We had a lot of fun! I am also very grateful to all chemists who patiently showed a mathematician how to do experimental work - without their help, this thesis would not have been realized in this way. Especially, I want to thank Dipl.-Ing.(FH) R. Richter, H. Höltzen, C. Thome and J. Kriesamer for their steady and excellent support concerning our technical equipment and S. Kern-Schumacher for her advice in all kind of administrative things.

I am very thankful to Martin Fengler for being my *mathematical support team* in any kind of situation - I learned so much from you! Thank you for encouraging me to start this thesis and for sharing 4 wonderful years together.

Furthermore, special thanks go to Michael Reiser for putting a lot of work and effort into reading and correcting the manuscript - thank you so much!

Last but not least, I thank my family and especially my grandfather for their steady support throughout my whole student and PhD-years.

Abstract

The main focus of this work lies in the mathematical modelling of quantitative composition activity relationships (QCARs) for heterogeneous catalysts tested for the oxidation of propene to acrolein. The use of combinatorial chemistry approaches together with high-throughput screening techniques also plays an important role here. One the one hand, the thesis tries to give a validation of the used synthesis and screening processes and on the other hand it is checked how well QCARs can be realized by the two applied interpolation techniques: Kriging and multilevel B-Splines.

Following a sol-gel procedure approximately 2400 catalysts have been prepared and tested with the help of high-throughput synthesis and screening approaches. The samples include two complete pentanary composition spreads (elements Cr, Mn, Co, Te and Ni) having 10-%-wise variations in composition together with further refinements. The screening for catalytic activity itself has been realized in a high-throughput reactor system for sequential testing. An indicative sign of a potentially good catalyst candidate has been a large GC signal for acrolein in the product gas composition.

For the analysis of data, new visualization techniques needed to be developed and introduced into the field of heterogeneous catalysis since common visualization approaches could not cope with more than three dimensional data sets.

Another challenge has been the calculation of activities of 5%-wise variations given the 10%-wise samples by Kringing and B-Splines. Since the underlying functional relationship between composition and catalytic activity is not known, a direct evaluation of both interpolation techniques cannot be easily given.

Zusammenfassung

Das Hauptaugenmerk dieser Arbeit richtet sich auf die mathematische Modellierung von Zusammensetzung-Aktivität-Beziehungen für heterogene Katalysatoren in der Oxidation von Propen. Der Einsatz von Methoden der Kombinatorischen Chemie und Hochdurchsatzansätzen sowohl zur Herstellung als auch zum Testen dieser Katalysatorproben spielt dabei eine entscheidende Rolle. Die Arbeit versucht, zum einen eine Validierung der verwendeten Synthese- und Screeningverfahren zu geben, zum anderen aber auch zu prüfen, wie gut eine Beschreibung von Zusammensetzung-Aktivität-Beziehungen mit zwei mathematischen Interpolationsverfahren (Kriging, multilevel B-Splines) möglich ist.

Im experimentellen Teil der Arbeit wurden rund 2400 Katalysatoren, zusammengesetzt aus den Elementen Cr, Mn, Co, Te und Ni, synthetisiert und auf ihre Aktivität hin getestet. Dabei wurden zwei komplette pentanäre Datensätze hergestellt (10%-ige Variation der Zusammensetzung). Alle Proben wurden in einer Hochdurchsatz-Screening Apparatur auf ihre Aktivität hinsichtlich der Oxidation von Propen (Zielprodukt Acrolein) untersucht. Als Kriterium für Aktivität wurden GC Signale aller interessanten Produkte aufgenommen. Hohe GC Signale entsprachen einer hohen Aktivität der Katalysatoren für das entsprechende Produkt.

Zum Auswerten der Daten waren neue Visualisierungskonzepte zu entwickeln, wie sie im Bereich der heterogenen Katalyse noch nicht verwendet wurden. Mit Hilfe des Kriging und B-Spline Ansatzes konnten die Aktivitäten von Katalysatoren mit engeren Rasterungen im Suchraum geschätzt und mit experimentellen Werten verglichen werden. Dies lieferte neben der Reproduzierbarkeitsanalyse auch einen Vergleich der zwei verwendeten Modellierungsmethoden, wie es ansonsten mangels Kenntnis des wahren zugrundeliegenden funktionalen Zusammenhangs nicht möglich ist.

Contents

Part I

Introduction

1 Introduction and Objectives

1.1 Combinatorial Chemistry

In the 1980s [157, 158, 97] the need to rapidly and inexpensively synthesize many chemical compounds spawned a new branch of chemistry known as *combinatorial chemistry*. Originally, the techniques of this fast growing field have been primarily developed by pharmaceutical companies and research in pharmaceutical sectors to find new candidate drugs. Traditionally, chemists in the pharmaceutical business had to synthesize possible drug candidates one by one before they could screen their activities. The basic idea of combinatorial chemistry is to synthesize rapidly large amounts of different compounds at the same time using a process that is supported by automation and computation. Combinatorial chemistry contrasts with the time-consuming and labor intensive method of traditional chemistry where compounds are synthesized individually, one at a time. A variety of successes of combinatorial approaches in pharmaceutical companies have been reported where drugs could have been developed within months instead of years. To create and screen thousands of compounds within days, hundreds of millions of dollars have been invested in these techniques and nowadays nearly all of the major pharmaceutical companies have their own combinatorial chemistry department [129]. Once the benefits of combinatorial chemistry became apparent, inorganic and polymer chemists started applying these methods, too.
Originally, combinatorial chemistry has been based on the premise that the probability of finding a hit in a random screening process is proportional to the number of samples considered. In the beginning, the main objective of combinatorial chemistry has been the simultaneous generation of a large variety of samples with an additional screening process. Following this approach, the success rate of finding new hits is greatly enhanced, while the time needed is considerably reduced. There are even numerous advantages coming along with combinatorial chemistry, these new approaches could not diminish the importance of "classical" techniques that still play a significant role. Figure 1.1 illustrates the principal characteristics of conventional and combinatorial strategies and also stresses the synergistic effects of combining both. The combinatorial approach should be considered to be a sort of completion of conventional research, but not in a competitive way. A detailed description of many advantages of conventional approaches has been intensively discussed by Schlögl [148]. Still, the conventional way of research plays the most important role of development and evaluation processes for new materials. Once a promising hit has been discovered by combinatorial techniques it needs to be validated and synthesized in laboratory scales, furthermore tested under conventional conditions to confirm the results found by high-throughput. The natural diversity of the periodic system bears too many possibilities to systematically synthesize and screen all of them. For example, thinking of 75 stable elements and neglecting the variety of possible structures, phases and compositions, 2,775 binary, 67,525 ternary and over

one million quaternary mixtures can be obtained, cf. [104]. Up to now, one has to be sure that we know only a very small part of all these possible materials. At the moment combinatorial chemistry can be considered to be the most promising concept to explore this vast variety of possibilities - although we must be aware of that a complete recording of chemical diversity appears to be impossible. Even combinatorial approaches cannot perform this. But they can improve the chance of discovering completely new materials having special or improved features.

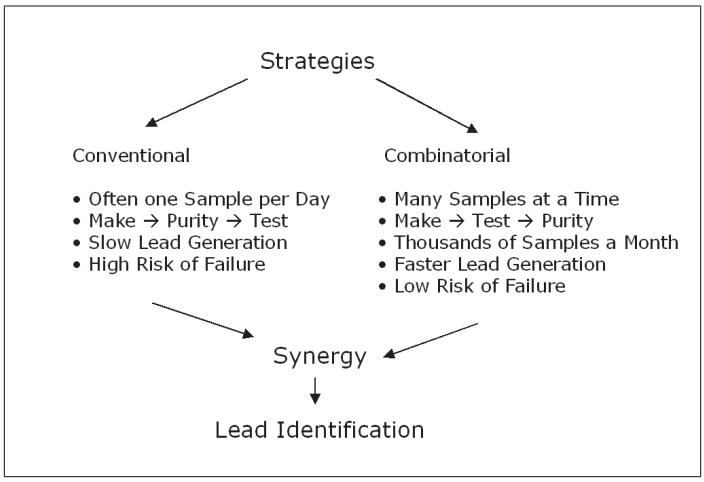


Figure 1.1: Principal characteristics of conventional vs. combinatorial strategy of drug discovery.

The main advantage of combinatorial strategies compared to conventional approaches is that they lead to samples that have been all obtained automatically under identical conditions and thus are highly reproducible. Automated equipment can perform operations more rapidly (and for 24h a day) and can also cope with very small amounts of reactants in a precise way.

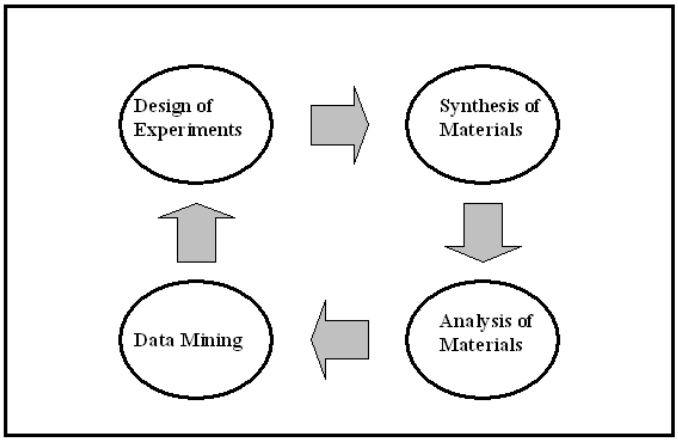


Figure 1.2: Schematical workflow of combinatorial materials science.

Besides of being faster, combinatorial methods are also safer and have a lower environmental impact, since they use only small quantities of reactants. Being more costly in equipment than conventional approaches, combinatorial chemistry can only be successful if hits are identified within short time periods. Therefore, the experiments need to be planned very carefully.
The combinatorial approach can be divided into four subfields, which are illustrated by Figure 1.2.

Design of Experiments (DOE)
Before combinatorial experiments are being started, the experiments need to be planned in an intelligent way. Intelligent means here, that as few experiments as possible should be run to gain maximal knowledge about a studied system. In literature, this approach is called "Design of Experiments" (DOE) and there exist vast amounts of publications and monographs on that topic. In context with combinatorial approaches, Cawse [21, 22] and Agrafiotis [2] published interesting ideas and realizations and we refer to references [13, 14] for examples of DOE realized in industry.

Synthesis of Materials
The synthesis of materials aims to synthesize and analyze as many samples in parallel as possible. Today, mainly automated equipment is used for high-throughput syntheses and techniques like physical vapor deposition (PVD) [28] or liquid phase syntheses (Sol-Gel-, Hydrothermal-, Precipitation-, Polymerisation Syntheses). The latter case often allows the elegant use of ink-jet printers. In 1997, a library containing new fluorescent materials has been created out of aqueous solutions of nitrates of La, Eu, Gd and Al using ink-jet technology by Sun and coworkers, [172]. With this technique, single volumes of 0.5 nl can be dosed exactly on desired positions. A combinatorial realization of hydrothermal syntheses has been reported by Klein et al. [94].

Analysis of Materials
Once the HT-synthesis of combinatorial libraries has been finished, the materials need to be characterized and screened for a desired function. Again, these screening processes should also be as automated as possible to save time and to increase accuracy. It is impossible to mention all possible HT-screening techniques here since there are too many. As the most important techniques one might think of video cameras to detect optical features or IR-cameras having a high temperature resolution for the detection of spatially resolved reaction heats during a chemical reaction, cf. [73]. Also spatially resolved mass spectrometry has been applied to analyze products of combinatorial experiments, cf. [24, 87, 174].

Data Mining
Automated processes applied in combinatorial experiments lead to vast amounts of data that need to be handled and evaluated. This data management cannot be done manually and finding trends and correlations within large data sets is not a trivial task.
The term *Data Mining (DM)* can be explained as the extraction of previously unknown information out of databases using statistics, machine learning techniques,

databases and visualization approaches [190]. In literature, also the term *Knowledge Discovery in Databases (KDD)* has been coined being some sort of generic term to data mining. According to Fayyad et al. [41], KDD is defined as *"the non-trivial process of identifying valid and potentially useful new patterns in vast amounts of data."*
The use of mathematical and statistical techniques in data mining processes already revealed many successes. The main techniques that are applied in data mining are factor analysis, principle component analysis, regression (non-linear and linear), classification methods (e.g. hierarchical clustering, k-means clustering), outlier tests, trend analysis and so on. They all aim to discover existing structures and correlations that might be somehow hidden or not directly accessible. The whole bundle of methods can be divided into two parts: those that do *prediction* and those that do *classification*. Hastie and Tibshirani [65] and Härdle [62] provide good insight in most of these techniques. For a detailed introduction to principle component analysis we refer to Joliffe [84]. Data mining also includes the technique of *automatic learning*, also called *statistical learning*. These methods deal with learning out of empirical data with the help of mathematically based algorithms. The special interest in these techniques for material science in general lies in the possibility to establish predictions of activity out of experimental data for samples with new compositions. This means that it is extremely important to extract as much information out of the data sets as possible to derive new knowledge about the correlation of composition, micro structure and function of a certain material. Possibilities and chances of data mining techniques are also summarized in [177]. Maier *et al.* [106] recently reviewed the state of the art in materials science also including an overview on the current status of the application of data mining techniques.
Within the combinatorial workflow, pieces of information that have been obtained by data mining can now be used for the planning of new experiments and thus, the circle in Figure 1.2 is closed.

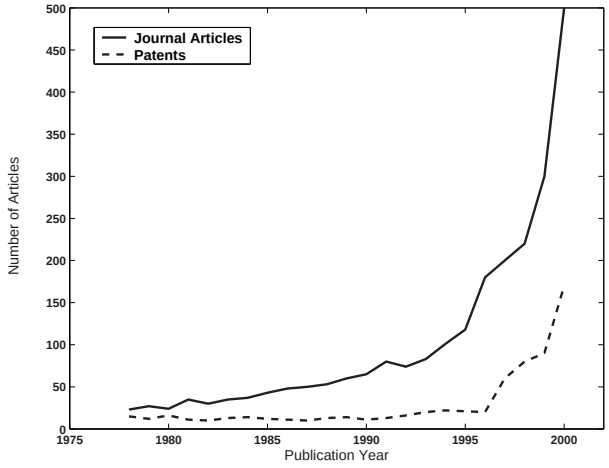


Figure 1.3: Growth of combinatorial literature and patents (non organic and non pharmaceutical)[1].

[1] http://www.istl.org/02-winter/refereed.html

Taking a look at the number of publications that have been published during the last two decades dealing with combinatorial approaches, it can be clearly seen that the growth in combinatorial literature is continuously expanding. Figure 1.3 gives an overview on the number of publications and patents that have been published within the last 20 years. Obviously, an ongoing growth among the number of publications can be observed from 1995. And the trend still continues until today indicating high research activities.
Especially during the last decade, combinatorial chemistry has been also finding other applications in various fields such as for example high-temperature superconductors [195], magnetoresists [15], synthesis of metal surfaces exhibiting a gradient in composition [9], in the synthesis of organic host compounds [143], optimization and discovery of luminescent metal oxide materials [28, 29, 181] and many other more. The following section is dedicated to the application of combinatorial chemistry approaches in heterogeneous catalysis.

1.2 Combinatorial Heterogeneous Catalysis

Today our high living standards are mainly influenced by optimized chemicals and fuels manufactured with the help of so-called *catalysts*. In general, catalysts are substances that decrease the activation energy of a chemical reaction by interfering into its reaction mechanism such that the rate constant of the conversion is increased (or decreased). At the end of the reaction, the catalyst has not been wasted or converted [136].
These catalysts strongly contribute to the world's economy since catalysts are used in the production of over 7000 compounds worth over $3 trillion globally. About 90% of chemical processes and 60% of chemicals production relie on the use of catalysts. Globally there are about 100 catalyst manufacturing companies and the world wide catalysts manufacture is estimated to $ 8.5 billion per year, cf. [156]. In the future, these figure are likely about to increase as a direct result of increasing pressure to develop environmental friendly manufacturing processes. From an economic point of view, the benefits that come along with the development and use of efficient catalysts are enormous: catalytic processes can be realized in a cheaper way, having lower operating costs, creating less by-products or lead to formation of products with higher purities. In this context it becomes clear that there is an always growing need for optimized and improved catalysts for every kind of application. However, the discovery of new catalysts continues to be a challenging and rather unpredictable trial-and-error process [130]. The conventional catalyst development uses a large variety of recipes and synthesis routes, which are in general very time consuming. Following the synthesis, the prepared samples need to be characterized and tested and further improved until no improvements can be obtained anymore. In addition, time scales here are given in months and years until a developed catalyst is commercialized and applied in industrial processes. Facing increasing competition, these time limits become narrower and narrower going along with increasing pressure on industry due to environmental aspects. This bottleneck can only be overcome by new, innovative ideas that would lead to shorter development times for highly demanded products and chemicals. But then also the development time for appropriate catalysts need to be drastically decreased. Another crucial point here

is the fact that the natural resources slowly run out and there are less resources available for research and development.
Approximately twenty years ago, these problems have been already recognized within the pharmaceutical industry which led to a strong spreading of combinatorial and high-throughput approaches within this community. Since 1999 the number of conferences and workshops on combinatorial catalysis has remarkably increased and this rather new research field of *combinatorial catalysis* has gained its acceptance in industry and academia. According to Senkan [156] the term *combinatorial catalysis* itself can be described as a *methodology or a set of tools where large diversities of solid-state materials libraries are prepared, processed, and tested for activity and selectivity in a high-throughput fashion.*

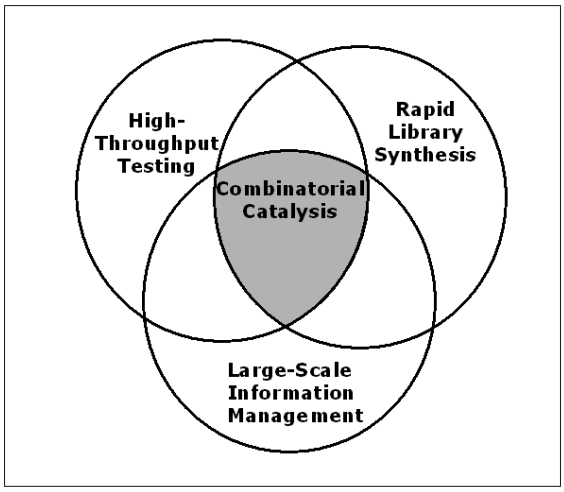


Figure 1.4: Components of combinatorial catalysis.

As already mentioned above, high-throughput screening techniques produce an enormous amount of data that need to be organized and processed. Therefore, combinatorial catalysis also has to come along with sophisticated data management systems. Here, the border between chemistry and informatics is more and more vanishing and in the future more and more interdisciplinary work will be needed to realize efficient combinatorial approaches. Figure 1.4 illustrates the main components that need to be incorporated into combinatorial catalysis. It is of essential importance here, that all parts are equally weighted to prevent the whole approach from bottlenecks limiting the throughput. Although this is not an easy task, the optimal integration of all parts should be the ultimate goal here.
The transfer of combinatorial methodologies to heterogeneous catalysis is a more challenging task than for other scientific fields of materials science [82]. There are many more aspects to consider arising from the complex and dynamic structure of catalysts:
Heterogeneous catalyst are in general solid materials with multiple discontinuities in structure and composition that cannot be easily controlled or systematically varied. Additionally, these structures directly depend on the processing and synthesis of a material and may change when the catalyst is exposed to reaction conditions. Here, good catalysts will exhibit less changes resulting in slower deactivation and longer life times. In contrast to other research areas there exist no defined descriptors that

could perfectly describe a heterogeneous catalyst [148] which makes research in this field a challenging task. Furthermore no theories are yet known that would allow to predict the catalytic performance of a material just out of its composition or structure. Within the last years, some research activities have been focused on so-called quantitative-activity relationships (QSARs) or quantitative composition-activity relationships (QCAR) trying to expand the "similarity principle" known in organic chemistry to solids [26, 92, 147, 162]. But this stays a difficult task.
Heterogeneous catalysts must further bind the reactants, lead them to bonding interactions and allow the products to depart again. This is assumed to happen on active sites situated on the surface of the catalyst particle or within its pores. Multitudes of active sites exit on solid surfaces but their distribution changes with time. These changes are caused by rates of solid-state diffusion and gas-solid interactions, both influenced by temperature and local composition. As a direct consequence, catalysts deactivate with time due surface reconstruction, sintering, poisoning, coke formation and volatilization [31]. But it has been also observed that some catalysts need some time "on stream" to reach their peak performance [38] which might be also induced by the events mentioned before (e. g. surface reconstruction). To bring some more light into the darkness here, new methodologies are needed that allow the realizations of a large number of systematic and parallel experiments helping to learn more about heterogeneous catalysis and its mechanisms on the molecular scale. That is why combinatorial catalysis may immensely profit from highly-parallel experimentation and screening methods coming along with all advantages mentioned above. Although software development for the organization of high-throughput experiments and automated data management still is in its infancy. The use of combinatorial approaches in heterogeneous catalysis may contribute to the general understanding of catalysis in different manners: Clearly, it may speed up the whole development process by increasing the chance of developing totally new and unexpected hits. Furthermore, the systematic application of sophisticated data mining techniques to large data bases may lead to the discovery of more trends and patterns for structure-activity relationships leading to new input for practical catalyst development.

History and State of the Art of Combinatorial Heterogeneous Catalysis
The idea of applying combinatorial techniques to heterogeneous catalysis is not totally new but can be dated back to the early years of the 20th century where Mittasch at the BASF company coordinated an impressive research program to find a catalyst for the synthesis of ammonia [117, 118]. For the systematic search for a catalytic material approximately 20.000 experiments have been run by many researchers in parallel, testing about 6.500 substances until they finally ended up with the optimized catalyst. This catalyst allowed the industrial production of ammonia according to the Haber-Bosch process and it is still in use until today changed by only slight modifications. This has been the first great success of a combinatorial approach realized in industrial research. According to Senkan, [156], as a consequence here, the birth of combinatorial inorganic chemistry should not be attributed to materials research [61], as one can often find in the literature, but to heterogeneous catalysis instead.
In the years after Mittasch's success, the combinatorial approach to catalyst optimization has not been further followed, probably due to high labor costs. For more than half a century, catalysis research has been mainly realized following the

conventional approach of synthesizing and testing one sample at a time. Thus, the development of new catalysts usually took several years. In 1970, this rather ineffective and slow approach has been criticized by Hanak [61] indicating that using parallel synthesis and testing methods (*'multi-sample-concept'*) would be more efficient and remarkably speed up the development process. But again, at that time, the ideas of Hanak have not been adapted by other research groups and it took some more years until the combinatorial approach has been brought to life by the pharmaceutical industry. As mentioned above, high-throughput and combinatorial techniques are widely spread within the pharmaceutical community and today often 10.000 samples can be tested per day.
Combinatorial libraries of potential catalysts can comprise organic, organometallic, inorganic or solid materials. Especially in heterogeneous catalysis, the active sites of the catalyst are embedded in a solid state material. Thus, the synthesis of combinatorial libraries can be realized by a variety of solution- based or vapor deposition methods. A pioneering work in this field has been carried out by Schultz et al. [195] in 1995, at the University of California. Here, the group applied vapor deposition methods to synthesize thin film libraries of superconducting and magnetoresistant materials. A short time period after these advances, researchers at Symyx Technologies produced high-density, thin film libraries of photoluminescent materials using improved vapor deposition methods [28, 29]. As one of the first applications of combinatorial approaches in catalysis the work of Weinberg et. al. [24] at the Symyx Company has to be mentioned. In 1999, the group published results on the catalytic oxidation of CO and the reduction of NO by metal alloy catalysts, since this reaction has been quite thoroughly studied and perfectly lend itself to test and validate combinatorial synthesis and screening techniques. Combinatorial chemistry with respect to heterogeneous catalysis also includes high-throughput synthesis techniques, high-throughput screening and data management. All three fields need to be perfectly linked together to avoid the "bottleneck" that would remarkably decrease the throughput and therefore the efficiency of the whole approach. Within the last decade, many publications have been published dealing with *combinatorial catalysis*. In 1998, Jandeleit et al. summarized and critically analyzed the development in the application of combinatorial methodologies to the discovery of homogeneous and heterogeneous catalysts between 1995 and 1998. For a detailed overview on the developments achieved within this time period the reader is referred to some review articles, e.g. [80, 184, 160, 19, 188, 128]. In 2001, Senkan [156] published a review on combinatorial heterogeneous catalysis giving an excellent overview on recent catalysis research up to 2001.
During the last five years research in heterogeneous catalysis has been extremely influenced by the use of combinatorial methods. It is impossible to give a detailed listing of all published works here, since this would go beyond the scope of this thesis. Nevertheless, we should mention some success stories here. Within the last 8 years the group around Maier published several successful results of new catalyst materials discovered by the use of high-throughput synthesis and screening methods. The group mainly worked with sol-gel techniques to synthesize the catalyst samples using automated dispensing systems and emissivity corrected IR-thermography [72, 73] for the detection of catalytically active species. Saalfrank [141, 142] discovered a new catalyst being free of precious metals for CO oxidation at room temperature. Kirsten [89] worked on combinatorial strategies to discover new catalysts with the help of

genetic algorithms. Paul [127] used high-throughput screening set-ups for the testing of antimony-rich selective oxidation catalysts. Furthermore, research of this group also focuses on the development of screening set-ups [174] and the development of data management systems [47].
Corma and coworkers widely use combinatorial approaches in their research for new catalyst materials in oil reforming, petrochemical or fine chemicals processes but also to find quantitative structure property relationships [25, 26]. The group around Baerns published several success stories dealing with the development of high-throughput synthesis and screening equipment for catalytic materials, e.g. [59]. They also used combinatorial techniques combined with data mining or learning approaches (genetic algorithms, artificial neural networks) to discover new heterogeneous catalysts [191, 192, 139]. A detailed overview on recent research activities in the field of applying genetic algorithms and artificial neural networks to heterogeneous catalysis is given by a contemporary review by Maier et al. [106].
Research work on quantitative structure activity relationships has also be reported by projects of the groups around Schüth and Mirodatos [92, 91]. At the moment, numerous industrial and academic research groups all over the world deal with combinatorial heterogeneous catalysis to optimize already known catalysts or to discover new materials. Within the last 10 years, nearly all big chemical companies invested in high-throughput equipment and there have also been several company foundations aiming at the development and sale of high-throughput set-ups (e.g. Symyx Technologies (USA), Avantium Technologies (The Netherlands), Chemspeed Ltd. (Switzerland) or the hte-AG (Germany).

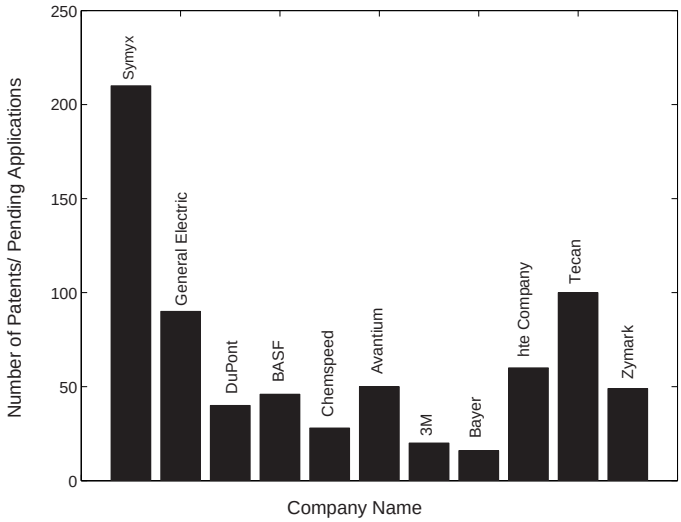


Figure 1.5: High-throughput patents and pending applications (non-bioscience and non-pharmaceutical) split by company, [30].

In the last ten years, the number of patents filed in this area has grown significantly. Figure 1.5 shows the number of patents published within the last years by the most common chemical companies using high-throughput technology. TecanInc. and Zymark Inc. that supply robotics systems useful for materials and chemical research have also made significant additions in their product offerings to the industry.

1.3 Parameter Spaces and Composition Spreads

1.3.1 Parameter Spaces

As mentioned above, dealing with heterogeneous catalysts is a multi-dimensional task. A large variety of parameters influence the performance of a heterogeneous catalyst and therefore the parameter space that has to be dealt with is extremely large. If it is assumed that there are 50 useful, stable and affordable elements in the periodic table that could be used as constituents for heterogeneous catalysts, 1225 binary, 19600 ternary, 230 000 quaternary and about 10^{10} decanary combinations can be constructed, following equation (1.1),

$$N(N_E, N_M) = \binom{N_E}{N_M} = \frac{N_E!}{N_M!(N_E - N_M)!} \tag{1.1}$$

where N_E denotes the number of possible elements, N_M denotes the number of elements chosen for the mixtures and N gives the total number of possible combinations. Even when neglecting different compositions and structural diversities this parameter space significantly becomes larger and larger with increasing N_M. If additionally n_c compositional increments are used the total number of combinations is given by equation (1.2)

$$N(N_M, n_c) = \binom{n_c + N_M - 2}{n_c - 1} = \frac{(n_c + N_M - 2)!}{(n_c - 1)!(N_M - 1)!}. \tag{1.2}$$

The following table gives the numbers of compositional combinations for multielement mixtures:

n_c	$N_M = 2$	$N_M = 3$	$N_M = 4$	$N_M = 5$	$N_M = 6$
6	6	21	56	126	252
11	**11**	**66**	**286**	**1001**	**3003**
15	15	120	680	3060	11628
21	**21**	**231**	**1771**	**10626**	**53130**
25	25	325	2925	20475	118755
31	31	496	5456	46376	324632
41	41	861	12341	135751	1221759

Table 1.1: Total number of compositional combinations of multielement mixtures.

The most important rows for our considerations are the 10%-wise compositional increments ($n_c = 11$) and 5%-wise compositional increments ($n_c = 21$). This means, for three elements we can synthesize 66 catalyst samples having 10% compositional variations while four elements yield 286 samples. Varying five samples in this way 1001 samples have to be considered. Using 5%-wise compositional variations of 5 elements we would end up with 10626 samples whose syntheses definitely exceed the capacities of conventional catalyst preparation within reasonable time. Within the scope of our work we have been able to prepare 1001 catalyst samples, which gives

a complete 10%-wise compositional variation of 5 elements. It can be clearly seen that the number of combinations that have to be tested and evaluated for increasing N_M and n_c drastically grows and quickly reaches the limit of practical realizability. As we have seen above, there nearly exist an infinite number of possibilities to combine elements out of the periodic table to develop new catalysts. Additionally, synthesis routes can be varied, also the concentrations of the used precursor solutions, the temperature programs, the pretreatment of the catalysts and so on. Thus, the systematic synthesis and testing of all possible catalysts within a finite time period is not possible. In the literature the term *curse of dimension* can be found to appear within this context to describe the explosion of new possible mixtures by just adding another element or using finer increments. Approaches to tackle this problem can be found in the field of *Design of Experiments (DOE)*, where maximal information should be obtained out of a minimal number of experiments. In the literature, there exist several theories about the optimal sampling of a search space and especially for high-throughput research DOE is quite essential (cf. Figure 1.2). Hamprecht and Agrell [60] considered the way sampling of a defined search space can be done to gain the clearest possible picture of it. Furthermore, they express the problem of sampling high-dimensional search spaces, where most of the points will lie on the surfaces and only few inner points are obtained using uniformly spaced samples. As an example, it can be mentioned here, that the sampling of a pentanary search space using 10% increments results in 1001 samples, but only 126 samples really consist of all five elements. All other samples can be interpreted as subsets of the whole search space lying on hypersurfaces. The methodology of DOE approaches can be found within several monographs and articles, e. g. [22, 60] and [106], and within the references therein.
The curse of dimension still prevents the possibility to systematically sample the parameter space for more than 10 considered elements. In view of this large dimensional parameter spaces, combinatorial approaches seem to be the most rational way for the discovery of new catalysts. The use of improved combinatorial synthesis and screening methods still cannot realize to completely screen this large search space but they can help to build appropriate, diverse combinatorial libraries that may lead to an increased probability of finding a hit for a chemical reaction of interest. This can only be realized by including chemical knowledge that has been obtained out of conventional, traditional catalysis research and experiments but also combined with chemical intuition or numerical simulation methods. These computational approaches still are in their infancy with respect to heterogeneous catalysis due to the high complexity of the topic. There are only few theories or well understood mechanisms available that would help to explain what really happens on the molecular scale during a catalyzed chemical reaction. Thus, it is very hard to work out structure activity or composition activity relationships for heterogeneous catalysts and only little successes have been achieved here during the last years (cf. above).

1.3.2 Composition Spreads

The term *Composition Spread* will be used constantly throughout this thesis and therefore it should be clarified here. As discussed above, the characterization of solid catalyst materials depends on many parameters (synthesis parameters, surface characteristics, pore structure, phases, etc.) and of course on the chemical composition

of the material. For the realization of search strategies, the parameter space needs to be efficiently sampled. This can be done by working out detailed experimental designs, varying different parameters at a time or by fixing all parameters except one and varying just this one key parameter. In this study, the chemical composition of a catalytic material has been chosen to be the key parameter. Thus, the dimensionality of the search space has been given by the number of contained elements of a material. Then a sampling can be done by considering samples having different compositions and therefore giving different points within this search space. Since the composition of a material can only consist of discrete values, a finite number of possible combinations is given. Following table 1.1, the variation of three elements using 10% increments produces 66 combinations, a so-called *ternary composition spread*. Using four elements with again 10% increments in composition leads to 286 combinations, a *quaternary composition spread*. In the case of solid materials with defined compositions, one speaks of *discrete composition spreads* since the increments cannot become infinitely small, due to fabrication limits. In the approach discussed here, composition spreads having 10%-wise variations in composition have been considered. Regions that appeared to be of special interest have been sampled by 5%-wise steps. In other fields, e.g. thin films, also continuous composition spreads can be obtained by sputtering techniques using sophisticated masks, cf. [196, 176]. Here, continuous gradients in composition can be achieved and a predefined search space can be completely screened, at least from a theoretical point of view. In practice, there are still problems arising from continuous composition spread approaches, for example, how to address a certain composition for screening.
The following figure shows, how discrete ternary and quaternary composition spreads are graphically presented within this thesis.

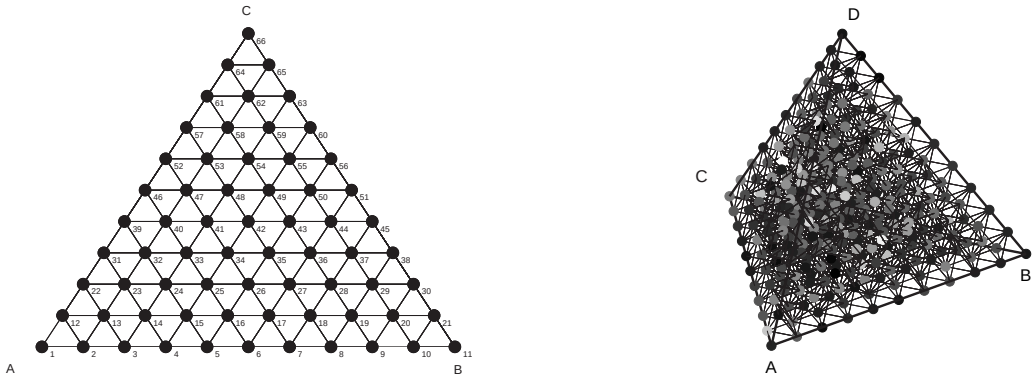


Figure 1.6: Discrete ternary (left) and quaternary (right) composition spread.

1.4 Preparation of Solid-State Libraries of Catalytic Materials

Techniques to combinatorially synthesize solid-state libraries of catalytic materials have to lend themselves well for parallelization and automation. Basically, the techniques can be grouped into two categories: *thin film deposition-based* methods and *solution-based* methods of synthesis. In 2001, Senkan [156] presented a summary of numerous combinatorial heterogeneous catalyst library synthesis methods and on the one hand this overview clearly showed the infancy of the field at that time. On

the other hand, it also illustrated that these approaches bear many research and development opportunities for the discovery and optimization of new catalysts.
Today, six years later, the use of the mentioned synthesis methods has further spread among industrial and academic research groups and combinatorial techniques could tighten their position.
In general, a crucial practical issue in library synthesis is the so-called *upscaling*. The most important goal of catalysis research is to develop new catalysts that can be produced in large scale to be ready for industrial applications. Therefore, it has to be checked, if each step of the combinatorial synthesis route can be realized in scaling-up. A well known problem here is the scaling-up of materials previously discovered as thin films since chemical and physical structures of those films may not be easily transferable to the corresponding bulk material.

Thin Film Deposition Based Methods
For the preparation of thin film combinatorial libraries sputtering techniques with physical masking are widely used. Weinberg and coworkers [24] for example applied sputtering techniques for the discovery and optimization of CO oxidation catalysts. The techniques used here have been based on the pioneering work of Hanak [61] and Xiang et al. [195]. These authors published appropriate methodologies for the parallel synthesis of spatially addressable solid-state materials libraries. One main advantage of sputtering techniques is the fact that they allow the quick synthesis of a large number of samples in parallel, i.e. they perfectly lend themselves for the preparation of very large library systems. Furthermore, these techniques allow the production of films having almost every desired and flexible composition. After the deposition, the films can be further processed, for example, calcined or reduced to produce catalytic materials. The most important thin film deposition methods are:

- Thermal chemical-vapor deposition, [85, 119]
- Plasma chemical-vapor deposition, [90, 178]
- Molecular Beam Epitaxy, [88, 113, 125]
- Pulsed Laser Deposition, [7, 66, 102].

For the application in heterogeneous catalysis, there are several limits: first, thin films only have very small geometrical surface areas. Second, interactions between support materials and active phases, effects caused by the particle size and transport phenomena cannot be studied at thin films. Thus, thin film techniques for the preparation of libraries of catalytic materials suffer from several drawbacks that cannot be neglected. Today, most commercial catalysts are prepared by solution-based methods.

Solution-Based Synthesis of Libraries
The application of solution-based syntheses for the preparation of solids bears the advantage that the difficult handling of solids, e.g. automated dosing, can be totally by-passed working with solutions. In general, solution-based synthesis techniques can be divided into two groups: *coprecipitation* and *impregnation*, [144, 135]. Furthermore, there exist several other methods that can be considered of being variants

of the two groups mentioned above: hydrothermal synthesis, complexation, gelation, crystallization, ion exchange, grafting, adsorption and deposition, [152].
For coprecipitation, two or more solutions are mixed, followed by precipitation, filtration, washing, drying, forming and activation of the catalytic material. For impregnation techniques, the solutions containing the catalytic compounds are brought together with a porous carrier material, followed by drying and activation steps. Here, the surface area and mechanical properties of the final catalytic material are mainly determined by the type of the carrier. To apply impregnation and coprecipitation techniques successfully in combinatorial high-throughput approaches, the processing steps need to be automated and miniaturized. Good results have been achieved by the use of dispensing robots for dosing the precursor solution on the carrier materials [137, 138]. Other research groups used modified ink jet printers to dose precursor solutions on carrier materials [34, 172, 121]. The use of hydrothermal synthesis techniques has already been reported several times, cf. [3, 94], and for detailed information the review article by Akporiaye and coworkers [4] might be interesting. Further success stories of applying these synthesis techniques in a combinatorial way have been reported and a good overview on different works can be found in reference [156]. Sol-gel preparation techniques lend themselves in a perfect way to the automated synthesis of catalyst libraries by the use of dispensing robot systems [145]. The precursor solutions are mixed in the desired ratios by the robot and the final catalytic materials (metal oxides) are then obtained by additional gelation, drying and calcination steps. All catalysts discussed within this thesis have been prepared following exactly this route and the following section is dedicated to the explanation of the sol-gel process from a mechanistical point of view.

1.5 Preparation of Solids by Sol-Gel Methods

A process that has, in the past years, gained much notoriety in the glass and ceramic field is the sol-gel process. Using this synthesis technique a large variety of inorganic networks from silicon or metal alkoxide monomer precursors can be produced. The first notification of sol-gel approaches can even be traced back in the mid 1800s where Ebelmen reported the first silica gelation from $SiCl_4$ in moist air [37]. Extensive studies followed by the early 1930s and a renewed interest evoked in the early 1970s when it became possible to form monolithic inorganic gels at low temperatures and convert them into glasses without the need of a high temperature melting process. Through the sol-gel procedure, homogeneous inorganic oxide materials with desirable properties such as hardness, optical transparency, chemical durability, tailored porosity, and thermal resistance, can be produced at room temperature, as opposed to the much higher melting temperatures required in the production of conventional inorganic glasses. Various material shapes can be generated in the gel state such that the ceramics and glasses produced in this way appear in the shape of monoliths, films, fibers, or monosized powders. Sol-gel techniques comprise a large variety of applications including optics, protective and porous films, optical coatings, window insulators, dielectric and electronic coatings, high temperature superconductors, reinforcement fibers, fillers, and, important here, *heterogeneous catalysts*, [48, 67]. The preparation of catalyst materials using the sol-gel technique comes along with a bundle of advantages, since the syntheses can be

well controlled by the researchers and individually adapted to "design" desired features of the new material. Following [53, 58, 95, 96], the sol-gel process allows to get

- a superior homogeneity and higher purity of a material
- an improved microstructural control of the metallic particles that serve as isolated active site of a catalyst
- large BET[2] surface areas
- an improved thermal stability of supported metals
- a well-defined, narrow pore size distribution
- an easy access to a large variety of metals via their precursors
- a porous material out of rather simple synthesis steps (no extra steps as filtration, reflux or distillation)

Derived from the name itself, the sol-gel process involves the evolution of inorganic networks through the formation of a colloidal suspension (sol) and gelation of the sol to form a network in a continuous liquid phase (gel), [86]. The formed gel is an interconnected network with pores of submicrometer dimensions that are filled with a solvent and polymeric chains whose average length is greater than a micrometer, [86]. For the metal precursors, numerous reactants can be used having the metal ion surrounded by appropriate, reactive ligands. Here, metal alkoxides are mostly used as precursors since they react readily with water. Widely used metal alkoxides are the alkoxysilanes, such as tetramethoxysilane (TMOS) and tetraethoxysilane (TEOS). However, other alkoxides such as aluminates, titanates, and borates are also commonly used in the sol-gel process, often mixed with TEOS.
Looking at the molecular level and the behavior of the functional groups during a sol-gel process, the overall reaction can be divided into three steps: *hydrolysis, alcohol condensation* and *water condensation*. A scheme of that can be found in Figure 1.7 where these three reaction steps are illustrated by alkoxysilanes.

[2] named after Brunauer, Emmett and Teller, who developed an elegant method to determine the surface area of a porous material by gas adsorption, [18].

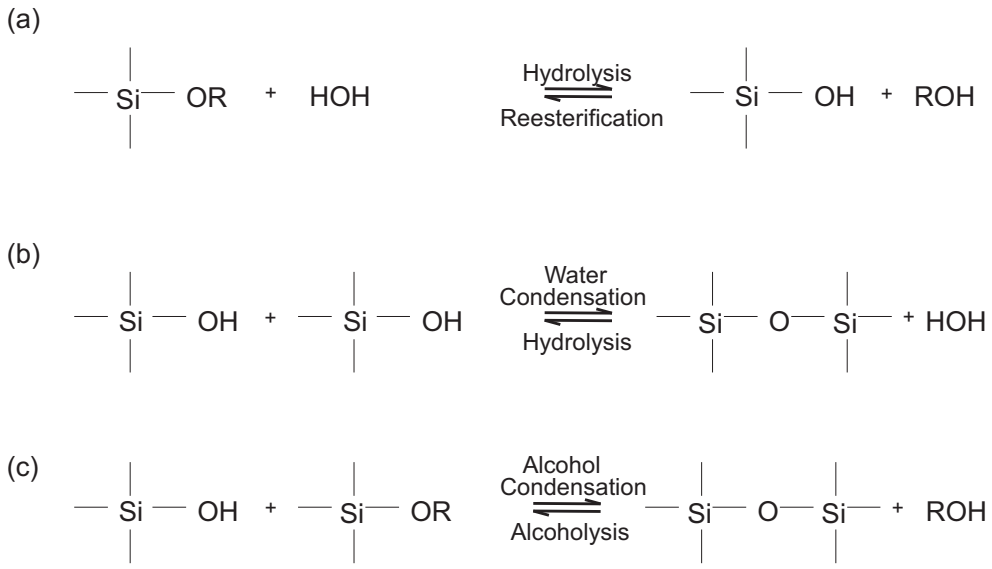


Figure 1.7: The three main reaction steps of the sol-gel process, (a) hydrolysis, (b) water condensation, and (c) alcohol condensation.

Step (a) is the hydrolysis reaction of the alkoxysilanes to form silanols (terminal alkoxide groups (OR) are replaced by hydroxyl groups). The subsequent water and alcohol condensation steps ((b) + (c)) involve the silanol groups (Si-OH) to produce siloxane bonds (Si-O-Si) together with either water or an alcohol as side products.
The properties and special features of a sol-gel network are influenced by a number of synthesis factors that control the rate of hydrolysis or condensation. As the most important factors, the pH-value, reaction temperature and time, concentration of the reactants, presence of catalysts and the molar ratio of H_2O/Si should be mentioned, [16]. Variations of these factors result in different structures and properties of the sol-gel network within certain ranges.
For the application in catalysis, it is important to end up with highly porous materials with defined active sites given by the metal centers together with large surface areas. Sol-gel synthesis makes these materials accessible in a rather convenient way. In general, porous materials are divided into three groups, according to their pore diameters. Microporous materials only contain pores having less than 2 nm pore diameter. The pore diameters of mesoporous materials range between 2 nm and 50 nm, while macroporous materials only contain pores with diameters larger than 50 nm.
By means of different gel drying the pore sizes can be controlled reliably. Drying by evaporation under normal conditions causes a shrinkage of the gel due to present capillary pressure. Then, the resulting dried gel, a so-called *xerogel*, is often reduced in volume compared to the original wet gel. Xerogels mostly consist of micropores. Drying under supercritical conditions (i.e. there is no interface between liquid and vapor) causes no capillary pressure and therefore the shrinkage is much smaller. As a product here, so-called *aerogels* are obtained, mostly consisting of air. In most cases, aerogels do contain meso- or macropores. Xerogels and aerogels both have high porosities resulting in large surface areas.
Using different metal alkoxides as precursors there is the danger that one species

will lead to higher hydrolysis and condensation rates than the other species. If hydrolysis and condensation rates of used precursors are too different, the formation of domains of a certain metal within the final material might be a direct consequence. To prevent that, one can either work with an excess of solvent in order to decrease the concentration of metal ions [57], or add complexing agents. These chemicals (organic acids, diols, β-diketons etc.) stabilize the metallic species by coordinating the metallic center causing a balancing effect of the hydrolysis and condensation rates. The stabilizing function of the complexing agent is based on the interaction between the donor electrons of its oxygen atoms and the empty orbitals of the metal cation. Thus, the metallic center is less accessible for the nucleophilic attack of H_2O molecules which decreases the hydrolysis rate.
The sol-gel procedure can be acidly, basically or fluoric catalyzed [173]. In combinatorial materials science, the acid-catalyzed sol-gel process has gained the most importance. In contrast to basically catalyzed sol-gel recipes, a larger variety of doping elements and elements in general can be used since many precursors of these elements would precipitate as hydroxides within a basic milieu. Furthermore, the role of the used solvent to prepare the precursor solutions is also important. In many cases, iso-propanol is used because of its low surface tension which prevents the channels from collapsing due to high capillary forces. Within the following figures we sketch the mechanism of the hydrolysis and condensation reaction of the acidly catalyzed sol-gel process for the well-studied case of tetraethoxysilane (TEOS).

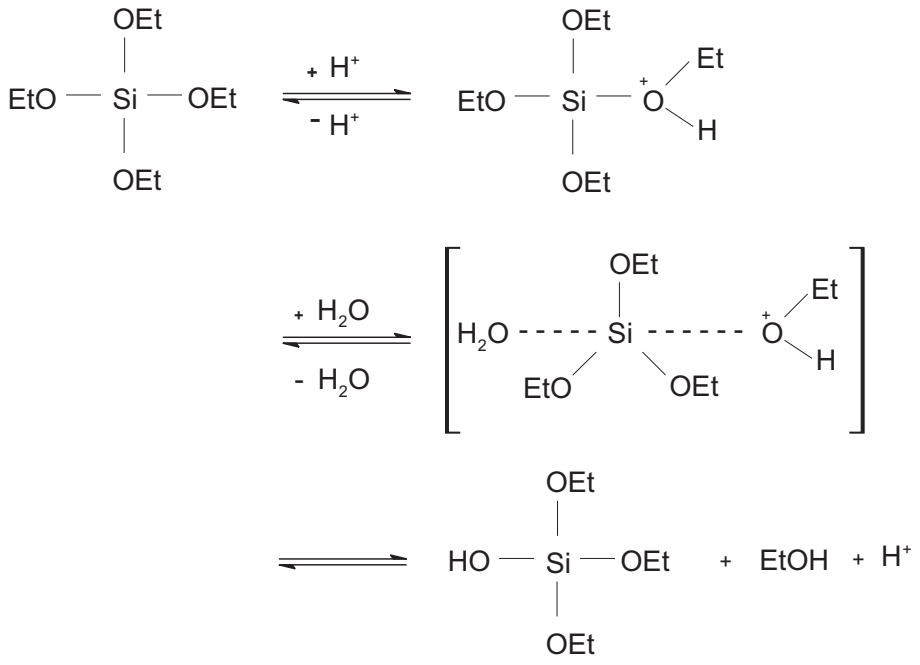


Figure 1.8: Hydrolysis reaction under acidic conditions.

By mixing the precursors (dissolved in a solvent, e.g. an alcohol), the acid and water, the ethoxide groups of the used metal ethoxides are protonated at the oxygen atom. Then a nucleophilic attack by a water molecule can take place at the silicon center with ethanol being the leaving group.

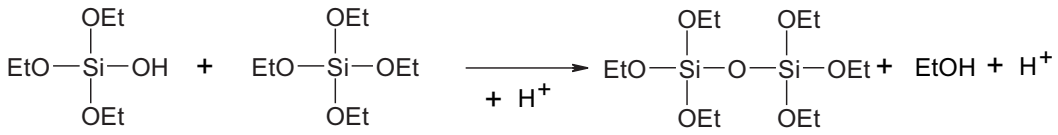


Figure 1.9: Condensation reaction under acid conditions.

Figure 1.9 shows how the hydrolysed metal ethoxides condensate and build chains by splitting of ethanol. Under acidic conditions, the hydrolysis and condensation of terminal ethoxide groups is favored such that in the beginning of the condensation process rather long chains are built up. Since the metal ion does not directly influence the condensation rate, a statistical distribution of metal ions according to their molar ratios within the ethoxide chains on the molecular level can be achieved, assuming that the used precursors possess similar alkoxide groups. Due to the growth of long chains, the network can enclose other metal atoms apart from the precursor ions [101]. In this way, a large variety of doping elements can be added if their precursor solutions do not lead to precipitation. As the chains are getting longer, the probability of crossed connections increases leading to a gel.
In general, under acidic conditions, the hydrolysis rate overtakes the condensation rate resulting in a gel having little branching. On the opposite, working under basic conditions will lead to highly branched gels since the condensation rate being much faster than the hydrolysis steps, cf. [16].
For a detailed insight into sol-gel chemistry and further reading, the reader is referred to the literature, e.g. [53, 105, 115, 123, 179, 182] and especially to the monograph [16]. The application of sol-gel techniques for catalyst preparation has been reported in numerous articles, e. g. in [93] and [103].

1.6 High-Throughput Screening Methods

Once the libraries have been prepared, the materials need to be screened for their desired properties. In catalysis research, the desired property is the ability of a material to catalyze a predefined reaction given by its activity and/or selectivity. Still, the screening of catalytic libraries is a challenging task mainly because of the time dependence of the catalytic performance of the materials. Many catalysts deactivate with time when exposed to reaction conditions by sintering, poisoning or coking but it might also happen, that catalysts need some time before reaching their full performance. So, new catalysts need to be tested for extended reaction times adding another bottleneck to the combinatorial workflow. Miniaturization and massive parallelization of reactor systems coupled with fast screening systems may counteract this bottleneck. Worldwide, several research groups work on these issues, cf. [70, 73, 126, 154].
The appropriate screening tool for each application has to be always carefully chosen and evaluated. In general, optical methods and mass spectrometry are the most common screening techniques used in catalyst screening because of their broad applicability and relative speed. In the field of high-throughput screening methods for porous heterogeneous catalytic materials a throughput of several hundreds per day is the upper limit - in contrast to thin film libraries where these numbers can be

easily topped. High-throughput screening set-ups for heterogeneous catalysts can be characterized by the following issues: catalyst samples are contained in specially shaped libraries and due to their synthesis techniques catalyst samples are often tested at mg-scale sample weight. Furthermore, the screening often yields results being of more qualitative nature (so-called *primary screening*) and the hits need to be verified in conventional experiments being closer to real operating conditions than high-throughput set-ups.
Screening methods can be basically divided into two groups: *sequential* and *parallel methods*.

Parallel Screening Methods

IR-thermography is widely used for the screening of catalytic activities because of its ease of application. This technique has been firstly introduced by Willson and co-workers as a screening tool for heterogeneous catalysts in 1996, [120], using hydrogen oxidation as model reaction. A short time later, Maier and co-workers [73] successfully refined this approach by also taking into account the different emissivities of materials making the tool a more quantitative one. IR-thermography allows the detection of very small differences in temperature, as they occur during catalytic reactions. It has already been applied as a time-resolved screening tool for catalytic reactions, [133]. Further parallel screening methods can be mentioned, namely:

- Laser Induced Fluorescence Imaging (LIFI):
  Detection of Fluorescent Species, [171].
- Resonance-Enhanced Multiphoton Ionization (REMPI):
  In-situ ionization of reaction products by UV lasers, followed by detection of photoions or electrons by spatially addressable microelectrodes, [153].
- Photoacoustics, [83]
- IR-Spetroscopy, [164]
- Color Indication Methods, [150].

Sequential Screening Methods

Numerous publications dealing with sequential screening methods have been published during the last years. The most important approaches to be mentioned here are spatially resolved mass spectrometry, cf. [24, 174, 126] and gas chromatography [186]. In several applications, the time needed to screen one sample lies in the range of 1-2 minutes. Unfortunately, these set-ups cannot be used to study the time dependent behavior of catalyst samples, e.g. deactivation, due to the relatively short reactant-catalyst contact. In most cases, these screening methods have been developed to identify different reaction products of partial oxidation reactions as a function of catalyst composition or formulations.
Identification of the time on-stream performance of a catalytic material can be approached by the use of array microreactors [155]. Here, it is possible to test large numbers of catalysts in parallel and for longer time periods. For a practical use in industrial applications, the lifetime and long-term stability of a catalyst is of essential importance.
For the screening of our catalyst samples, a sequential gas chromatography set-up

has been used. A detailed description of the reactor system and the analytical parts is given within section 7.3 of this thesis. Furthermore, the used screening system has already been described in literature, cf. [87, 126, 174, 185].

1.7 Selective Oxidation of Propene to Acrolein

For the screening of our pentanary composition spreads the oxidation of propene has been chosen as model reaction. Among all possible products that can be obtained by the oxidation of propene we exclusively focused on acrolein as our product of interest. By using GC-analytics also other by-products have been detected during the primary screening. A listing of these by-products can be found within the experimental part of this thesis, cf. section 7.3.

1.7.1 Acrolein

Acrolein (also called propenal, acrylaldehyde), CH_2 =CH-CHO, is a colorless, toxic and volatile substance (melting-point: -88 °C, boiling-point: 52.7 °C). Emission of acrolein into the atmosphere is mainly done by automotive exhaust gases, by pyrolysis of tobacco while smoking, by overheating of fat, e. g. by deep-frying. Acrolein has been discovered in 1843 by Redtenbacher and at the beginning of the 20th century, its economic importance was rather small. Until that time, ist has been synthesized out of glycerine. In 1942, Degussa developed the first technical process using the condensation of formaldehyde and acetaldehyde at 300-320 °C in the gas phase. The yields have been 70-80%. Since the used catalyst showed only short stability of approximately 150 h it had to be regenerated by air and water for 24 h at 500-550 °C. The first catalytic process using propene has been realized by Shell in 1958 with the help of a Cu_2O-catalyst supported on SiC. By the addition of gaseous promotors (hydrogen halides, halogenated hydrocarbons) an acrolein yield of up to 65% could have been achieved. Until today, the production of acrolein is nearly entirely realized by the heterogeneously catalyzed oxidation of propene in the gas phase according to the following reaction, [32]

$$CH_2\text{=CH-}CH_3 + O_2 \rightarrow CH_2\text{=CH-CHO} + H_2O, \qquad \Delta H_R = \text{-369.2 kJ mol}^{-1}$$

1.7.2 Catalysts

The selective oxidation of propene to acrolein in air is done using a heterogeneous mixed metal oxide catalyst, mainly consisting of molybdenum and bismuth, [8, 54, 122]. For catalysts containing Bi and Mo only a small dependence of selectivity from yield has been reported and due to their excellent stability, these catalysts can be used for several years. The discovery of appropriate catalysts for the selective oxidation of propene to acrolein has been realized by putting much effort into the optimization of the original Bi-Mo or Bi-Mo-P system by including further elements like Co, Fe, Ni but also alkaline metals and alkaline earth metals, cf. [32] and the corresponding patents therein. Usually, the catalysts are stabilized by adding silica or alumina as support or framework materials. The composition of the used catalyst can be summarized by the following formula:

$$(Mo_{12}Bi_aFe_b(Co_c + Ni_d)_eP_fM_gO_y)_z(SiO_2)_{(1-z)}$$

with M being additives like Na, K, Cs, Mg, Ca, Cr, Sm, Pt, etc..

These multi-functional mixed metal oxide catalysts are used as extrudate or shell-catalysts in the fixed bed of tube bundle reactors.
Some words about the mechanism of this selective oxidation:
The mechanism of the main reaction has been proven by isotope marking experiments with ^{18}O and pulsing experiments in the absence of oxygen in the gaseous phase to be of Mars-van Krevelen kind. The oxidation of propene to acrolein takes place at two different active centers of the catalyst. At the first center, propene is oxidized by the lattice oxygen of the catalyst. By delivering this oxygen the catalyst is reduced. The oxygen from the reaction atmosphere is adsorbed at the second active site followed by spilling over to vacant anionic positions where it is integrated into the lattice by re-oxidation of the catalyst. An efficient catalysis can only take place, if the transport of electrons but also of the oxygen species between the catalytic centers is quick enough. Here, the electronic structure of the catalyst is decisive. By investigations with IR- & Raman-spectroscopy, several types of lattice oxygen species could have been discovered for different Bi-molybdates. By adding redox pairs like Cr^{2+}/Cr^{3+}, Fe^{2+}/Fe^{3+} or Ce^{3+}/Ce^{4+} the mobility of the lattice or the reoxidation of the lattice can be accelerated, respectively. Here, the rate-determining step seems to depend on the catalyst system. On the one hand, pulsing experiments have shown that taking up the oxygen into the lattice for the re-oxidation is the rate-determining factor here. On the other hand, in the work of Glaeser et al. [51], the diffusion of oxygen on the catalyst surface is thought to be the slowest part of the re-oxidation process. Using kinetic modelling of the reaction in a catalytic wall reactor, the dependence of the rate-determining step from the reaction temperature could have been shown, cf. [131, 132]. At a reaction temperature below 370 °C, the reoxidation of the catalyst is the rate-determining step. Here, the presence of water supports the reoxidation process and reduces total oxidation. At temperatures above 370 °C, the reduction of the catalyst by propene is rate-determining. There are still discussions going on concerning this mechanism that can be followed in the literature, e.g. [189].
Now as before, the economical importance of the heterogeneously catalyzed selective oxidation of propene can be observed in many research activities spread over the whole world. The world's leading chemical companies like BASF, Degussa, Nippon Shokubai Inc., Nippon Kayaku Inc. and Standard Oil (Sohio) but also research laboratories intensively deal with that topic.

1.7.3 Industrial Processing

For the industrial realization of the propene oxidation as described above, the feed gas consists of propene (5-10 Vol.-%), air (40-64 Vol.-%) and inert additives also including water. The inert additives are needed to keep the reaction gas mixture out of the explosive range. With a gas hourly space velocity (GHSV) of 1300 - 2600 h^{-1} , reaction temperatures of 300-450 °C and a pressure of 1-3 bar the yields lie between 93-98% with an acrolein selectivity of 90% (w.r.t. propene). As by-products, mainly 4-10% acrylic acid and 2-6% CO_2 are obtained. Furthermore, acetaldehyde, formaldehyde and acetic acid are obtained in minor amounts. At the reactor exit, the product is quenched to prevent further reacting. Being an exothermic process, the resulting energy of -369.2 kJ mol^{-1} can be used to heat up

the feed gas, as it is already done in modern plants.
To get the acrolein out of the process, the reaction gas needs to be further treated. In a first step, acrylic acid and other high boiling chemicals are washed out. The remaining gas mixture is then led through an absorber to extract acrolein in cold water. Here, parts of the gas mixture (e.g. propene, CO_2, O_2, N_2) are recycled back into the reactor. Following the absorber, a desorption step is placed to get acrolein out of the sorbate. In a next step, low boiling substances (e.g. acetaldehyde) are removed by distillation such that acrolein is obtained with a purity higher than 95%. In all steps, it is important to prevent acrolein from polymerisation. This is done by adding an inhibitor during the processing and storage, e.g. hydroquinone, [32].

1.7.4 Recent Developments

The selective oxidation of propene has been intensively studied within the last years, [12, 49, 107, 168, 197]. Research focuses on the development of new and improved catalysts and other reactor concepts. In one approach, propane is dehydrated to give propene in a first step, then it is oxidized to acrolein. Economically more interesting is the realization of this reaction in one step by using a bifunctional catalyst that catalyzes the dehydration of propane to propene and oxidation of propene. At the moment, the selectivity of this process to acrolein is still too small for commercialization. To increase the selectivity, membrane reactors have been used successfully at laboratory scale. Apart from the use of classical tube bundle reactors, also catalytic wall reactors are applied. The advantages here lie in an improved heat control of the process, i. e. an isothermal reaction procedure within the catalyst bed can be obtained. This prevents the appearance of so-called *hot-spots*, that may lead to undesired side reactions. That is one reason why the development of a riser-reactor working with a fluidized bed has been realized. Another research field to obtain better catalysts for this reaction is the application of aerosol-synthesis approaches, e.g. high temperature aerosol decomposition, HTAD. By this technique, nanoscaled Bi-Mo catalysts could have been generated having high crystallinity. Another synthesis route being tried is the partial oxidation of ethane by oxygen to give acrolein and acetaldehyde. By aldol condensation of formaldehyde, acrolein can be obtained. The corresponding literature belonging to all processes described above can be found within the references of [32].

1.7.5 Application of Acrolein

Possessing an aldehyde group and a conjugated double bond system, acrolein can exhibit numerous reactions. The most important chemicals that are produced out of acrolein are:

- Allylacohol being further processed to glycerine
- Methionine, an essential amino acid ($> 0.5 \cdot 10^6$ t per year, Degussa, Novus, Rhodia)
- acrylic acid and acrylates (approx. $3.5 \cdot 10^6$ t per year, BASF, Röhm & Haas, Union Carbide)
- Pyridine, 3-Picoline (reaction with NH_3 in the gaseous phase, Daicel)

- 1,3-Propandiol (Adisseo, Degussa)

In industrial applications, the reaction of acrolein to acrylic acid and acrylates is the most important, realized by approx. $3.5 \cdot 10^6$ t per year. Furthermore, acrolein is used to be transformed to the essential amino acid methionine ($> 0.5 \cdot 10^6$ t production on the world market). The worldwide demand for acrylic acid and methionine is steadily growing such that the need for acrolein will also increase. 1,3-Propandiol is essentially used in drug industry and in the production of polyester fibers. For more information on acrolein and its industrial application we refer to the literature, e.g. [16, 187].

1.8 Scopes of this Thesis

The scopes of this thesis are of different kinds and cover experimental but also theoretical aspects. The following questions can be considered to be of essential interest:

1. Is it possible to construct quantitative composition activity relationships (QCARs) for heterogeneous catalysts based on primary high-throughput screening data?
2. Does the chosen synthesis route allow the identical preparation of two complete pentanary composition spreads (1001 samples each)?
3. Can a new sol-gel recipe be found to include Te as an additional element in combination with Cr, Mn, Co and Ni?
4. How reliable does the chosen high-throughput synthesis work?
5. Where lie the problems with high-throughput screening data?
6. How good is the reliability of the screening set-up?
7. How good is the data quality? What about false positives/ false negatives?
8. Which mathematical approaches/techniques can be used to model these relationships?
9. Is it possible to estimate activities of heterogeneous catalysts for a predefined model reaction using Kriging and B-Splines?
10. After prediction has been done, is it possible to validate the results in practice (by synthesizing and testing the calculated compositions)?
11. What are the main differences between the two mathematical approaches?

1.9 Outline

For the convenience of the reader, the thesis starts with an introductional part. The first chapter is dedicated to combinatorial chemistry in general and its application to heterogeneous catalysis. The relatively new research area “combinatorial chemistry” is introduced and an overview on some historical aspects but also the state of the art in combinatorial heterogeneous catalysis is given. Furthermore, the most

important high-throughput preparation and screening techniques for solid catalyst materials are addressed. Since all our materials considered here have been prepared by a sol-gel method, cf. pp. 130, also a short introduction into this preparation technique is given.

The second chapter is dedicated to Kriging, an interpolation technique known from geostatistics. Chapter 2 gives a detailed introduction into Kriging with main focus on setting up the Kriging system to solve a given estimation problem. At the end of this chapter we discuss how the given high-throughput screening data are transformed into appropriate coordinate representations such that Kriging and the B-Spline approach can be applied. It is also described, how composition spreads can be elegantly represented in a mathematical way and how the composition of a catalyst corresponds to a point in a three- up to five-dimensional search space.

Chapter 3 covers a short introduction into the theoretical aspects of the chosen B-Spline approach.

In the second part of this thesis describes the obtained results.
Chapter 4 summarizes the most successful screening results of the catalyst libraries and shows, if regions of interest found by a primary screening can be reproduced in a second generation of catalysts. Additionally, the reliability of the synthesis route and the high-throughput screening set-up are intensively considered and discussed.

Dealing with high-dimensional data sets, there is always the question of an elegant and effective visualization. Chapter 5 is therefore dedicated to different visualization approaches to visualize our raw screening data having more than three dimensions. But also interpolated and therefore "continuous data" are visualized to indicate regions of interest found by the primary screening. Chapter 5 presents some visualization approaches known in other scientific fields with a first application to heterogeneous catalyst screening data.

In Chapter 6 the numerical results are discussed and the estimated activities obtained by the Kriging and the B-Spline Model can be compared to those of re-synthesized samples. Furthermore, the best catalyst compositions for the selective propene oxidation are mentioned according to each model.

The final part of this thesis includes a detailed description of all experimental work that has been done to synthesize and screen the whole set of catalysts.
Finally, Chapter 8 summarizes the most relevant results. Moreover, we sketch aspects that appear to be of great relevance for further research.

2 Kriging

2.1 Introduction

This chapter introduces an interpolation technique that is known as *Kriging*. Belonging to the field of *Geostatistics*, Kriging is used to approximate or interpolate spatial data, i.e. to reconstruct phenomena over a domain based on observations taken at a limited number of points (e.g. meteorological data, geological data, etc.). Geostatistics itself deals with statistics applied to geology or perhaps more generally to problems in the earth sciences.
Within the late 40's of the last century the gold prices drastically decreased due to increasing amounts of Russian gold brought to the world market leading the South-African gold mines into huge difficulties. Daniel G. Krige, a young mining engineer of the South-African Witwatersrand mines thought of developing a possibility to convey more gold with less operating expense. To know about the gold content of ore and therefore to judge about the chance to find gold within a certain area needs experience and many test drills. Krige had the following idea: If there is the possibility to draw as much information out of as few drills as possible this would certainly help to increase efficiency. Today, the term Geostatistics includes many methods and approaches originally developed by Krige: tools to gain knowledge about surfaces or even volumes out of pointwise measurements.
Starting in the mid-60's and especially in the mid-70's Geostatistics became much more closely affiliated with the work of George Matheron and this connection might be still the prevailing one today. In his early works "Traité de géostatistique appliquée" [109] and "La théorie des variables régionalisées, et ses applications" [111], he formalized the ideas of Krige. It took several years until the impact of his works has been accepted within the geoscience community since the fundamental problem of deducing information on surfaces out of discrete, pointwise data not only occurs in geology but also in other fields of science.
With increasing calculating capacities and widely-used applications of geographic information systems (GIS) the interest in Geostatistics spread also to other scientific communities.
In the remaining parts of this thesis a compact introduction to Kriging is given together with an illustration how this approach has been applied to high-throughput data in heterogeneous catalysis. With the help of Kriging, the activities of a priori non-tested catalyst samples have been estimated such that these results could be compared to those of synthesized compositions afterwards.

Applications
Geostatistics is very much an applied discipline and its development has been the work of mining engineers, petroleum engineers, hydrologists, soil scientists, geologists as well as statisticians. There are applications in epidemiology, plant pathology or entomology as well as forestry, atmospheric sciences, global change and geogra-

phy. Apparent is the overlap with GIS and spatial statistics in general.

Problems and Objectives
In one respect Geostatistics might be viewed as simply a methodology for interpolating data on an irregular grid but this seems to be too simplistic. A number of interpolation methods and algorithms were already well known when Geostatistics began to become popular: Inverse Distance Weighting and Trend Surface Analysis as well as the Nearest Neighbor Algorithm.
First of all, Geostatistics is concerned with **spatial** data. That means that each data point is associated with a location in space and there is at least an implied connection between the location and the data value. Here, location can be thought of to have two meanings: one is simply a point in space (which only exists in an abstract mathematical sense) and secondly with an area or volume in space. For example, let $x, y, ..., w$ be points (not just coordinates) in 1, 2, or 3 dimensional space and $Z(x), Z(y), ...$ denote observed values at these locations. These might be the grade of copper, temperature, concentration of a pollutant etc.. Now, let t be a location that is not "sampled". The objective then is to estimate or predict the value $Z(t)$. If only this information is given then the problem is *ill-posed*, i.e., it does not have a unique solution. One way to obtain a unique solution to this problem is to introduce a model. There are two ways to do this: one is deterministic and the second is stochastic or statistical. Both approaches must somehow incorporate the idea that there is uncertainty associated with the estimation or prediction step. The value at the non-sampled location is not itself random but the knowledge of it is uncertain. One approach then is to treat $Z(x), Z(y), ..., Z(t)$ as being the values of random variables. If the joint distribution of these random variables were known then the "best" estimator (with best meaning unbiased and having minimal variance of the error of estimation) would be the conditional expectation of $Z(t)$ given the values of the other random variables. However, the data consists of only one observation of the random variables $Z(x), Z(y), ...$ and none of the random variable $Z(t)$, hence it is not possible to estimate or model this distribution using standard ways of modelling or fitting probability distributions.

Kriging
The interpolation technique *Kriging* has been named by Matheron [109, 111] after Daniel G. Krige. It is a linear estimation technique that yields the **best linear unbiased estimator (BLUE)** of a non-observed value at an non-sampled location t. During following sections it will become clear how the Kriging approach tackles the handicap of ill-posed problems mentioned above.

2.2 Preliminaries

In this section, the elementary concepts and definitions of mean, variance, covariance, random functions, regionalized variables, expectation etc. are presented (following [23] and [180]) that serve as fundamental prerequisites for introducing Kriging. For a more detailed introduction to Kriging the reader is referred to some standard text books and monographs, for example [180, 23, 79, 27].

2.2.1 Definitions

Notation
Throughout this thesis a point in n-dimensional space is denoted by x. This means, in 3D x stands for the coordinates (x_1, x_2, x_3). In most cases, no explicit notation for the coordinates of a point will be needed so that, except when stated otherwise, $x_1, x_2, ...$ will stand for distinct points in $\mathbb{R}^n$.

Random Function
Let $\mathcal{D} \subset \mathbb{R}^n$ be a domain with a positive volume and $(\Omega, \mathcal{A}, P)$ be a probability space. A *random function* (RF) is a function of two variables $Z(x, \omega)$ such that for each $x \in \mathcal{D}$ the section $Z(x, \cdot)$ is a random variable on $(\Omega, \mathcal{A}, P)$. Each function $Z(\cdot, \omega)$, defined on $\mathcal{D}$ as the section of the RF at $\omega \in \Omega$, is a realization of the RF. Shortly, the RF is denoted by $Z(x)$ and a realization by the lowercase $z(x)$. A random function is also called a *stochastic process* when x varies in a 1D space (e.g. x is interpreted as time) and a *random field* when x varies in a space of more than one dimension.

Regionalized Variable
A variable that denotes the spatial distribution of a factor is called a *regionalized variable* or *spatially dependent*. It is denoted by $z(x)$ with x giving the coordinates of the location (in most applications $x \in \mathbb{R}^3$). Again, $z(x)$ is considered to be a realization of a parent RF $Z(x)$. Matheron [110] coined the term *regionalized variable* to designate a numerical function $z(x)$ depending on a location x, and combining large irregularities of detail with spatial correlation. Figure 2.1 illustrates this connection: the regionalized variable $z(x)$ is one realization of the random function $Z(x)$. A regionalized value $z(x_0)$ at a specific location x_0 is a realization of a random variable $Z(x_0)$ which is itself a member of an infinite family of random variables, $\{Z(x) : x \in \mathcal{D}\}$, the random function $Z(x)$. Here, the point x_0 is an arbitrary point of region $\mathcal{D}$ which may or may not have been sampled.

Random Function	$Z(x)$	$\longrightarrow$	$Z(x_0)$	The random variable's
$\Downarrow$	$\downarrow$		$\downarrow$	realization is a
Regionalized Variable	$z(x)$	$\longrightarrow$	$z(x_0)$	regionalized value

Figure 2.1: Regionalized variables as a realization of a random function [180].

Distribution Function
Let z be a value randomly drawn from a set of values Z. Again, each value z can be considered as a realization of a random variable Z. For a random variable Z that takes values in $\mathbb{R}$, the *probability distribution function* $F(z)$ is defined as

$$F(z) = P(Z < z) \qquad \text{with} \quad -\infty < z < \infty \tag{2.1}$$

The distribution function indicates the probability P that a value of the random variable Z is below a given value z. If z is partitioned into intervals of infinitesimal

length dz, the probability that a realization of Z belongs to such an interval is $F(dz)$. If the distribution function is differentiable, the *density function* $p(z)$ can be defined as

$$F(dz) = p(z)\,dz, \tag{2.2}$$

i.e. the density function $p(z)$ is given by the derivative of the distribution function.

Mean, Expectation
Let $z(x_1), ..., z(x_n)$ be different realizations of the random variable $Z(x)$. Then the *mean value* $\bar{z}$ of $z(x_1), ..., z(x_n)$ is given by

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z(x_i). \tag{2.3}$$

The idealization of the concept of the mean value is the *mathematical expectation* or *expected value*. The expected value of Z, also called *first moment* of the random variable, is defined as the integral over the realizations z of Z weighted by the density function

$$E[Z] = \int_{z\in\mathbb{R}} zp(z)\,dz = m \tag{2.4}$$

Since the expectation is a linear operator, it holds by definition (2.4):

$$E[a] = a, \tag{2.5}$$
$$E[bZ] = bE[Z] \tag{2.6}$$
$$E[a + bZ] = a + bE[Z] \tag{2.7}$$

for constants $a, b \in \mathbb{R}$.
The second moment of the random variable is the expectation of its squared value

$$E[Z^2] = \int_{z\in\mathbb{R}} z^2 p(z)\,dz \tag{2.8}$$

and the k-th moment is defined as the expected value of the k-th power of Z

$$E[Z^k] = \int_{z\in\mathbb{R}} z^k p(z)\,dz. \tag{2.9}$$

When Z is a discrete random variable with states z_i, the integral in the definition (2.4) of the mathematical expectation is replaced by a sum

$$E[Z] = \sum_i z_i p_i = m, \tag{2.10}$$

where $p_i = P(Z = z_i)$ is the probability that Z takes the value z_i.

Variance
The variance σ^2 of the random variable Z, called the theoretical variance, is defined as

$$\text{var}(Z) = E\left[(Z - E[Z])^2\right] = E[(Z - m)^2] = \sigma^2. \tag{2.11}$$

After expanding the binomial expression it yields

$$\text{var}(Z) = E[Z^2 + m^2 - 2mZ] \tag{2.12}$$

and by the linearity of the expectation,

$$\text{var}(Z) = E[Z^2] - (E[Z])^2. \tag{2.13}$$

The variance can thus be expressed as the difference between the second moment and the squared first moment.

Sample Variance
The sample variance of a finite number of values always has a finite expectation and it characterizes the dispersion of the values around their common mean. Let $z(x_1), ..., z(x_n)$ be a particular realization of the random function $Z(x)$. If the n values are known the sample variance can be calculated using the mean value $\bar{z}$:

$$s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(z(x_i) - \bar{z})^2 \qquad \text{with} \quad \bar{z} = \frac{1}{n}\sum_{i=1}^{n} z(x_i) \tag{2.14}$$

Covariance
The theoretical covariance σ_{ij} between two random variables Z_i and Z_j is defined as

$$\begin{aligned} \text{cov}(Z_i, Z_j) &= E\left[(Z_i - E[Z_i]) \cdot (Z_j - E[Z_j])\right] \\ &= E\left[(Z_i - m_i) \cdot (Z_j - m_j)\right] = \sigma_{ij} \end{aligned} \tag{2.15}$$

where m_i and m_j are the means of the two random variables. The covariance of Z_i with itself is the variance of Z_i,

$$\sigma_{ii} = E[(Z_i - m_i)^2] = \sigma_i^2 \tag{2.16}$$

The covariance divided by the square root of the variances is called the theoretical *correlation coefficient*

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2 \sigma_j^2}} \tag{2.17}$$

Two random variables are said to be *uncorrelated* if their covariance is zero.

Experimental Covariance
Let $Z_1(x)$ and $Z_2(x)$ be two random variables having n realizations $z_1(x_1), ..., z_1(x_n)$ and $z_2(x_1), ..., z_2(x_n)$, respectively. Furthermore, let $\bar{z}_1$ and $\bar{z}_2$ be the corresponding mean values. Then the *experimental covariance* is defined by

$$s_{12} = \frac{1}{n}\sum_{i=1}^{n}(z_1(x_i) - \bar{z}_1)(z_2(x_i) - \bar{z}_2). \tag{2.18}$$

For two random variables, the experimental covariance gives the dispersion of the data cloud around its center of mass (defined by $(\bar{z}_1,\bar{z}_2)$). The difference between a value of one variable and the mean value is called a *residual*. The covariance of $Z_1(x)$ and $Z_2(x)$ is positive, if the residuals of $Z_1(x)$ tend to have the same sign than the residuals of $Z_2(x)$, on average. If the residuals show opposite signs on average, the covariance is negative. If $Z_1(x)$ is identical with $Z_2(x)$, the covariance is equal to the experimental variance.
In many applications, the covariances of pairs of variables need to be compared. Sometimes, the units of the variables can not be easily compared because they are of different type, e.g. cm, %, kg, °C, etc.. Then it is necessary to standardize each variable $Z(x)$ by subtraction of its mean value and dividing by its *standard deviation* s, which is the square root of the sample variance s^2.

$$\tilde{z}_i(x) = \frac{z_i(x) - \bar{z}}{s} \tag{2.19}$$

The standardized variable has a variance that is equal to 1. The covariance of two standardized variables $\tilde{z}_1$ and $\tilde{z}_2$ is called the experimental *correlation coefficient*, denoted by r_{12}. In general, it is bounded by

$$-1 \leq r_{12} \leq 1 \tag{2.20}$$

Variance-Covariance Matrix
Let $Z_1(x), ..., Z_N(x)$ be N different random variables with n realizations $z_{1i}(x), ..., z_{ni}(x)$, $i = 1, ..., n$ each. These realizations can be arranged in a matrix Z:

$$Z = \begin{pmatrix} z_{11} & \dots & z_{1j} & \dots & z_{1N} \\ \vdots & & \vdots & & \vdots \\ z_{i1} & \dots & z_{ij} & \dots & z_{iN} \\ \vdots & & \vdots & & \vdots \\ z_{n1} & \dots & z_{nj} & \dots & z_{nN} \end{pmatrix} \tag{2.21}$$

Let further M be a $n \times N$ matrix having the repeated mean values of each variable in each column:

$$M = \begin{pmatrix} m_1 & \dots & m_j & \dots & m_N \\ \vdots & & \vdots & & \vdots \\ m_1 & \dots & m_j & \dots & m_N \\ \vdots & & \vdots & & \vdots \\ m_1 & \dots & m_j & \dots & m_N \end{pmatrix} \tag{2.22}$$

A matrix Z_c of centered variables is obtained out of Z by subtracting M from the raw data matrix

$$Z_c = Z - M \tag{2.23}$$

The *variance-covariance matrix* is the then defined by

$$\begin{aligned} V &= \frac{1}{n} Z_c^\top Z_c \\ &= \begin{pmatrix} \operatorname{var}(Z_1) & \dots & \operatorname{cov}(Z_1, Z_j) & \dots & \operatorname{cov}(Z_1, Z_N) \\ \vdots & & \vdots & & \vdots \\ \operatorname{cov}(Z_i, Z_1) & \dots & \operatorname{var}(Z_i) & \dots & \operatorname{cov}(Z_i, Z_N) \\ \vdots & & \vdots & & \vdots \\ \operatorname{cov}(Z_N, Z_1) & \dots & \operatorname{cov}(Z_N, Z_j) & \dots & \operatorname{var}(Z_N) \end{pmatrix} \\ &= \begin{pmatrix} s_{11} & \dots & s_{1j} & \dots & s_{1N} \\ \vdots & & \vdots & & \vdots \\ s_{i1} & \dots & s_{ii} & \dots & s_{iN} \\ \vdots & & \vdots & & \vdots \\ s_{N1} & \dots & s_{Nj} & \dots & s_{NN} \end{pmatrix} \end{aligned} \tag{2.24}$$

The variance-covariance matrix is always symmetric and nonnegative definite. The nonnegative definiteness of the covariance-variance matrix plays an important role in theory and also has practical implications. For the exact definition together with three criteria of nonnegative definiteness the appendix section in [180] is referred to.

Probability Distributions

According to Figure 2.1, the random mechanism $Z(x_0)$ acting at some given point x_0 of a region generates realizations that follow a probability distribution F

$$P(Z(x_0) \leq z) = F_{x_0}(z). \tag{2.25}$$

Here, P stands for the probability that an outcome of Z at the point x_0 is lower than a fixed value z.

Let $Z(x_1)$ and $Z(x_2)$ be two random variables at two different locations x_1 and x_2. A *bivariate distribution function* for these two random variables is

$$P(Z(x_1) \leq z_1, Z(x_2) \leq z_2) = F_{x_1,x_2}(z_1, z_2), \tag{2.26}$$

where P is the probability that simultaneously an outcome of $Z(x_1)$ is lower than z_1 and an outcome of $Z(x_2)$ is lower than z_2.

Analogously, multiple distribution functions for n random variables located at n different points can be defined:

$$F_{x_1,\dots,x_n}(z_1, ..., z_n) = P(Z(x_1) \leq z_1, ..., Z(x_n) \leq z_n) \tag{2.27}$$

This approach of defining multiple distribution functions provides a very general model to describe processes in nature or technology. Often in practice, only few data from one or several realizations of the random function can be obtained and it is impossible to deduce all the mono- and multivariate distribution functions for a certain set of points. This means, that some kind of simplification is needed here which is provided by the idea of stationarity.

Strict Stationarity
In general, stationarity means that characteristics of a random function stay the same when shifting a given set of n points from one part of the region to another. This phenomenon is called *translation invariance*.
The definition of *strict stationarity* is the following: A random function is called *strictly stationary* if for any set of n points $x_1, ..., x_n$ (where n is an arbitrary positive integer) and for any vector h

$$F_{x_1,...,x_n}(z_1, ..., z_n) = F_{x_1+h,...,x_n+h}(z_1, ..., z_n). \tag{2.28}$$

In words this means that a translation of a point configuration in a predefined direction does not cause any changes to the multiple distribution or that a phenomenon is homogeneous in space and repeats itself in the whole space.
There is a wide range of several degrees and types of stationarity ranging from non-stationary random functions whose characteristics change at any time and location to the concept of a strictly stationary function whose distribution functions are the same always and everywhere. A weaker type of stationarity is defined by the so-called *second-order* or *weak stationarity*.

Second-Order Stationarity
For a strictly stationary random function holds, that its moments, if they exist, are also invariant under translations. If only stationarity of the first two moments is considered, this is called *second-order stationarity* or *weak stationarity*. Then, for points x and $x + h$ of $\mathcal{D} \subset \mathbb{R}^n$ this means:

$$\begin{cases} E[Z(x)] = m & \text{for all} \quad x \in \mathcal{D} \\ E[Z(x) - m][Z(x + h) - m] =: C(h) & \text{for all} \quad x, x + h \in \mathcal{D} \end{cases} \tag{2.29}$$

where m is the mean (constant) and $C(h)$ is called the *covariance function* which will be discussed more closely in the following sections. $C(h)$ depends only on the *separation* h, also called *lag* h. A second-order (stationary) random function (SRF) is called *isotropic* if its covariance function only depends on the length $|h|$ of vector h and not on its orientation. The opposite behavior is called *anisotropic*.
As strict stationarity requires the specification of the multipoint distribution (2.28) for any set of points $x_1, ..., x_n$ a relaxation of this approach can be to consider only pairs of points x_1, x_2 in the domain $\mathcal{D}$ and try to deduce only the first two moments, not the full distribution. For the case of the Gaussian distribution, this strategy is ideal since the first two moments entirely characterize the distribution. This relaxation yields quite good results in practice as long as the data histogram does not show too heavy tails. Assuming the stationarity of the first two moments of the variable is one possibility. Furthermore, there is the possibility to assume

the stationarity of the first two moments of the difference of a pair of values at two points, which is called *intrinsic stationarity* and leads to the notion of the variogram.

2.3 The Variogram

To solve interpolation or filtering problems the theory of stochastic processes and random function has been applied for a relatively long time. The proposed methods are mainly based on the first two moments of the random functions. For real world applications, these moments are never known before and there must be a way to determine them. One strength of geostatistics lies in the fact that it proposes a methodology to identify the features of the stochastic model but also a methodology for the interpolation itself. Spatial phenomena are in general unique, non-reproducible, often defined in a two- or three-dimensional domain, obtained from unevenly distributed sample points and, in most cases, too complex for a precise deterministic description. These phenomena are regionalized variables $\{z(x) : x \in \mathcal{D} \subset \mathbb{R}^n\}$ that are regarded as realizations of random functions. A random function (RF) $\{Z(x) : x \in \mathbb{R}^n\}$ is characterized by its finite-dimensional distribution (also called here *spatial distribution*) which is the set of all multidimensional distributions of k-tuples $(Z(x_1), Z(x_2), ..., Z(x_k))$ for all values of k and all configurations of the points $x_1, x_2, ..., x_k$. In practice, it is rather impossible to calculate sample multidimensional distributions for more complex k-tuples, even if a very large number of realizations of a random function is available. The common case is that one has only one single realization and therefore the distributions cannot be determined, except under an assumption of stationarity which describes some sort of repetition in space. With the help of stationarity two point configurations can be considered that are identical up to a translation to be statistically equivalent. Assuming an uneven distribution of sample points, the only (approximately) identical configurations that can be found are pairs of sample points. To study these phenomena, the main tool will be the *variogram* showing the behavior of the dissimilarity between $Z(x)$ and $Z(x+h)$ depending on distance h. The following section distinguishes among three main definitions of the variogram, namely

- Theoretical Variogram
- Sample (Experimental) Variogram
- Regional Variogram

2.3.1 Variogram and Covariance Function

Dissimilarity versus Separation
The variability of a regionalized variable $z(x)$ at different scales can be measured by computing the dissimilarity between pairs of data values z_1 and z_2 located at points x_1 and x_2 in a spatial domain $\mathcal{D}$. Let $\gamma^\star$ be the measure for dissimilarity of two values defined by

$$\gamma^\star_{12} = \frac{(z_1 - z_2)^2}{2} \tag{2.30}$$

i.e. half of the squared difference between the two values.
The two points x_1, x_2 in space can be linked by a vector $h = x_2 - x_1$, as shown in Figure 2.2.

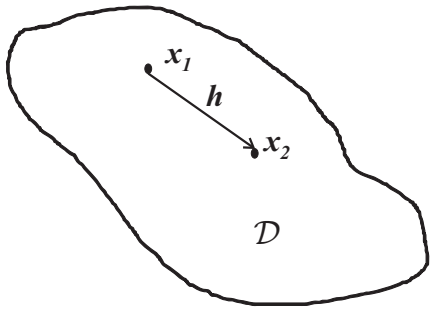


Figure 2.2: Vector h linking two points in 2D space.

Furthermore, the dissimilarity $\gamma^\star$ depends from the distance and the orientation of the point pair described by vector h, i.e.

$$\gamma^\star(h) = \frac{1}{2}\left(z(x_1 + h) - z(x_1)\right)^2 . \tag{2.31}$$

Using all sample pairs in a data set (usually up to a distance of half the diameter of the region) and plotting the dissimilarities $\gamma^\star$ against the spatial separation h (the absolute values of vector h) leads to a so-called *variogram cloud*. Figure 2.3 illustrates a schematic example. The dissimilarity often increases with distance as near samples tend to be more alike than more distant samples.

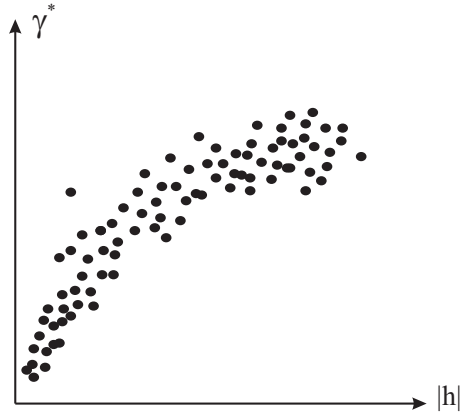


Figure 2.3: Plotting the dissimilarities $\gamma^\star$ against the spatial separation h results in a variogram cloud.

As the dissimilarity is a squared quantity, the sign of vector h, i.e. the orientation in which the points x_1 and x_2 are considered, does not effect the results. The dissimilarity is symmetric with respect to h:

$$\gamma^\star(-h) = \gamma^\star(+h) \tag{2.32}$$

For the exploration of features of spatial data, the variogram cloud provides an excellent tool. Values of the variogram cloud can, for example, be linked to the position of the sample pairs on a map representation to study their spatial behavior. Anomalies, inhomogeneities can be detected by having a closer look at high dissimilarities at short distances. Haslett et al. [64] were the first to introduce the use of the variogram cloud in combination with other views of the data together with many examples. The variogram cloud can show different behaviors along different directions of the separation $h = x_2 - x_1$ in the case of anisotropy. To detect anisotropic behavior, the variogram cloud needs to be calculated for several directions. In 2D, at least four equally spaced directions are usually considered (e.g. the coordinate axes and the diagonals). In 3D, the use of regular polyhedra yields equally spaced directions.

Experimental Variogram
Considering a variogram cloud it is possible to extract the following information:

- *The sample variogram.* This is the curve giving the mean of the halved squared differences as a function of distance. For practical applications, the sample variogram is obtained from the variogram cloud by subdividing it into classes of separation and computing an average for each class.
- *Any other characteristic of the cloud* which is again calculated by classes of distance (median, quartiles, etc..)
- *Box Plots.* They present several characteristics of the variogram cloud in a single figure, usually the mean, the median and some other quantiles of the halved squared distances for each class of separation.

The sample variogram is calculated and displayed by classes of direction since, like the variogram cloud, it can also be anisotropic. The sample variogram is defined by:

$$\hat{\gamma}(h) = \frac{1}{2N_h} \sum_{x_1 - x_2 \approx h} [z(x_2) - z(x_1)]^2 \tag{2.33}$$

where N_h denotes the count of pairs of points that are separated (approximately) by the lag h. It is calculated for discrete values of h and in practice, the length of vector h is often kept inferior to half the diameter of the region. For pairs of samples with vectors h of a length almost equal to the diameter of the region, the corresponding samples are located near the border. Vector classes formed with such pairs will have no contribution from samples at the center of the region and thus are not representative of the whole data set. Figure 2.4 gives an illustration of that situation.
Usually one can observe that the average dissimilarity between values increases when the spacing between the pairs of sample points is increased. For larger spacings the variogram sometimes reaches a sill which can be equal to the variance of the data. The definition of the sample variogram can be generalized to more than two dimensions (sometimes called *variogram map*). By the sample variogram, a decomposition of the sample variance into distances is achieved. The average of all terms $\hat{\gamma}(h)$ for all possible lags including the lag $h = 0$, weighted by N_h, is the mean of $\frac{1}{2}(z(x_2) - z(x_1))^2$

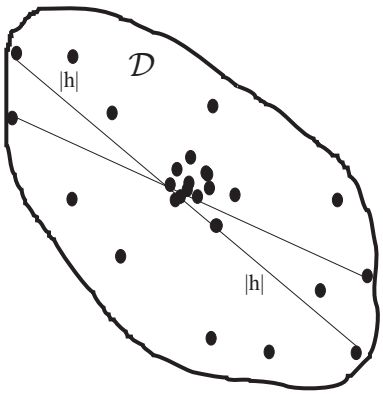


Figure 2.4: Example of pairs of points lying more than half of the diameter of set $\mathcal{D}$ apart.

for all pairs (x_1, x_2) including those for which $x_1 = x_2$, and it coincides with the sample variance of the data. The sample variogram provides us with an efficient tool for the analysis of spatial data based on dissimilarity.

Regional Variogram
The experimental variogram of samples $z(x_i)$ is the sequence of dissimilarities for different distance classes. Assuming that there are samples given for the whole domain $\mathcal{D}$ then the variogram can be computed for every possible pair in the domain. Let $\mathcal{D}(h)$ be defined as the intersection of the domain $\mathcal{D}$ with a translation $\mathcal{D}_{-h}$ of itself. Then $\mathcal{D}(h)$ describes all points x having a counterpart $x + h$ in $\mathcal{D}$.
The *regional variogram* $g_r(h)$ is the integral over the squared differences of a regionalized variable $z(x)$ for a given lag h:

$$g_R(h) = \frac{1}{2|\mathcal{D}(h)|} \int\limits_{\mathcal{D}(h)} (z(x+h) - z(x))^2 \, dx \tag{2.34}$$

where $|\mathcal{D}(h)|$ is the area (or volume) of the intersection $\mathcal{D}(h)$.
In practice, the regionalized variable $z(x)$ might only be known at a few locations and generally it is not possible to approximate $z(x)$ by a simple deterministic function. Thus, $z(x)$ is considered to be a realization of a random function $Z(x)$. The associated regional variogram

$$G_R(h) = \frac{1}{2|\mathcal{D}(h)|} \int\limits_{\mathcal{D}(h)} (Z(x+h) - Z(x))^2 \, dx \tag{2.35}$$

is a randomized version of $g_r(h)$ and its expectation defines the theoretical variogram $\gamma(h)$ of the random function model $Z(x)$ over the domain $\mathcal{D}$.

$$\gamma(h) = E\left[G_R(h)\right] \tag{2.36}$$

Theoretical Variogram
To express the variation of a random function $Z(x)$ in space, the differences between values at pairs of points x and $x+h$ can be considered. These differences $Z(x+h)-Z(x)$ are called *increments*. The theoretical variogram $\gamma(h)$ is defined by the so-called *intrinsic hypothesis*, which stands as a short form for "a hypothesis of intrinsic stationarity of order two". This hypothesis that characterizes the stationarity behavior of the random function includes the following two assumptions:

- The mean $m(h)$ of the increments, also called the *drift*, is invariant for any translation of a given vector h within the domain. Furthermore, the drift is supposed to be zero independent from the position of h in the domain.
- The variance of the increments is finite with a finite value $2\gamma(h)$, depending on the length and the orientation of a given vector h, but again being independent of the position of h in the domain. This can be formally expressed as:
  For any pair of points $x, x+h \in \mathcal{D}$ it holds

$$E[Z(x+h)-Z(x)] = m(h) = 0 \tag{2.37}$$
$$\text{var}[Z(x+h)-Z(x)] = 2\gamma(h) \tag{2.38}$$

With the help of these two properties of an intrinsic stationary random function the definition of the theoretical variogram follows:

$$\gamma(h) = \frac{1}{2}E\left[(Z(x+h)-Z(x))^2\right] \tag{2.39}$$

The existence of expectation and variance of the increments does not imply the existence of the first two moments of the random function itself: Although the variance of the increments may be finite for any vector h, an intrinsic random function can have an infinite variance.
Some properties of the variogram are:

- At the origin, the value of the variogram is zero by definition:

$$\gamma(0) = 0 \tag{2.40}$$

- The values of the variogram are positive:

$$\gamma(h) \geq 0 \tag{2.41}$$

- The variogram is an even function:

$$\gamma(-h) = \gamma(h) \tag{2.42}$$

- The variogram grows slower than $|h|^2$, i.e.:

$$\lim_{|h|\mapsto\infty} \frac{\gamma(h)}{|h|^2} = 0 \tag{2.43}$$

Covariance Function
For the definition of the *covariance function* the stationarity of the first two moments (mean and covariance) of the random function is essential.

$$\begin{cases} E[Z(x)] = m & \text{for all} \quad x \in \mathcal{D} \\ E[Z(x) - m][Z(x+h) - m] =: C(h) & \text{for all} \quad x, x+h \in \mathcal{D} \end{cases} \tag{2.44}$$

Some properties of the covariance function are:

- The covariance function is bounded and its absolute value does not exceed the variance

$$|C(h)| \leq C(0) = \text{var}(Z(x)) \tag{2.45}$$

- Similar to the variogram, it is an even function

$$C(-h) = C(+h). \tag{2.46}$$

  But unlike the variogram it can also take negative values.

- The covariance function divided by the variance is called the *correlation function*

$$\rho(h) = \frac{C(h)}{C(0)}, \tag{2.47}$$

  which is bounded by

$$-1 \leq \rho(h) \leq 1 \tag{2.48}$$

- Furthermore, the variogram function can be deduced from a covariance function by

$$\gamma(h) = C(0) - C(h). \tag{2.49}$$

In general, the reverse is not true, because the variogram is not necessarily bounded. Thus, the hypothesis of second-order stationarity is less general than the intrinsic hypothesis (for the monovariate case) and unbounded variogram models do not have a covariance function counterpart.

- A covariance is a positive definite function. This means that the use of a covariance function $C(h)$ for the computation of the variance of a linear combination of $n+1$ random variables $Z(x_i)$ ($i \in I$ for any index set I, describing any subset sampled from a second-order stationary random function) must be positive. It is necessarily linked to a positive semi-definite matrix C of covariances

$$\text{var}\left(\sum_{i=1}^{n} w_i Z(x_i)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C(x_i - x_j) = w^\top C w \quad \geq \quad 0 \tag{2.50}$$

- **Conditionally Negative Definite Function**
  The variogram is a *conditionally* negative definite function. To guarantee the positivity of the variance of any linear combination of $n+1$ random variables being a subset of an intrinsic random function, the $n+1$ weights $w_i, i = 0, ..., n$ need to sum up to zero. In that case, the variance of a linear combination of intrinsically stationary random variables is given by

$$\begin{aligned} \text{var}\left(\sum_{i=0}^{n} w_i Z(x_i)\right) &= \sum_{i=0}^{n}\sum_{j=0}^{n} w_i\, w_j\, \gamma(x_i - x_j) \\ &\geq 0, \qquad \text{if} \quad \sum_{i=0}^{n} w_i = 0. \end{aligned} \tag{2.51}$$

It is important to notice here that the properties described above are not sufficient to totally characterize a covariance or variogram function. An additional and necessary condition for a function to become a covariance or a variogram function is that all variance calculations have to lead to a nonnegative result.

2.3.2 Interpretation of a Variogram

Plotting a sample variogram $\hat{\gamma}(h)$ against $|h|$ for a given direction of h generally results in the following properties:

1. It always starts at zero, for $h = 0$, $z(x+h) - z(x) = 0$.
2. It increases with $|h|$.
3. It continues to increase, or else stabilizes at a certain level.

Range and Sill
The amount of the variogram increase reflects the degree of dissimilarity of more distant samples. For the case the variogram increases infinitely, the variability of the phenomenon has no limit at large distances. If, conversely, the variogram reaches a limit, this value is called the *sill*. This means that there is a distance beyond which $Z(x)$ and $Z(x+h)$ are considered to be uncorrelated. This distance is usually called the *range*. For a better understanding the range can also be interpreted as an *area of influence* of a sample.

Behavior near the Origin - Nugget Effect
Sill and range describe the variogram behavior at large distances. To examine the continuity and spatial regularity of a regionalized variable it is important to consider the variogram's behavior near the origin. The type of continuity of the regionalized variable might be differentiable, continuous but not differentiable, or discontinuous. The last case, when the variogram is discontinuous at the origin, this phenomenon is called *nugget effect* indicating that the values of the variable change abruptly at a very small scale. Formally, this means:

$$\lim_{h \to 0} \hat{\gamma}(h) \neq 0 \tag{2.52}$$

The term *nugget effect* itself originally arises from gold deposits where gold commonly occurs as nuggets of pure metal that are much smaller than the sampling size. This may result in a strong grade variability among adjacent samples and causes a discontinuity of the variogram at the origin. In general, the nugget effect is caused by

- a microstructure or so-called "geological noise" which describes a component of the phenomenon with a range shorter than the sampling distance,
- a structure with a range smaller than the smallest distance between sample points,
- and measurement or positioning errors.

How measurement or positioning errors influence the Kriging procedure shall be discussed in the following sections.

When the average dissimilarity of the values of a variogram is constant for all spacings of h, there is no spatial structure in the data. Conversely, a non zero slope of the variogram near the origin indicates structure. An abrupt change in slope indicates the passage to a different structuring of the values in space. If the sampling points are too far apart, it is impossible to tell from the variogram what is the exact behavior at the origin. Furthermore, knowledge about the physics of the problem is essential for modelling the variogram.

2.3.3 Different Variogram Models

The use of a variogram in a Kriging procedure requires continuous variogram values for every distance $|h|$. Of course, this cannot be provided by the experimental variogram since only discrete measurements can be realized in practice. Fitting the experimental variogram by an appropriate variogram function helps to overcome this problem. Using a theoretical variogram also guarantees that the variance of any linear combination of sample values is positive. This is important for setting up a *Kriging system* where the values of an experimental variogram can lead to negative Kriging variances.
When fitting a theoretical variogram function to the sequence of average dissimilarities (experimental variogram) it is important to notice that the fit implies an interpretation of both the behavior at the origin and at large distances, beyond the range of the experimental variogram. In most applications, the fit is merely done by eye, because it is generally not so relevant how well the variogram function fits the sequence of points. The important point here is that the right type of continuity is assumed for the regionalized variable together with the stationarity hypothesis associated to the random function. These assumptions will lead to the choice of an appropriate variogram function which has a greater impact on the Kriging results than the way the theoretical function is fitted. A detailed discussion of this context can be found in [112].

There are several reasons to favor the variogram instead of the covariance function: The variogram is a more general tool than the covariance. Another reason is more of

practical interest: The variogram, unlike the covariance function, does not depend on the existence of a mean value. In practice, the mean is not known in most cases and has to be estimated out of the data, which also adds a bias. Therefore, the variogram is often preferred to the covariance function.
In the following some of the most common covariance functions are presented. They are mainly defined for isotropic (i.e. rotation invariant) random functions. For the graphical representations the covariance functions are plotted as variograms using relation (2.49):

Nugget-effect Model
The *nugget-effect model* describes a covariance function $C(h)$ that models a discontinuity at the origin:

$$C_{nug}(h) = \begin{cases} b & \text{for} \quad |h| = 0 \\ 0 & \text{for} \quad |h| > 0 \end{cases} \tag{2.53}$$

where b is a positive value. In variogram representation, the variogram is zero at the origin and has the value b for $h \neq 0$. An example of a corresponding variogram function can be seen in Figure 2.5. The nugget-effect is used to model a discontinuity at the origin of the variogram, i.e. when

$$\lim_{|h| \to 0} \gamma(h) = b. \tag{2.54}$$

The phenomenon of the nugget-effect can be seen as an equivalent to the concept of *white noise* in signal processing.

Spherical Model
Another widely spread covariance function is the *spherical model*

$$C_{sph}(h) = \begin{cases} b \left(1 - \frac{3}{2}\frac{|h|}{a} + \frac{1}{2}\frac{|h|^3}{a^3}\right) & \text{for} \quad 0 \leq |h| \leq a \\ 0 & \text{for} \quad |h| > a \end{cases} \tag{2.55}$$

The range of the spherical covariance is indicated by parameter a, i.e. the covariance vanishes when the range is reached. The parameter b represents the maximal value of the covariance: the spherical covariance steadily decreases, starting from the maximum value b at the origin and vanishes beyond range a.
The nugget-effect model can be interpreted as a special case of the spherical model with an infinitely small range. But there is one crucial difference between these two models: $C_{nug}(h)$ describes a discontinuous phenomenon, whose values change abruptly from one location to the other, while $C_{sph}(h)$ represents a continuous phenomenon being not differentiable. A corresponding spherical variogram can be found in Figure 2.6. In this example, the sill $b = 1$ is reached at a range of $a = 3$.

Exponential Covariance Function
The exponential covariance function model falls off exponentially with increasing

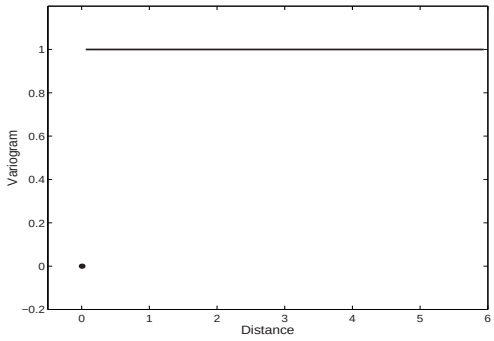


Figure 2.5: A nugget-effect variogram: its value is zero at the origin and $b = 1$ elsewhere.

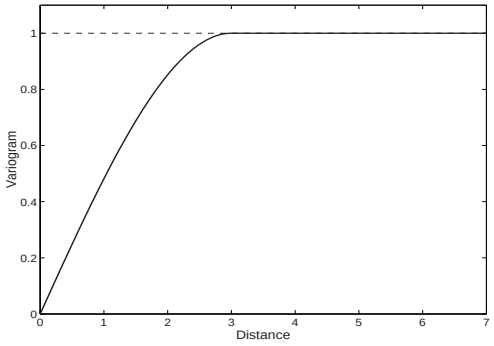


Figure 2.6: A spherical variogram with sill $b = 1$ and range $a = 3$.

distance

$$C_{exp}(h) = b \exp\left(-\frac{|h|}{a}\right) \quad \text{with } a, b > 0. \tag{2.56}$$

Parameter a determines how quickly the covariance falls off. For a value of $h = 3a$ the covariance function has decreased by 95% of its value at the origin. Thus, this distance has been termed the *practical range* of the exponential model. This model is continuous but not differentiable at the origin. It drops asymptotically towards zero for $|h| \to \infty$. An illustration of an equivalent variogram can be found in Figure 2.7.

Gaussian Model

The *Gaussian Model* with scale parameter $a > 0$ is defined by

$$C_{gauss}(h) = \exp\left(-\frac{|h|^2}{a^2}\right) \tag{2.57}$$

is a covariance in $\mathbb{R}^n$ for any n. Its practical range is about $1.73a$, as can be seen in Figure 2.8 showing the corresponding variogram.

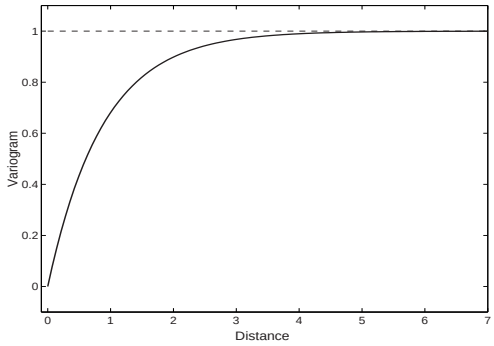


Figure 2.7: Exponential Variogram Model with sill $b = 1$ and range parameter $a = 1$.

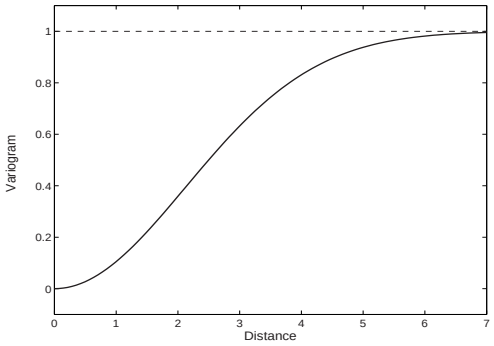


Figure 2.8: A Gaussian Variogram Model with $a = 3$.

Matérn Class of Covariance Functions

Covariance functions belonging to the so-called Matérn class (named after Bertil

Matérn [108]) possess an additional parameter ν describing the differentiability or smoothness. In most applications, a necessary differentiability can be deduced out of the physical context of a problem. The mathematical representation of this class of covariance functions depending on smoothing parameter ν is given by:

$$C(h) = \begin{cases} 0 & \text{for} \quad h = 0 \\ \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\frac{h}{a}\right)^{\nu} K_{\nu}\left(\frac{h}{a}\right) & \text{for} \quad h > 0 \end{cases} \tag{2.58}$$

for $a > 0$, $\nu \geq 0$ and K_{ν} being a modified Bessel function of the second kind of order ν (cf. [1]), $\Gamma(\cdot)$ denoting the usual gamma function and a giving the range. The variogram representation of this model is the following:

$$\gamma(h) = \begin{cases} 0 & \text{for} \quad h = 0 \\ C\left(1 - \frac{1}{2^{(\nu-1)}\Gamma(\nu)} \left(\frac{|h|}{a}\right)^{\nu} K_{\nu}\left(\frac{|h|}{a}\right)\right) & \text{for} \quad h > 0 \end{cases} \tag{2.59}$$

with C being the sill.
Depending on the choice of parameter ν, this class of variogram functions also contains the most common types like the exponential or the gaussian model.

$$\nu = \begin{cases} \frac{1}{2} & \text{Exponential Type} \quad \gamma(h) = b\left(1 - \exp\left(-\frac{|h|}{a}\right)\right) \\ \infty & \text{Gaussian Type} \quad \gamma(h) = b\left(1 - \exp\left(-\frac{|h|^2}{a^2}\right)\right) \end{cases} \tag{2.60}$$

This model can have any behavior near the origin and Figure 2.9 illustrates the most common cases. For $\nu \to \infty$, the model converges towards the gaussian type.

Anisotropy
In practice, calculating the experimental variogram might lead to quite different results depending on the direction of vector h. This phenomenon is called *anisotropy*. For the applications concerning heterogeneous catalyst samples, always an isotropic behavior is assumed, i.e. the variogram does not change with direction of h but only with distance among sample points. Changes in composition of the catalytic samples are assumed to cause a rotationally symmetric change around each sample. For a detailed discussion of anisotropy, hole effects, periodicities and how such effects are addressed in the context of variogram models, see standard textbooks (e.g. [23, 180]).

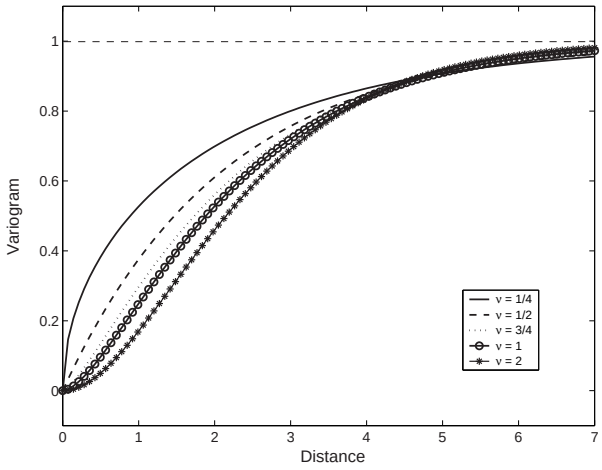


Figure 2.9: Depending on parameter ν, variograms show different shapes near the origin.

2.4 Simple Kriging

In general, Kriging is a geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when estimating values in unknown areas. A kriged estimate is always a weighted linear combination of the known sample values around the point to be estimated. Kriging leads to weights that result in optimal and unbiased estimates. It attempts to minimize the error variance and sets the mean of prediction errors to zero so that there are no over- or under-estimates. Furthermore, a unique feature of Kriging is that it also gives an estimation of the error at each interpolated point, providing a measure of confidence for the results.

Suppose data are given for one type of measurement at various locations $x_1, ..., x_n$ of a spatial domain $\mathcal{D}$. For each location $x_i, i = 1, ...n$ a random variable $Z(x_i)$ is considered. Let x_0 be an additional location in $\mathcal{D}$ with random variable $Z(x_0)$. Furthermore, it is assumed that these random variables are a subset of an infinite collection of random variables, namely the random function $Z(x)$, defined at any location of x of the domain $\mathcal{D}$. Let the random function be second-order stationary, as has been defined by equation (2.29). This means that both the expectation and the covariance are translation invariant over the domain. Then, the expected value $E[Z(x)] = m$ is the same at any point x of the domain and the covariance between a pair of locations depends only on the vector h separating them.
Now, the goal is to calculate a weighted average to estimate the unknown value at location x_0 using information at points $x_1, ..., x_n$. This situation is illustrated in Figure 2.10. The estimation procedure for location x_0 is based on the knowledge of the covariances between the random variables at the points sampled.
The estimation approach called *Simple Kriging* (SK) is closely related to multiple regression considered within a spatial context. One main aspect here is the fact, that simple Kriging assumes a known, constant mean of the data which is often not available in practice. In this case, it is estimated out of the sample data. Sometimes, simple Kriging is also referred to as *Kriging with known mean.*

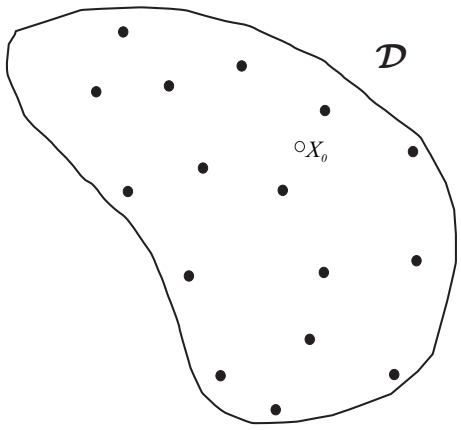


Figure 2.10: Sample Points within domain $\mathcal{D}$ and the estimation point x_0.

The weighted average for this estimation is defined as

$$Z^\star(x_0) = m + \sum_{i=1}^{n} w_i(Z(x_i) - m) \tag{2.61}$$

where w_i are weights belonging to the residuals $(Z(x_i) - m)$. Unlike in multiple regression, the mean m is the same for all locations following the stationarity assumption.
For the estimation error at location x_0 the difference between the estimated and the true value is to be considered:

$$Z^\star(x_0) - Z(x_0) \tag{2.62}$$

An estimator is said to be *unbiased* if the estimation error is zero on average:

$$\begin{aligned} E[Z^\star(x_0) - Z(x_0)] &= m + \sum_{i=1}^{n} w_i \left(E[Z(x_i)] - m\right) - E[Z(x_0)] \\ &= m + \sum_{i=1}^{n} w_i(m - m) - m \\ &= 0. \end{aligned} \tag{2.63}$$

The variance of the estimation error, called the *estimation variance* σ_E^2, can be expressed as

$$\sigma_E^2 = \mathrm{var}(Z^\star(x_0) - Z(x_0)) = E[(Z^\star(x_0) - Z(x_0))^2]. \tag{2.64}$$

Expanding this expression leads to

$$\begin{aligned} \sigma_E^2 &= E[(Z^\star(x_0))^2 + (Z(x_0))^2 - 2Z^\star(x_0)Z(x_0)] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i\, w_j C(x_i - x_j) + C(x_0 - x_0) - 2\sum_{i=1}^{n} w_i\, C(x_i - x_0). \end{aligned} \tag{2.65}$$

where the last equation has been computed using

$$\text{cov}[Z(x_i), Z(x_j)] = C(x_i - x_j), \tag{2.66}$$

which directly follows from the stationarity assumption (2.29): the spatial covariances only depend on the difference vector between the points.
Minimal estimator variance is obtained by setting the first partial derivatives to zero

$$\frac{\partial \sigma_E^2}{\partial w_i} \qquad \text{for } i = 1, ..., n \tag{2.67}$$

Explicitly, this means that for each $i = 1, ..., n$ this yields,

$$2 \sum_{j=1}^{n} w_j \, C(x_i - x_j) - 2C(x_i - x_0) = 0 \tag{2.68}$$

and the system of equations for simple Kriging can be written

$$\sum_{j=1}^{n} w_j \, C(x_i - x_j) = C(x_i - x_0) \qquad \text{for } i = 1, ..., n \tag{2.69}$$

In matrix notation, this leads to

$$\begin{pmatrix} C(x_1 - x_1) & C(x_1 - x_2) & \dots & C(x_1 - x_n) \\ C(x_2 - x_1) & C(x_2 - x_2) & \dots & C(x_2 - x_n) \\ \vdots & \vdots & \ddots & \vdots \\ C(x_n - x_1) & C(x_n - x_2) & \dots & C(x_n - x_n) \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} C(x_1 - x_0) \\ C(x_2 - x_0) \\ \vdots \\ C(x_n - x_0) \end{pmatrix} \tag{2.70}$$

The left hand side of the equation system contains the matrix of data-to-data covariances between the locations. On the right hand side the covariances between each data location and the point to be estimated are calculated. Solving this system yields the desired weights $w_i, i = 1, ..., n$. The system has a unique solution if the matrix containing the covariances is nonsingular. This is always the case if the used covariance function is strictly positive definite and all sample points are distinct (which is the case for all the following applications).

Simple Kriging also provides a variance for each location x_0. It is obtained by substituting the left hand side of the kriging system by its right hand side in the first term of the expression of the estimation variance σ_E^2. Then the variance of simple kriging is given by

$$\begin{aligned} \sigma_{SK}^2 &= \sum_{i=1}^{n} w_i \, C(x_i - x_0) + C(x_0 - x_0) - 2 \sum_{i=1}^{n} w_i \, C(x_i - x_0) \\ &= C(0) - \sum_{i=1}^{n} w_i \, C(x_i - x_0) \end{aligned} \tag{2.71}$$

Properties of $Z^\star(x_0)$:
The estimator $Z^\star(x_0)$, the kriging estimator received in this way, is an exact interpolator. If the point x_0 coincides with a sample point, e.g. x_1, then $Z^\star(x_0)$ is equal to $Z(x_1)$. Furthermore, $Z^\star(x_0)$ is the *best* estimator as it minimizes the error variance (in this case the variance is exactly zero).

2.5 Kriging the Mean

This section describes a procedure to estimate the mean value of samples that possess some kind of spatial correlation. The mean value can also be estimated by the arithmetic mean which is the simplest approach. If the spatial correlations of the samples should be included into estimating the mean, a weighted average has to be constructed.

Let $Z(x_i)$ be a number of samples at positions $x_i, i = 1, ..., n$ irregularly spread in a domain $\mathcal{D}$. An estimator $m^\star$ for the mean value m should be calculated which is assumed to be a weighted average:

$$m^\star = \sum_{i=1}^{n} w_i \, Z(x_i) \tag{2.72}$$

Now, the weights w_i need to be calculated. It is assumed that the mean exists for all samples of domain $\mathcal{D}$, i.e. $E[Z(x)] = m$ for all x in $\mathcal{D}$. The estimator should be unbiased which implies $E[m^\star - m] = 0$. Then it holds that

$$E[m^\star - m] = E\left[\sum_{i=1}^{n} w_i Z(x_i) - m\right] = m \sum_{i=1}^{n} w_i - m \stackrel{!}{=} 0. \tag{2.73}$$

It follows that the weights w_i need to sum up to 1 to satisfy the unbiasedness condition.

$$\sum_{i=1}^{n} w_i = 1 \tag{2.74}$$

Further $Z(x)$ need to be second order stationary with an existing covariance function $C(h)$ which describes the correlation between any pair of points x and $x + h$ in the spatial domain

$$C(h) = \mathrm{cov}(Z(x), Z(x+h)) = E[Z(x) \cdot Z(x+h)] - m^2. \tag{2.75}$$

Now, the variance of the estimation error in this unbiased procedure is the average of the squared error

$$\mathrm{var}(m^\star - m) = E[(m^\star - m)^2] - (\underbrace{E[m^\star - m]^2}_{=0}). \tag{2.76}$$

The variance of the estimation error can be expressed in terms of the covariance

function

$$\begin{aligned}\mathrm{var}(m^\star - m) &= E[m^{\star 2} - 2mm^\star + m^2] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i\, w_j E[Z(x_i)Z(x_j)] - 2m\sum_{i=1}^{n} w_i \underbrace{E[Z(x_i)]}_{=m} + m^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i\, w_j C(x_i - x_j) \qquad (2.77)\end{aligned}$$

The "best" weights will be the weights that minimize $\mathrm{var}(m^\star - m)$ and respect the unbiasedness condition (2.74) which gives the following minimization problem:

$$\boxed{\text{Minimize} \quad \mathrm{var}(m^\star - m) \qquad \text{subject to} \quad \sum_{i=1}^{n} w_i = 1} \qquad (2.78)$$

The minimum can be found by setting the first order partial derivatives to zero. To include the unbiasedness constraint into this problem, the minimization is solved using the method of *Lagrange*, which can be found in detail e.g. in [169]. The Lagrange approach solves a constrained minimization problem by converting it into an unconstrained minimization of the Lagrange function.
This function φ is defined out of the original objective function plus a term containing a *Lagrange multiplier* μ,

$$\varphi(w_i, \mu) = \mathrm{var}(m^\star - m) - 2\mu\left(\sum_{i=1}^{n} w_i - 1\right). \qquad (2.79)$$

For ease of subsequent computation, the Lagrange multiplier is multiplied by factor 2 here.
To solve the optimization problem the partial derivatives of the objective function $\varphi(w_i, \mu)$ are set to zero

$$\frac{\partial\varphi(w_i,\mu)}{\partial w_i} = 0 \quad \text{for } i = 1, ..., n \qquad (2.80)$$

$$\frac{\partial\varphi(w_i,\mu)}{\partial \mu} = 0 \qquad (2.81)$$

This yields a linear system of $n+1$ equations. Solving this system leads to the optimal weights w_i for the estimation of the mean by a weighted average.

$$\begin{cases} \displaystyle\sum_{j=1}^{n} w_j\, C(x_i - x_j) - \mu = 0 \quad \text{for} \quad i = 1, ..., n \\ \displaystyle\sum_{j=1}^{n} w_i = 1. \end{cases}$$

Summarized, this results in the system for *Kriging the mean (KM)*.
The minimal estimation variance σ^2 is computed by using the equations for the optimal weights

$$\sum_{j=1}^{n} w_j C(x_i - x_j) = \mu \quad \text{for } i = 1, ..., n \tag{2.82}$$

and in the expression of the estimation variance

$$\begin{aligned} \sigma^2 &= \operatorname{var}(m^\star - m) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C(x_i - x_j) \\ &= \sum_{i=1}^{n} w_i \mu \qquad = \mu \underbrace{\sum_{i=1}^{n} w_i}_{=1} = \mu. \end{aligned} \tag{2.83}$$

This means that the variance of Kriging of the mean is directly given by the Lagrange multiplier μ.
If there is no spatial correlation between points in space, the arithmetic mean is the best linear unbiased estimator and also the corresponding estimation variance is recovered. In general, it can be said that Kriging of the mean is a generalization of the arithmetic mean estimator to the case of spatially correlated samples.

2.6 Ordinary Kriging

Ordinary Kriging (OK) is the most widely applied Kriging method. It can be used to estimate a value at a point of a region for which a variogram is known. As mentioned above, the variogram is computed with the help of data points lying in the neighborhood of the estimation location or by using the whole data set. Ordinary Kriging can also be used to estimate a block value.

Ordinary Kriging Problem
The value at location x_0 should be estimated as illustrated in Figure 2.11 using the

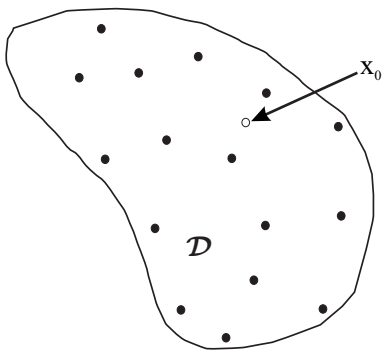


Figure 2.11: A domain with irregularly sampled points and a location x_0 of interest for estimation.

data values from n surrounding sample points $x_i, i = 1, ..., n$. The estimator $Z^\star_{OK}$ should be a linear combination of the points x_i with weights w_i.

$$Z^\star_{OK}(x_0) = \sum_{i=1}^{n} w_i Z(x_i). \tag{2.84}$$

In contrast to simple Kriging, ordinary Kriging constrains the weights to sum up to one because in the particular case when all data values are constant, the estimated value should also be equal to this constant. Again, it is assumed that the data are part of a realization of an intrinsic random function $Z(x)$ with a variogram $\gamma(h)$. Furthermore, the unbiasedness is achieved by unit sum weights:

$$\begin{aligned} E[Z^\star(x_0) - Z(x_0)] &= E\Big[\sum_{i=1}^{n} w_i Z(x_i) - Z(x_0) \cdot \underbrace{\sum_{i=1}^{n} w_i}_{=1}\Big] \\ &= \sum_{i=1}^{n} w_i E[Z(x_i) - Z(x_0)] = 0, \end{aligned} \tag{2.85}$$

with the expectations of the increments being zero.
For the estimation variance $\sigma^2_E = \mathrm{var}(Z^\star(x_0) - Z(x_0))$ it holds that it is the variance of the linear combination

$$Z^\star(x_0) - Z(x_0) = \sum_{i=1}^{n} w_i Z(x_i) - 1 \cdot Z(x_0) = \sum_{i=0}^{n} w_i Z(x_i), \tag{2.86}$$

with a weight w_0 equal to -1 and

$$\sum_{i=0}^{n} w_i = 0. \tag{2.87}$$

It has to be noted here that the variogram is authorized for ordinary Kriging, but not for simple Kriging because simple Kriging does not constrain the weights to sum up to 1. Condition (2.87) directly corresponds to the one found in expression (2.51) of a conditionally negative function. Using that connection, the estimation variance for ordinary Kriging can be expressed as

$$\begin{aligned} \sigma^2_E &= E[(Z^\star(x_0) - Z(x_0))^2] \\ &= -\gamma(x_0 - x_0) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i\, w_j\, \gamma(x_i - x_j) + 2\sum_{i=1}^{n} w_i \gamma(x_i - x_0). \end{aligned} \tag{2.88}$$

Minimizing the estimation variance due to the constraints on the weights results in the *ordinary Kriging system* (OK)

$$\begin{pmatrix} \gamma(x_1 - x_1) & \dots & \gamma(x_1 - x_n) & 1 \\ \gamma(x_2 - x_1) & \dots & \gamma(x_2 - x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n - x_1) & \dots & \gamma(x_n - x_n) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1^{OK} \\ w_2^{OK} \\ \vdots \\ w_n^{OK} \\ \mu_{OK} \end{pmatrix} = \begin{pmatrix} \gamma(x_1 - x_0) \\ \gamma(x_2 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \\ 1 \end{pmatrix}, \tag{2.89}$$

with w_i^{OK} being the weights that have to be computed and μ_{OK} being the Lagrange parameter. On the left hand side of the system, the dissimilarities between all data points are coming into play while the right hand side shows the dissimilarities between each data point and the estimation point x_0.
The ordinary Kriging system can also be written in the form

$$\begin{cases} \sum_{j=1}^{n} w_j^{OK} \gamma(x_i - x_j) + \mu_{OK} = \gamma(x_i - x_0) & \text{for } \alpha = 1, ..., n \\ \sum_{j=1}^{n} w_j^{OK} = 1. \end{cases} \tag{2.90}$$

The estimation variance of ordinary Kriging is

$$\sigma_{OK}^2 = \mu_{OK} - \gamma(x_0 - x_0) + \sum_{i=1}^{n} w_i^{OK} \gamma(x_i - x_0). \tag{2.91}$$

The ordinary Kriging approach yields *exact interpolations*: if x_0 is identical with another sample point x_i, then the estimated value is the same as the data value at that point, i.e.

$$Z^\star(x_0) = Z(x_i) \quad \text{if} \quad x_0 = x_i. \tag{2.92}$$

Looking at the Kriging system, this becomes clear: With x_0 being one of the sample points, the right hand side of the Kriging system is equal to one column of the left hand side matrix. Then the vector w containing a weight equal to 1 for that column and all other weights equal to zero (also $\mu_{OK} = 0$) is a feasible solution of this system. Actually, it is the only solution as the left hand matrix is not singular.

2.6.1 Ordinary Kriging Applied to Heterogeneous Catalyst Data

Within the scope of this thesis more than 2400 catalyst samples have been prepared and screened for their catalytic activity for the oxidation of propene. The main focus has been laid on the correlation between chemical composition of the catalysts and their activity. With the help of Kriging, activities of non-synthesized compositions can be estimated given the activities of "surrounding" samples. Due to the construction of the Kriging estimator, estimation values larger than the largest measured value can be obtained, which is not always the case for estimation techniques. In applying the Kriging approach to real catalyst data, this feature plays the most

important role. Calculating the Kriging estimator of any desired catalyst composition within the sampled search space allows the construction of an appropriate model of the functional correlation of composition and activity (QCAR, quantitative composition activity relationship). This allows a deeper understanding of the phenomena within the sample space such that the most active compositions can be easily identified and further investigated. Another advantage of the Kriging approach compared to other interpolation or approximation techniques is the fact that it is not restricted to three dimensions as some other interpolation techniques are. For Kriging, the dimensionality of the data set plays no role since the only "critical" calculations deal with distance measures (e.g. euclidian distances) defined in $\mathbb{R}^n, n \in \mathbb{N}, n > 0$. In practice, the size of the Kriging system is given by the number of screened samples. The combinatorial sampling of a pentanary composition spread using 10%-wise increments yields 1001 samples which appears to be the actual limit with respect to time and synthesis parameters. In this case, the Kriging matrix has the size of 1002×1002 that can be easily handled by a common PC. Sampling a six elemental search space using 10%-wise increments the number of samples to be synthesized and tested moves to 3003, being a cumbersome task within reasonable time. So, the limiting factor of applying Kriging to heterogeneous catalyst data lies not in Kriging itself but in the availability of appropriate combinatorial screening data. To show the efficiency of the estimation procedure with Kriging, a complete pentanary composition spread has been prepared and screened automatically including 10 ternary and 5 quaternary composition spreads. The results of the Kriging estimation procedure applied to the most promising ternary and quaternary data sets together with the complete pentanary are described in detail in Chapter 6.

2.7 Data Transformation

To work with data resulting from high-throughput screening experiments, appropriate transformations in a mathematically useful coordinate system has to be performed. Originally, each data set consists of composition information and activity values assigned to each sample. This means, for each catalyst, its elemental composition is considered together with its activity for the oxidation of propene to a certain product. As already mentioned above, within this thesis the focus exclusively lies on the activities of catalysts for acrolein among all the possible by-products. The catalysts are mixed metal oxides composed of n different metals in various contents and oxidation states. Therefore the content of oxygen within the materials may differ depending on the preparation techniques, pretreatment or during the reaction and is not further specified. The total metal content (defined as 100%) is indeed uniquely determined if the contents of $n-1$ metals are specified.
Mathematically, a catalyst consisting of n different metals can be described as a point lying in the n-dimensional space $\mathbb{R}^n$ with each coordinate axis corresponding to an element. Within the following sections ternary, quaternary and pentanary composition spreads are dealt with separately and it is discussed how the data transformation has been performed.

2.7.1 Ternary Composition Spreads

Ternary composition spreads $A_aB_bC_cO_x$ with $a, b, c = 0, ..., 1$ represent the simplest case within this study. With a 10%-wise variation in composition, in total 66 different samples can be considered. Using the Cartesian coordinate system, ternary compositions lie on the diagonal of the unit cube if the composition values between 0 and 1 are interpreted as Cartesian coordinates. Figure 2.12 illustrates this:

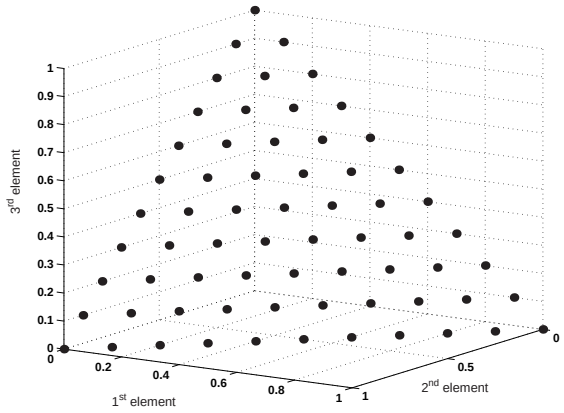


Figure 2.12: Ternary system in 3D lying on the diagonal of the unit cube.

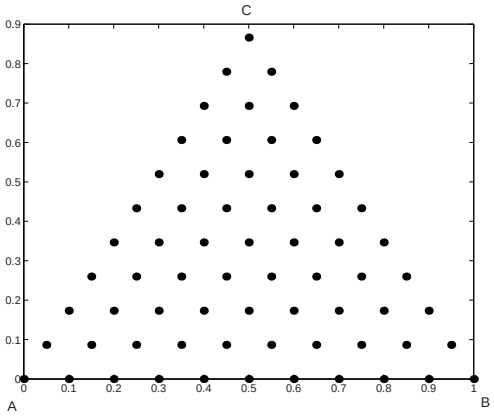


Figure 2.13: Ternary system transferred into 2D representation within the Cartesian coordinate system.

Having the constraint that the proportions of metals sum up to 1 (100%), information about the nth dimension is redundant such that a ternary system can also be presented in 2D using an adequate coordinate transformation. For the aimed purposes the data have been transformed into a 2D representation as shown in Figure 2.13.
Without loss of generality, the vertices of the triangle have been fixed to the coordinates $(0, 0)$, $(1, 0)$ and $(0.5, \sqrt{3}/2)$. Mathematically this has been done by multiplying the original composition values by matrix M_T. Let $p = (a, b, c)$ be a point in $\mathbb{R}^3$ describing a catalyst with composition $A_aB_bC_cO_x$. The coordinates of this point p in 2D representation (x_{2d}, y_{2d}) can be obtained by multiplying (b, c) by matrix M_T.

$$\underbrace{\begin{pmatrix} 1 & 0.5 \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}}_{M_T} \cdot \begin{pmatrix} b \\ c \end{pmatrix} = \begin{pmatrix} x_{2d} \\ y_{2d} \end{pmatrix} \tag{2.93}$$

This convention results in a representation of the ternary system with the first element corresponding to the left vertex, the second element corresponding to the right vertex and the third element is given by the top of the triangle. Thus, points belonging to a ternary composition spread lie in a triangular shaped plane within $\mathbb{R}^2$. Their activity values form the third dimension such that the interpolation problem that needs to be solved here by Kriging is given in $\mathbb{R}^3$, cf. Figure 2.14. Calculating the Kriging estimators corresponds to the calculation of a triangular surface providing the activity for any composition within the parameter space.

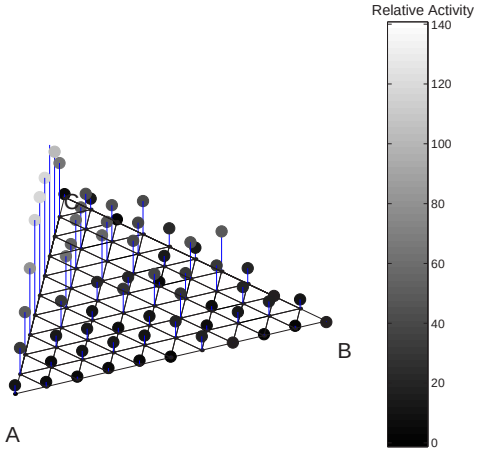


Figure 2.14: Ternary Composition Spread in 2D with activity giving the 3rd dimension.

2.7.2 Quaternary Composition Spreads

A quaternary composition spread $A_aB_bC_cD_dO_x$ with $a, b, c, d = 0, ..., 1$ in 10%-wise variations yields 286 samples. As already discussed above, quaternary points are originally lying in $\mathbb{R}^4$. Since an activity value belongs to each point there are actually five dimensions to deal with. The compositional data can be transformed into 3D representation analogously as discussed for the ternary case. This gives an even tetrahedron in 3D. Without loss of generality, this tetrahedron is defined based on the triangle defined for the ternary case and the fourth element sits on top of the tetrahedron. Then the coordinates of the vertices in 3D are given by:

$$\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\\frac{\sqrt{3}}{2}\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\\frac{1}{3}\frac{\sqrt{3}}{2}\\\sqrt{\frac{2}{3}}\end{pmatrix} \tag{2.94}$$

Let $p = (a, b, c, d)$ be a point in $\mathbb{R}^4$ describing a catalyst with composition $A_aB_bC_cD_dO_x$. The corresponding coordinates (x_{3D}, y_{3D}, z_{3D}) for a representation in 3D can be obtained by a multiplication with matrix M_T (rotation of the coordinate system):

$$\underbrace{\begin{pmatrix}1 & 0.5 & 0.5\\0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6}\\0 & 0 & \sqrt{\frac{2}{3}}\end{pmatrix}}_{M_T} \cdot \begin{pmatrix}b\\c\\d\end{pmatrix} = \begin{pmatrix}x_{3D}\\y_{3D}\\z_{3D}\end{pmatrix} \tag{2.95}$$

After this transformation the complete quaternary composition spread yields an even tetrahedron in 3D as can be seen in the following figure.

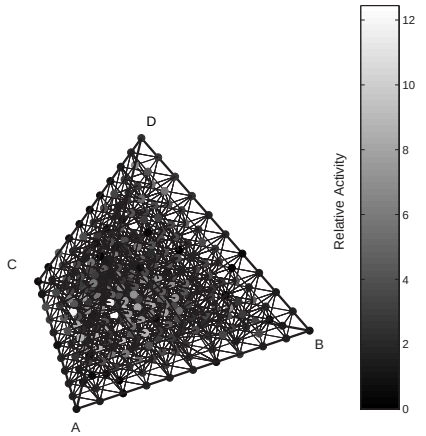


Figure 2.15: Example of a quaternary composition spread $A_aB_bC_cD_dO_x$ plotted in 3D. The grey scale corresponds to the measured activity giving a fourth dimension.

2.7.3 Pentanary Composition Spreads

The last case to consider here is a pentanary composition spread $A_aB_bC_cD_dE_eO_x$. Together with the activity values corresponding to each point there are six dimensions to deal with. Following the transformation approach described above, the composition information can be represented in 4D yielding some sort of generalized tetrahedron, a convex polyhedron in 4D. Then, the coordinates of the vertices of this polyhedron have to be determined applying the following procedure:
Consider the coordinates of the vertices of the tetrahedron in 3D (4 vectors in $\mathbb{R}^3$) and expand them by one component. The fifth vertex in $\mathbb{R}^4$ should satisfy the condition that its distance to all other vertices must be the same as it is always the case for equilateral polyhedra. Without loss of generality and following the schemes explained above this distance is set to 1 unit. Furthermore, all components of the vectors should be positive since they directly correspond to the composition values of materials (negative solutions have no applicable meaning here).
Setting up the system of equations to be solved:
Consider the known coordinates of the vertices of a tetrahedron in $\mathbb{R}^3$ having edge length 1:

$$\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\\frac{\sqrt{3}}{2}\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\\frac{1}{3}\frac{\sqrt{3}}{2}\\\sqrt{\frac{2}{3}}\end{pmatrix} \tag{2.96}$$

To come to a polyhedron in $\mathbb{R}^4$ having five vertices, the dimension need to be expanded by one component.

$$\mathbf{e_1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e_2} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e_3} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \\ 0 \end{pmatrix}, \mathbf{e_4} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3}\frac{\sqrt{3}}{2} \\ \sqrt{\frac{2}{3}} \\ 0 \end{pmatrix} \tag{2.97}$$

To calculate the fifth vertex $x = (x_1, x_2, x_3, x_4)$ it should hold: Vector x must have the same distance from all other vertices and this distance should be 1 using the euclidian distance measure. Then,

$$\|\mathbf{x} - \mathbf{e_i}\| = 1 \qquad \text{for} \quad i = 1, \ldots, 4. \tag{2.98}$$

This leads to the following system of non-linear equations:

$$\begin{array}{rcl} x_1^2 + x_2^2 + x_3^2 + x_4^2 & = & 1 \\ x_1^2 - 2x_1 + x_2^2 + x_3^2 + x_4^2 & = & 1 \\ x_1^2 - x_1 + x_2^2 - \sqrt{3}x_2 + x_3^2 + x_4^2 & = & 1 \\ x_1^2 - x_1 + x_2^2 - \frac{\sqrt{3}}{3}x_2 + x_3^2 - 2\sqrt{\frac{2}{3}}x_3 + x_4^2 & = & 1 \end{array} \tag{2.99}$$

A positive solution for the fifth vertex is:

$$\mathbf{x} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{6} \\ \frac{\sqrt{6}}{12} \\ \frac{1}{2}\sqrt{\frac{5}{2}} \end{pmatrix} \tag{2.100}$$

Analogously, this approach can also be applied to find the sixth vertex in $\mathbb{R}^5$ and so on. Solving the corresponding equation system leads to

$$\mathbf{x} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{6} \\ \frac{\sqrt{6}}{12} \\ \frac{\sqrt{10}}{20} \\ \sqrt{\frac{3}{5}} \end{pmatrix} \tag{2.101}$$

Summarizing, the calculation of coordinates for convex polyhedra in high dimensional spaces can be done using the following transformation matrices:

- Tetrahedron in $\mathbb{R}^3$ for quaternary composition spreads:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6} \\ 0 & 0 & \sqrt{\frac{2}{3}} \end{pmatrix}$$

- Polyhedron in $\mathbb{R}^4$ for pentanary composition spreads:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} \\ 0 & 0 & \sqrt{\frac{2}{3}} & \frac{\sqrt{6}}{12} \\ 0 & 0 & 0 & \frac{1}{2}\sqrt{\frac{5}{2}} \end{pmatrix}$$

- Polyhedron in $\mathbb{R}^5$ for composition spreads with six elements:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} \\ 0 & 0 & \sqrt{\frac{2}{3}} & \frac{\sqrt{6}}{12} & \frac{\sqrt{6}}{12} \\ 0 & 0 & 0 & \frac{1}{2}\sqrt{\frac{5}{2}} & \frac{\sqrt{10}}{20} \\ 0 & 0 & 0 & 0 & \sqrt{\frac{3}{5}} \end{pmatrix}$$

Within this thesis the data have always been transformed into the corresponding spaces to apply the interpolation methods even though Kriging can easily cope with multi-dimensional data. For the sake of a good visualization the reduction of dimension by 1 comes along with a bundle of advantages for application.

3 Multilevel B-Splines

This chapter is dedicated to the theory and methodology of *multilevel B-Spline interpolation*, a technique that is also used next to Kriging for interpolating catalyst data. Within the next section a short overview on the most common interpolation techniques is given together with the introduction of a multilevel B-Spline approach to interpolate scattered data following the work of Lee and coworkes [100].

3.1 Scattered Data Interpolation

The problem of scattered data interpolation can be described as fitting a smooth surface through data samples that are nonuniformly distributed. In general, the 2D interpolation problem can be formulated as follows: Let $P = \{p_1, ..., p_N\} \subset \mathbb{R}^2$ and values of function f at these points, $f(p_1), ..., f(p_N)$. Then an interpolation (approximation) function $\hat{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$ has to be found that describes the given data points in the 'best' manner. The definition of 'best approximation' or 'best interpolation' cannot be easily answered since the real underlying function is not known. Thus, no quality criteria (e.g. least squares distances etc.) or approximation errors can be calculated.
For practical applications this problem plays an important role in many scientific and engineering fields where data points are often sampled sparsely and only at distinct positions. As the main goal of any interpolation technique the reconstruction of an underlying function is intended. Then this function can be evaluated at any desired position which allows the propagation of information associated with the scattered data onto all positions in the domain. There are many sources of scattered data but among the most common are measurements of physical quantities, results from experiments and computational values. Scientific and engineering fields that mostly yield scattered data are geology, meteorology, oceanography, cartography and especially mining. Scattered experimental data is also often produced in chemistry, physics and engineering and nonuniformly spaced computational values may result from finite element solutions of partial differential equations, computer graphics or computer vision. A great interest of scattered data interpolation techniques is indeed the gained information at non-sampled points. Having reconstructed the function, a lot of useful applications, such as better visualization of data, can be performed. For example, in geography or cartography, the reconstruction of a terrain surface from scattered altitude values is exactly one application for interpolation approaches.

3.1.1 Application to Heterogeneous Catalyst Data

Within this thesis, scattered data interpolation has been used to construct the activity surface of scattered catalyst activity measurements. From experiments, the activity values of predefined compositions of mixed metal oxide catalysts have been

given. The interesting question here is to find an underlying function connecting composition and activity such that activity values of non-synthesized catalysts can be estimated. Once the function is calculated, it can be evaluated at any desired composition values being equivalent to calculating an estimator by the Kriging approach.

3.1.2 Interpolation Techniques

There has been a lot of activity within this field of research during the last two decades, but scattered data interpolation remains a quite difficult and computationally expensive problem. There is a vast amount of literature devoted to this subject describing various approaches with different limitations in smoothness, time complexity or allowable data distributions. For a detailed overview the reader is referred to several surveys on this topic [10, 45, 74, 151]. In the following, a short description of the most common interpolation techniques is given with an explicit focus on the multilevel B-Spline interpolation approach presented by Lee *et al.* [100].
Most interpolation methods present $\hat{f}$ as a weighted sum of the known values $z_k = f(p_k)$, i.e.

$$\hat{f}(x,y) = \sum_{i=1}^{N} w_i z_i \quad , \qquad (x,y) \in \mathbb{R}^2 \tag{3.1}$$

and differ in the method to calculate the weights.
In one of the simplest interpolation methods - the *Inverse Distance Weighted Method* - the interpolation function $\hat{f}(x,y)$ at an arbitrary point $(x,y) \in \mathbb{R}^2$ is defined as a weighted sum having weights proportional to the k^{th} power of the inverse distance between the point (x,y) and points with known values. This method suffers from several well known drawbacks: The estimates are bounded by the extrema in the samples values and thus this method is not appropriate for the application to the catalyst interpolation problem. The probability that one gets the most active compositions within a composition spread is extremely small such that the interpolation technique must also be able to yield larger activity values than the maximum measured value. The Inverse Distance Weighted Method uses weights of every scattered point, as also Kriging does require all weights to be recomputed if any data point is added or removed or just somehow modified. For $k = 1$ the interpolation function is C^0-continuous and has many cusps and corners. For $k > 1$ derivatives are zero on P.
Shepard [159] proposed a variation in the inverse power, with two different weighting functions using two separate neighborhoods based on the inverse distance weighted method. Franke and Nielson [46] published a modified quadratic Shepard's method to improve these drawbacks and to produce C^1-continuous interpolation functions.
Delaunay triangulation of P combined with subsequent linear or cubic interpolation within each triangle [40, 134] has also been applied to construct an interpolation surface.
Another popular technique to interpolate scattered data points is to define the interpolation function as a linear combination of radially symmetric basis functions, each centered at a data point. The unknown coefficients for the linear combinations

have to be obtained by solving a linear equation system. For large data sets, the coefficient matrices often suffer from poor condition numbers and need to be preconditioned, cf. [35, 36]. There are several possibilities for choosing the appropriate basis functions, e. g. Gaussian, multiquadratics originally suggested by Hardy [63], thin plate splines or inverse multiquadratics [45].
In image processing, scattered data interpolation or approximation plays an essential role in the reconstruction of images from nonuniform samples and therefore within this community there exist by far more techniques and further improved methods. The reader is referred to [140] for a survey on the application of radial basis functions for image warping and [52] for a survey on nonuniform reconstruction techniques.
The method described by Lee *et al.* [100] is based on a B-Spline approximation technique that has been proposed for image morphing [98, 99] constructing C^2-continuous interpolation functions from arbitrarily scattered data. Thus, this technique has been chosen to be applied to the catalyst data. Data resulting from screening complete composition spreads are not randomly spread within the search space but follow a predefined, regular grid. This gives a special case of scattered data. The algorithm of Lee *et al.* works with a hierarchy of control lattices to generate a sequence of approximation functions which sum up to the desired function. The "heart" of the algorithm is composed of an effective and fast B-Spline approximation technique applied to the different control lattice resolutions. The method is also applicable to multivariate data sets and therefore catalyst data comprising more than 3 elements can also be interpolated. Within the following section, the methodology of the used multilevel B-Spline interpolation algorithm is explained in detail.

3.2 A Multilevel B-Spline Algorithm

Although this interpolation method applies to multivariate data the mathematical background is introduced here for the bivariate case. This leads to the ordinary surface interpolation problem already described above. Let z_k be data values sampled at an arbitrary set of positions (x_k, y_k). Now one aims at finding an appropriate interpolation function $f(x, y)$ such that $f(x_k, y_k) = z_k$ and an evaluation of f at any arbitrary position (x*,y*) becomes available. Together, the triplet (x_k, y_k, z_k) comprises the given scattered data.
Let $\Omega = \{(x, y) | 0 \leq x < m, 0 \leq y < n\}$ be a rectangular domain in the xy-plane. A set of scattered points $P = \{(x_i, y_i) \,|\, 0 \leq x_i < m, 0 \leq y_i < n,\ i = 1, ..., N\} \subset \Omega$ with $z_i = f(x_i, y_i)$ is given. The approximation or interpolation function $\hat{f}$ is formulated as a uniform bicubic B-Spline function defined by a control lattice Φ overlaid on domain Ω. Without loss of generality it is assumed that Φ is an $(m + 3) \times (n + 3)$ lattice that spans the integer grid in Ω, cf. Figure 3.1 page 62. Let ϕ_{ij} be the value of the ij-th control point on lattice Φ with location (i, j) for $i = -1, 0, ..., m + 1$ and $j = -1, 0, ..., n + 1$. Then, the interpolation function $\hat{f}$ can be defined using these control points by

$$\hat{f}(x, y) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(s) B_l(t) \phi_{(i+k)(j+l)} \tag{3.2}$$

where $i = \lfloor x \rfloor - 1, j = \lfloor y \rfloor - 1, s = x - \lfloor x \rfloor$ and $t = y - \lfloor y \rfloor$. Functions B_k and B_l are uniform cubic B-Spline basis functions defined as

$$\begin{aligned} B_0(t) &= (1-t)^3/6 \\ B_1(t) &= (3t^3 - 6t^2 + 4)/6 \\ B_2(t) &= (-3t^3 + 3t^2 + 3t + 1)/6 \\ B_3(t) &= t^3/6 \end{aligned}$$

where $0 \leq t < 1$. The basis functions serve to weight the contribution of each control point to $\hat{f}(x, y)$ based on the distance to (x, y). Then, the problem of finding function $\hat{f}$ is reduced to solving for the control points in Φ that best approximate the data points.

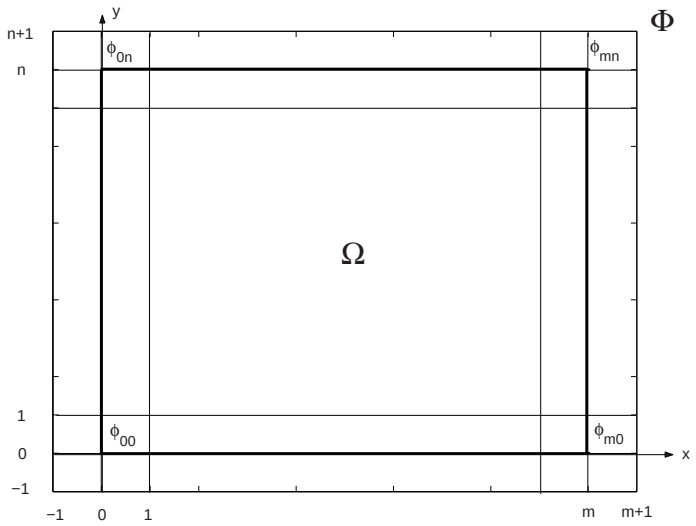


Figure 3.1: Configuration of control lattice Φ defined over domain Ω, cf. [100].

The missing values of control points ϕ_{ij} are computed along the following equations, giving the least square solution to weights ϕ_{ij}, cf. [75, 100].

$$\phi_{ij} = \frac{\sum_{p \in P} w_p^2 \phi_p}{\sum_{p \in P} w_p^2} \tag{3.3}$$

$$\phi_p = \frac{w_p z_p}{\sum_{a=0}^{3} \sum_{b=0}^{3} w_{ab}^2} \tag{3.4}$$

where $w_p = w_{kl} = B_k(s) B_l(t)$, $k = (i+1) - \lfloor x_p \rfloor$, $l = (j+1) - \lfloor y_p \rfloor$, $s = x_p - \lfloor x_p \rfloor$ $t = y_p - \lfloor y_p \rfloor$.

Having all weights ϕ_{ij} of control lattice Φ calculated, one immediately gets the value of function $\hat{f}$ at any location of domain Ω. The smoothness and accuracy of the

approximation and the contribution of each data point to the value of $\hat{f}(x, y)$ at given locations depend primarily on the size and fineness of control lattice Φ. As Φ becomes finer, the influence of a data point is constrained to a smaller neighborhood. This leads to a closer approximation of P but, as a consequence, $\hat{f}$ tends to contain sharp local peaks at the data points. Using this approach, a tradeoff between the smoothness and the accuracy of the approximation function $\hat{f}$ has to be found. This is done by applying a multilevel B-Spline approach with a hierarchy of control lattices generating a sequence of approximation functions $\hat{f}_k$ that result in the desired function $\hat{f}$ when summed up. In this sequence, a function resulting from a coarse control lattice provides a rough approximation which is further refined in its approximation accuracy by functions derived by finer lattices. With Φ being sufficiently fine compared to the data distribution, the scattered points are interpolated by $\hat{f}$.
Let $\hat{f}_1, ..., \hat{f}_K, K \in \mathbb{N}, K > 1$ be a sequence of approximation functions resulting from K different control lattices $\phi_1, ..., \phi_K$ with ϕ_i being of size $(n^i + 3) \times (m^i + 3)$ corresponding to domain Ω_i of size $n^i \times m^i$. Then, the resulting approximation function $\hat{f}$ is defined by

$$\hat{f}(x, y) = \sum_{i=1}^{K} \hat{f}_i(x, y) \tag{3.5}$$

where K is the number of hierarchy levels.
Summarizing, the multilevel B-Spline Approximation (MBA) Algorithm can be formulated as follows [100]:

MBA Algorithm:

1. Let $i = 1, z_P^{(1)} = z_p$ for all $p \in P$ and domain Ω_1 have size $n \times m$.
2. Compute control lattice Φ_i for domain Ω_i of size $n^i \times m^i$ using data points $z_p^{(i)}$ and $\hat{f}_i(x, y)$.
3. Compute residuals $z_p^{(i+1)} = z_p^{(i)} - f_i(x_p, y_p), \qquad p \in P.$
4. Let $i = i + 1$.
5. Repeat process from step 2 for $i = 2, ..., K$.

Part II

Results

4 High-Throughput Screening

In this chapter the results of the high-throughput screening of all synthesized catalyst samples are presented. In total, 2,400 samples have been screened for propene oxidation. Two identical pentanary composition spreads have been studied consisting of elements Cr, Mn, Co, Te and Ni (10% increment) together with 400 catalysts being a refinement (5% increment) around the most active regions of the pentanaries. All catalysts and their compositions are given in the Appendix (cf. pp. 161) together with a unique sample number.

4.1 Sequential High-Throughput Screening of Catalyst Libraries

All the catalysts have been tested under identical conditions in a high-throughput screening reactor system that has been recently developed within the research group [174]. The whole set-up is explained and illustrated in detail within section 7.3. For the screening process, the temperature has been fixed at 350 °C and the composition of the feed gas has been adjusted with 28.4 vol-% propene and 71.6 vol-% synthetic air. Among all possible products that can be monitored after the reactions by GC measurements (cf. p. 135) the focus is laid on the formation of acrolein. Indication for a high activity of a catalyst to form acrolein are large acrolein GC signals (peak areas), in the following referred to as *activity values*.
Throughout the discussion of these results each sample is referred to by its sample number enhanced by 'a' or 'b' corresponding to data set 'A' or 'B'. The two identical pentanary composition spreads A and B have been synthesized sequentially using the same precursor solutions and dispensing robot. The interesting point in doing a replication has been to check the reliability of the synthesis route and the screening system. This chapter addresses the following issues:

1. To validate the whole screening set-up a slate library containing only reference material (Hopcalite) and empty wells has been measured several times to study the reliability of the system.

2. To check the distribution of temperature on a slate library, the temperature in every filled well of the reference library has been recorded by a thermocouple.

3. Due to the limited number of wells on the slate plate (206 wells), the samples of the pentanary composition spread had to be split over 5 different slate libraries and hence, the whole data set is put together by 5 single measurement runs, performed at different days. It is to be checked if this influences the samples behavior and if screening results from these five separate measurements can be put together in one data set.

4. Within the reactor, a temperature gradient of about 20 °C has been observed which surly influences the activity of the catalysts according to their position

on the plate. Higher temperatures should lead to more active samples, which is also a problem in finding 'real' hits.

5. How good has the reproducibility of data sets A and B been? Do the results lie within the same range?

Having discussed all the above mentioned influences, the most active regions within the 6-dimensional search space needed to be determined to reveal the best catalyst for acrolein. This is not a trivial task since six dimensions need to be handled at the same time. So, one challenge has been to find and to develop appropriate visualization approaches to cope with 5D data (6D with activity for acrolein). Chapter 5 is dedicated to the visualization of multivariate screening data.
The main focus throughout this thesis is to validate the two modelling approaches: Kriging and Multilevel B-Splines. It should be checked how predicted activities coincide with the experimental values for 5%-wise compositions. Therefore, the two complete data sets together with the 400 samples of the refinement pose an invaluable prerequisite to discover composition-activity relationships for heterogeneous catalysts. The synthesis of these two pentanary composition spreads also played an important role since such complete data sets from combinatorial high-throughput screening have not been available up to now. It has to be noticed again that the data sets exclusively focus on the composition of the catalysts, all other parameters such as synthesis route, calcination times, pretreatment of the precursors etc. have been kept fixed. This is not always the case when dealing with high-throughput screening data and that makes these data sets so valuable.

4.2 Screening Results of the Reference Library

The reference library has been alternatingly filled by Hopcalite, always leaving empty wells in between. In total, 108 wells containing Hopcalite and 98 empty wells have been tested. Figure 4.1 illustrates the layout of this library. Figure 4.2 (cf. page 67) contains the mean values of the screening indicated by color. Red colored samples show rather high activity for acrolein while blue and dark blue samples correspond to low activities and empty wells that should contain at most traces of acrolein since the wells are not totally isolated for gas diffusion. The difference between Hopcalite samples and empty wells is clearly recognizable, indicating that the screening method works quite well for this reaction. From Figure 4.2 a kind of overall trend can be observed: the most active samples have been screened on inner library positions, indicating that these positions are somehow advantaged. What causes these trends shall be studied more closely within the next sections. In total, three screenings of the library have been performed within eight months.

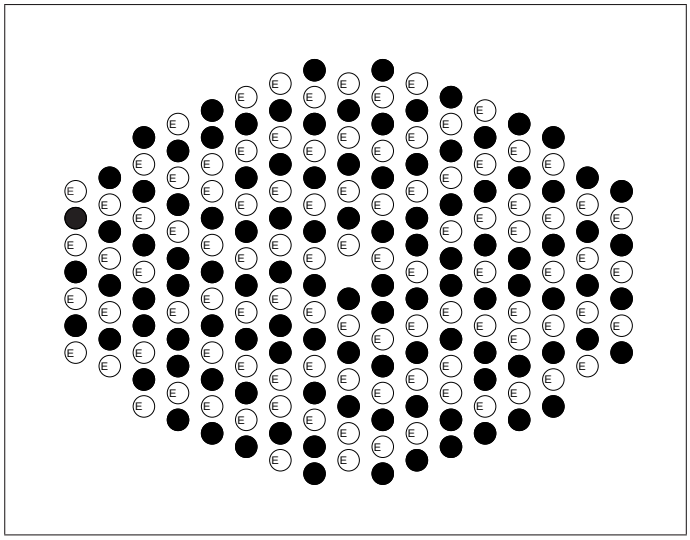


Figure 4.1: Layout of the Reference Library: 108 wells filled with Hopcalite (black wells), 98 empty wells (marked by *E*).

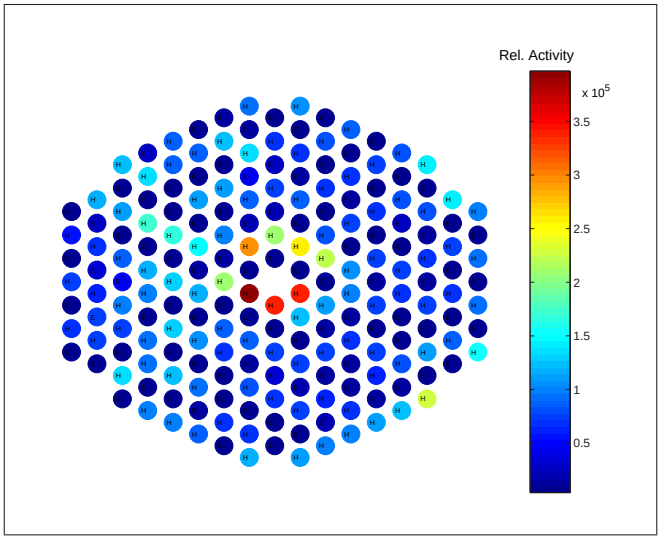


Figure 4.2: Results of the Reference Library (mean values taken over three screenings).

Table 4.1 summarizes the statistics of the measurements. The positive trend toward the inner positions explains the quite large variability between maximum and minimum value leading to large standard deviation values. Nevertheless, the reproducibility of the screening results has been extremely satisfactory, especially for the last two measurement runs. A slight decrease of activity within eight months might be due to some sintering effects caused by the reaction conditions or aging behavior of the catalyst. The advantaging of inner library positions might be due to a higher temperature within this area of the plate. Therefore, the temperature distribution of the reference library has been studied in more detail as shown in the next section. The repeated screening of the reference library also provided enough valuable data to determine a 95% -confidence interval for the Hopcalite values following the theory

of observation errors (Student's t distribution) . For the three measurements this gives:

	10/27/06	05/24/06	06/01/06
Minimal Value	$7.34 \cdot 10^4$	$5.00 \cdot 10^4$	$2.96 \cdot 10^4$
Maximal Value	$3.66 \cdot 10^5$	$4.12 \cdot 10^5$	$4.23 \cdot 10^5$
Median	$1.20 \cdot 10^5$	$7.70 \cdot 10^4$	$7.42 \cdot 10^4$
Mean $\bar{X}$	$1.37 \cdot 10^5$	$9.82 \cdot 10^4$	$9.25 \cdot 10^4$
Standard Deviation s_x	$6.36 \cdot 10^4$	$5.98 \cdot 10^4$	$5.81 \cdot 10^4$
Number of Measurements: n	108	108	108
Deviation from Mean: $\frac{s_x}{\sqrt{n}}$	$6.11 \cdot 10^3$	$5.75 \cdot 10^3$	$5.59 \cdot 10^3$
t-factor for $n = 108$	1.98	1.98	1.98
Deviation:$1.98 \cdot \frac{s_x}{\sqrt{n}}$	$1.21 \cdot 10^4$	$1.13 \cdot 10^4$	$1.10 \cdot 10^4$
Deviation [%]	8.83	11.40	11.89
Interval	$1.37 \cdot 10^5$ +/- $1.21 \cdot 10^4$	$9.82 \cdot 10^4$ $+/- 1.13 \cdot 10^4$	$9.25 \cdot 10^4$ $+/- 1.10 \cdot 10^4$

Table 4.1: Measurement statistics of the reference library.

Table 4.1 demonstrates how large the deviation among the Hopcalite samples has been. Several statistics for each screening containing the two extrema, median and mean value and the standard deviation are considered. Since only 108 samples have been measured and no normal distribution of the measurement values can be assumed the sample mean value does not represent the measurements in a correct way. Therefore, the stochastic deviation of the measurements u_{rand} is calculated to give the borders of a 95% - confidence interval. Following the theory of observation errors (Student's t distribution) [1], the random deviation among the measurements is given by the quotient:

$$u_{rand} = t_{p,n} \frac{s_x}{\sqrt{n}}$$

where s_x being the standard deviation, n being the number of measurements and p the desired statistical confidence level. In this case, $t_{0.05,108} = 1.98$, cf. page 1123 in [17] and the corresponding confidence intervals can be determined as has been done in table 4.1.
Furthermore the table also contains the percentage deviation, ranging from 8.83% for the first screening of the library to 11.89% for the third screening. That means, the actual activity values of a sample can be determined with a deviation of approximately 10% inaccuracy. For a high-throughput screening technique this result is acceptable.

4.3 Temperature Profile of the Reference Library

To establish a temperature profile of the filled reference library the temperatures in each filled well have been monitored by a thermocouple. Figure 4.3 shows the results. Obviously, the temperature is not constantly distributed over the library plate, it increases from left to right with a definite maximum around the inner positions. The maximum measured temperature of 356 °C has been reached around the center, while the minimal measured temperature of 334 °C appears at the left margin. This means, a temperature gradient of 22 °C has been observed spread over the library. This may be caused by the architecture of the reactor heating system or an unsteady heat transfer between slate and steel, while the draft of the hood may contribute, too. The lower temperatures at the margins obviously are due to insufficient insulation. This definitely has an influence on the activity of a catalyst sample since the speed of a catalytic reaction directly depends on the temperature.

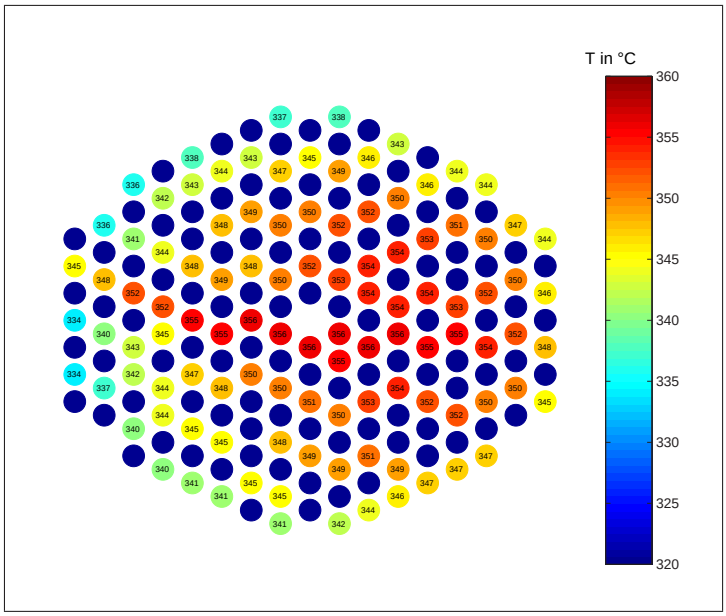


Figure 4.3: Temperature profile of the reference library (dark blue wells correspond to empty wells with no temperature measurement).

This means within the screening data it can be observed that many of the most active samples are situated on inner library positions giving the danger of too many "false positives". Therefore, a strategy to filter the data according to that influence is proposed. This can be done by multiplying each activity value with an adequate factor depending on the position and therefore the temperature of the sample. After this correction, it might happen that the most active samples "loose" some activity while other samples, less active, gain activity, especially those samples measured at lower temperatures situated at outer positions.

Strategy: Correction of Temperature/Position Influence

For the correction of trends within the screening data a temperature value for each well is needed. Since the temperature could have been only measured within the

filled wells appropriate temperature values for the empty positions are required. This has been done by using a Kriging model to calculate the temperatures at the empty well positions. Then, each well/position of the library corresponds to a certain temperature value. Figure 4.4 shows the results of the kriged temperature distribution.

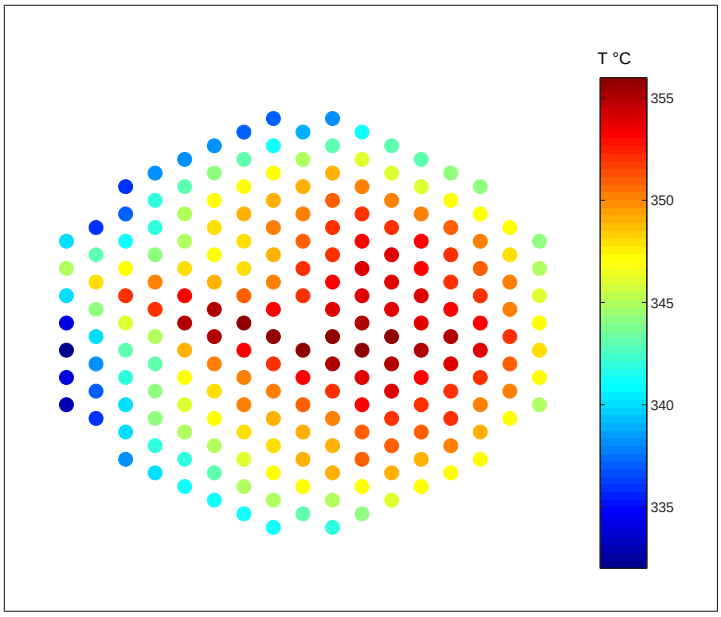


Figure 4.4: Results of kriging the missing temperature values.

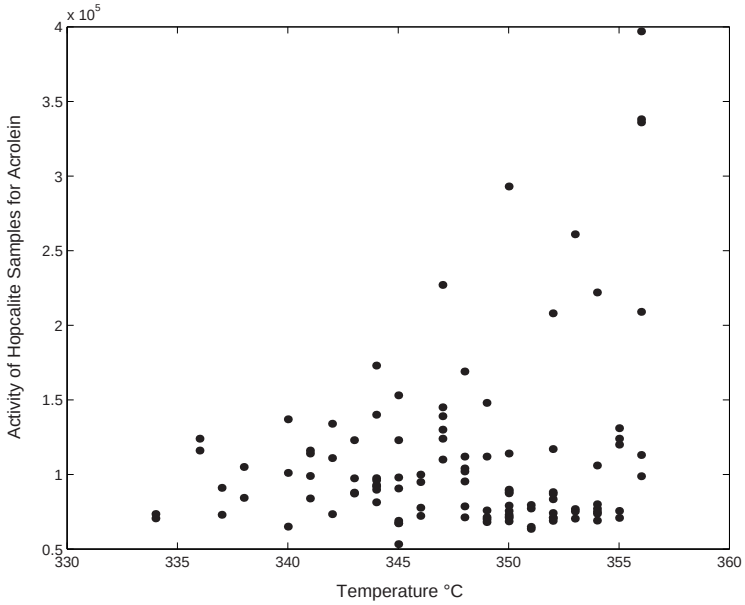


Figure 4.5: Activity versus temperature of the reference library.

As a main aspect, the dependence of activity on temperature need to be studied and the two plots above give an idea of this correlation.

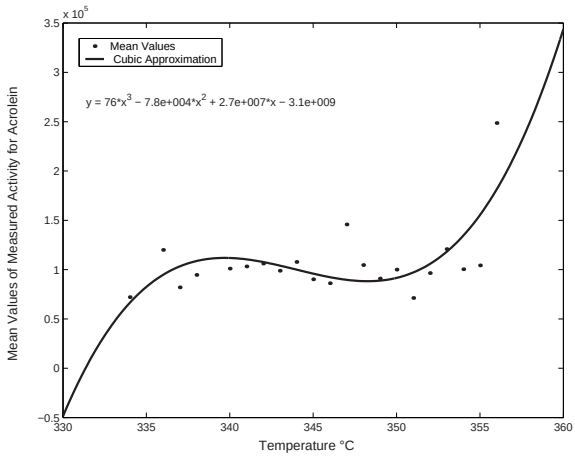


Figure 4.6: Mean values of activity versus temperature of the reference library fitted by a cubic polynomial.

It can be observed that there is a positive trend towards higher temperatures indicating that the activity of the samples is positively influenced by higher temperature values. Since some temperatures have been measured more often than others it was decided to work with the mean activity values of the measured activities, see Figure 4.6. Then, an activity value for each empty well can be calculated by a fitting polynomial and a complete library is obtained with 206 activity and temperature values based on the temperature and activity measurements extended by modelled values. The data can then be normalized to the mean activity value at the desired temperature of 350 °C by assigning a scaling factor to each well/position. This factor will be larger than 1 for positions with low temperatures or low activity values and less than 1 for regions with high measured activities due to high temperatures. After this normalization, all the samples of the Hopcalite library show the same activity as should be theoretically the case. When screening the pentanary composition spreads also a trend has been observed that many active samples occurred near the innermost positions indicating the consequences discussed above. Thus, this normalization strategy has also been applied to the screened catalyst libraries to correct the temperature influence of inner positions. Of course, it has to be assumed here that all synthesized materials show a similar heat transfer than the Hopcalite material does although materials may behave differently. To calculate the normalization factors according to the positions seems to be a good approach since all libraries have been measured in the same way such that every position of the library exactly sits at the same place in the reactor. Thus, in every run the screening has been realized under the same conditions yielding the same errors and influences.
It can be also shown that the use of steel libraries leads by far to smaller temperature gradients than slate, even at operating temperatures above 400 °C.

4.4 Screening Identical Libraries

To study the reliability of the screening system for the oxidation of propene a pentanary composition spread has been synthesized twice under exactly the same condi-

tions. All samples have been sequentially synthesized within three days keeping all other parameters fixed. For a complete pentanary composition spread extended by reference samples five different slate plates have been needed to place all the catalyst powders. In total, 10 (2 × 5) libraries have been filled to perform the screening at 350 °C. The exact distribution of the samples on the slate plates can be found in the Appendix (see pp. 173). The following figures show the unscaled results (raw data) of screening the 10 libraries for propene oxidation. For these plots, the same scaling has been applied such that very active samples are colored in dark red and can be identified at once. Another important aspect here was to check the reproducibility.

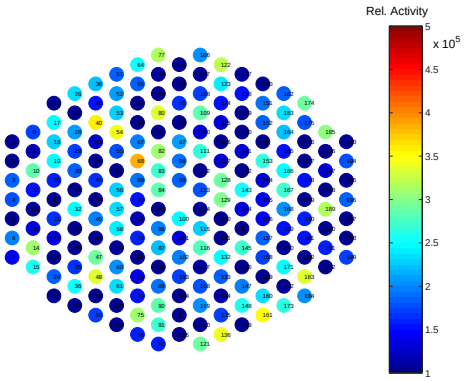


Figure 4.7: Results of ht-screening of library 1a (unscaled).

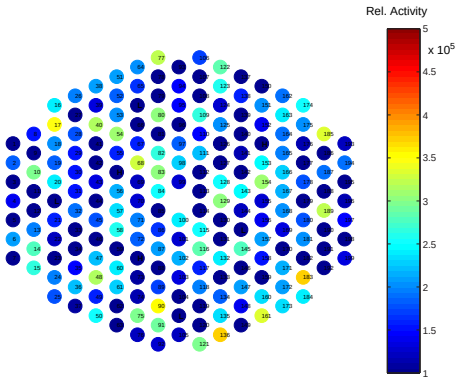


Figure 4.8: Results of ht-screening of library 1b (unscaled).

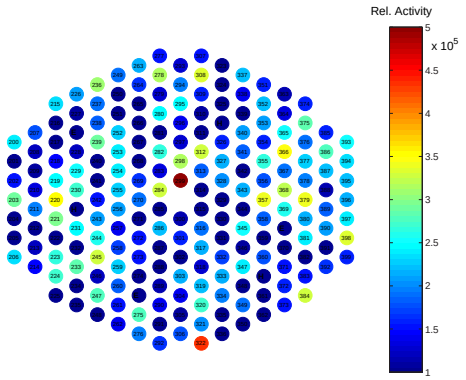


Figure 4.9: Results of ht-screening of library 2a (unscaled).

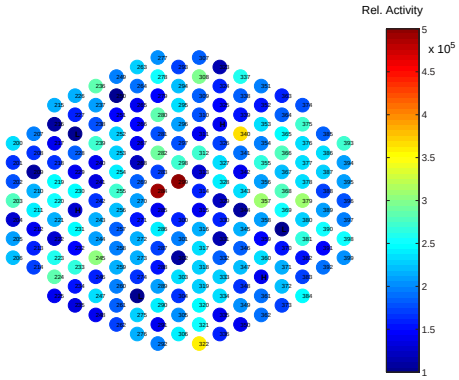


Figure 4.10: Results of ht-screening of library 2b (unscaled).

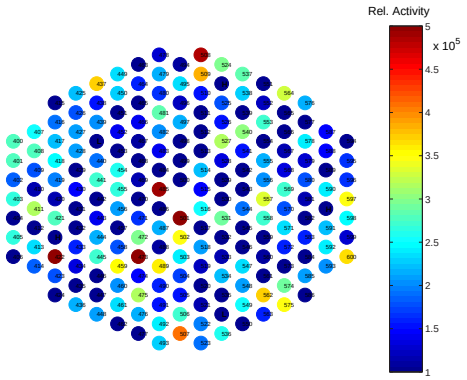


Figure 4.11: Results of ht-screening of library 3a (unscaled).

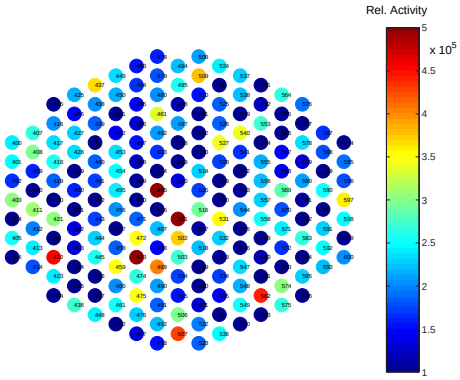


Figure 4.12: Results of ht-screening of library 3b (unscaled).

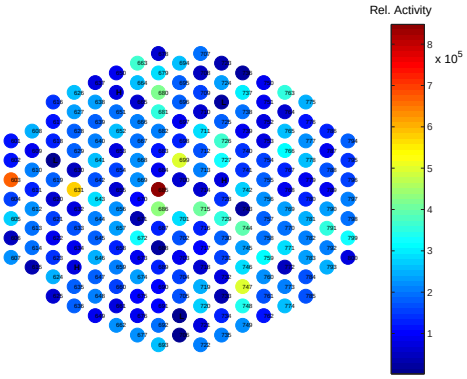


Figure 4.13: Results of ht-screening of library 4a (unscaled).

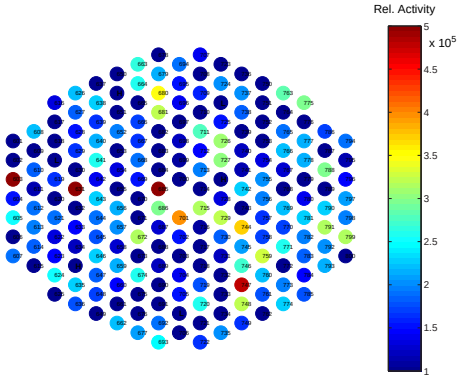


Figure 4.14: Results of ht-screening of library 4b (unscaled).

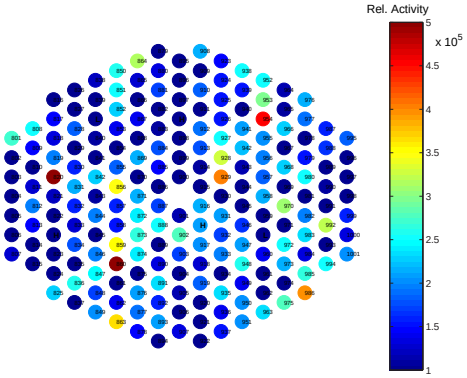


Figure 4.15: Results of ht-screening of library 5a (unscaled).

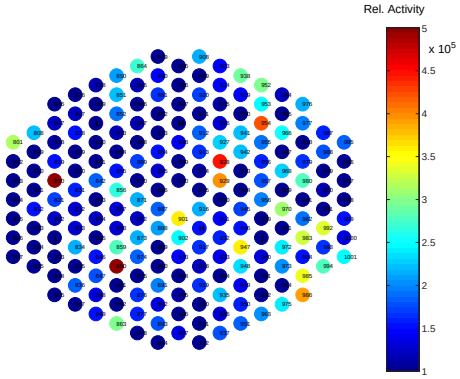


Figure 4.16: Results of ht-screening of library 5b (unscaled).

A pairwise comparison of the libraries showed in general good reproducibiloity.

- Libraries 1a and 1b:
  Sample 68 achieved higher GC signals on plate 1a, the results of other samples on these plates could be well reproduced.

- Libraries 2a and 2b:
  The reproducibility of libraries 2a and 2b has also been satisfactory. For example, samples 322 and 299 showed in both cases comparably good activities.

- Libraries 3a and 3b:
  These two libraries showed the best reproducibility amongst all pairs. Except for sample 508 nearly no discrepancies were observed. All very active samples of library 3a also showed high activities in their reproduction on plate 3b. Sample 501 showed reproducibly the highest activity of all libraries. Sample 473 appeared to be the third-best catalyst in both data sets.

- Libraries 4a and 4b:
  The leading samples yielded high activity values on both libraries. Sample 685 appeared to be the second-best catalyst in both cases. Remarkably here are the good activity values of sample 603 situated at an outer position where a lower temperature value may cause rather a decrease in activity. In general, the overall reproducibility among both plates is well succeeded.

- Libraries 5a and 5b:
  The results of the leading samples could have been reproduced and libraries 5a and 5b show a good reproducibility. Sample 901 shows less activity on library 5a than on library 5b but nearly all other samples follow the same pattern.

The following two tables illustrate the results of the 25 best catalysts (data without scaling) from both screenings given their composition and position on the plates.

Ranking	**Sample Number**	**X**	**Y**	**Cr**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein Signal**
1	**501**	9	10	0.1	0	0.2	0.4	0.3	$9.84 \cdot 10^5$
2	**685**	8	8	0.1	0	0.4	0.3	0.2	$8.48 \cdot 10^5$
3	**473**	7	12	0	0.2	0.7	0.1	0	$8.02 \cdot 10^5$
4	860	6	12	0	0.1	0.8	0.1	0	$6.95 \cdot 10^5$
5	820	3	8	0	0.2	0.6	0.1	0.1	$6.73 \cdot 10^5$
6	603	1	8	0	0.3	0.5	0.1	0.1	$6.67 \cdot 10^5$
7	299	9	8	0.1	0.2	0.4	0.3	0	$6.04 \cdot 10^5$
8	631	4	8	0	0.2	0.5	0.1	0.2	$5.80 \cdot 10^5$
9	422	3	12	0	0.1	0.7	0.1	0.1	$5.61 \cdot 10^5$
10	747	12	13	0	0.3	0.3	0.1	0.3	$4.96 \cdot 10^5$
11	699	9	7	0.9	0	0	0.1	0	$4.92 \cdot 10^5$
12	485	8	8	0.1	0	0.3	0.6	0	$4.89 \cdot 10^5$
13	508	10	1	0.1	0	0.3	0.4	0.2	$4.87 \cdot 10^5$
14	954	13	5	0	0.3	0.6	0.1	0	$4.54 \cdot 10^5$
15	686	8	9	0.1	0.4	0	0.3	0.2	$4.38 \cdot 10^5$
16	322	10	16	0.1	0.2	0.6	0.1	0	$4.34 \cdot 10^5$
17	680	8	3	0.1	0.1	0.7	0.1	0	$4.14 \cdot 10^5$
18	507	9	16	0	0.4	0.2	0.1	0.3	$4.11 \cdot 10^5$
19	929	11	8	0	0.1	0.6	0.1	0.2	$4.04 \cdot 10^5$
20	715	10	9	0	0.3	0.4	0.1	0.2	$4.00 \cdot 10^5$
21	986	15	14	0.1	0	0.4	0.4	0.1	$3.96 \cdot 10^5$
22	744	12	10	0.1	0.3	0.3	0.3	0	$3.92 \cdot 10^5$
23	509	10	2	0	0.4	0.3	0.1	0.2	$3.85 \cdot 10^5$
24	68	7	7	0	0.1	0.5	0.1	0.3	$3.82 \cdot 10^5$
25	437	5	3	0	0.2	0.4	0.1	0.3	$3.70 \cdot 10^5$

Table 4.2: Best catalysts of data set A after ht-screening (unscaled data).

Comparing the unscaled screening results of data set A and B it can be observed that the leading group, i.e. the best three catalysts are *exactly* the same samples in both cases. That illustrates the good reliability of the synthesis route together with the screening system. Among the best 25 catalysts of each data set, 17 catalysts are contained in both leading groups, indicating again a good reproducibility of the measurement results.

Ranking	Sample Number	X	Y	Cr	Mn	Co	Te	Ni	Acrolein Signal
1	**501***	9	10	0.1	0.0	0.2	0.4	0.3	$9.44 \cdot 10^5$
2	**685***	8	8	0.1	0.0	0.4	0.3	0.2	$7.86 \cdot 10^5$
3	**473***	7	12	0.0	0.2	0.7	0.1	0.0	$7.60 \cdot 10^5$
4	485*	8	8	0.1	0.0	0.3	0.6	0.0	$7.34 \cdot 10^5$
5	299*	9	8	0.1	0.2	0.4	0.3	0.0	$7.33 \cdot 10^5$
6	284	8	8	0.1	0.1	0.2	0.5	0.1	$7.21 \cdot 10^5$
7	603*	1	8	0.0	0.3	0.5	0.1	0.1	$6.37 \cdot 10^5$
8	631*	4	8	0.0	0.2	0.5	0.1	0.2	$5.53 \cdot 10^5$
9	820*	3	8	0.0	0.2	0.6	0.1	0.1	$5.44 \cdot 10^5$
10	860*	6	12	0.0	0.1	0.8	0.1	0.0	$5.11 \cdot 10^5$
11	340	12	5	0.0	0.4	0.4	0.1	0.1	$4.84 \cdot 10^5$
12	747*	12	13	0.0	0.3	0.3	0.1	0.3	$4.80 \cdot 10^5$
13	322*	10	16	0.1	0.2	0.6	0.1	0.0	$4.79 \cdot 10^5$
14	422*	3	12	0.0	0.1	0.7	0.1	0.1	$4.56 \cdot 10^5$
15	928	11	7	0.2	0.0	0.1	0.4	0.3	$4.49 \cdot 10^5$
16	562	13	14	0.1	0.3	0.4	0.1	0.1	$4.41 \cdot 10^5$
17	507*	9	16	0.0	0.4	0.2	0.1	0.3	$4.28 \cdot 10^5$
18	954*	13	5	0.0	0.3	0.6	0.1	0.0	$4.22 \cdot 10^5$
19	489	8	12	0.1	0.3	0.0	0.4	0.2	$4.01 \cdot 10^5$
20	701	9	10	0.0	0.0	0.0	0.3	0.7	$3.97 \cdot 10^5$
21	929*	11	8	0.0	0.1	0.6	0.1	0.2	$3.95 \cdot 10^5$
22	986*	15	14	0.1	0.0	0.4	0.4	0.1	$3.89 \cdot 10^5$
23	379	15	9	0.1	0.2	0.4	0.1	0.2	$3.88 \cdot 10^5$
24	357	13	9	0.1	0.4	0.0	0.4	0.1	$3.88 \cdot 10^5$
25	509*	10	2	0.0	0.4	0.3	0.1	0.2	$3.79 \cdot 10^5$

Table 4.3: Best catalysts of data set B after ht-screening (unscaled data). Samples marked by * are also contained among the best 25 samples of data set A.

In a next step, the same ranking of the catalyst performance was done after having removed the temperature influence in the data. Tables 4.4 and 4.5 show the ranking after the data have been multiplied by the scaling factor according to their position. After scaling, sample 473 (originally on third position) takes the leading position in both data sets. Comparing the unscaled with the scaled data sets one notices slight changes within the ranking of the top 25 catalysts. For data set A, 11 catalysts remain among the leading 25 samples after scaling while 14 catalysts are kept for data set B. Considering the best three catalysts before and after the scaling, a group of 6 samples can be identified being the most active ones: 473, 685, 860, 820, 501 and 299.

Ranking	Sample Number	X	Y	Cr	Mn	Co	Te	Ni	Acrolein Signal	Scaling Factor
1	473	7	12	0	0.2	0.7	0.1	0	$8.02 \cdot 10^5$	1.00
2	860	6	12	0	0.1	0.8	0.1	0	$7.75 \cdot 10^5$	1.12
3	820	3	8	0	0.2	0.6	0.1	0.1	$7.64 \cdot 10^5$	1.14
4	685	8	8	0.1	0	0.4	0.3	0.2	$7.19 \cdot 10^5$	0.85
5	603	1	8	0	0.3	0.5	0.1	0.1	$5.97 \cdot 10^5$	0.89
6	299	9	8	0.1	0.2	0.4	0.3	0	$5.71 \cdot 10^5$	0.95
7	401	1	7	0	0.1	0.1	0.4	0.4	$5.18 \cdot 10^5$	1.88
8	747	12	13	0	0.3	0.3	0.1	0.3	$5.11 \cdot 10^5$	1.03
9	422	3	12	0	0.1	0.7	0.1	0.1	$5.02 \cdot 10^5$	0.89
10	631	4	8	0	0.2	0.5	0.1	0.2	$4.98 \cdot 10^5$	0.86
11	748	12	14	0.2	0.1	0.4	0.3	0	$4.92 \cdot 10^5$	1.43
12	411	2	9	0.2	0	0.2	0.4	0.2	$4.81 \cdot 10^5$	1.54
13	75	7	15	0.1	0.4	0.2	0.1	0.2	$4.63 \cdot 10^5$	1.49
14	508	10	1	0.1	0	0.3	0.4	0.2	$4.63 \cdot 10^5$	0.95
15	954	13	5	0	0.3	0.6	0.1	0	$4.54 \cdot 10^5$	1.00
16	189	16	9	0.2	0	0.3	0.4	0.1	$4.48 \cdot 10^5$	1.41
17	437	5	3	0	0.2	0.4	0.1	0.3	$4.40 \cdot 10^5$	1.19
18	863	6	15	0.2	0.5	0	0.3	0	$4.37 \cdot 10^5$	1.19
19	763	14	3	0.1	0.4	0.1	0.3	0.1	$4.33 \cdot 10^5$	1.23
20	970	14	9	0.1	0.4	0.4	0.1	0	$4.33 \cdot 10^5$	1.41
21	14	2	11	0.2	0.2	0	0.3	0.3	$4.27 \cdot 10^5$	1.37
22	368	14	8	0.1	0.2	0	0.3	0.4	$4.26 \cdot 10^5$	1.33
23	759	13	12	0.1	0.2	0	0.4	0.3	$4.25 \cdot 10^5$	1.35
24	605	1	10	0.2	0.2	0	0.4	0.2	$4.22 \cdot 10^5$	1.51
25	562	13	14	0.1	0.3	0.4	0.1	0.1	$4.16 \cdot 10^5$	1.13

Table 4.4: Best catalysts of data set A after ht-screening, scaled data. Sample 473 yielded the best activity value.

It is interesting to notice that the second and third best samples of data set B (after scaling) have a scaling factor below 1 which means that their original activity values are decreased by a certain amount since they are located on inner plate positions. Nevertheless, they possess very good activity values. Furthermore sample 473 has a factor of 1 causing no activity change such that the result of this catalyst stays very good. In the scaled data set A, the second and third best catalysts (860 and 820) have been assigned a scaling factor larger than one leading to a slight improvement of their activity values due to their plate positions. Another interesting results has been obtained for sample 603 situated at an outer position of the library. This sample yielded very good activity values during the screening and can be found among the best 10 catalyst within all tables. Analyzing the data sets in this way helped to get a first impression of the best catalyst samples. Fortunately, removing the temperature trends out of the data did not totally disturb the original ranking and a leading group of the 6 catalysts mentioned above could be identified.

Ranking	Sample Number	X	Y	Cr	Mn	Co	Te	Ni	Acrolein Signal	Scaling Factor
1	473*	7	12	0	0.2	0.7	0.1	0	$7.60 \cdot 10^5$	1.00
2	299*	9	8	0.1	0.2	0.4	0.3	0	$6.93 \cdot 10^5$	0.95
3	685*	8	8	0.1	0	0.4	0.3	0.2	$6.67 \cdot 10^5$	0.85
4	485	8	8	0.1	0	0.3	0.6	0	$6.22 \cdot 10^5$	0.85
5	820*	3	8	0	0.2	0.6	0.1	0.1	$6.18 \cdot 10^5$	1.14
6	284	8	8	0.1	0.1	0.2	0.5	0.1	$6.11 \cdot 10^5$	0.85
7	860*	6	12	0	0.1	0.8	0.1	0	$5.70 \cdot 10^5$	1.12
8	603*	1	8	0	0.3	0.5	0.1	0.1	$5.70 \cdot 10^5$	0.89
9	928	11	7	0.2	0	0.1	0.4	0.3	$5.61 \cdot 10^5$	1.25
10	947	12	11	0.1	0.2	0.3	0.3	0.1	$5.18 \cdot 10^5$	1.45
11	562*	13	14	0.1	0.3	0.4	0.1	0.1	$4.97 \cdot 10^5$	1.13
12	747*	12	13	0	0.3	0.3	0.1	0.3	$4.95 \cdot 10^5$	1.03
13	340	12	5	0	0.4	0.4	0.1	0.1	$4.84 \cdot 10^5$	1.00
14	411*	2	9	0.2	0	0.2	0.4	0.2	$4.80 \cdot 10^5$	1.54
15	631*	4	8	0	0.2	0.5	0.1	0.2	$4.74 \cdot 10^5$	0.86
16	401*	1	7	0	0.1	0.1	0.4	0.4	$4.65 \cdot 10^5$	1.88
17	748*	12	14	0.2	0.1	0.4	0.3	0	$4.63 \cdot 10^5$	1.43
18	189*	16	9	0.2	0	0.3	0.4	0.1	$4.56 \cdot 10^5$	1.41
19	489	8	12	0.1	0.3	0	0.4	0.2	$4.52 \cdot 10^5$	1.13
20	759*	13	12	0.1	0.2	0	0.4	0.3	$4.47 \cdot 10^5$	1.35
21	203	1	9	0.2	0	0	0.5	0.3	$4.46 \cdot 10^5$	1.36
22	368*	14	8	0.1	0.2	0	0.3	0.4	$4.45 \cdot 10^5$	1.33
23	75*	7	15	0.1	0.4	0.2	0.1	0.2	$4.42 \cdot 10^5$	1.49
24	437*	5	3	0	0.2	0.4	0.1	0.3	$4.36 \cdot 10^5$	1.19
25	970*	14	9	0.1	0.4	0.4	0.1	0	$4.35 \cdot 10^5$	1.41

Table 4.5: Best catalysts of data set B after ht-screening, scaled data. Samples marked by * are also contained among the best 25 samples of the scaled data set A.

How good the reproduction of the whole data set has been can be seen in Figure 4.17 where the activity results of all 1001 samples are plotted. The x-axis denotes the activity values of data set A, the y-axis the activity values of data set B. In case of a perfect matching this plot should results in a bisecting line of an angle. The best catalyst samples for the unscaled case are 501, 685 and 473 which can be easily identified in the plot lying in the outermost right corner. It is remarkable that catalyst ranking among the best 25 samples show a good reproducibility and appear close to the bisecting line. The overall trend of this plot confirms the results already discussed above in an elegant way. How the data behave after the scaling process is illustrated in Figure 4.18. Here, it can be observed that the best sample after the scaling process is catalyst 473. Its position in the outmost right corner of the coordinate system indicates the best activity values in both data sets. Also the leading groups of the best catalyst show a very good reproducibility and differ noticeably from the rest of the samples. Sample 501 lost its good position due to an inner library position and a small scaling factor of 0.3. It fell back to position 129 (data set A) and 153 (data set B) in the overall ranking.

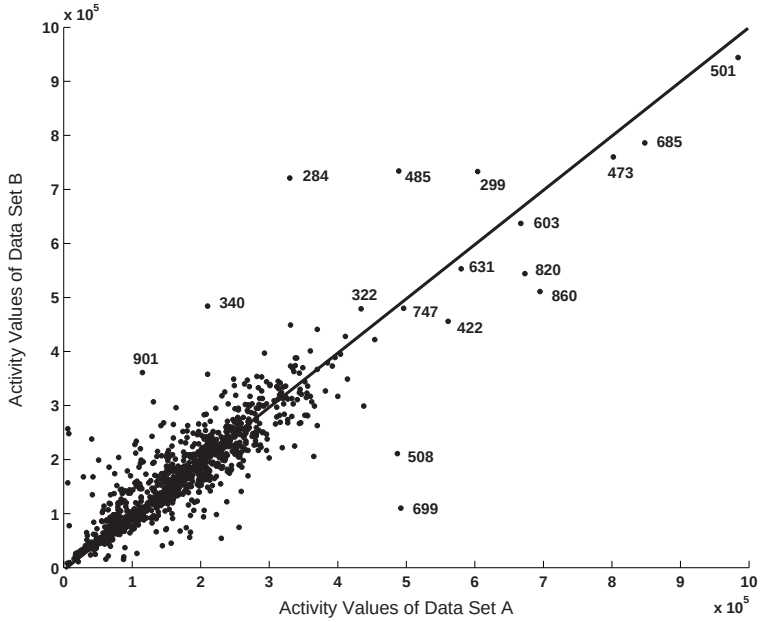


Figure 4.17: Comparison of the unscaled data sets A and B with respect to the measured activity values. The leading group with samples 501, 685 and 473 clearly catches the eye.

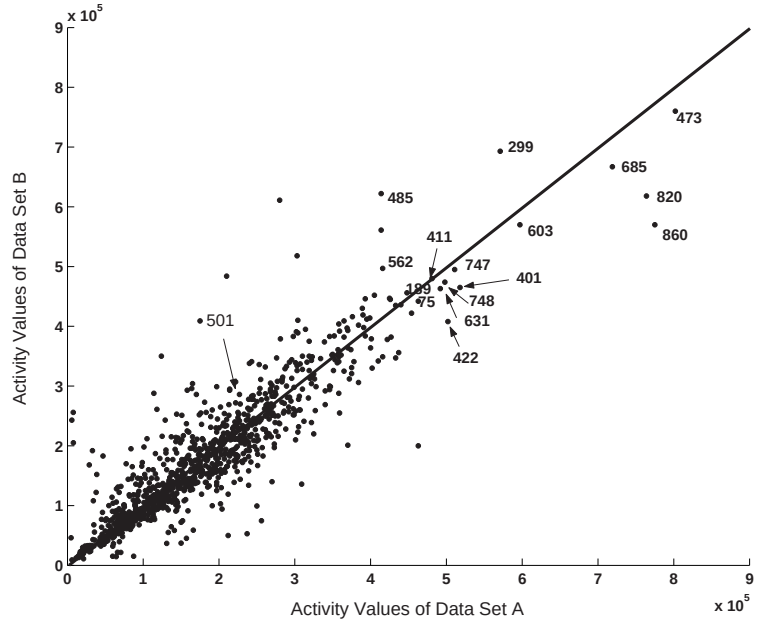


Figure 4.18: Comparison of the scaled data sets A and B with respect to the measured activity values. The best sample 473 appears in the outmost right corner. Sample 501 lost its good position due to a scaling factor of only 0.3.

Other very good catalysts that got factors smaller than one due to their inner library positions still remain among the best samples. Furthermore, the leading group has been supplemented with samples from outer positions due to scaling factors larger

than one. An interesting result has been obtained for sample 603 lying at the margin position (1,8). This sample has been already very active in both original data sets and by its factor of 1.56 it even gains more activity by the scaling. Thus, sample 603 promises to be a really good candidate.

4.5 Splitting the Samples onto Five Libraries

To screen a complete pentanary composition spread consisting of 1001 samples five slate plates were needed. Thus, the measurements have been carried out sequentially during several days which might have influenced the results. To check this the screening results of the reference samples located at each library have been considered and it was controlled if their results lie within the same range. The following two figures illustrate the measured results by box-plots.

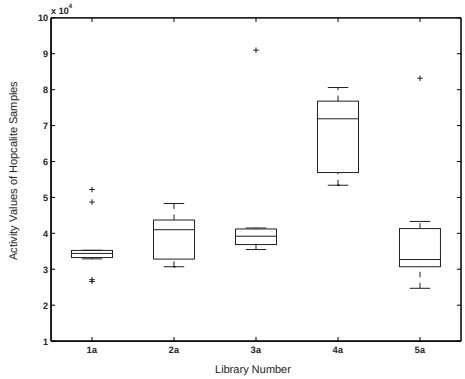


Figure 4.19: Box plot of reference samples of libraries 1a-5a.

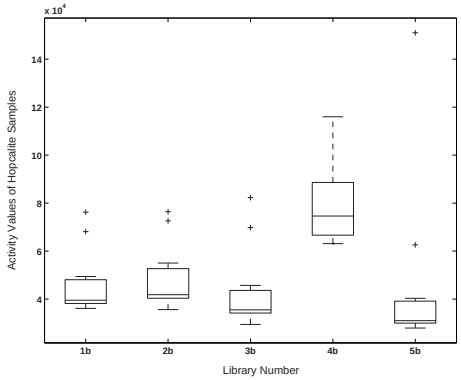


Figure 4.20: Box plot of reference samples of libraries 1b-5b.

The screening results of the Hopcalite samples on different libraries showed good accordance among each other. The median values of libraries 1a, 2a, 3a and 5a appeared within the same range, a satisfying result. The same can be noted with respect to libraries 1b, 2b, 3b and 5b. It has been interesting to see that the screening of the Hopcalite samples of both libraries 4a and 4b yielded higher activity values that lie far outside the 95%-confidence interval. Why this is the case for these two libraries cannot be explained. But in 8 out of ten cases the Hopcalite samples yielded results within the desired range such that a weighting of the results according to the library is neglected.

4.6 Screening Results of the Ternary Composition Spreads

Within this section the most interesting screening results among the ten ternary composition spreads that can be extracted out of the complete pentanary system are presented. The left figure always corresponds to data set A while the right figure contains the results out of data set B. For the following figures, the color corresponds to the measured GC signals for acrolein, giving a sort of "relative activity" for the catalyst samples. The most active regions have been finer sampled (5%-wise increments) by a second generation of catalysts, to validate the modelling approaches. These compositions have been marked by small black points.

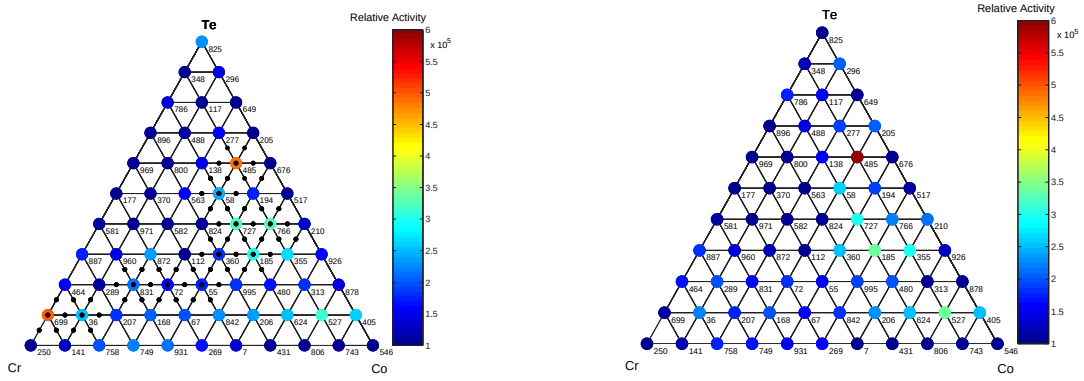


Figure 4.21: Screening results of the ternary system CrCoTe A (left) and CrCoTe B (right) without scaling.

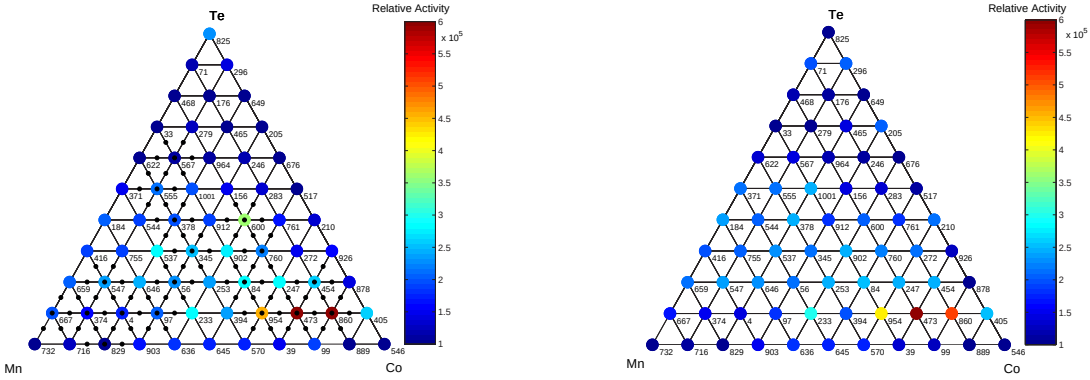


Figure 4.22: Screening results of the ternary system MnCoTe A (left) and MnCoTe B (right) without scaling.

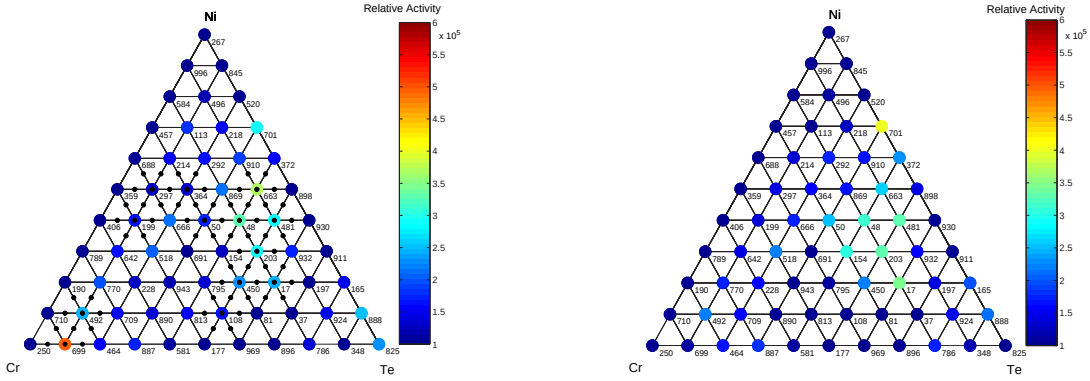


Figure 4.23: Screening results of the ternary system CrTeNi A (left) and CrTeNi B (right) without scaling.

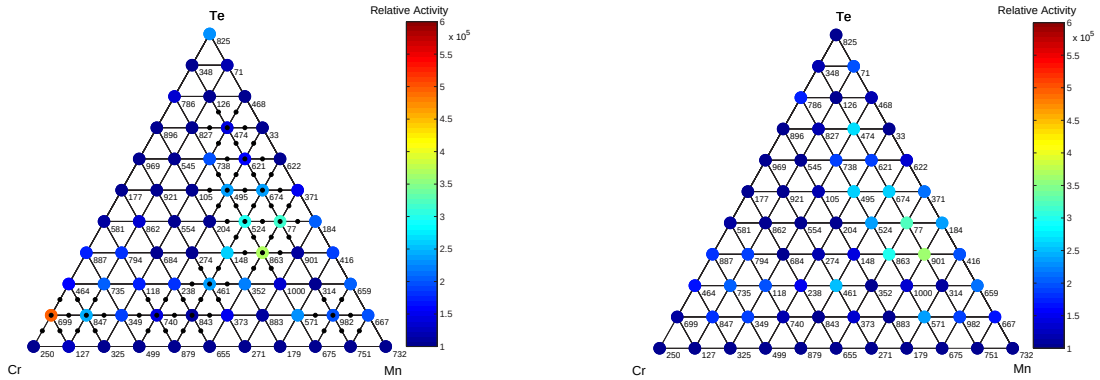


Figure 4.24: Screening results of the ternary system CrMnTe A (left) and CrMnTe B (right) without scaling.

According to Table 4.2 the ternary catalysts 473 and 860 show very good activities. It is very interesting that these samples are neighbors within the same system MnCoTe, cf. Figure 4.22. This means, the region around samples 473 and 860 appears to be a region of interest for further investigations. The same can be deduced for samples 699 and 485, although the high activity value of sample 699 only occurred in data set A. The activity of 485 has been verified in data set B. These representations are as important as the statistical treatment since they visualize the consistency of the activity-composition dependence. Outliers are identified by significant differences to the next neighbors, while active materials should be surrounded by other active samples. Avtivity is assumed to change gradually.

4.7 Screening Results of the Quaternary Composition Spreads

Analogously to the section above the screening results of the five quaternary systems are presented within this section. Quaternary composition spreads can be represented by tetrahedra, as has been described within section 2.7.2. In this representation, the ternary systems that build the equilateral triangles at the four faces of the tetrahedra have been removed. Then, only real quaternary compositions are drawn that are normally covered by ternaries. Nevertheless, it is not as easy as with the ternary case to determine the composition of an active sample by just looking at the plots - this requires some practice. Using appropriate software tools, these drawbacks of a fixed representation as tetrahedra can be decreased. Within the next chapter it will be explained, how an interactive representation of quaternary composition spreads may facilitate the interpretation. For the following figures, the results from data set A (left sides) and B (right sides) are again presented in direct comparison.
Figure 4.25 shows a quaternary system consisting of the elements Mn, Co, Te, and Ni. Again, the high reproducibility of the samples can be noted and especially samples 747, 631, 422, 820 and 603 showed comparable high activities in both data sets. For the preparation of a second library generation this area has been identified as a region of interest to be sampled finer by a second generation of catalysts.

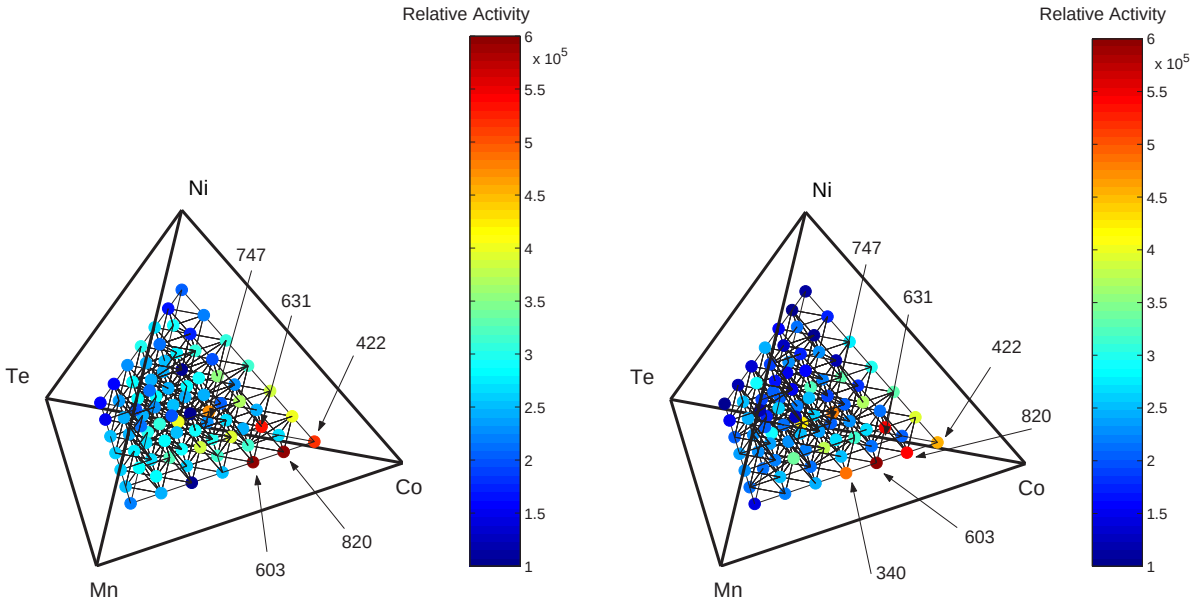


Figure 4.25: Screening results of the quaternary system MnCoTeNi A (left) and MnCoTeNi B (right) without scaling.

Within Figure 4.26 the most active samples of both data sets can be clearly identified verifying the good reproducibility of the synthesis method and the screening. Samples 501 and 685 together with a small neighborhood of surrounding points have been synthesized again to check the power of the modelling approaches. These results will be discussed in Chapter 6.

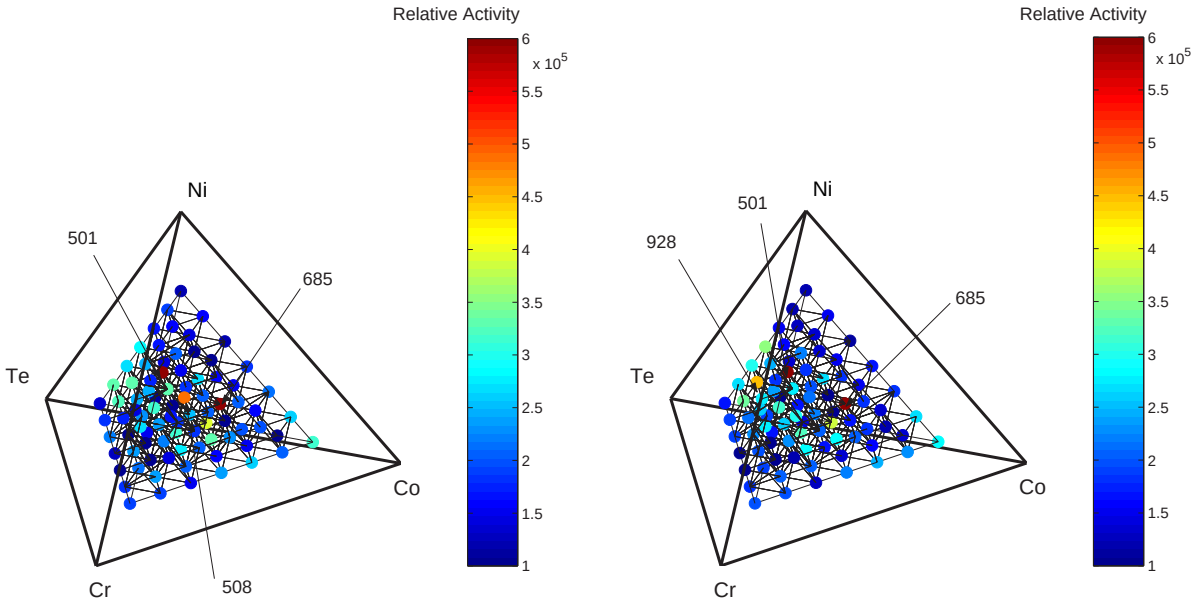


Figure 4.26: Screening results of the quaternary system CrCoTeNi A (left) and CrCoTeNi B (right) without scaling.

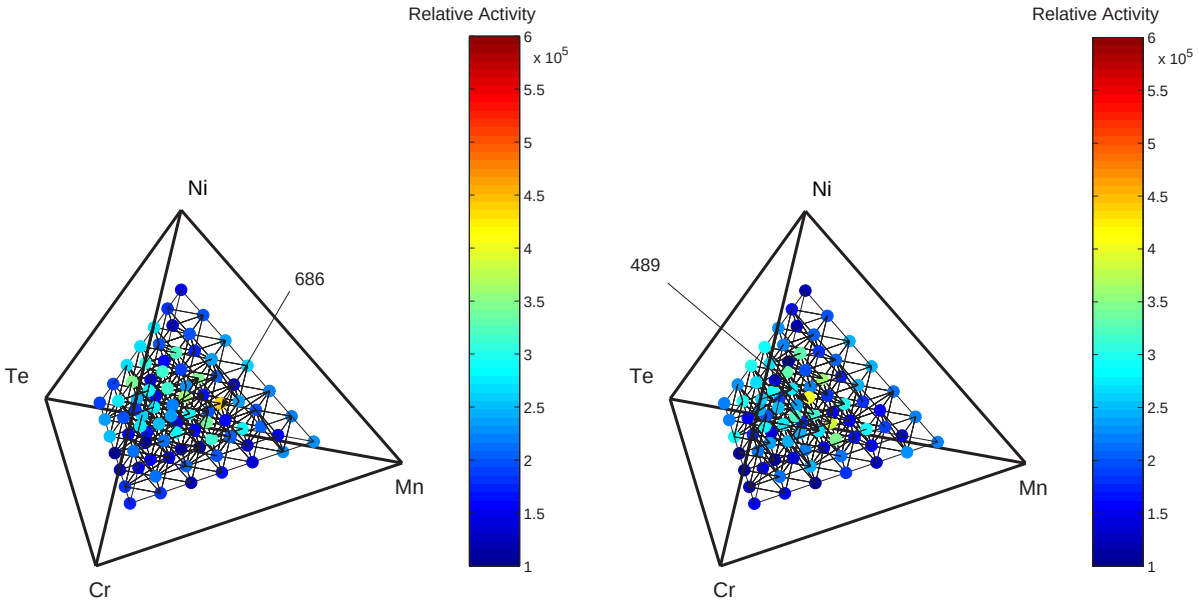


Figure 4.27: Screening results of the quaternary system CrMnTeNi A (left) and CrMnTeNi B (right) without scaling.

In Figure 4.27 the system CrMnTeNi is illustrated. In this system, there has been a noticeable difference between the two data sets. Sample 686 that proved to be quite active in data set A did not show the same behavior within data set B. There, sample 489 gained the best screening result.

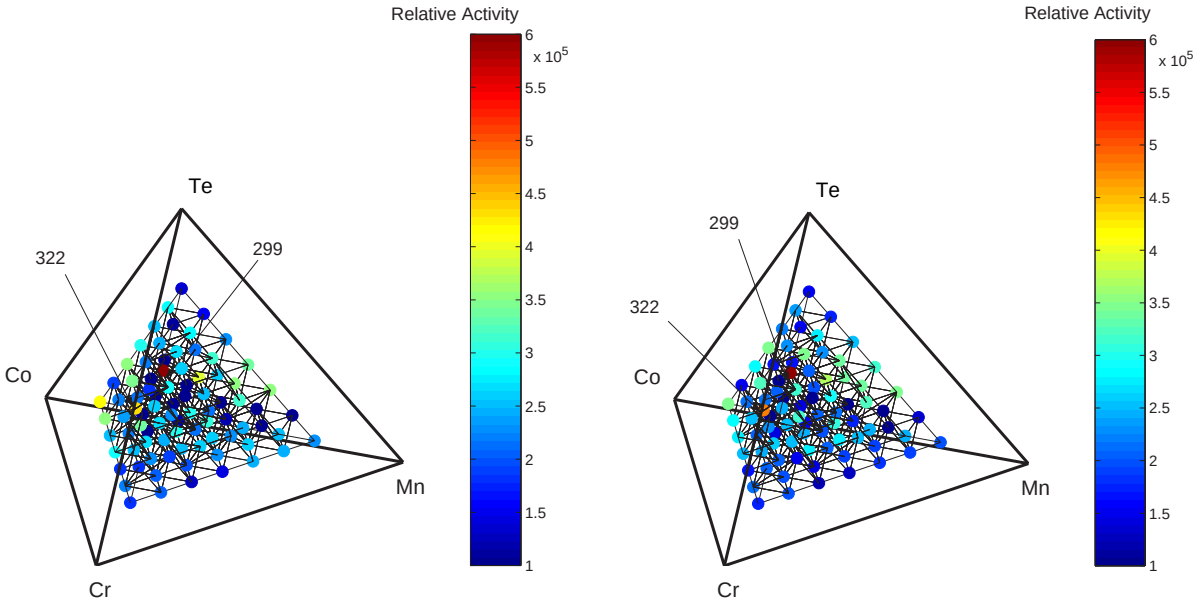


Figure 4.28: Screening results of the quaternary system CrMnCoTe A (left) and CrMnCoTe B (right) without scaling.

In Figure 4.28 the screening results of the system CrMnCoTe are illustrated. In both cases, the most active catalyst samples have been samples 299 and 322 indicating the good reproducibility.

4.8 Screening Results of the 2nd Generation of Catalysts

Within this section the screening results of a second generation of catalysts are identified that has been synthesized according to the regions of interest identified after the first screening process. In total, again 400 catalyst samples have been prepared lying in the 5% neighborhood of the most active samples. The active samples themselves have also been synthesized repeatedly to manifest their performances. The exact compositions of these 2nd generation catalysts can be found within a table given in the Appendix (pp. 169) together with the layouts of libraries 6 and 7. To test whether the results can be compared to those of the previous screenings the reference samples (Hopcalite) that have been placed on each library were considered first. The screening results of the references should lie within the same range than before. The following table gives the mean values of Hopcalite activity for each measured library.

Library	Mean Value of Hopcalite Activity ($\times 10^4$)
1a	3.60
2a	5.71
3a	4.91
4a	6.83
5a	5.27
1b	4.60
2b	4.80
3b	4.27
4b	7.99
5b	4.55
6	6.11
7	7.04

Table 4.6: Mean values of Hopcalite activity of all measured libraries.

From this table it can be seen that the Hopcalite samples screened on libraries 6 and 7 do not noticeably differ from the other results. A mean value of $6.11 \cdot 10^4$ for library 6 has also been reached on library 4a and the result for library 7 has been topped on library 4b, such that there have been no remarkable anomalies. It can be clearly seen that the results of the Hopcalite samples possess a large variance, ranging from $3.60 \cdot 10^4$ on library 1a to $7.99 \cdot 10^4$ on library 4b for catalytic samples having the same composition. This may have certain reasons, ranging from positional influences on the plates to sintering or aging effects. Since all samples are taken out of the same batch, it is also an indication for the general experimental error (noise). Furthermore, the set-up is thought to be a *first screening* and thus the chance of detecting false positives or negatives cannot be excluded. For materials in general it has to be kept in mind that the systematic variation of the chemical composition throughout the whole composition space may also lead to the formation of different phases. The transformation from amorphous mixed metal oxides to phase separations is assumed to be most likely the main reason for the observed experimental deviations. Having only 66 measurement values taken at discrete points, the probability of a phase transformation between single measured samples is quite large. It

has also been observed during our studies, that adjacent samples often show totally different activity values, lying far away from each other even though their chemical composition shows small variations.
This has to be paid attention to when working with the predictive models. Since the input data are indeed very noisy it will not be possible to estimate the activity value of a sample exactly. All that can be done is to predict the activity to lie within a certain range as shall be further discussed within Chapter 6.
Next, the most active samples of the second screening are summarized. It appears again that the best active sample is situated on an inner library position. Therefore, also the corrected data have been considered to remove this apparent influence of temperature gradients on the plates.

Ranking	**Sample**	**X**	**Y**	**Cr**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein Signal**
1	1287*	8	9	0.0	0.0	0.15	0.4	0.45	$7.06 \cdot 10^5$
2	1396	17	7	0.0	0.3	0.5	0.1	0.1	$5{,}50 \cdot 10^5$
3	1372*	14	11	0.0	0.35	0.40	0.1	0.15	$5.37 \cdot 10^5$
4	1196*	17	8	0.45	0.0	0.0	0.45	0.1	$5.35 \cdot 10^5$
5	1286*	8	8	0.0	0.0	0.15	0.35	0.5	$5.22 \cdot 10^5$
6	1352*	12	15	0.0	0.2	0.55	0.1	0.15	$5.20 \cdot 10^5$
7	1363*	13	14	0.0	0.3	0.45	0.1	0.15	$5.15 \cdot 10^5$
8	1354*	13	4	0.0	0.15	0.5	0.1	0.25	$5.11 \cdot 10^5$
9	1353*	13	3	0.0	0.15	0.6	0.1	0.15	$5.10 \cdot 10^5$
10	1360*	13	10	0.0	0.2	0.55	0.10	0.15	$5.10 \cdot 10^5$

Table 4.7: Best catalysts of the secondary screening without scaling. Samples marked with * are not contained within the first generation of catalysts.

Within table 4.7, very active samples appeared out of the quaternary system Mn-CoTeNi that has been already identified to contain promising samples lying in the area around samples 820, 603, 631 and 422 (cf. Figure 4.25).

Ranking	**Sample**	**X**	**Y**	**Cr**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein Signal**
1	1008*	1	12	0.0	0.15	0.75	0.1	0.0	$7.67 \cdot 10^5$
2	1354*	13	4	0.0	0.15	0.5	0.1	0.25	$7.08 \cdot 10^5$
3	1356	13	6	0.0	0.2	0.5	0.1	0.2	$6.93 \cdot 10^5$
4	1371*	14	9	0.0	0.3	0.45	0.1	0.15	$6.81 \cdot 10^5$
5	1360*	13	10	0.0	0.2	0.55	0.1	0.15	$6.76 \cdot 10^5$
6	1391*	16	9	0.0	0.35	0.25	0.1	0.3	$6.58 \cdot 10^5$
7	1367*	14	5	0.0	0.25	0.6	0.1	0.05	$6.32 \cdot 10^5$
8	1362*	13	12	0.0	0.25	0.5	0.1	0.15	$6.28 \cdot 10^5$
9	1351*	12	14	0.0	0.25	0.45	0.1	0.2	$6.11 \cdot 10^5$
10	1003*	1	7	0.0	0.05	0.85	0.1	0.0	$6.04 \cdot 10^5$

Table 4.8: Best catalysts of the secondary screening with scaling. Samples marked with * are not contained within the original data sets.

So, this area has been verified to be an area of interest for this reaction. Furthermore, good samples have also be reproduced within the ternary systems CoTeNi and CrTeNi. Applying the scaling procedure to this second generation of catalysts led to the top ten catalysts shown in Table 4.8. Here, most of the samples again originate from the quaternary system MnCoTeNi while two samples contain Mn, Co and Te. The scaling procedure confirmed the trend observed out of the unscaled data: the system MnCoTeNi yielded the best results. Due to the screening results lying very close for those catalysts of the second generation and the experimental error, there have been only two samples appearing among the best ten samples before and after the scaling. Nevertheless, the strong trend towards the quaternary system MnCoTeNi has not been affected. In the following, the screening results of areas that have been finer sampled are presented and compared to the first generation screening data.

Ternary Systems

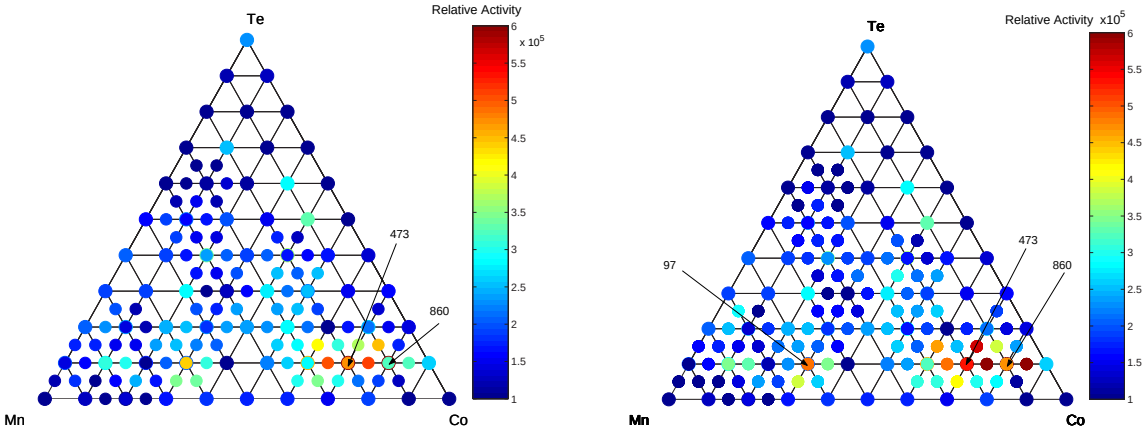


Figure 4.29: Results of system MnCoTe with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data.

From Figure 4.29 it can be observed that the finer sampling could verify most of the excellent performance out of screening 1. As mentioned above, the noise contained within the data also causes some deviations here. Especially the best sample 473 yielded a lower result in the second screening that has been further decreased by the scaling factor. Nevertheless, the region around sample 473 has been detected as the most active region among all ternary composition spreads.
In Figure 4.30, the most interesting sample seems to be 699. The original impression out of Figure 4.21 was that of 699 being an outlier while the refinement showed that the neighborhood also yielded active samples. In contrast to that the activity of sample 485 could not be confirmed by its neighborhood, so this might be a false positive.

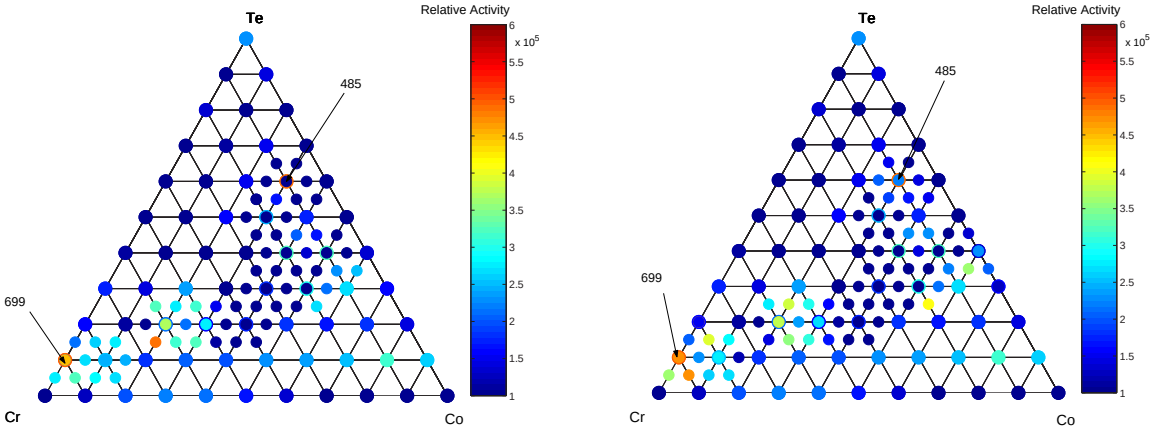


Figure 4.30: Results of system CrCoTe with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data.

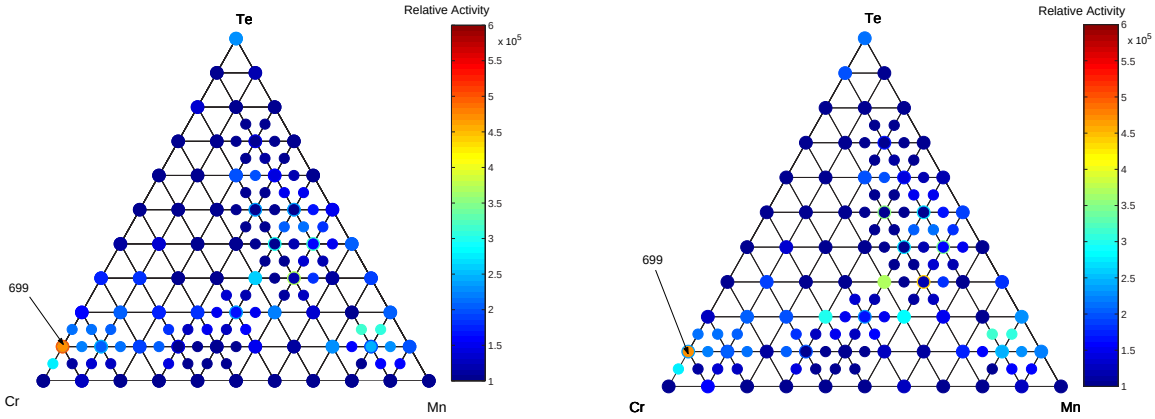


Figure 4.31: Results of system CrMnTe with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data.

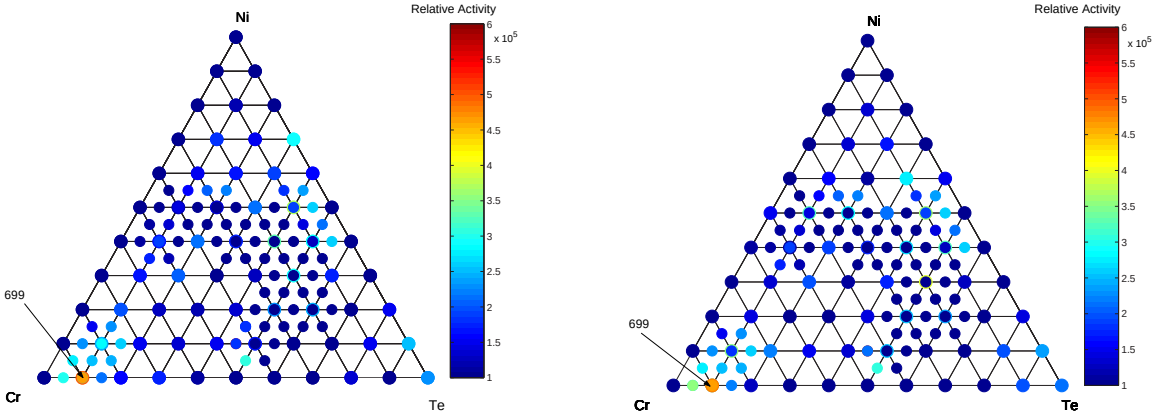


Figure 4.32: Results of system CrTeNi with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data.

In Figure 4.30, the most interesting sample seems to be 699. The original impression out of Figure 4.21 was that of 699 being an outlier while the refinement showed that

the neighborhood also yielded active samples. In contrast to that the activity of sample 485 could not be confirmed by its neighborhood, so this might be a false positive.
Within Figure 4.31 again sample 699 is contained. The neighborhood towards Mn shows a lower activity than for Co in the figures above. But still sample 699 is not totally isolated. Further active samples within the unscaled composition spread could not be verified. The application of the scaling procedure causes an increase of the activity of these samples that did not appear among the refinement here. In total, this composition spread is rather inactive but the refinement could verify some trends of the original data set.
For the last ternary system CrTeNi considered here, things have almost been the same than with CrMnTe. The only active region appeared around sample 699 being the most active one but similar to CrCoTe, the region around sample 699 shows quite good activity values, also for increasing Ni content. The region of interest stayed also active after the scaling procedure.

Quaternary Systems

For the refinement of a quaternary composition spread, the system MnCoTeNi has been chosen due to its good results in the primary screening. The following figure presents the results of the secondary screening compared to the first screening.

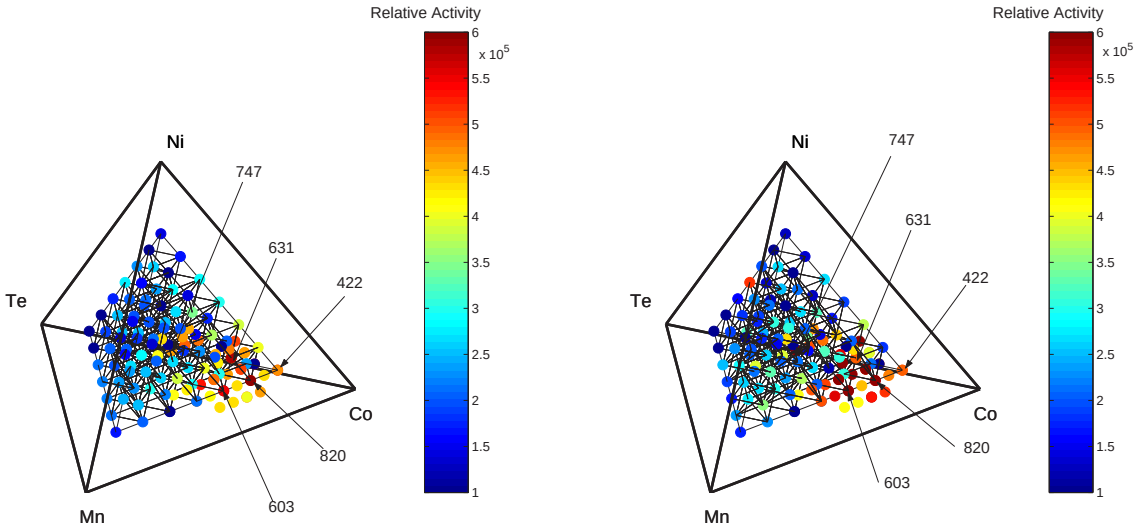


Figure 4.33: Results of system MnCoTeNi with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data

These results show that the screening results of the first generation have been reproduced in an excellent way. In both cases, the areas of interest around the marked samples have been confirmed and the refined catalyst samples perfectly fit into the original screening results.
Among all five quaternary systems, this system showed the best results of the refinement measurements such that the considerations in the quaternary case are restricted to this composition spread.

Pentanary System

The screening results of the pentanary samples cannot be as easily visualized as the ternary and quaternary ones. Therefore, the screening results of the pentanary compositions are summarized within the following tables, again differentiating between data set A and B. Among the 1001 samples of the complete pentanary composition spread, only 126 samples have exactly five components. Here, the best 10 catalyst among these 126 are mentioned.

Ranking	Sample	X	Y	Cr	Mn	Co	Te	Ni	Acrolein Signal
1	562	13	14	0.1	0.3	0.4	0.1	0.1	$3.70 \cdot 10^5$
2	726	11	6	0.1	0.2	0.1	0.3	0.3	$3.62 \cdot 10^5$
3	40	5	5	0.1	0.2	0.1	0.4	0.2	$3.60 \cdot 10^5$
4	681	8	4	0.1	0.1	0.1	0.4	0.3	$3.55 \cdot 10^5$
5	791	16	10	0.1	0.1	0.2	0.4	0.2	$3.54 \cdot 10^5$
6	763	14	3	0.1	0.4	0.1	0.3	0.1	$3.52 \cdot 10^5$
7	502	9	11	0.1	0.3	0.2	0.3	0.1	$3.49 \cdot 10^5$
8	729	11	10	0.2	0.2	0.1	0.3	0.2	$3.47 \cdot 10^5$
9	856	6	8	0.1	0.3	0.1	0.3	0.2	$3.46 \cdot 10^5$
10	379	15	9	0.1	0.2	0.4	0.1	0.2	$3.38 \cdot 10^5$
Means				0.11	0.23	0.18	0.29	0.19	

Table 4.9: Best pentanary catalysts in data set A.

Ranking	Sample	X	Y	Cr	Mn	Co	Te	Ni	Acrolein Signal
1	284	8	8	0.1	0.1	0.2	0.5	0.1	$7.21 \cdot 10^5$
2	562	13	14	0.1	0.3	0.4	0.1	0.1	$4.41 \cdot 10^5$
3	379	15	9	0.1	0.2	0.4	0.1	0.2	$3.88 \cdot 10^5$
4	136	11	16	0.1	0.3	0.1	0.4	0.1	$3.74 \cdot 10^5$
5	245	5	12	0.1	0.2	0.2	0.3	0.2	$3.73 \cdot 10^5$
6	502	9	11	0.1	0.3	0.2	0.3	0.1	$3.70 \cdot 10^5$
7	183	15	13	0.1	0.2	0.2	0.4	0.1	$3.69 \cdot 10^5$
8	308	10	2	0.1	0.3	0.3	0.1	0.2	$3.63 \cdot 10^5$
9	947	12	11	0.1	0.2	0.3	0.3	0.1	$3.58 \cdot 10^5$
10	531	11	10	0.1	0.1	0.4	0.3	0.1	$3.53 \cdot 10^5$
Means				0.1	0.22	0.27	0.28	0.13	

Table 4.10: Best pentanary catalysts in data set B.

From Tables 4.9 and 4.10 it can be seen that among the leading group sample 562 achieved nearly identical ranking in data set A and B with similar activity values. This result again indicates that the synthesis and screening route allows a replication of samples and their activities. Furthermore, samples 502 and 379 can be also found within the leading group of data sets A and B achieving comparable screening results. The rest of the catalysts appeared either in leading group A or leading group B,

indicating some sort of deviations among the pentanaries between the two data sets. Compared to the screening results of the ternaries and quaternaries presented above the samples of the pentanary system are less active. Except sample 284 of data set B that achieved a very large activity value of $7.21 \cdot 10^5$. Since this sample does not lie within the range of very good catalysts in data set A the possibility of sample 284 being an outlier has to be kept in mind. It is also interesting that only one sample (729) among the best 20 catalysts considered within these two tables contains 20% Cr while all other samples contain 10% Cr - a clear trend.
The elements Mn, Co and Te seem to have a positive influence on the performance of a catalyst here having mean values larger than 0.2 which would be the evenly spread value. Cr and Ni apparently best contribute to the catalytic activity with a content of 10-20 %.
Analogously to the ternary and quaternary systems some areas of the composition space have been finer sampled for a screening of a second generation of catalysts. As for higher dimensions the number of neighbors of samples grows very fast it has not been possible to synthesize all interesting neighborhoods here. In total, 14 pentanary samples have been synthesized half a year later to check their activities (4 new pentanaries with 5% increment and 10 (2×5) reproductions of samples with 10% increment). The following table gives an overview on the performance of this second generation of pentanaries:

Sample	**X**	**Y**	**Cr**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein Signal**
1375	14	14	0.10	0.20	0.05	0.40	0.25	$2.38 \cdot 10^5$
1376	15	4	0.10	0.20	0.05	0.35	0.30	$4.18 \cdot 10^5$
1377*	15	5	0.10	0.10	0.20	0.40	0.20	$9.37 \cdot 10^4$
1378	15	6	0.10	0.15	0.05	0.40	0.30	$1.01 \cdot 10^5$
1379*	15	7	0.10	0.10	0.30	0.40	0.10	$2.28 \cdot 10^5$
1380*	15	8	0.20	0.30	0.10	0.30	0.10	$5.27 \cdot 10^5$
1381*	15	9	0.10	0.20	0.20	0.40	0.10	$3.62 \cdot 10^5$
1382	15	10	0.05	0.20	0.05	0.40	0.30	$3.25 \cdot 10^5$
1383*	15	11	0.20	0.10	0.30	0.30	0.10	$5.52 \cdot 10^4$
1384*	15	12	0.10	0.10	0.20	0.40	0.20	$8.29 \cdot 10^4$
1385*	15	13	0.10	0.10	0.30	0.40	0.10	$9.79 \cdot 10^4$
1386*	15	14	0.20	0.30	0.10	0.30	0.10	$2.90 \cdot 10^4$
1387*	16	5	0.10	0.20	0.20	0.40	0.10	$1.76 \cdot 10^5$
1388*	16	6	0.20	0.10	0.30	0.30	0.10	$4.56 \cdot 10^4$

Table 4.11: Second generation of pentanary catalysts. Samples marked with * have been already synthesized in the first generation.

These measurements will become valuable to validate the Kringing and B-Spline model, discussed more closely within Chapter 6.

4.9 Summary

Summarizing, it can be said that the high-throughput screening of data set A and data set B has been in good accordance. The most active samples could be reproduced in data set B and their activity values have been confirmed, at least for the ternary and quaternary case. These samples and the neighborhood around them are studied more closely within Chapter 6, where Kriging and multilevel B-Splines are used to estimate activity values of non-synthesized samples but also of samples, that have been re-synthesized or have 5%-wise compositional increments.
The global maximum of activity performance has been achieved by the catalyst sample **501** ($Cr_{0.1}Co_{0.2}Te_{0.4}Ni_{0.3}$), followed by samples **685** ($Cr_{0.1}Co_{0.4}Te_{0.3}Ni_{0.2}$) and **473** ($Mn_{0.2}Co_{0.7}Te_{0.1}$) for the unscaled case. These three candidates showed excellent performances within both data sets and even their internal ranking could be exactly reproduced. Working with scaling factors sample **473** got the best results in both data sets while sample **501** completely vanished out of the leading group. Actually this sample received an extremely small scaling factor due to its inner library position. Nevertheless, this catalyst should also be evaluated once again in the search of the optimal catalyst for the reaction studied. The pentanary mixtures in general received lower activity values.

Chapter 6 will show how well the performance of catalysts can be predicted by a Kriging approach and a multilevel B-Spline approximation algorithm. Furthermore, it is discussed whether the scaling of data has a large influence on the estimation results. It will be also analyzed if the estimation procedures can lead to new compositions that might show larger activity values than the global screening maximum.
Dealing with screening data from high-throughput combinatorial experiments often more than three dimensions have to be dealt with. This means, that, for example, ordinary scatter plots are not the best way to illustrate the results and correlations in an elegant and informative way. Looking at the tetrahedra presented it can be clearly seen what challenges lie in a good visualization technique for five-dimensional data sets. Compositional information needs to be correlated with the screening results and for quaternary systems this yields a five dimensional data set, for pentanary systems one ends up with six dimensions. The following chapter is dedicated to this problem: visualizing more than four-dimensional data.

5 Visualization

This chapter focuses on a selection of visualization techniques that have been developed and applied to visualize the screening results of quaternary and pentanary composition spreads.
Newly developed, automated synthesis techniques but also the increasing use of combinatorial chemistry approaches and high-throughput screening (HTS) lead to an explosion of generated data in the laboratories - data that are quite useless without appropriate analyzing tools generating knowledge out of them. Within our work an essential part has been the visualization of our high-throughput screening and prediction results. Especially the visualization of quaternary systems needed to be updated and reconsidered since the simple presentation of a quaternary composition spread as a regular tetrahedron suffers from several drawbacks. Furthermore, visualization of the pentanaries has been another challenge. Here, more than 5 dimensions have to be considered and illustrated.
Within the next sections several possibilities to visualize high-throughput screening data are discussed. An emphasis is laid on quaternary and pentanary composition spreads since the usual way of presenting a ternary composition spread as an equilateral triangular pattern seems to be most elegant.
The development of different techniques, strategies and applications to visualize data belongs to a certain research area of computer graphics, the so-called *information visualization* [20, 42, 183]. Here, visualization is most often directly connected to data mining processes but apart from that it also plays an essential role within all research fields, where lots of data have to be handled. Within the last years the spreading of the so-called *visual data mining* could have been observed, a research area within the information visualization especially adapted to the needs of data mining, [6, 165]. In contrast to visualization in research, information visualization not aims less at the illustration of chemical or physics measurement data or simulations, but more on the visualization of correlations, pattern and information. Within the last decade, intensive research has been going on within this field and the survey of Wong and Bergeron [194] can be studied for a detailed description of the historical development.
Out of the vast variety of visualization techniques, especially for large, multivariate data sets, some techniques have been chosen that lend themselves very well for the visualization of high-throughput screening data.

5.1 Types of Data and Dimensionality

Large data sets like screening data or results of combinatorial experiments often contain large amounts of single entries, so-called *records*, together with a set of variables, the *dimensions*. One record in our case consists of one measured sample with the single variables corresponding to the composition of the material (contents of Cr, Mn, Co, Te and Ni) and the produced acrolein signal. In information visualization,

the number of variables also gives the dimensionality of the data set. According to Shneiderman [161], data sets can be one, two- or multi-dimensional and also contain more complex data types such as text, hypertext, hierarchies, graphs or algorithms. Alternatively, these data sets are also called uni-, bi- or multivariate.

One-dimensional Data Sets
Typical representatives of one-dimensional data sets are time dependent data. Each point on the time axis corresponds to one or several measurement values. In chemistry, measuring the time dependence of a concentration or reaction often results in a one-dimensional data set. In this application with high-throughput screening, there have been no one-dimensional data.

Two- and Three-dimensional Data Sets
These data sets consists of two or three defined variables. In combinatorial chemistry two dimensional data sets can be binary composition spreads while three-dimensional data sets correspond to ternary composition spreads. Usually, there is a screening value adapted to each composition such that the dimensions are increased by one here. Most of the time those lower dimensional data sets can be illustrated by 2D and 3D scatter plots. Although this illustration techniques seems to be quite simple, it can become more and more confusing for very large data sets or a high density of data.

Multi-dimensional Data Sets
In most cases, the data sets consist of more than three dimensions and therefore cannot be illustrated by simple 2d or 3d scatter plots. Those multi-dimensional data sets can easily contain thousands of records as they are generated by automated high-throughput systems or combinatorial set-ups, often directly stored within relational databases. These data can only be illustrated by elaborated visualization techniques since an effective projection onto 2D or 3D is not a trivial task. For this thesis, the visualization of more than four dimensions as they occur with quaternary and pentanary composition spreads plays the central role.

5.2 Visualization Techniques

Within this section some visualization techniques discovered within other research areas or other application fields are presented. Furthermore it is discussed how the tetrahedral plots can be improved with respect to readability and validity. Most of the plotting work has been realized in MATLAB.

5.2.1 Parallel Coordinates

Parallel coordinates were proposed by Inselberg [76] in 1981 as a new way to represent multi-dimensional information. Since the original proposal, much subsequent work has been added, e.g. [77, 78]. In the traditional approach with Cartesian coordinates, all axes are mutually perpendicular and the visualization is restricted to three dimensions. In parallel coordinates all axes are parallel to one another and equally spaced. By drawing the axes parallel to one another, one can represent points, lines and planes in more than 3 dimensions. A vertical line is used for the

projection of each dimension or variable, arranged in a way that the maximum and minimum values of each dimension are scaled to the upper and lower boundaries on those vertical lines. A connecting line made up of $n-1$ lines at the appropriate dimensional values connects the axes to represent an n-dimensional point. The following figure presents a parallel coordinate representation of the screening results of the quaternary system MnCoTeNi for acrolein.

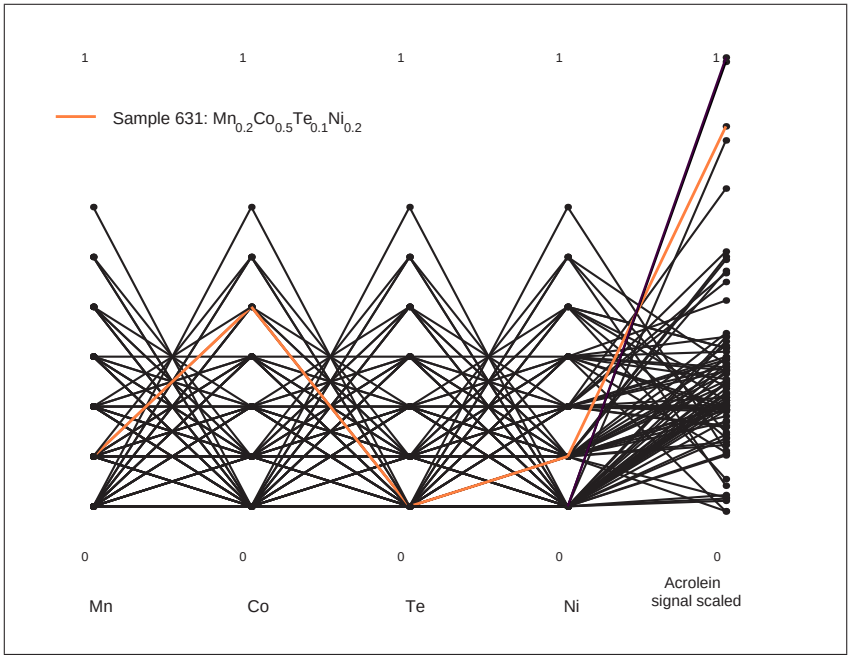


Figure 5.1: Parallel coordinates plot of the screening results (acrolein signal) of the quaternary system MnCoTeNi. As an example, sample 631 is marked in color such that its composition becomes easily accessible.

In this plot all quaternary compositions can be illustrated together with their achieved acrolein signal. To give a better representation, the acrolein signals have been normalized to the interval [0,1]. In a non-colored presentation it would be quite hard to identify the connections belonging to one catalyst sample but in most software solutions (e.g. Spotfire®) these kind of plots are interactively accessible such that clicking on a line highlights the corresponding connections. This makes it very easy to find the best compositions or to identify trends within the data. Here, sample 631 has been highlighted to reveal its composition and screening results ranging among the best catalysts of this data set. Parallel coordinates plots are an elegant tool to illustrate high-dimensional data sets with different variables and objectives.

5.2.2 RadViz

By the RadViz method, cf. [5, 6], n-dimensional data sets are nonlinearly projected onto 2 dimensions using the idea of springs being arranged around the perimeter of a circle. The term *RadViz* is derived from radial visualization and the whole approach can be considered to be quite similar to parallel coordinates, belonging to the field of so-called *lossless* visualization techniques. The key to understand the way RadViz

visualization works lies in a rather physical way of thinking: It is assumed that n-dimensional data points are laid out as points equally spaced around the perimeter of a circle. Furthermore, the ends of n springs are attached to these n perimeter points while the other ends of the springs are attached to a data point. The spring constant K_i, $i = 1, \ldots, n$ equals the values of the ith coordinate of the fixed point. Then, each data point is displayed at the position where the sum of the spring forces equals 0. In general, the data point values are usually normalized to have values between 0 and 1 (in the case of catalyst compositions this is already the case since all entries sum up to 1). For example, if all n coordinates of a n-dimensional data point have the same value, this point will appear in the exact center of the circle. If the point is described by a unit vector, then that point will appear exactly at the fixed point on the edge of the circle (where the spring for that dimension is fixed). Of course it can happen that several points are mapped onto the same position, i.e. the mapping is not a unique transformation. It is also a non-linear transformation of the data that allows to preserve certain patterns and to obtain a quite intuitive representation of the data.

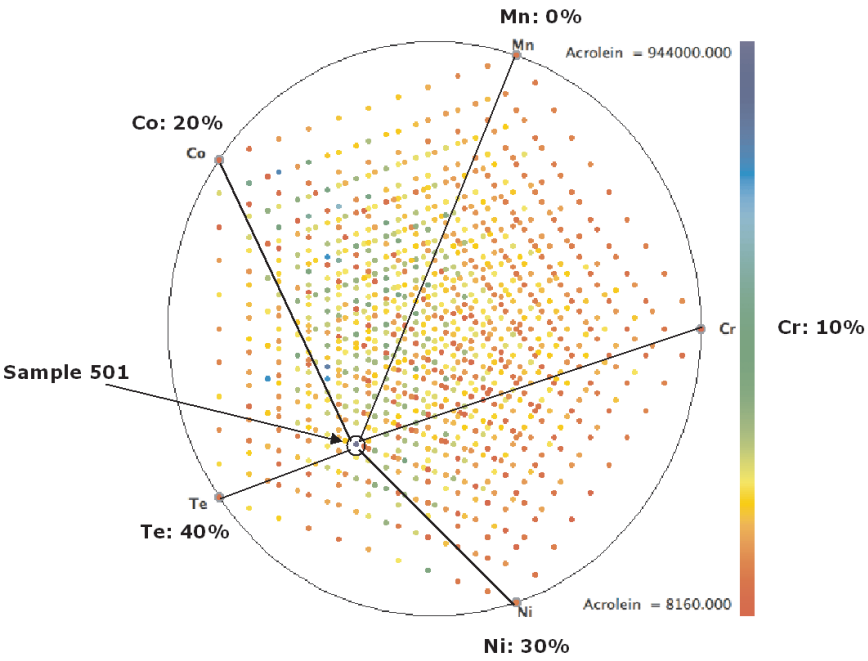


Figure 5.2: Presentation of the pentanary data set in RadViz. The color scheme corresponds to the activity values measured for acrolein (raw data).

In the plot, sample 501 is explicitly highlighted to demonstrate how RadViz works. It is not easy to directly get the composition of the sample, but only qualitative trends. The longer the connecting line from the sample to the element on the boundary of the circle, the less is its content in the sample. Sample 501 contains no Mn, thus, this connection is the longest. Te holds the largest content, thus this connecting line is shortest. In general, the main features of this visualization method are:

- Points having approximately equal coordinate values will appear close to the center of the circle.
- Points having similar values whose dimensions are placed opposite each other on the circle will also appear close to the center of the circle.
- Points having one or two coordinates greater than the others lie closer to those dimensions.
- An n-dimensional line will be mapped to a line.
- A sphere will be mapped to an ellipse.
- An n-dimensional plane maps to a bounded polygon.

Figure 5.2 illustrates the pentanary data set A (1001 catalyst samples with activity for acrolein) in radviz presentation. The color indicates the peak area of the acrolein signal measured by gas chromatography. The larger the peak areas, the more acrolein has been formed and the better the catalysts. Due to the regular sampling of the composition space using 10% increment, the obtained pattern in radviz appears as an equilateral polygon.
The visualization of the pentanary composition spread in this way allows an immediate investigation of trends and correlations. In this case, it can be seen that the best catalyst (dark blue point) contains more Co and Te than Mn or Cr. Indeed, the best sample of this data set has been catalyst 501 ($Cr_{0.1}Mn_{0.0}Co_{0.2}Te_{0.4}Ni_{0.3}$). By the distribution of the blue points in this radviz representation we can easily derive qualitative correlations of metal contents and activity of the catalysts. As another trend, samples containing larger amounts of Cr or Mn have not been very active for acrolein formation. Thus, for the studied reaction the elements Te, Co and Ni seem to play a more important role. Having a look at the sorted catalyst samples at table 4.2 on page 75, these observations can be confirmed.
Summarizing, it can be said that the radviz visualization technique provides a quick and elegant tool for a lossless illustration of high-dimensional data sets. The obtained trends and correlations are more of qualitative kind since an exact determination of the composition of a catalyst is not trivial. For software realizations this drawback can be solved, for example, by providing the possibility to get the compositional information by clicking at the sample.

5.2.3 Heat Maps

For the representation of multidimensional data sets, the use of so-called *heat maps* can also be an efficient tool. Heat maps are colored schemes that "translate" numeric values into color. This means, heat maps are color-coded tables. In the present application, the complete pentanary data set consists of 1001 rows (catalysts) and 6 columns (5 elements & activity values). Still, the main interest with that data set is the search for correlations and trends between composition and activity of a catalyst sample. Figure 5.3 displays a heat map representation of 100 catalysts taken out of the pentanary data set. For this illustration, the data table has been transposed such that the rows contain the composition (rows 1-5) and measured activity values (row 6).

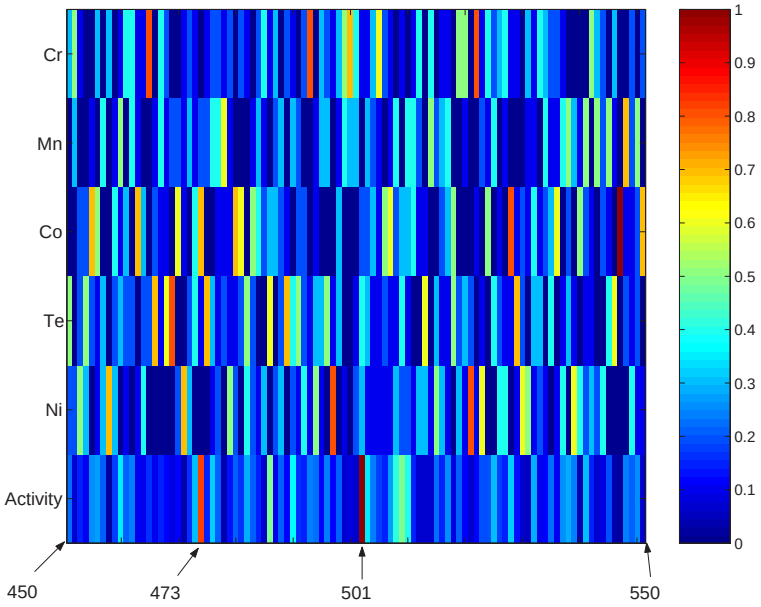


Figure 5.3: Representation of catalyst samples 450-550 by a heat map. The color corresponds to the composition of the samples but also to the normalized activity values.

The columns then correspond to the catalyst samples. In this example, only 100 samples are considered due to readability reasons. To yield a better color-scheme, the activity values have been standardized to the most active sample (sample 501, cf. p. 75) such that this sample gets an activity value of 1 and all other catalysts lie within the interval [0,1].

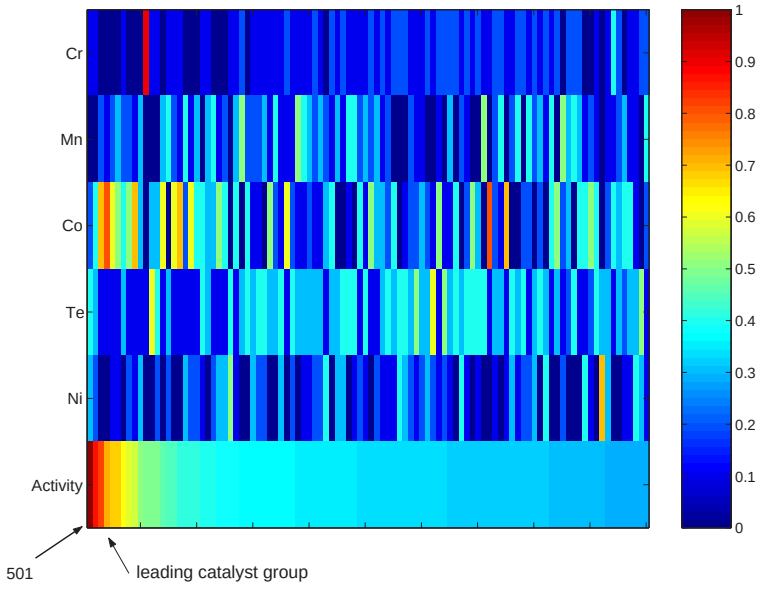


Figure 5.4: Representation of the 100 best catalyst samples of data set A (unscaled data) by a heat map. The color corresponds to the composition of the samples but also to the normalized activity values.

The representation of the data by a heat map allows a quick visual inspection of the performance of the catalysts. The range between samples 450 and 550 has been chosen for this example and the best sample 501 can be clearly identified by its activity value of 1 given by a dark red line. Furthermore, the composition of sample 501 ($Cr_{0.1}Mn_{0.0}Co_{0.2}Te_{0.4}Ni_{0.3}$) can be determined out of the color scheme. Using too many samples within one heat map may become quite difficult to interpret since the compositions of the catalysts cannot be uniquely determined by visual inspection. Here, sorting of the data according to a certain column or row may improve the results. This has been done in Figure 5.4 showing the best 100 catalyst samples (unscaled data).
Sorting the data and illustrating them by a heat map can reveal certain trends concerning the composition of the catalysts. In the second example, it can be observed that the leading group of catalysts contains larger amounts of Co compared to the content of the other four elements.
In general, looking at the complete data set in this way can reveal information about the distribution of good samples. In the case of an evenly sorted data set, the appearing of red clusters in the activity row can be a hint that the compositions of these samples are quite promising.
Summarizing, it can be said that heat maps provide another convenient visualization tool to illustrate large multidimensional data sets. The application of normalization and sorting the data should be used to get an optimized representation with respect to the used color scheme. For each data set, these settings should be individually set.

5.2.4 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) [84] is a projection technique widely used to handle multivariate data containing interrelated variables. The application of PCA can reduce the information dimensionality in a way that the loss of information is minimized, by constructing uncorrelated axes that finally lead to a transformation of the original coordinate system. Multivariate data sets then can be visualized by their projections in 2D or 3D with a minimal loss of information.
In the studied application when considering only acrolein as product of interest, 5 out of 6 used data columns (namely the columns containing the composition information) are orthogonal to each other and thus uncorrelated. They all contain identical entries, only in a shuffled way (since a *complete* and evenly sampled composition spread is considered generated by mixture design). Due to this, none of these 5 columns is redundant and a reduction of dimensionality can at most be achieved by one. If we added other products of interest, e.g. other columns to this data set, PCA would be useful to detect trends and pattern within the data by removing redundant information. For more information on how this approach has been realized using high-throughput screening data in a nine-dimensional data set, the reader is referred to some former work in this field, cf. [163]. Furthermore, this work shortly introduces the theory of PCA but also mentions recently published applications of PCA in the chemical sector.

5.3 Developed MATLAB Environments

To visualize the catalyst data several graphical user interfaces (GUIs) have been developed in MATLAB providing a convenient access to its visualization power.

5.3.1 CatVis

This GUI enables the user to plot ternary and quaternary composition spreads where color schemes, sizes of markers or additional circles around the markers (*Circles*-button) can be arranged. It is also possible to add a bounding triangle for ternary compositions if only parts of a ternary composition spread should be plotted and the vertices of the triangle are missing (*Triangle*-button). Furthermore, the user can scale the colorbar according to certain needs such that the color scheme is to be slightly varied (Colorbar Min, Colorbar Max). The color scheme itself is chosen by the context menu *Colormap* out of several schemes available (default layout: 'jet'). With *Units Colorbar* and *Label Colorbar* text can be added to the colorbar. If no text should appear in the plot, blanks should be placed here to prevent errors. For quaternary composition spreads illustrated as a tetrahedron additionally the viewing position can be adjusted (3D view). The Azimuth value gives the vertical rotation while the Elevation value gives the height of the view (in degrees). Here, the default values are -37.5 (az) and 30 (el) yielding the standard 3D view.

Characters	Corresp. Element	Position in Plot (Ternary/Quaternary)
1-2	1	left corner/ front corner
3-4	2	right corner/ right back corner
5-6	3	top of triangle/left back corner
7-8	4	top of tetrahedron

Table 5.1: Correlation between characters of the elements box and their position in the plot.

Data are loaded in via the standard dialogue box in form of *.txt files. As a convention here the *.txt files should be organized in such a way that the first column corresponds to the composition values of the 1st element, the second column corresponds to the 2nd element and so one till the last column including the screening results or any other measurement values. It is important that a point is used instead of a comma as decimal separator. The *Elements* box should be used to add element names to the figure. In both cases (ternary and quaternary) always 8 characters need to be spent to name the corners. For ternary plots, only the first 6 characters are used and table 5.1 explains where the labels are placed in the plot (for 3D: in default view position).

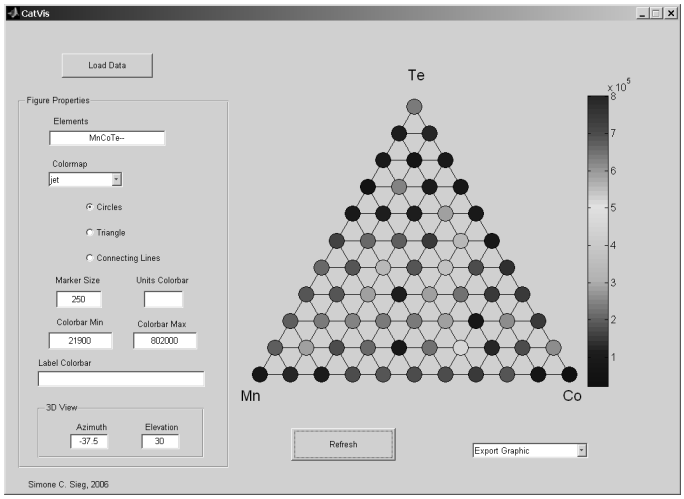


Figure 5.5: Screenshot of the CatVis GUI.

5.3.2 TetraView

The GUI TetraView has been developed to scroll through a 3D tetrahedral plot layer by layer. In contrast to the complete representation done in CatVis, this representation might be helpful to look at the screening results of inner points of the tetrahedron.

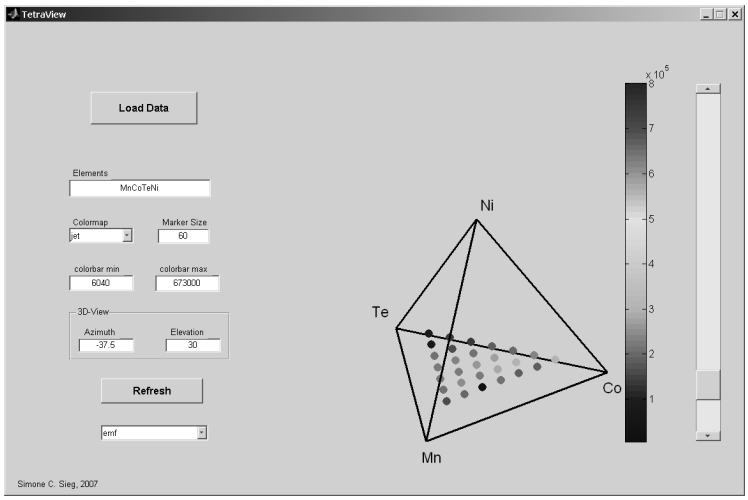


Figure 5.6: Screenshot of the TetraView GUI.

Loading in the data is again done via a *.txt-file as described above. Analogously to CatVis, the corners of the tetrahedron can be labelled using the *Elements* edit box. All other features that can be changed are similar to CatVis. By default, the colorbar appears at the right side.

5.3.3 View4d

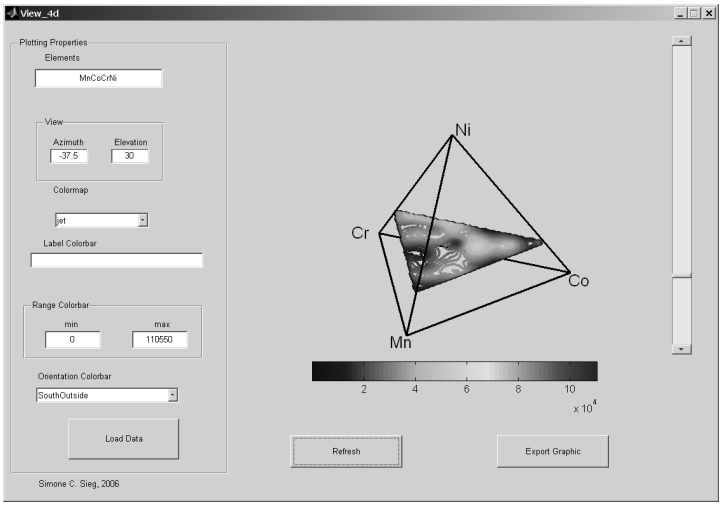


Figure 5.7: Screenshot of the View4d GUI. The colorbar has been placed at the south margin (SouthOutside).

This GUI has been developed to illustrate the modelling results of a quaternary system by using a sliced representation. Having the results from the Kriging or B-Spline model the tetrahedron not only consists of 286 discrete points but of many more such that each layer is more closely packed with points. Illustration of these points in form of lines connecting points with identical activity values (similar to contour lines) yields an elegant way of visualization. This representation can help to identify regions of interest more easily than the discrete, pointwise representation can achieve.
The operation of the GUI is similar to the ones described above with self-explaining edit boxes. The only difference here lies in the way the input data need to be processed. For this representation no *.txt-file can be imported but a *.mat-file is needed containing three arrays X, Y, Z with coordinate information in 3D and the according activity information stored in an array V. These data can be obtained by the MATLAB routine plot_cube.m. The operation of this m-file is documented quite well by its comments and the internal help function.

5.3.4 View5d

This GUI View5d has been designed to enable the visual inspection of a five dimensional composition spread helping the search for regions of high activity values.

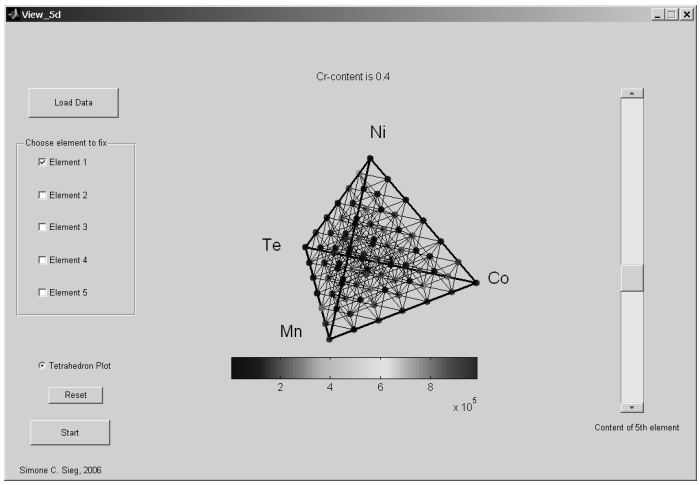


Figure 5.8: Screenshot of the View5d GUI. The colorbar has been placed at the right margin (EastOutside).

Following the illustration used for the quaternary case (cf. section 2.7.3) we would come to a generalized polygon in 4D with five corners which cannot be represented in 3D.
The idea of View5d lies in the fact that always one element is kept fixed to a certain composition value, to look then at all points having that fixed content. In the case of 0% of an element, one ends up with a subspace of a quaternary composition spread that can be easily represented by a tetrahedron. Scrolling through the content of one element this tetrahedral representation is kept while the number of points is decreased. This is a logical consequence since with increasing content of one element the number of mixtures including exactly this content decreases. The visible mixtures contain all samples with fixed 5th element content but not only pentanary mixtures but also binaries, ternaries and quaternaries. Choosing a fixed elemental content of 90% results in a plot with only four points remaining. One difficulty for this representation lies in the fact that the corner of the presented tetrahedra do not correspond to pure mixtures but to binary mixtures of the fixed element and the element represented by the corner.

6 Numerical Results and Modelling

In this chapter the results of the prediction models applied to the most active composition spreads are discussed. As presented in Chapter 4 the best performing ternary composition spread was MnCoTe containing samples 473, $Mn_{0.2}Co_{0.7}Te_{0.1}$ and 860, $Mn_{0.1}Co_{0.8}Te_{0.1}$. In the quaternary case, the composition spread including the best samples is MnCoTeNi. As already seen in Chapter 4 screening of a library of identical samples yielded a certain variance among the results. This means the data contain a certain noise caused by synthesis or screening mistakes. Thus it is clear, that the predictive power of the models cannot go beyond this noise. For practical applications this means that only an interval of each estimated activity value can be determined lying in the range of the observed variance of approximately $1.14 \cdot 10^4$. This variance has to be subtracted from an estimated result or added to the result to obtain the corresponding interval.

6.1 Kriging Model

6.1.1 The Kriging GUI

The application of Kriging to estimate the activity of non-synthesized samples can also be realized by a MATLAB GUI that has been developed. The following figure shows a screenshot of this application.

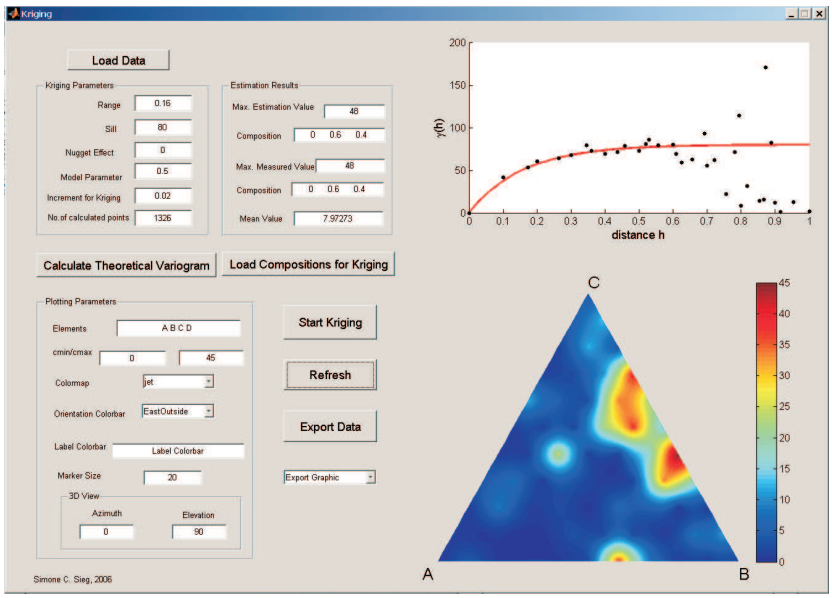


Figure 6.1: Screenshot of the Kriging GUI. The operation is described in the text.

Operating the Kriging GUI
Similar to other GUIs, the application of the Kriging GUI starts with loading in the data via *.txt-files. These files consist of four (ternaries), five (quaternaries) or six (pentanaries) columns (decimals separated by points (!)) with the last column containing the measured activity values or other experimental data. After loading in the data, the experimental variogram is immediately calculated and plotted in the coordinate system in the upper right corner of the GUI. Additionally, the composition of the best sample together with its screening result is displayed. To fit a theoretical variogram function the appropriate parameters can be typed in (*Kriging Parameters*). According to the chosen increment for Kriging the number of calculated points is returned. The smaller the chosen increment value, the more points within the composition spread are kriged and visualized. Quite good results can be already achieved for an increment of 0.02; for smaller increments, the calculation time steadily increases. The button *Calculate Theoretical Variogram* yields the theoretical variogram function. This function can be changed and further adjusted by changing the Kriging parameters. Furthermore, the model parameter determines the type of Kriging model used for calculation (0.5: Exponential Type, ∞: Gaussian type, cf. Figure 2.9). The button *Calculate Theoretical Variogram* always yields the refreshed variogram function. If only compositions defined a priori should be kriged, they can be loaded in via a *.txt-file by *Load Compositions for Kriging*. If all Kriging parameters are fixed, the Kriging process is started by *Start Kriging* and the results will be displayed. For the ternary case a modelled triangular surface is obtained, in the quaternary case, the original tetrahedron is illustrated together with the estimated points. In this case a representation of the Kriging results in TetraView might be more informative. To change the layout of the plot several parameters can be adjusted like the colormap, orientation of the colormap or the 3D view. With *cmin/cmax* the range for the coloring can be individually changed such that very active regions are highlighted. Input into the boxes *Elements* and *Label Colorbar* labels the corners of the plots (cf. Table 5.1, p. 100) and the colorbar. The field *Marker Size* controls the size of the markers in the tetrahedral plot (a value between 60 and 80 is a good choice here). With *Refresh* the set parameters can be applied to the plot. Pressing *Export Data* the results of the Kriging procedure can be exported in a *.txt-file (compositions, Kriging estimator, Kriging variance). As a final step, the graphic can be exported into certain file formats.

6.1.2 Kriging of the Ternary System MnCoTe

Within this section the performance of the Kriging model is tested when applied to screening data of the ternary composition spread MnCoTe. For this composition spread a second generation consisting of 101 catalysts has been prepared and screened to refine the most active region of the search space. It has been an interesting task to check whether the estimated activity values lie in accordance with the measured results. Therefore the Kriging estimator has been calculated for each sample of the refinement generation. Furthermore it was to be checked if estimated activity values can surpass the measured ones and lead to new active compositions. Within this composition spread, samples **473** ($Mn_{0.2}Co_{0.7}Te_{0.1}$) and **860** ($Mn_{0.1}Co_{0.8}Te_{0.1}$) got the best activity values for acrolein. For the refinement of this composition spread 101 samples have been synthesized and screened again

under the same synthesis and screening conditions as the parent data sets. Here, two ternary starting systems are considered taken out of data set A, the unscaled and scaled case. As shown in Figure 4.22, the reproduction of the samples show excellent similarities between ternaries A and B such that it is sufficient to consider the data obtained out of data set A only. Due to the temperature gradient observed during the screening process the behavior of the scaled data is also considered. It will be discussed in the following sections, whether the scaling procedure helped to improve the results of the prediction.
First, the complete ternary composition spread MnCoTe taken out of data set A (unscaled) is considered. Figure 6.2 illustrates the kind of input data used for prediction calculations.

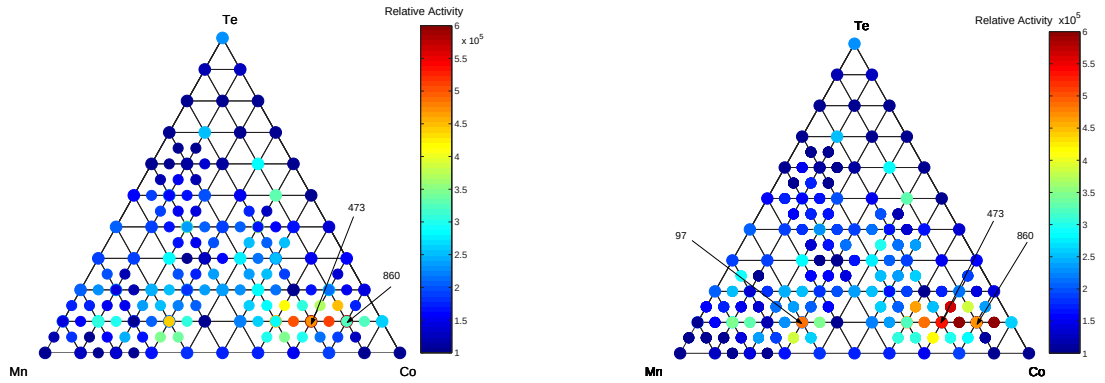


Figure 6.2: Screening results of the system MnCoTe together with finer sampling of the regions of interest. Left: unscaled data, Right: scaled data

It can be clearly seen that the most active catalyst samples are 473 (peak area: 802,000) and 860 (peak area: 695,000) situated next to each other within a defined *region of interest*. The rest of the composition spread appears to be rather inactive for the propene oxidation studied. Unfortunately, the excellent performance of sample 860 could not be confirmed by the second screening and reproduction. The screening results for this sample appeared to be significantly lower. The neighborhood around sample 473 could be verified as an active region, indicating that the good performance of sample 473 is not a measurement error.
Now, the goal has been the determination of a so-called *activity surface* that describes the dependence between chemical composition of the catalysts and their activities. This means, quantitative composition activity relationships (QCARs) have to be established.
To apply the Kriging model, an experimental variogram has to be calculated (cf. pp. 36) out of the measured data. Then, a theoretical variogram function must be found that *best* fits the experimental variogram cloud. There exist many different theoretical variogram functions and according to each problem, an appropriate function has to be chosen. For this case, a variogram function with a parabolic behavior (approximately quadratic) near the origin has been chosen, i.e. $\gamma(|h|) \approx c|h|^2$ for $h \to 0$, being continuous and differentiable at $h = 0$. This kind of variogram functions is mostly used to describe regular phenomena. For catalysts, a regular variation of activity with composition is assumed and the experiments have been planned such that the compositions of the catalysts vary with a predefined, constant increment (10%-wise variation). The following figure exhibits the corresponding variogram

models according to the experimental data.

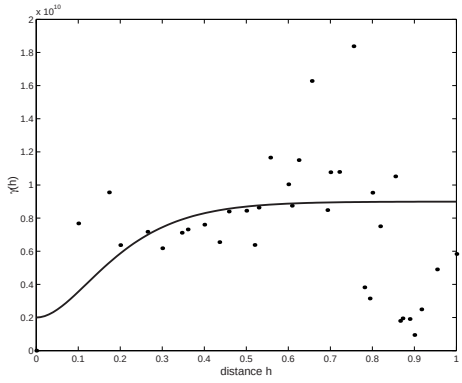


Figure 6.3: Experimental Variogram of MnCoTe taken out of data set A (no scaling used) together with the fitted variogram function.

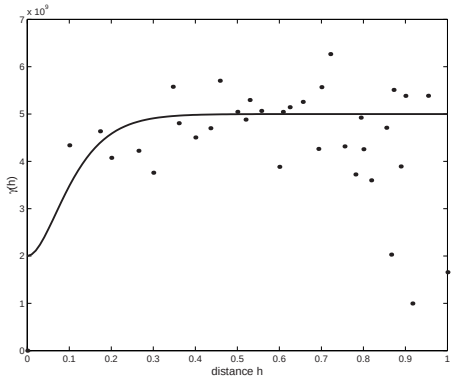


Figure 6.4: Experimental Variogram of MnCoTe taken out of data set A with applied scaling together with the fitted variogram function.

Furthermore, the chosen theoretical variogram functions are characterized by a nugget effect that indicates the presence of possible variability between samples with very similar compositions (cf. p. 40). The used variogram functions are given by the Matérn class of variogram functions (see equation 2.59, p. 44), using the parameters summarized in the following table. To check the performance of the chosen Kriging models, the activity of 101 catalyst samples synthesized in a second generation has been estimated.

	Unscaled Data	Scaled Data
Range	0.09	0.06
Sill	$0.7 \cdot 10^{10}$	$3 \cdot 10^9$
Nugget Effect	$0.2 \cdot 10^{10}$	$2 \cdot 10^9$
Smoothing Parameter	2	2

Table 6.1: Chosen Kriging parameters for the ternary system MnCoTe (unscaled and scaled case)

Using these theoretical variogram functions, ordinary Kriging has been applied to estimate the activity values of non-synthesized samples. Figure 6.5 illustrates the results for the unscaled case.

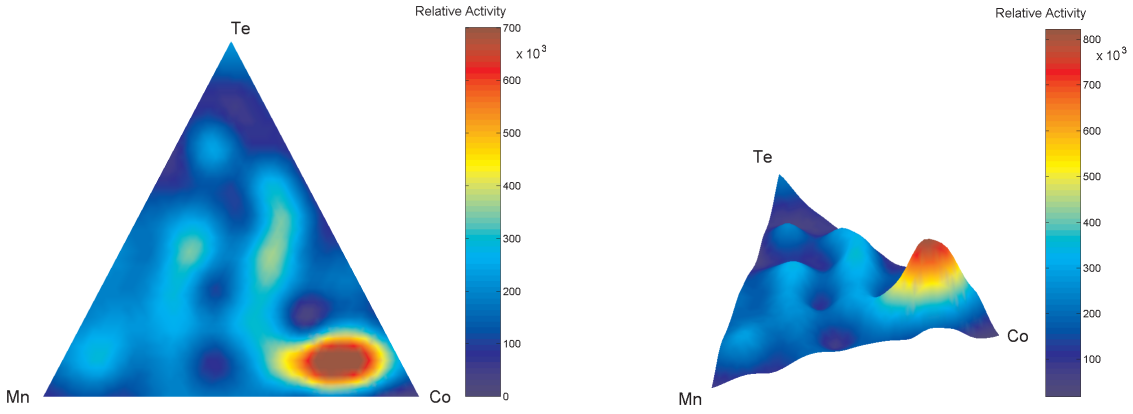


Figure 6.5: Kriging results for the unscaled system MnCoTe.

The Kriging estimator for the activities of 20,301 compositions has been calculated within the ternary composition space which corresponds to an increment of 0.5%. Then, the correlation between the composition and the activity can be illustrated by a smooth surface. Within this unscaled data set it has been quite interesting to see that a coherent part of the triangle appeared as an active region. The most active samples 473 and 860 also have been obtained in this region. Since these two samples are neighbors in composition the focus has been laid on the compositions around them.
The Kriging model can also be applied to predict the activities of any desired composition within the considered search space. In this example, the most active composition calculated by the model has been ($Mn_{0.18}Co_{0.73}Te_{0.09}$) with a calculated activity value of 835,894 (Kriging variance: 54,360). In a second generation of catalysts, the surrounding compositions (5%-wise variations) have been synthesized to check whether the good performance of the first generation could be confirmed. In the following figure the calculated results of this second generation of catalysts are illustrated. The case of an ideal prediction would be given, if the points appeared lying on the bisecting line of the coordinate system. In this representation it can be observed that there are larger differences between measured and calculated activity values, especially within the active region. This can be explained by the fact that the catalyst samples of the second generation in general yielded lower activity values. Unfortunately, the excellent screening result of sample 860 could not be reproduced by the second generation. In both cases, the unscaled and scaled data set, this sample yielded lower activity values. Since the Kriging model uses the first generation screening results as input data, the activity value of sample 860 is estimated to its first screening value and therefore this calculated value exceeds the measured one. Why this sample gained less activity during the second generation cannot be properly explained and might be due to experimental issues. A significant deviation of the second screening results from the first ones in the region of high activity has been observed that can not be properly explained. As a positive result here, very active samples have been obtained within the direct neighborhood of sample 860 that could verify the trend of the first screening.
To evaluate the quality of the considered models, a *value of fit* has always been calculated, which can be described as the mean of the relative deviation between

measured and calculated activities. Let therefore x_i and y_i be the calculated and measured activity value of sample $i, i = 1, \ldots, n$. Then, the value of fit (VoF) is given by

$$VoF = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - x_i|}{y_i} \tag{6.1}$$

This means, for each model a VoF can be obtained and the smaller this value, the better correspond experimental and predicted activity values.
The VoF for the reproduction of samples taken out of the unscaled ternary system MnCoTe is **0.32**.

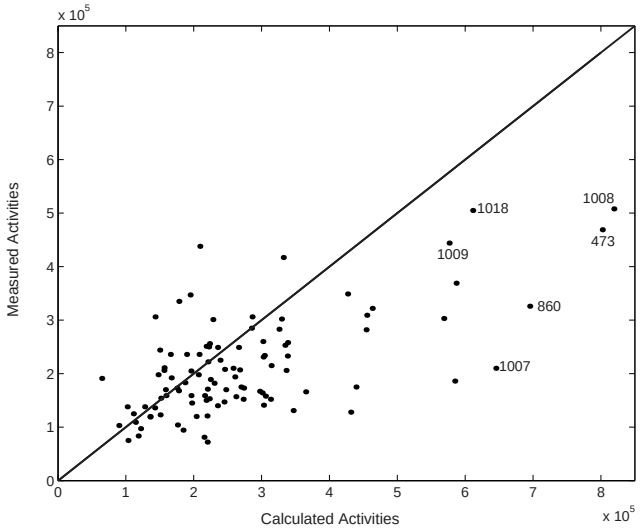


Figure 6.6: Comparison of calculated and measured activity values for the ternary system MnCoTe (unscaled data).

The samples with 5%-wise variation around samples 473 and 860 showed also good activities, cf. Figure 6.2, which confirmed the successful first screening.
In the following the scaled data set is considered to check whether the correction of the data improves the prediction results. Figure 6.2 also illustrates the screening results of the scaled data. It can be seen that some samples in the active region have slightly reduced activity due to their library position while the activity of other samples has increased. In general, the scaling procedure did not disturb the trends observed in the unscaled case: the active region also stayed active. In contrast to the unscaled data, another quite active region appeared around sample 97, $Mn_{0.6}Co_{0.3}Te_{0.1}$. The VoF for the ternary system MnCoTe using scaled data can be calculated at **0.31**, which is nearly identical to the unscaled case.

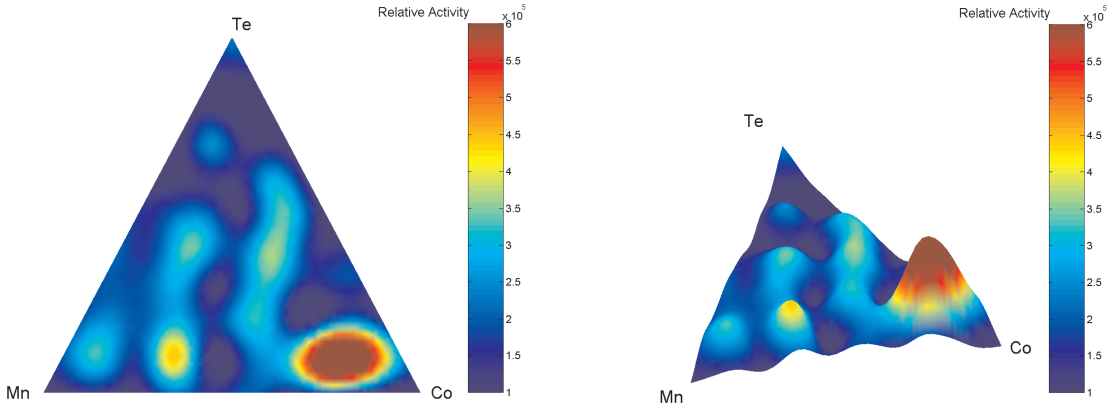


Figure 6.7: Kriging results for the scaled system MnCoTe.

Figure 6.7 gives the results of Kriging applied to the scaled data. For the calculations, the Kriging parameters already given in table 6.1 have been used and the corresponding experimental and fitted variogram are shown in Figure 6.4.
According to the Kriging model, the most active sample has been again estimated to lie within the active region at $Mn_{0.16}Co_{0.74}Te_{0.1}$ with a calculated activity value of 793,491 (Kriging variance: 52,610), lying below the maximum of the unscaled case. The following figure displays the comparison of measured and calculated activity values. In general it can be said that the model overestimated most samples of the second generation. Since the measured activity values of the second generation have been significantly smaller than those of the first generation the behavior of the model becomes clear. Why these deviations in activity occurred between first and second generation cannot be cleared. In fact, the values for the Hopcalite samples placed on the libraries reached comparable results, cf. section 4.8 page 85.

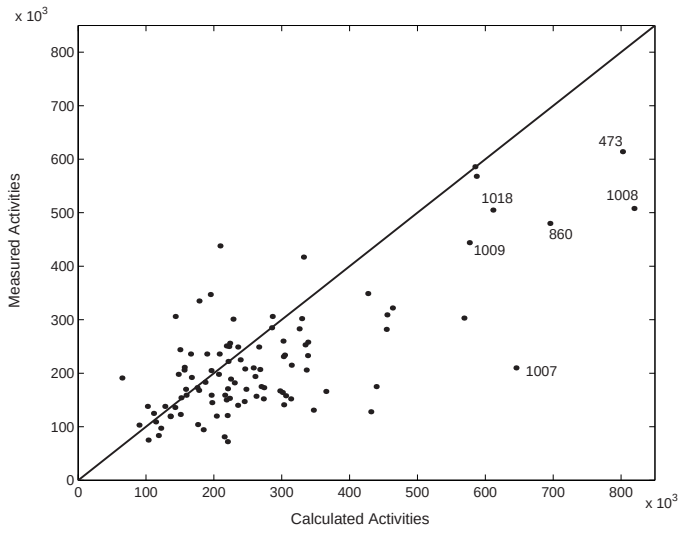


Figure 6.8: Comparison of calculated and measured activity values for the ternary system MnCoTe (scaled data).

The most active samples of the first generation 473 and 860 also reached less activity during the second screening and due to the conditions of the model the estimated values are exactly those measured in the first screening, the input data. From figure 6.8 this can be clearly observed. An explanation for this behavior might be that samples belonging to the second generation have been synthesized half a year later using new precursor solutions while the Hopcalite samples have been again taken out of the same charge. Furthermore, this indicates once again the difficulty of comparing samples being not synthesized out of the same precursor solutions. As discussed in Chapter 4 the reproduction of two identical libraries out of the same precursor solutions yielded excellent results. Thus, the synthesis route and the screening system themselves are working reliably.
Figure 6.2 also indicates that several samples lying in the close neighborhood of samples 473 and 860 showed quite large activities such that the positive results within this region of interest could be verified by the reproduction. For an intensive search of a new active catalyst for the oxidation of propene, the researcher should focus on the close neighborhood of samples 473 and 860.

6.1.3 Kriging of the Quaternary System MnCoTeNi

Similar to the discussion of the ternary system the quaternary system MnCoTeNi should be studied more closely. This system again exists twice and the good reproducibility among these two twin systems has been already reported within section 4.7. Thus, only the composition spread obtained out of data set A is dealt with here. The following tables give the best 10 samples for the unscaled and the scaled case, respectively.

Ranking	**Sample**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein ($\times 10^5$) Signal**
1	473	0.2	0.7	0.1	0.0	8.02
2	860	0.1	0.8	0.1	0.0	6.95
3	820	0.2	0.6	0.1	0.1	6.73
4	603	0.3	0.5	0.1	0.1	6.67
5	631	0.2	0.5	0.1	0.2	5.80
6	422	0.1	0.7	0.1	0.1	5.61
7	747	0.3	0.3	0.1	0.3	4.96
8	954	0.3	0.6	0.1	0.0	4.54
9	507	0.4	0.2	0.1	0.3	4.11
10	929	0.1	0.6	0.1	0.2	4.04

Table 6.2: Best catalysts of the system MnCoTeNi (unscaled data).

Table 6.2 shows that the two best samples here are ternary mixtures lying on the face of the tetrahedron. These two samples 473 and 860 have been already studied in the previous section. Therefore, this section exclusively focuses on the best quaternaries, starting with sample 820.

Ranking	**Sample**	**Mn**	**Co**	**Te**	**Ni**	**Acrolein ($\times 10^5$) Signal**
1	473	0.2	0.7	0.1	0.0	8.02
2	860	0.1	0.8	0.1	0.0	7.75
3	820	0.2	0.6	0.1	0.1	7.64
4	603	0.3	0.5	0.1	0.1	5.97
5	401	0.1	0.1	0.4	0.4	5.18
6	747	0.3	0.3	0.1	0.3	5.11
7	422	0.1	0.7	0.1	0.1	5.02
8	631	0.2	0.5	0.1	0.2	4.98
9	954	0.3	0.6	0.1	0.0	4.54
10	437	0.2	0.4	0.1	0.3	4.40

Table 6.3: Best catalysts of the system MnCoTeNi (scaled data).

The scaling procedure did not change the leading group of this composition spread and the ternary samples 473 and 860 stayed the most active ones. Again, sample 820 is the best quaternary here. In general, applying the scaling procedure to this data set did only cause slight changes among the best catalysts.
For the calculation of non-synthesized samples by the Kriging model, the experimental and theoretical variogram have to be calculated. In the following table, the chosen variogram parameters are given leading to the fitted variogram function illustrated by the figures on the following page.

	Unscaled Data	Scaled Data
Range	0.07	0.06
Sill	$5.8 \cdot 10^9$	$7.8 \cdot 10^9$
Nugget Effect	$2 \cdot 10^9$	$2 \cdot 10^9$
Smoothing Parameter	2	2

Table 6.4: Chosen Kriging parameters for the quaternary system MnCoTeNi (unscaled and scaled case)

Again, the experimental variograms have been fitted by variogram functions that possess a parabolic behavior near the origin together with a nugget effect since the same assumptions apply as for the ternary case. The nugget effect again models a possible variability among very similar samples (measurement noise) while the parabolic behavior near the origin stands for very regular variations (smooth variation of the activity in dependence of the chemical composition is assumed). Both theoretical variograms are again given by the Matérn class of variogram functions (see equation 2.59, p. 44), using the parameters summarized in Table 6.4.

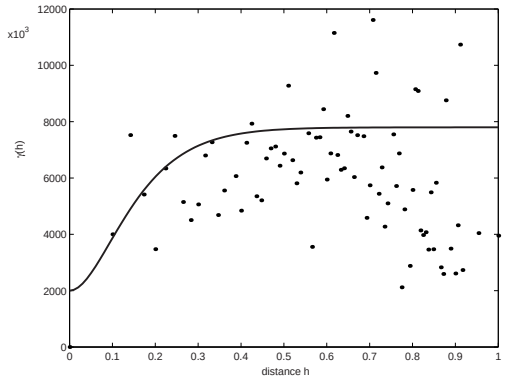


Figure 6.9: Experimental Variogram of MnCoTeNi taken out of data set A (no scaling used) together with the fitted variogram function.

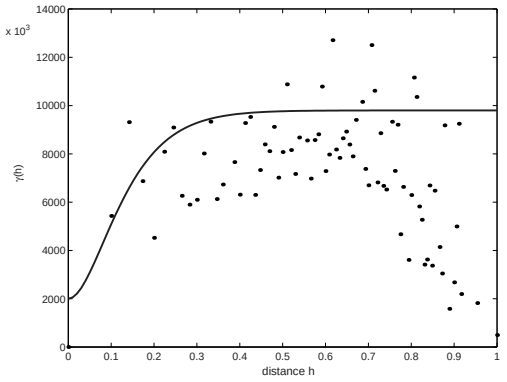


Figure 6.10: Experimental Variogram of MnCoTeNi taken out of data set A (scaling used) together with the fitted variogram function.

In the following figure, the result of Kriging applied to the unscaled quaternary data set MnCoTeNi is shown. It can be clearly seen, that the most active region appears around samples 473 and 860 that lie on one face (the triangle MnCoTe) of the tetrahedron, being ternary mixtures. Here and in the following representations the tetrahedron is turned around in such a way, that the triangle given by MnCoTe containing the most active samples comes to the front (Te content is plotted in upward direction). For this calculation, the tetrahedron has been sampled in 2%-increments giving 23,426 kriged points. The best composition calculated by the Kriging procedure lies in the ternary system MnCoTe that has been already studied in the previous section: $Mn_{0.18}Co_{0.72}Te_{0.1}$ with an estimated activity value of 822.489(Kriging variance: 64,327) while the maximum measured activity value is 802,000 at $Mn_{0.1}Co_{0.7}Te_{0.1}$). The calculation itself, performed on a common laptop computer, took about 5 minutes being completely implemented in MATLAB. Since the tetrahedron is a 3D object, looking inside is not possible in this illustration. To get a better understanding of what happens inside the tetrahedron, a more sophisticated illustration technique has been applied by slicing the tetrahedron in its upright direction. In this representation it is possible to look "inside" this 3D object searching for active regions. As already described in Chapter 5, the use of GUI View4d enables the scientist to interactively explore the tetrahedron by scrolling through all its calculated layers. In this studied example, it can be clearly observed that there is again a significant active region around samples 473 and 860. From the previous study of the corresponding ternary system it is already known that compositions around these samples are very active. In the quaternary case one can additionally observe that the addition of Ni as a fourth component also yielded quite active samples (cf. 2nd tetrahedron with Te: 10%). The active region extends towards the inner samples of the tetrahedron. However, the most active samples are still the ternaries that lie on the face MnCoTe of the tetrahedron containing no Ni.

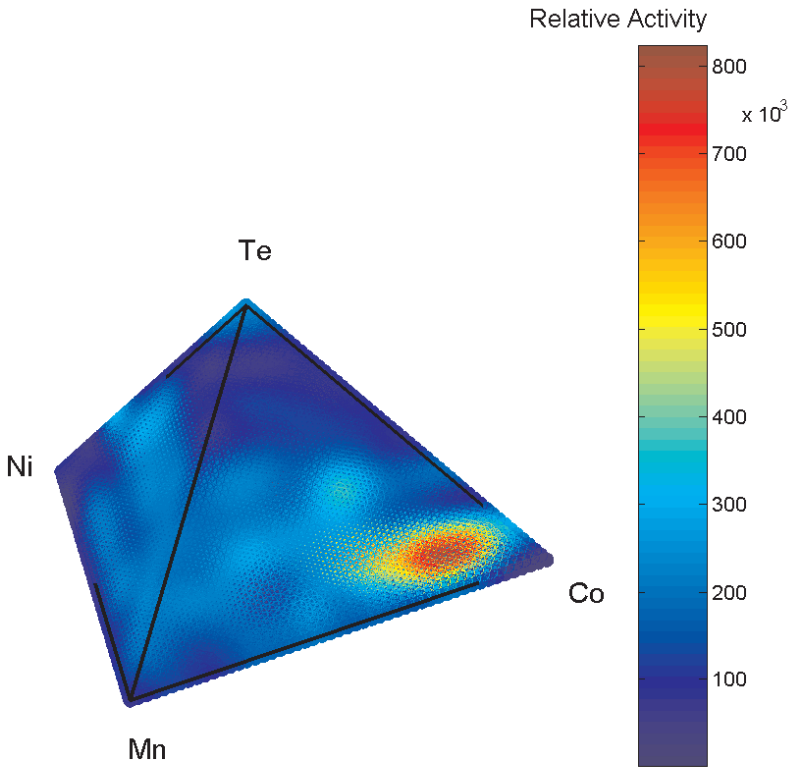


Figure 6.11: Result of Kriging the MnCoNiTe system at 2%-wise variation of composition.

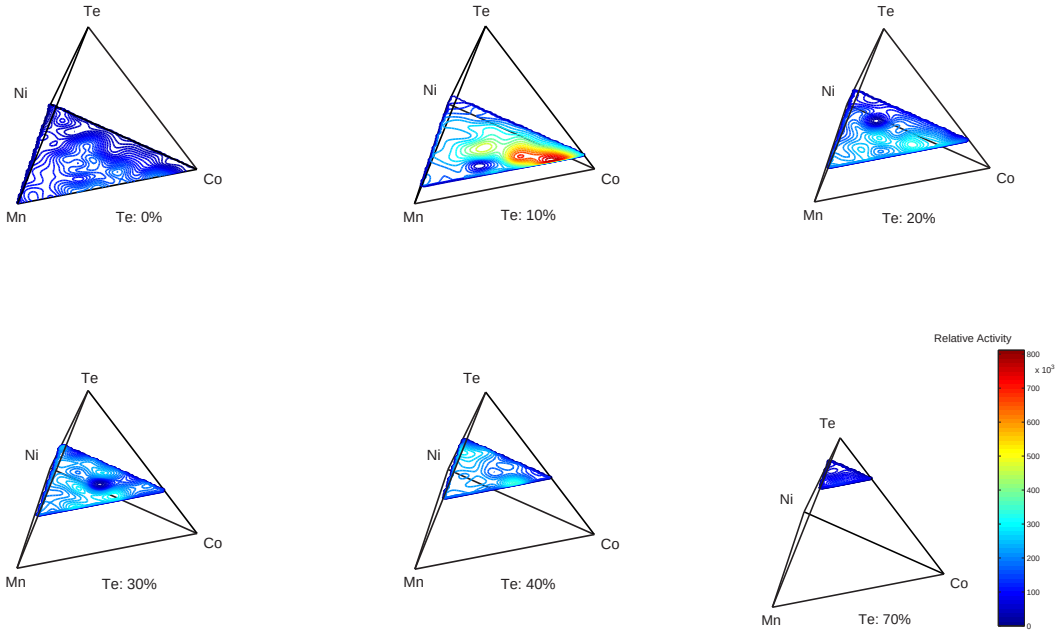


Figure 6.12: Results of Kriging the quaternary system MnCoTeNi illustrated in the slice plot representation.

Around the most active samples a second generation of catalysts has been prepared to verify this region of interest but also to compare the Kriging results with measured values. In total, for this example 38 catalysts have been re-synthesized being a refinement of the region of interest but also containing 10%-wise variations, cf. Figure 4.33 on page 89. For the comparison of measured and calculated activity values for these 38 re-synthesized samples, measured activity values have been again plotted against calculated ones (see below).
Also the scaled data system MnCoTeNi has been considered here. According to Table 6.3 the scaling procedure did only cause slight changes in the leading group of the most active samples. That is why the graphical considerations are restricted to the unscaled case, but the comparison between measured and estimated activity values is given for both cases.
Figure 6.13 shows the comparison of estimated and measured activity values for 38 quaternary samples prepared in the second generation. The VoF has been calculated to be **0.25**. It can be derived out of this figure, that the Kriging model tends both to over- and underestimate the activity here. Very good prediction have been obtained for catalyst 747 ($Mn_{0.3}Co_{0.3}Te_{0.1}Ni_{0.3}$). This sample showed very good performance again during its re-synthesis in the second generation. The other catalysts that have been considered more closely (631, 603, 820, 422) have been slightly overestimated by the Kriging model indicating that their activities measured in the second screening lie below the yielded results of the first screening. This trend has been already noticed when considering the ternary case MnCoTe.

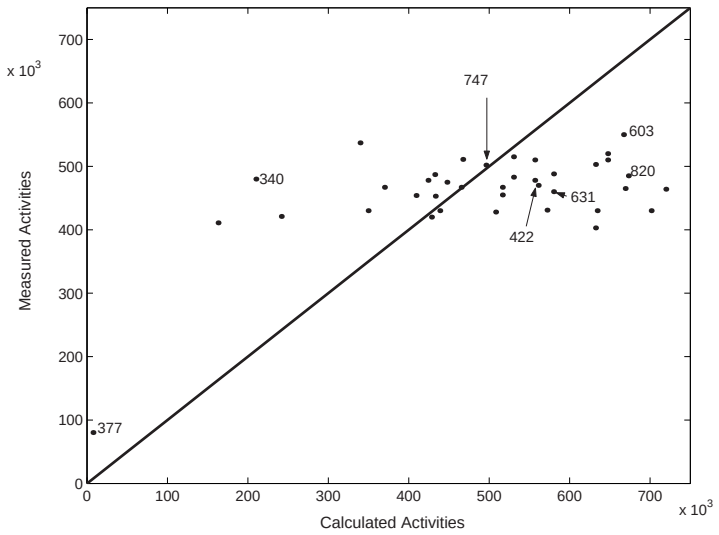


Figure 6.13: Comparison of calculated and measured activity values for the quaternary system MnCoTeNi (unscaled data).

The activity values of sample 340 and 377 have been underestimated here indicating that this sample gained a larger measurement result in the second generation than for the first screening. Summarizing, it can be said that in this case the re-synthesized samples did not follow a clear trend of over- or underestimation. The better the results of the second screening fit to the results of the first screening, the better can the Kriging model predict those activity values.

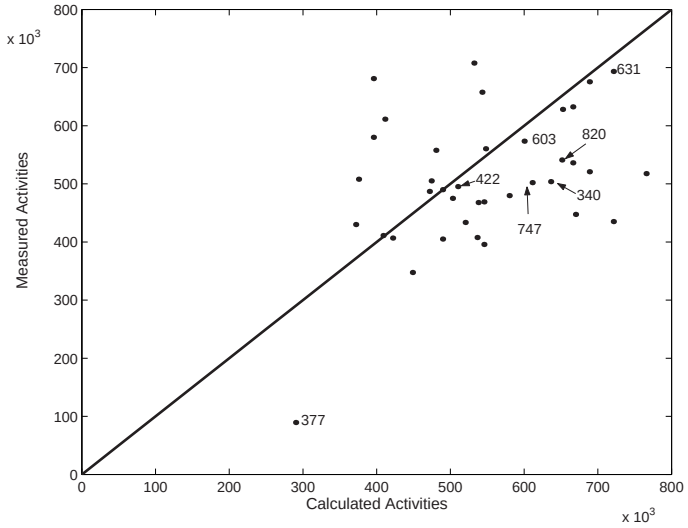


Figure 6.14: Comparison of calculated and measured activity values for the quaternary system MnCoTeNi (scaled data).

In this quaternary system, the reproduction yielded better results than those obtained in the ternary case studied previously.
The application of the scaling procedure for this quaternary system improved the predictive power for all marked samples except for sample 747. Samples 340 and 377 both changed from underestimation to overestimation, due to the scaling factors applied to the input data for the Kriging model. In this case, the VoF factor also lies at **0.25**. In general, the scaling procedure did not remarkably influence the predictive power here since the reproduction results of the second generation have only been slightly affected by it. In general it can be said that the Kriging model yielded better results in the quaternary case than in the ternary case due to a larger amount of input data. In the ternary case 66 measured catalyst samples served as input while the Kriging procedure can be performed for the quaternary case based on 286 samples.

The Quaternary system CrCoTeNi

The catalyst that reached the overall global maximum of activity in both data sets is the quaternary sample 501 ($Cr_{0.1}Co_{0.2}Te_{0.4}Ni_{0.3}$) followed by sample 685 ($Cr_{0.1}Co_{0.4}Te_{0.3}Ni_{0.2}$). Therefore also the neighborhood of these samples contained in the tetrahedron CrMnCoTe, cf. page 83 needs to be considered more closely. The following figure shows the slice plot representation of the tetrahedron containing samples 501 and 685 calculated by the Kriging procedure (range: 0.05, sill: 6 $\cdot 10^9$, nugget: 2 $\cdot 10^9$, smoothing parameter: 2).

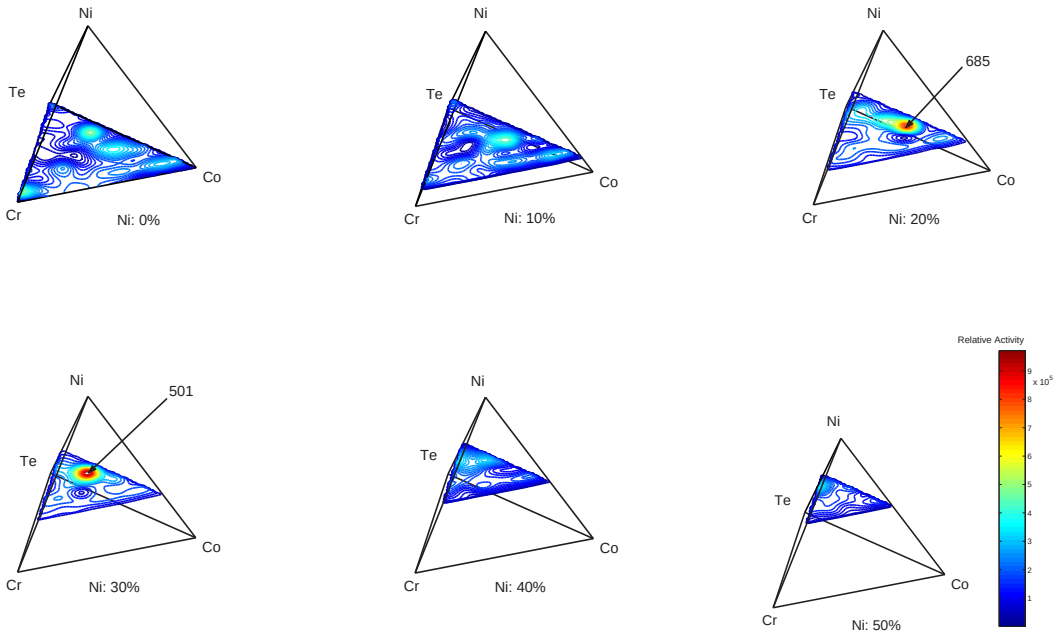


Figure 6.15: Results of Kriging the quaternary system CrCoTeNi (unscaled data) illustrated in the slice plot representation.

In this representation it might look as if the active region was only spread around the very active samples 501 and 685. Using the interactive visualization tool View4d (explained in Chapter 5) allows a continuous moving through the slices also displaying the layers between Ni: 20% and Ni: 30%. Then it can be observed that the active region is spread between the Ni range of 20% to 40% in vertical direction and not only just around those two hot spots. These regions should be examined in further studies more closely.

6.1.4 Kriging of the Complete Pentanary System CrMnCoTeNi

The best pentanary catalysts have been already listed within Chapter 4, cf. page 90. In this section the Kriging procedure is applied to the complete pentanary composition spread. It is interesting to see whether the model proposes another global activity maximum or a maximum lying close to samples 501 or 685. Furthermore the predictive power of the model can be checked when comparing the experimental screening results of 14 catalyst samples synthesized in a second generation. Again, only data set A is considered but also the scaled case. The following figure gives the experimental variograms (having a remarkably nice shape here) together with the fitted variogram functions. In contrast to previous experimental variograms more

points are available as input for the pentanary case. Again, it was decided to choose variogram functions having a parabolic behavior near the origin.

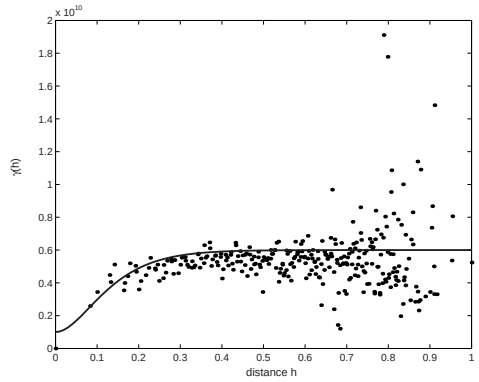


Figure 6.16: Experimental Variogram of CrMnCoTeNi taken out of data set A (no scaling used) together with the fitted variogram function.

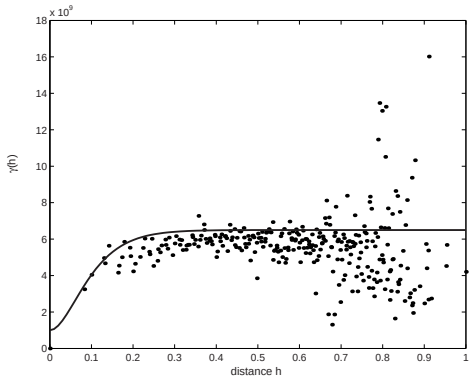


Figure 6.17: Experimental Variogram of CrMnCoTeNi taken out of data set A (scaling used) together with the fitted variogram function.

The chosen Kriging parameter are summarized in the following table:

	Unscaled Data	Scaled Data
Range	0.07	0.045
Sill	$5.2 \cdot 10^{10}$	$0.55 \cdot 10^{10}$
Nugget Effect	$0.1 \cdot 10^{10}$	$0.1 \cdot 10^{10}$
Smoothing Parameter	2	2

Table 6.5: Chosen Kriging parameters for the complete system CrMnCoTeNi (unscaled and scaled case)

For the complete data set 14 compositions have been re-synthesized and screened again to check the predictive power. Figure 6.18 summarizes the results obtained. In the unscaled case, the calculated activity values of samples 1382, 1381 and 1379 matched the measured activity values very well. Unfortunately, all other samples are overestimated by the model. Again, this is caused by the situation that the re-synthesized samples yielded lower activity values in the screening than during the first generation - a fact that has been observed throughout the whole studies and cannot be fully explained. Since the reproducibility has been excellent among the first generation of catalysts (all produced out of the same precursor batches), this deviation of screening results might be caused by the use of other precursor batches half a year later. In this study the calculated VoF has been **3.27** being extremely large compared to other studied systems.

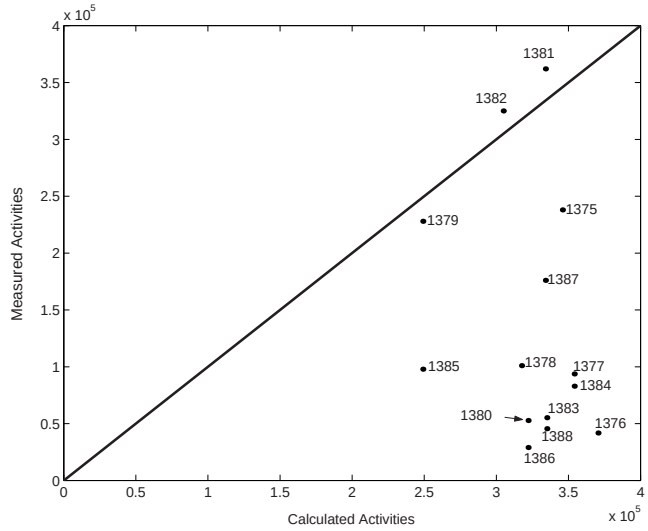


Figure 6.18: Comparison of calculated and measured activity values for the complete pentanary system CrMnCoTeNi (unscaled data).

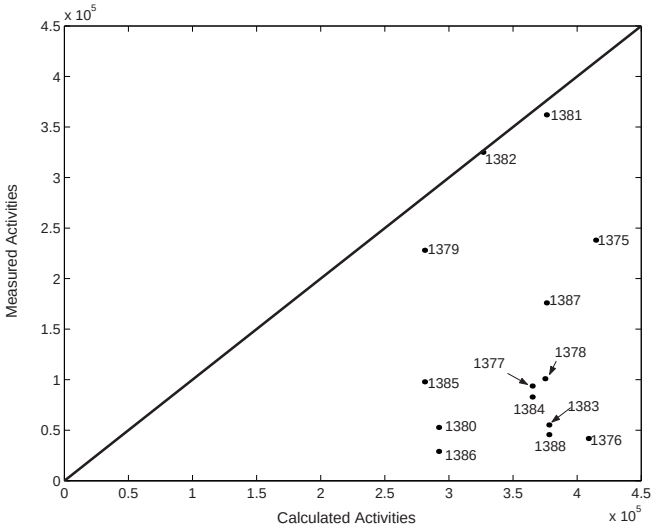


Figure 6.19: Comparison of calculated and measured activity values for the complete pentanary system CrMnCoTeNi (scaled data).

The behavior of calculated and measured activities in the scaled case is similar - samples 1382, 1381 and 1379 behave acceptably. The rest is again overestimated as already discussed above. For this prediction, a VoF of **3.46** has been calculated such that the scaling procedure could not help to improve the prediction results here. The Kriging estimators have been also calculated for a 2.5%-wise sampling of the pentanary composition space yielding 135,751 calculated compositions. The activity value of the most active catalyst, sample 501, has not been surpassed by the model. Unfortunately, it has not been possible to give a good visual presentation of the 135,751 calculated points. This has been too much for visualization via View5d since the program could not cope with the calculation of all connecting lines.

6.2 The B-Spline Model

In this section the interpolation results obtained by the B-Spline model are given. Again, the same examples chosen in the previous sections are discussed here to get a better feeling of similarities and differences of the two modelling procedures. First, the ternary system MnCoTe should be considered. An application of this procedure to heterogeneous catalyst data has been already published in [162].

6.2.1 The B-Spline Model Applied to the Ternary System MnCoTe

According to the B-Spline algorithm described in detail in Chapter 3, the fineness of control lattice Φ has a large influence on the final interpolation function. As Φ becomes finer, the influence of a data point is limited to a smaller neighborhood (comparable to a variogram model having a small range). This enables the data set to be more closely approximated, although the approximation surface will tend to contain more local peaks near the data points. The following figures will help to explain this behavior. In the figures it can be clearly observed how finer lattice sizes lead to sharp local maxima at the measured points. The calculated activity values at these points are overestimated and by far too large. It can also be seen how a finer lattice spacing causes larger (local) approximation values. Due to the construction of the B-Spline algorithm a finer lattice size leads to an interpolation of the data while coarser lattice sizes yield an approximation of the points. This can also be observed by taking a look at the ranges of the colorbar: the finer the control lattice the larger the estimated maximal activity value. For this case, only a lattice of 16×16 provides a good modelling result with a maximum activity value of 852,000 at the composition $Mn_{0.6}Co_{0.27}Te_{0.13}$ which exceeds the largest measured value of 820,000 at $Mn_{0.7}Co_{0.20}Te_{0.10}$. In total, 29,403 points have been calculated for these representations.

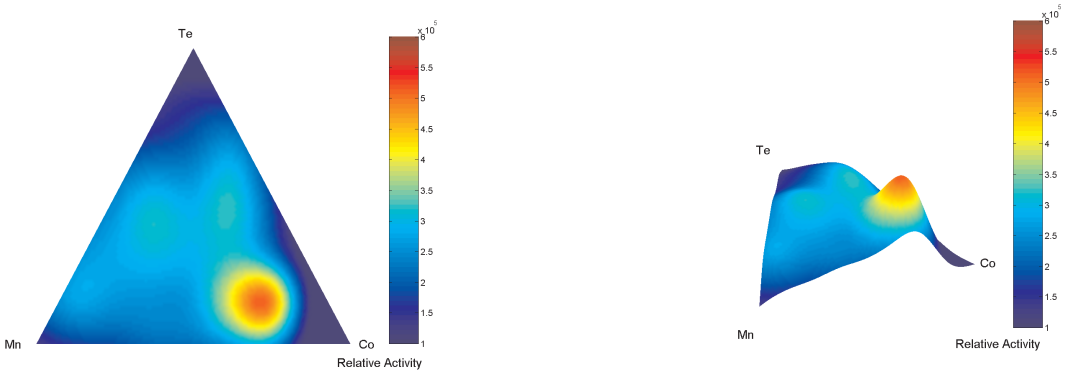


Figure 6.20: Modelling results of the B-Spline model applied to MnCoTe (unscaled data), lattice size 8×8. Left: 2D representation, Right: 3D view

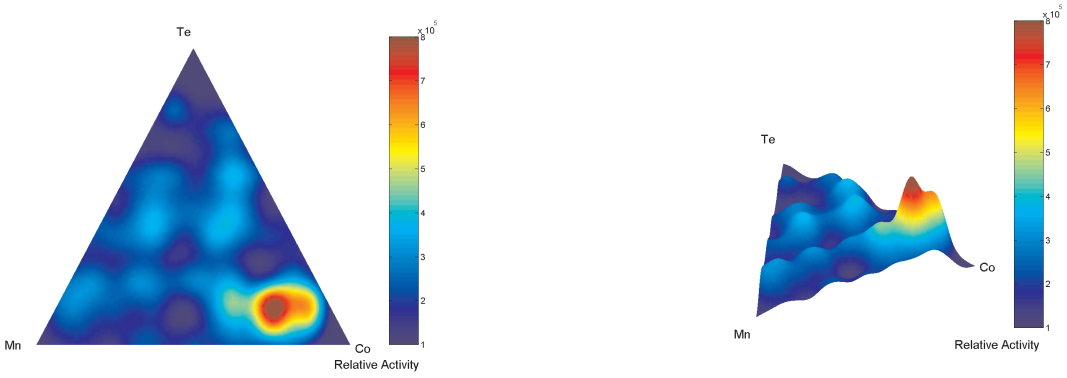


Figure 6.21: Results of the B-Spline Model with lattice size: 16×16

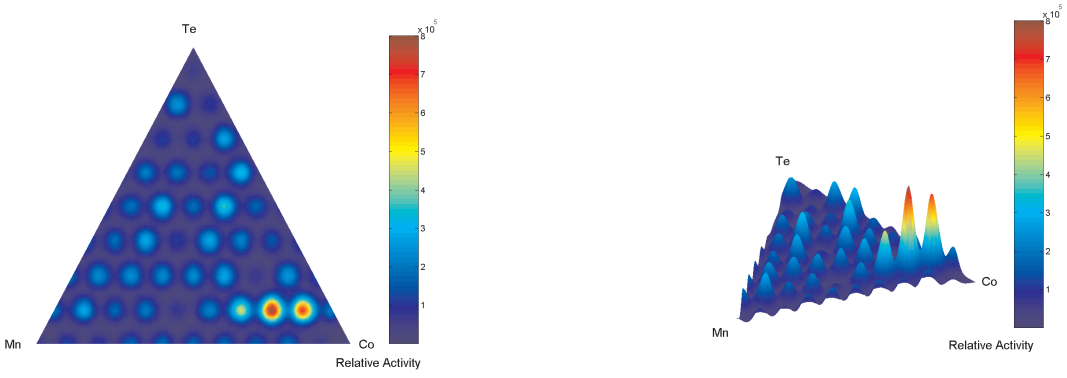


Figure 6.22: Results of the B-Spline Model with lattice size: 32×32

As already mentioned in section 6.1.2, 101 catalysts have be re-synthesized for this system using a finer increment of 5% but also containing 10%-wise original mixtures. The following figures illustrate the comparison of calculated and measured activities for the scaled and unscaled case.

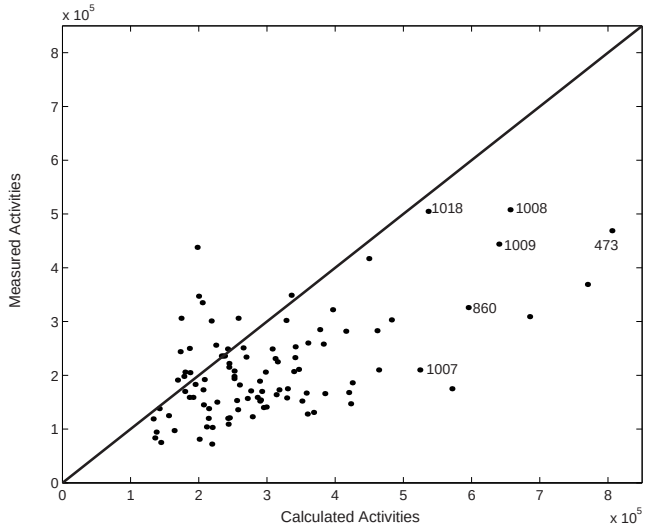


Figure 6.23: Comparison of calculated and measured activity values for the system MnCoTe (unscaled data). Lattice size: 16 × 16.

For the unscaled case the VoF has been calculated to be **0.61**. The same system modelled with the Kriging model yielded a VoF of **0.32**. The B-Spline model also tends to overestimate the activity values of the samples belonging to the second generation. Again, the low activity of the screening results of these 101 re-synthesized samples may be responsible. Nevertheless, for sample 1018 the predicted activity lies in good accordance with the experimental results. Especially, the most active sample of the first generation, sample 473, has been overestimated by the model since its screening result could not be confirmed during the second screening and re-synthesis. Applied to the same ternary data set MnCoTe, the Kriging model caused a smaller VoF such that this model describes the data in a better way here.

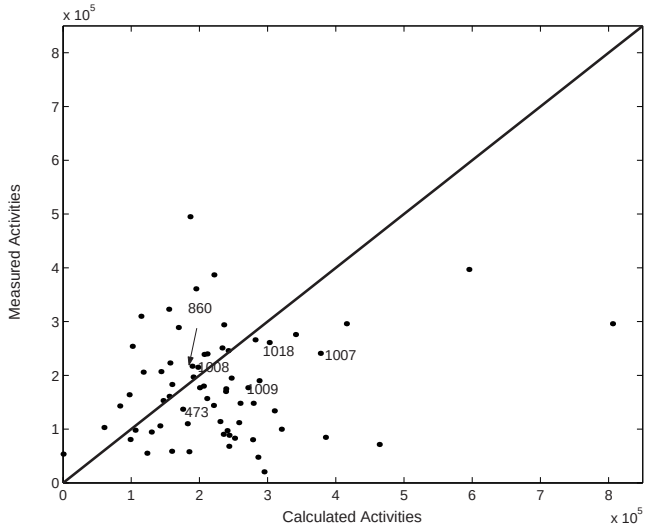


Figure 6.24: Comparison of calculated and measured activity values for the system MnCoTe (scaled data). Lattice size: 16 × 16.

Working with the scaled data, the VoF has been calculated to give **0.67**, even larger than for the unscaled case. Apart from two outliers in Figure 6.24 that cause the large VoF, the difference between measured and predicted activity values has been quite small, especially for samples 1008 and 860. Sample 860 already yielded a very good screening result in first generation. The corresponding application of the Kriging model yielded a VoF of **0.31**, nearly identical to the unscaled case. But still the Kriging procedure exceeds the predictive power of the B-Spline model by far. Taking a look at the results of the B-Spline modelling procedure applied to the scaled data (cf. Figure 6.2) the following situation arises, illustrated in Figure 6.25.

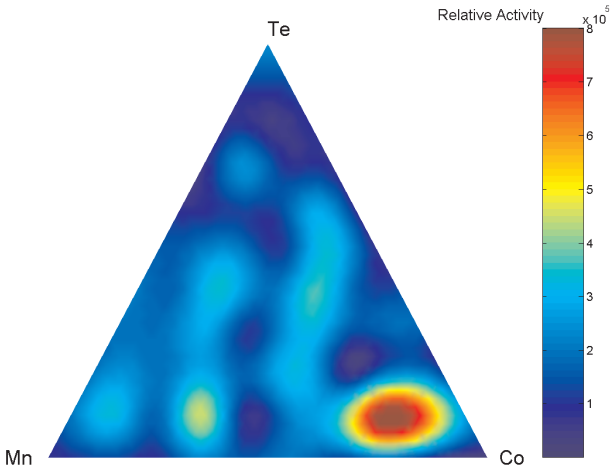


Figure 6.25: Results of applying the B-Spline model (lattice size: 16 × 16) to the scaled system MnCoTe.

The modelling result of the scaled case looks quite similar to the unscaled case. The maximal measured activity value in this data set has been 495,000 at $Mn_{0.2}Co_{0.5}Te_{0.3}$. The model estimates the maximal activity value for the composition $Mn_{0.17}Co_{0.46}Te_{0.36}$ at 479,000, lower than the largest measured activity value.

6.2.2 The B-Spline Model Applied to the Quaternary System MnCoTeNi

Similar to the Kriging approach, the quaternary system MnCoTeNi has been also modelled using the B-Spline algorithm. The maximal measured activity value in this system (unscaled case) has been obtained at $Mn_{0.2}Co_{0.7}Te_{0.1}Ni_{0.0}$ with an activity value of 802,000. This means, the best catalyst here is sample 473, a ternary mixture. In this application, a lattice resolution of 16 × 16 yielded the best results. The best catalyst according to the model has been calculated at $Mn_{0.17}Co_{0.68}Te_{0.15}Ni_{0.0}$ having an estimated activity value of 798,314 (+/− 11.400) staying below the maximal measured value. The following figure shows the results of the B-Spline procedure for the quaternary system MnCoTeNi (unscaled case) using the slice representation. The slices from 0% to 50% Te content have been chosen for illustration here leading to a good overview on this system. Again it must be noted here that a static illustration of discrete slices cannot cope with the interactive illustration of these

slice plots on a computer screen where the researcher can easily scroll through all possible levels.

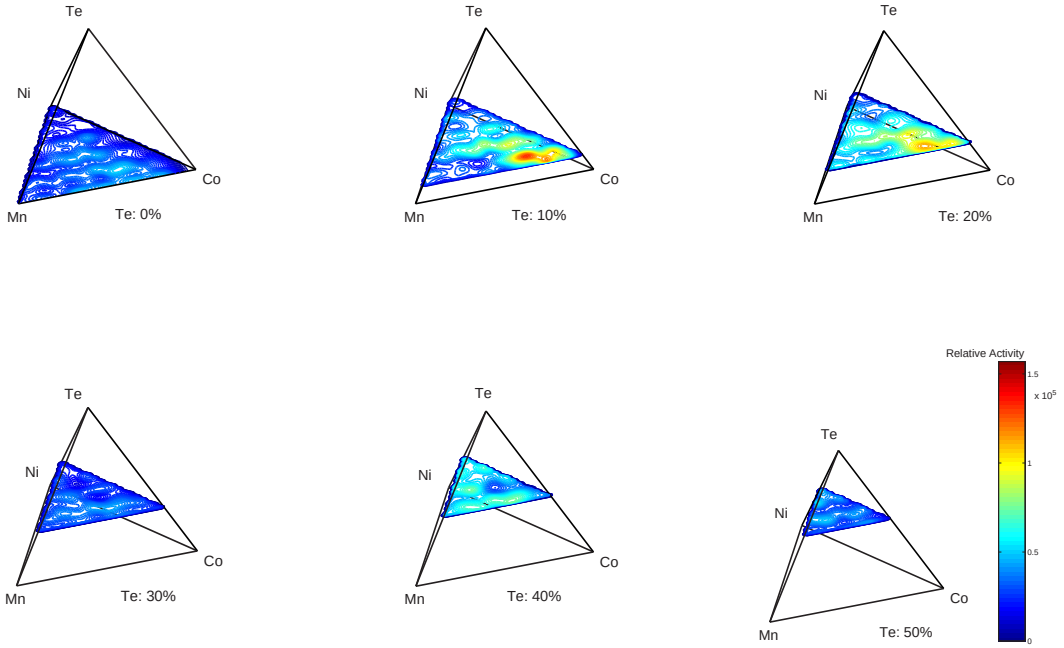


Figure 6.26: Results of modelling the quaternary system MnCoNiTe by B-Splines (unscaled data) illustrated in the slice plot representation.

Compared to the results of the Kriging model the B-Spline approach yielded nearly the same results. Just by visual inspection it is hard to say, where the differences between both procedures exactly are. The main difference can be observed taking a look at the range of the colorbar. For this data set, the Kriging system in general produced larger activity values.

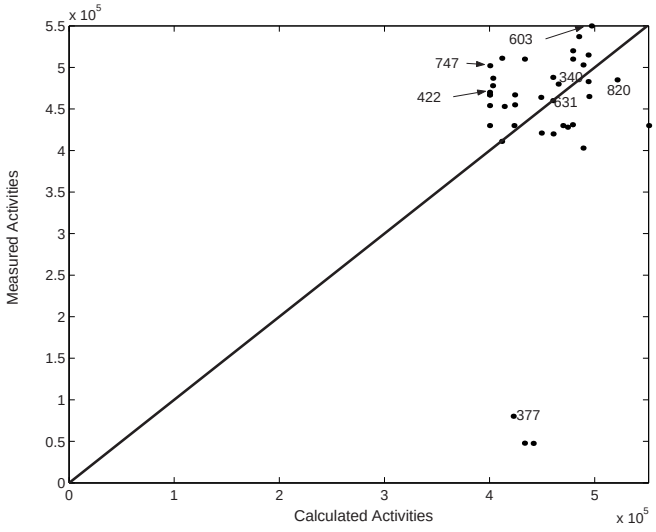


Figure 6.27: Comparison of calculated and measured activity values for the system MnCoTeNi (unscaled data) by B-Splines. Lattice size: 16 × 16.

The B-Spline procedure has been also used to check the differences between experimental and predicted activity values for 38 re-synthesized catalyst samples. Figures 6.27 and 6.28 contain the results:
The unscaled case resulted in a VoF of **0.63**. Except sample 377 together with two samples lying in its neighborhood the predictive power within this example has been quite good. The three outliers cause the large VoF here. In general, the trend of the data set could be captured.

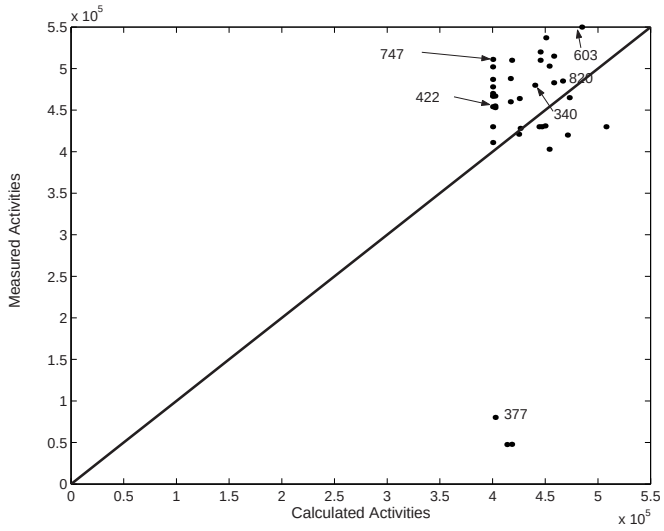


Figure 6.28: Comparison of calculated and measured activity values for the system MnCoTeNi (scaled data) by B-Splines. Lattice size: 16 × 16.

Considering the scaled case, only slight variations were observed. Again, around sample 377 two samples yielding very bad estimation results appear. The VoF value is calculated at **0.61** being once again enlarged by three outlying points. Compared to the Kriging results, the B-Spline model again resulted in larger VoF values indicating less predictive power.

6.2.3 The B-Spline Model Applied to the Complete Pentanary System CrMnCoTeNi

Unfortunately, the extension of the B-Spline algorithm to data sets in more than three dimensions could not be realized so easily. Since this approach could not cope with the predictive power of Kriging the time needed for a complete implementation has been used to improve the Kriging model and to develop several GUIs that allow a convenient access to the implemented MATLAB routines.

Part III

Experimental Part

7 Experimental Work

This chapter is dedicated to the experimental work that was done to synthesize a complete pentanary composition spread with 1001 samples containing the elements Mn, Co, Cr, Te and Ni. In the following sections, it is explained which metal precursors were employed and used in the acid-catalyzed sol-gel procedure to build the gels on the way to the final catalyst powders. Furthermore, it is explained in detail how the precursor solutions giving the later catalysts were mixed by a dispensing robot, calcined and transferred to the slate library plates to be accessible for the high-throughput screening. An explanation of the high-throughput screening reactor together with the entire setup is given within section 7.3.

7.1 Precursors

All the tested catalysts were prepared by the same synthesis route, following an acid-catalyzed sol-gel process that has been initially explained in section 1.5. Inspired by the work of Schmidt [149] who synthesized a complete pentanary composition spread out of the elements Mn, Co, Mo, Cr and Ni a quite similar approach was realized. Since a large amount of catalyst samples was planned to be synthesized the replacement of Mo obtained by the very expensive precursor Molybdenum-isopropoxide was necessary. This was realized by the use of Te instead of Mo after an adequate precursor for Te had been found. For the new compositions the recipe needed to be slightly modified. The incidental subsections deal with the synthesis of the metal propionates as metal ion precursors following the prescriptions described by Saalfrank [141].

7.1.1 Preparation of Cr(III)-Propionate

20,00 g (50.02 mmol) $Cr(NO_3)_3 \cdot 9H_2O$ was suspended in 300 ml propionic acid (4.0 mol) in a 500 ml one-necked round-bottomed flask. The reaction mixture was heated under reflux at 150 °C and constantly stirred. Then, the reflux was removed and the mixture was boiled at 150 °C for several days while the formation of nitrous oxides was observed (brown gases, toxic!). If the formation of NO_x became less (approx. after 2-3 days), the excessive propionic acid was removed by a rotary evaporator operated at 30 mbar and 50 °C. As a last step, the product was further dried at high vacuum (< 1 mbar) at 50 °C for several h to end up as a dark green powder which was further ground in a mortar. To check the composition of the synthesized chromium-propionate, a CHN-analysis was performed. The results are summarized in the following table.
The deviation of the theoretical results from the measured results can be explained by impurities such as amounts of water and traces of nitrate that remained in the powder and could not be removed totally.

	C	H	N
Theoretical Content [%]	39.88	5.58	0
Measured Content [%]	36.98	5.91	0.081

Table 7.1: Results of the CHN-Analysis of Cr(III)-Propionate.

7.1.2 Preparation of Co(II)-Propionate

The synthesis of Co(II)-propionate was realized following the description given by Saalfrank [141] and Spinolo [166].

$$CoCO_3 + 2CH_3 - CH_2 - COOH \rightleftharpoons Co(CH_3CH_2COO)_2 + H_2O + CO_2 \uparrow \quad (7.1)$$

In a 500 ml one-necked round-bottomed flask, 15 g (126.11 mmol) $CoCO_3$ were suspended in 250 ml propionic acid. The reaction mixture was continuously stirred and heated under reflux at 150 °C for 4 h. After cooling down, the dark violet suspension was filtered to remove remaining $CoCO_3$. Afterwards, excessive propionic acid was removed by a rotary evaporator operated at 30 mbar and 50 °C. Finally, the product was further dried in high vacuum (< 1 mbar) at 50 °C for several h to end up as a violet solid. Further grinding in a mortar decreased the grain size. To check the composition of the synthesized chromium-propionate, a CHN-analysis was performed and the corresponding results can be found in the following table:

	C	H	N
Theoretical Content [%]	35.14	4.91	0
Measured Content [%]	36.15	5.21	0.04

Table 7.2: Results of the CHN-Analysis of Co(II)-Propionate.

7.1.3 Preparation of Ni(II)-Propionate

The synthesis of Ni(II)-propionate was realized following the description given by Saalfrank [141].

$$Ni(OH)_2 + 2CH_3 - CH_2 - COOH \rightleftharpoons Ni(CH_3CH_2COO)_2 + 2H_2O \quad (7.2)$$

In a 500 ml one-necked round-bottomed flask, 10 g (107.89 mmol) $Ni(OH)_2$ were suspended in 100 ml propionic acid. The reaction mixture was continuously stirred

and heated under reflux at 150 °C for several h (transparent green solution). After cooling down, the excessive propionic acid was removed by a rotary evaporator operated at 30 mbar and 50 °C. Finally, the product was further dried in high vacuum at 50 °C for several h to give a light green solid. Further grinding in a mortar decreased the grain size. To check the composition of the synthesized nickel-propionate, a CHN-analysis was performed and the corresponding results are given in the following table:

	C	H	N
Theoretical Content [%]	35.18	4.92	0
Measured Content [%]	35.04	5.14	0.078

Table 7.3: Results of the CHN-Analysis of Ni(II)-Propionate.

7.1.4 Preparation of Mn(II)-Propionate

The preparation of Mn(II)-propionate was done using MnO as starting material.

$$MnO + 2CH_3 - CH_2 - COOH \rightleftharpoons Mn(CH_3CH_2COO)_2 + H_2O \tag{7.3}$$

In a 500 ml one-necked round-bottomed flask, 8.33 g (117.46 mmol) MnO were suspended in 200 ml propionic acid. The reaction mixture was continuously stirred and heated under reflux at 150 °C for several hours (transparent light brown solution). After cooling down, the excessive propionic acid was removed using by rotary evaporator operated at 30 mbar and 50 °C. Finally, the product was further dried in high vacuum (< 1 mbar) at 50 °C for several h to give a brownish solid. Further grinding in a mortar decreased the grain size. To check the composition of the synthesized nickel-propionate, a CHN-analysis was performed and the corresponding results can be found in the following table:

	C	H	N
Theoretical Content [%]	35.84	5.01	0
Measured Content [%]	35.23	4.92	0.075

Table 7.4: Results of the CHN-Analysis of Mn(II)-Propionate.

As the fifth metal precursor, telluric acid (H_6TeO_6) was used. All solid metal precursors were dissolved in Methanol in a concentration of 0.5 mol/l. To synthesize

all catalyst by a sol-gel approach, also propionic acid and a complexing agent (4-hydroxy-4-methyl-2-pentanone) were employed. The propionic acid route described by Saalfrank [141] was applied in a slightly modified way to synthesize all catalysts.

7.2 High-Throughput Synthesis of a Pentanary Composition Spread

The goal of this synthesis was the preparation of a complete pentanary composition spread sampling a five dimensional search space. The five dimensions correspond to the five used elements Mn, Co, Cr, Te and Ni. For the design of experiments, an equally spaced sampling of the whole composition search space in 10%-wise variation was applied. Then, the number of possible catalysts consisting of five metals is given by equation (1.2), introduced on page 11.

$$N(N_M, n_c) = \binom{n_c + N_M - 2}{n_c - 1} = \frac{(n_c + N_M - 2)!}{(n_c - 1)!(N_M - 1)!}$$

In the studied case, $N_M = 5$ (five elements) and $n_c = 11$, since the content of each element can take 11 different values (0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0) and this results in

$$N(5, 11) = \binom{11 + 5 - 2}{11 - 1} = \binom{14}{10} = \frac{14!}{4!\,10!} = 1001$$

samples in total. The 1001 prepared catalysts, their compositions together with a unique sample number are listed in the appendix of this thesis in section C.1 (pp. 161). To check the reliability of the synthesis and the robot system the composition spread was synthesized twice.

High-Throughput Synthesis of Catalysts

Two high-throughput syntheses of 1001 catalysts each were carried out with the help of a conventional dispensing robot (*Lissy* by Zinsser Analytic, cf. Figure 7.1).
The robot performed the mixing of the metal precursor solutions that were manually prepared before. For the mixing and further processing steps, sol-gel recipes were used that had been individually developed within several years by the Maier group. These recipes were implemented in the programming language *Python* to be accessible by the software *Plattenbau* [146]. This software tool was also developed by former members of the research group Maier to speed up the synthesis work of sol-gel catalyst preparations by dispensing robots. With the help of this software tool pipetting lists are created that can be directly exported to the robot. Furthermore, the software allows the design of a variety of rack layouts for the robot and library syntheses.
For the synthesis of the catalysts four metal precursors (Mn, Cr, Co and Ni) were used as propionates (cf. section 7.1) dissolved in Methanol (0.5 mol/l). For the fifth metal precursor Te, H_6TeO_6 was dissolved in Methanol (0.5 mol/l). The sol-gel recipe used in this approach was a slight modification of the propionic acid route described by Saalfrank [141]. To synthesize a mixed metal oxide of the composition

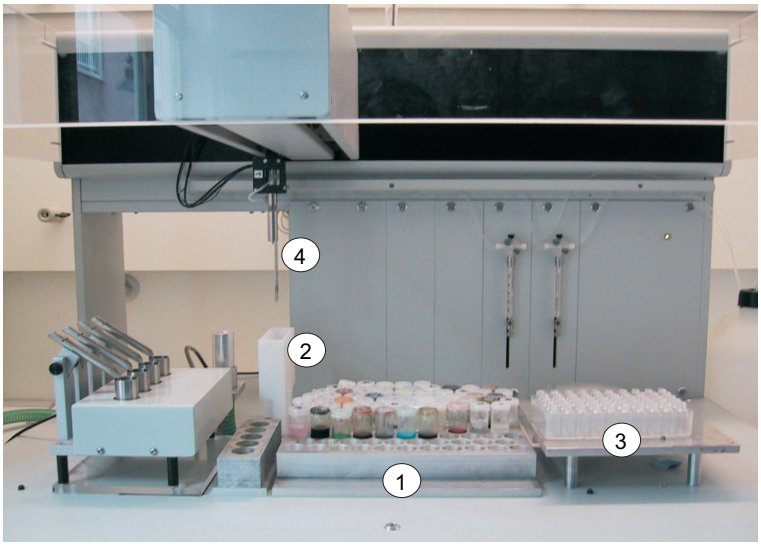


Figure 7.1: Zinsser dispensing robot used for the high-throughput synthesis: (1)Rack holding the precursor solutions (2) Washing station for needles (3) Synthesis rack holding the synthesis vials (4) Robot head with two needles

$A_aB_bC_cD_dE_eO_x$ the recipe can be characterized as follows:

$$\begin{gathered} \text{aA : bB : cC : dD : eE} \\ \vdots \\ \text{6(a+b+c+d+e)CA} \\ \vdots \\ \text{0.3(a+b+c+d+e)PA} \\ \vdots \\ \text{47.5(a+b+c+d+e)PS} \end{gathered}$$

where

- A, B, C, D, E denote the elements Mn, Co, Cr, Ni, Te.
- a, b, c, d, e denote the amounts of element A,...,E
- CA: complexing agent, 4-hydroxy-4-methyl-2-pentanone
- PA: Propionic Acid
- PS: Methanol, pure solvent

The precursor solutions were placed in 10 ml vials on a so-called *reagent rack* to be accessible for the needles of the robot. For the synthesis itself the liquids were transferred to small 2 ml vials that were arranged in a 50 sample layout on the *synthesis rack*. This means that within one run 50 samples were completed. Then the synthesis rack needed to be replaced. The order of pipetting the different solutions was always done in the same manner: complexing agent as first step, then propionic acid was added and finally all five metal precursor solutions were added in the following order: Cr, Mn, Co, Te and Ni. To intense the mixing the samples were

put on a vibrating plate for 30 min as soon as the pipetting has been terminated. Afterwards the samples were placed into a hood for one week to perform gelation. Then, the gels were calcined at 400 °C for five h to remove all remnants of organic compounds (solvent, propionic acid). In all cases, the produced catalyst amount was set to 400 μmol per sample. After the calcination, the catalyst powders needed to be stirred with a glass rod to improve their homogeneity in corn sizes. Finally, the powders were manually transferred into a library plate (containing 206 wells, thickness 6 mm, diameter 99 mm), Figure 7.2.

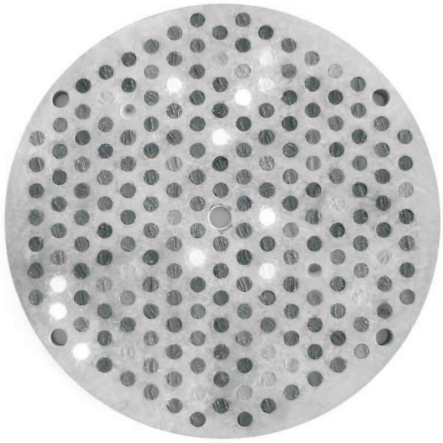


Figure 7.2: Picture of a slate library filled with catalyst samples.

The choice of the library material (slate or steel) depends on the screening method since emissivity corrected IR-Thermography (ecIR-Thermography) [72] requires materials that possess poor reflection properties together with a low reflectivity of infrared radiation. For the high-throughput screening setup that has been used in the primary screening process of all catalyst samples steel libraries were used in general. Here, slate libraries were used to allow further studies of a complete pentanary composition spread with ecIR-Thermography. Figure 7.2 illustrates a slate library filled with catalyst powders, Hopcalite and some empty wells for reference and background measurements, respectively. The used reference material Hopcalite is an industrial catalysts mainly consisting of MnO_2 and CuO in varying ratios often enhanced by CoO or Ag_2O. It is mostly applied for the oxidation of toxic CO to CO_2 in breathing apparatuses. The name *Hopcalite* is derived from the John Hopkins University ('Hop') and the University of California ('Cal') that discovered this catalyst family doing fundamental studies of the CO-oxidation during the first world war and in the 1920s.

7.3 High-Throughput Screening of Catalysts

The catalyst samples contained in the slate libraries were screened for propene oxidation (cf. section 1.7) in an open-well, high-throughput reactor system that has been described previously, cf. [87, 126, 174, 185]. In this reactor system, the catalyst library is placed in a custom-made reactor equipped with a heating system and an insulation. The cross-section of the assembled setup together with a magnification

of the reaction chamber is illustrated in Figure 7.3. During the screening the library is covered by a 15 mm thick ceramic mask made of Macor® that provides additional reaction volume on the top of each well. At the same time the mask serves as an insulation to guarantee a constant library temperature. At the top of this setup, a guiding mask is placed to help the threading of the sampling needle into each well. Inside the sampling needle a capillary bundle containing both the educt gas supply together with the product gas sampling system is arranged. The needle itself can be placed sequentially into each well of the library plate such that the space between catalyst powder and needle serves as reaction chamber for the studied reaction. Since the sampling needle is fixed, the whole reactor is moved by an *xyz*-stage to reach every single position of a well on the plate. Due to the open structure of the reactor, it cannot be avoided that moisture or oxygen from ambient air enters the reaction chambers. For the oxidation of propene, the educt gas total flow (5ml/min) consists of a mixture of 71.6 vol-% synthetic air and 28.4 vol-% propene (air/C_3H_6: 2.51) such that ambient air pouring towards the reaction chamber does not cause any problems.

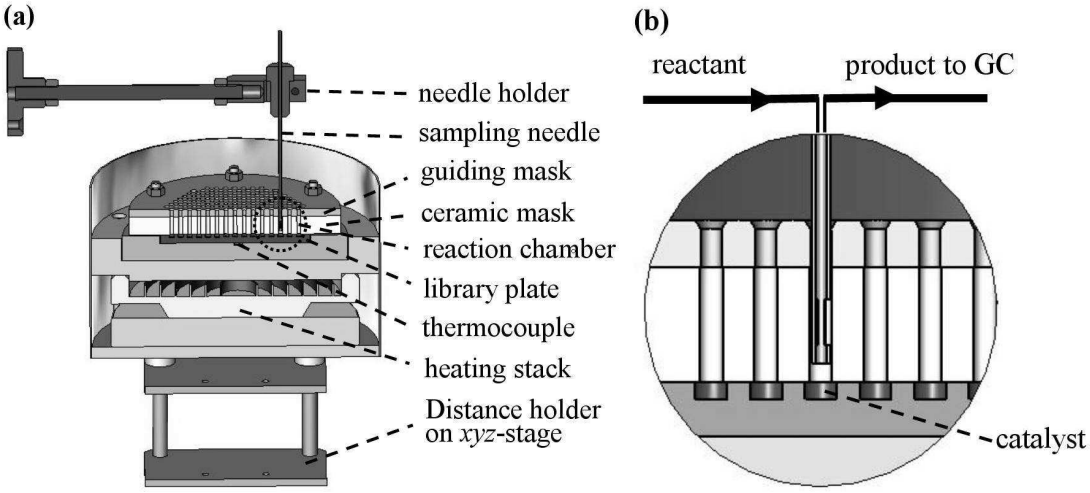


Figure 7.3: (a) Cross-section of the high-throughput reactor system (b) Magnification of the reaction chamber (dotted-circle in a)), [87].

During the reaction, the needle is placed directly over the catalyst sample such that the feed gas mixture can overflow the catalyst powder. The gas composition of products is analyzed by a micro-gas chromatograph (micro GC, model CP 4900; Varian), equipped with a thermal conductivity detector. Table 7.5 shows the used GC columns and the detectable chemicals.
The high-throughput screening system is run automatically controlled by the software TestRig [174] such that up to 200 catalyst samples can be screened within 8-10 hours. In a general experiment, it takes about 120 s to determine the catalytic activity of a catalyst by a single GC run (100 s reaction time, 10 s waiting outside hole, 10 s moving to the next position). The experimental setup parameters were determined by Schmidt [149] and needed not be modified since also the oxidation of propene was studied. With this setup, the complete screening process takes approximately 8 h.

GC CP 4900 Varian Columns	Chemicals
Porapak Q , 10m	1,5-hexadiene, propylene oxide, propionic aldehyde acetone, acrolein, benzene, H_2O, allyl alcohol
52CB, 4m (heated)	air, CO_2, propene

Table 7.5: Used GC columns and the corresponding detectable chemicals.

The sequential GC analysis is controlled by the GC software CP-Maitre Elite while the control of the *xyz*-stage is done via TestRig [174]. As model reaction, the oxidation of propene at 350 °C was studied. Therefore, the reactor was gradually heated up to 350 °C keeping the temperature fixed at 100 °C for about 30 min to remove possible water from the catalyst powders. The results of the screening were written into a *.txt file by the GC software including the retention times and peak areas of the detected chemicals. The feed gases propene and synthetic air were also detected during each GC run to provide an additional control of the reaction. Figure 7.5 shows the oxidation of propene and the reaction products that were monitored by gas chromatography.

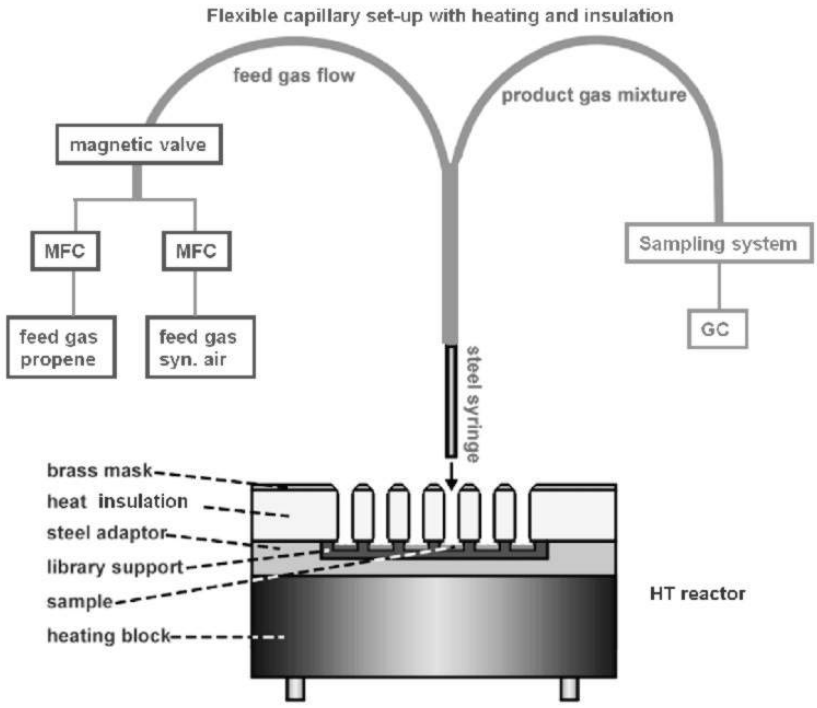


Figure 7.4: Scheme of the high-throughput screening reactor system and the applied GC analytic.

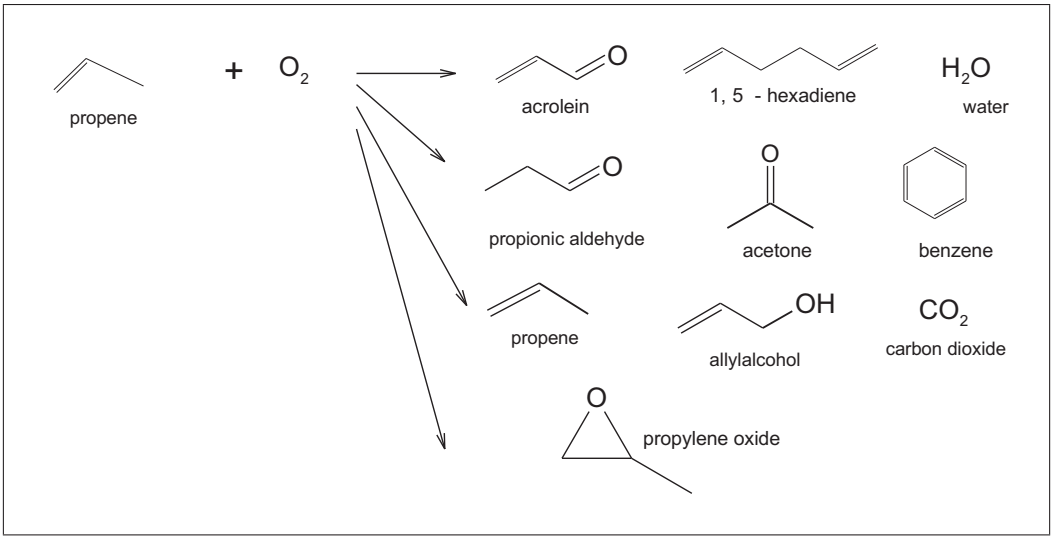


Figure 7.5: Detected products of propene oxidation at 350 °C.

Throughout this thesis the focus exclusively lies on the formation of acrolein out of propene. This means, the catalysts were validated by the amount of their acrolein production. Since all of the products mentioned above were monitored by the GC analytics, the collected data also include information about other products that can be further investigated. All synthesis results are contained in the data base *HT-dat* to allow an efficient access for further studies.

8 Summary and Outlook

8.1 Summary

Within the scope of this thesis several main issues have been addressed. First, the focus has been laid on the experimental work of synthesizing two identical pentanary composition spreads yielding 1001 samples each. Therefore, different precursors for Tellurium needed to be tested that allow the mixing with metal propionates without any precipitations. Good results using Telluric acid have been obtained here. The most active regions within this composition space have been sampled more closely by a second generation of catalysts consisting of 400 samples. In total, approximately **2,400** catalysts have been synthesized and screened using high-throughput synthesis techniques (a sol-gel approach realized by the use of pipetting robots) and an automated screening setup for the sequential testing of the samples with respect to their performance for the oxidation of propene. Especially the experimental work and experience in the laboratory helped to understand how the data are obtained and what the main error sources are. Different aspects of measurement accuracy became clear during the screening and testing of catalysts. The identical synthesis of two complete pentanary data has been realized yielding excellent screening results. Thus, the synthesis route using a sol gel procedure works reliably. It has been observed, that samples synthesized with the same precursor batches at the same time reached a high level of reproducibility. Unfortunately, this positive effect has not been observed for samples resynthesized with new precursor batches half a year later. This means that the screening results is strongly dependent on the quality of chemicals and precursors used. The reliability of the screening set up by itself has been checked by a library alternately filled with Hopcalite and empty wells. The variation of screening results ranged between 8-11% as a common result for primary high-throughput screening approaches. Also the existence of false positives or wrong negatives has been observed and has to be kept in mind when dealing with high-throughput screening data. The final decision about the active power of a catalyst can only be found in conventional tests and characterizations.
The main part of this thesis has been the application of two mathematical interpolation techniques, Kriging and multilevel B-Splines, to high-throughput screening data. One goal has been to check whether it is possible to predict activities of a priori non-synthesized samples out of the activities of neighboring samples. Therefore, the original data sets contained samples in 10%-wise variation of composition while the samples of the second generation have been synthesized in 5%-wise variations. One major problem that has been observed during the experimental and theoretical work has been the reproducibility of the high-throughput screening results. Two complete pentanary composition spreads have been synthesized within the shortest time period possible using the same precursor solutions. It has been possible to reproduce each library in an excellent way and the best catalysts of data set A also reached the best results in data set B. Unfortunately, the reproduction and refinement of regions

of interest of the composition space in a second generation of catalysts in general yielded lower activity values. This means that the best samples taken out of the first generation did not reach the expected results in the second generation. Why this has been the case can not be completely explained but only assumed. Since always a reference catalyst has been spread on each library plate the performance of these *known* samples could provide a control mechanism for the screening set-up. The reproducibility of the screening results of the Hopcalite samples always was acceptable, also for the second generation such that it can be excluded that the screening procedure itself caused the problems here. The deviating results between the first and second generation of catalyst might be due to the different precursor solutions or slight variations in the synthesis conditions. Although all synthesis parameters have been kept fixed and no changes in operating the samples have been performed. Since lower activities were obtained for the second generation, the predictive power of the Kriging and B-Spline model could not yield satisfying results for all examples studied. In most cases, the measured activity values of second generation catalysts have been overestimated by the models due to larger activity values of input data. Nevertheless, Kriging and multilevel B-Splines lend themselves in a very good way to model smooth correlations as it is assumed for the relation between catalytic activity and composition. In this application, the Kriging model yielded better estimation results due to its larger flexibility while the B-Spline model tends to overestimate the data leading to local maxima for finer lattice spacing.
The most important results of the high-throughput screening can be summarized as follows:
The catalyst that yielded the best results for the oxidation of propene is sample **501** $Cr_{0.1}Co_{0.2}Te_{0.4}Ni_{0.3}$. This catalyst has been synthesized three times yielding always excellent screening results. Also very good results have been observed for samples **685** ($Cr_{0.1}Co_{0.4}Te_{0.3}Ni_{0.2}$), **473** ($Mn_{0.2}Co_{0.7}Te_{0.1}$) and **860** ($Mn_{0.1}Co_{0.8}Te_{0.1}$), although the screening results for samples 473 and 860 in the second generation yielded lower results.
The best catalyst according to the Kriging model should have the composition $Mn_{0.18}Co_{0.73}Te_{0.09}$ (unscaled case) and $Mn_{0.18}Co_{0.73}Te_{0.09}$ for the scaled case. This means, the Kriging model always yielded a ternary mixture (MnCoTe) as global maximum.
For the B-Spline model, the facts are:
Only in the ternary case MnCoTe the predicted catalyst activity exceeded the maximal measured activity for the proposed composition $Mn_{0.6}Co_{0.27}Te_{0.13}$.
In general it can be said that several strong indications have been observed to assume a smooth dependence between composition and activity of a heterogeneous catalyst. The "similarity principle" already known in organic chemistry and biology seems to be also present for heterogeneous catalysts - but in a more complex way. It is clear that the composition plays an important role here but even more descriptors and factors need to be considered to get more complex models. Furthermore, more work need to be put in the development of reliably working synthesis and screening setups, to speed up the throughput for the search of new optimized catalysts.
All implementation work needed for this thesis has been done in MATLAB. In total, more than 10.000 lines of source code have been implemented to realize the calculations and extended visualization challenges. Especially the development of the GUIs took a large part of time and effort.

Within the scope of this thesis different tools have been developed that improve the handling of high-dimensional screening data. On the one hand, much work has been done to develop and apply several new visualization techniques and to provide convenient access to these techniques via the developed GUIs. On the other hand, the implemented Kriging and B-Spline approaches help to check for active regions in a data set and provide information on the regions between discrete data points. This information can then be very important for further experiments and catalyst generations.

8.2 Outlook

Throughout the whole thesis acrolein has been considered as product of interest among other products of the oxidation of propene. Since the GC signals of up to ten different products have been monitored, the whole discussion and analysis technique could also be applied to these data. To have a look at all monitored products would have gone beyond the scope of this thesis, so acrolein has been the product of choice. Nevertheless, the complete pentanary data sets contain large amounts of valuable information on all other products and the application of further data mining methods surly can reveal more trends and pattern. Researchers also dealing with selective propene oxidation can take the data sets and look for promising element combinations according to these primary screening results. The preparation of two identical pentanaries allows further interesting work on visualization, especially in higher dimensions. With the available data, visualization studies in more than 12 dimensions can be easily supported.
Since the Kriging model in general yielded better estimation results, further work can be focused on the optimization of the Kriging approach. An idea here could be to include locational uncertainty into the consideration of Kriging. This means that the Kriging model should be optimized by a term describing the deviation of input data from their real location. In the special case studied, the location of a data point has been given by its composition leading to a certain coordinate in space. Due to slight deviations between synthesized composition and theoretical composition (being always the case due to synthesis inaccuracies, only limited accuracy of the pipetting robot, mistakes in the concentrations of the precursor solution etc.) the assumed locations are not completely correct and these deviations could be included into the model. There exist some theories how these problems can be corrected and tackled by improved Kriging models (e.g. [50]).

Bibliography

[1] Abramowitz, M., Stegun, I. A.: *Handbook of Mathematical Functions.* Dover Publications, Inc., New York, 1972

[2] Agrafiotis, D. K.: *The Diversity of Chemical Libraries.* The Encyclopedia of Computational Chemistry, John Wiley & Sons UK, 742-761, 1998

[3] Akporiaye, D. E., Dahl, I. M., Karlsson, A., Wendelbo, R.: *Hydrothermalsynthesen von Zeolithen mit Methoden der kombinatorischen Chemie.* Angew. Chem., 110, 629-631, 1998

[4] Akporiaye, D., Wendelbo, R., Vistad, O., Karlsson, A., Plassen, M., Lillerud, K. P., McCoy, J.: *Combinatorial Hydrothermal Synthesis.* Zeoraito, 20, 18-24, 2003

[5] Ankerst, M., Keim, D., Kriegel, H.-P.: *Circle Segments: A Technique for Visually Exploring Large Dimensional Data Sets.* in: Proceedings of the IEEE Visualization Conference, 1996

[6] Ankerst, M.: *Visual Data Mining.* PhD-Thesis, Ludwig-Maximilians University Munich, Department of Mathematics and Informatics, Munich, 2000 dissertation.de - Verlag im Internet GmbH, Berlin, 2001

[7] Arai, N., Kim, T. W., Kubota, H., Matsumoto, Y., Koinuma, H.: *Combinatorial Fabrication and Cathodoluminescence Properties of Composition Spread $MHfO_3$:Tm^{3+}-films.* Appl. Surf. Sci. 197-198, 402-405, 2002

[8] Arntz, D., Böck, W., Burkhardt, W., Prescher, G.: EP Pat. 417722, Degussa, 1991

[9] Baker, B. E., Kline, N. J., Treado, P. J., Natan, M. J.: *Solution-Based Assembly of Metal Surfaces by Combinatorial Methods.* J. Am. Chem. Soc., 118, 8721, 1996

[10] Barnhill, R.: *Representation and Approximation of Surfaces.* in: Rice, J. R. (Ed.) Mathematical Software III, pp. 68-119, Academic Press, New York, 1977

[11] Bein, T.: *Effiziente Assays für kombinatorische Methoden zur Entdeckung von Katalysatoren.* Angew. Chem., 111, Nr. 3, 335-338, 1999

[12] Bettahar, M. M., Costentin, G., Savary, L., Lavalley, J. C.: *On the Partial Oxidation of Propane and Propylene on Mixed Metal Oxide Catalysts.* Appl. Catal. A: General, 145, 1, 1996

[13] Bingham, D., Sitter, R. R.: *Minimum-Aberration Two-Level Fractional Factorial Split-Plot Designs.* Technometrics, 41, 62-70, 1999

[14] Bisgaard, S., Sutherland, M.: *Split-Plot Experiments: Taguchi's Ina Tile Experiment Reanalyzed.* Quality Engineering, 16, 159-166, 2003

[15] Briceño, G., Chang, H., Sun, X., Schultz, P.G., Xiang, X.D.: *A Class of Cobalt Magnetoresistance Materials Discovered with Combinatorial Synthesis.* Science, 270, 273, 1995

[16] Brinker, C. J., Scherer, G. W.: *Sol-Gel Science.* Academic Press, Boston, 1990

[17] Bronstein, I. N., Semendjajew, K. A., Musiol, G.: *Taschenbuch der Mathematik.* Deutsch (Harri), 5. Aufl., 2001

[18] Brunauer, S., Emmett, P. H., Teller, E.: *Adsorption of Gases in Multimolecular Layers.* J. Am. Chem. Soc., 60, 309-319, 1938

[19] Burgess, K., Porte, A. M.,: *Accelerated Syntheses and Screening of Stereoselective Transition Metal Catalysts.* in: Doyle, M. P. (ed.), Advantages in Catalytic Processes, pp. 69-82, JAI Press, Greenwich, 1997

[20] Card, S. K., Mackinlay, J. D., Shneiderman, B.: *Readings in Information Visualization - Using Vision to Think.* Morgan Kaufman Publishers, San Francisco, 1999

[21] Cawse, J. N.: *Experimental Strategies for Combinatorial and High-Throughput Materials Development.* General Electric Company, 2000, Technical Report

[22] Cawse, J. N. (edt.): *Experimental Design for Combinatorial and High-Throughput Materials Development.* John Wiley & Sons, New Jersey, 2003

[23] Chilès, J. P., Delfiner, P.: *Geostatistics: Modeling Spatial Uncertainty.* Wiley, New York, 1999

[24] Cong, P., Doolen, R. D., Fan, Q., Giaquinta, D.M., Guan, S., McFarland, E. W.,Poojary, D. M., Self, K., Turner, H. W., Weinberg, W. H.: *High-Throughput Synthesis and Screening of Combinatorial Heterogeneous Catalyst Libraries.* Angew. Chemie, 111, 508-512, 1999; Angew. Chem. Int. Ed., 38, 484-488, 1999

[25] Corma, A., Serra, J. M.: *Heterogeneous Combinatorial Catalysis Applied to Oil Refining, Petrochemistry and Fine Chemistry.* Catalysis Today, 107-108, 3-11, 2005

[26] Corma, A., Serra, J. M., Serna, P., Moliner, M.: *Integrating High-Throughput Characterization into Combinatorial Heterogeneous Catalysis: Unsupervised Construction of Quantitative Structure/Property Relationship Models.* Journal of Catalysis, 232, 335-341, 2005

[27] Cressie, N. A. C.: *Statistics for Spatial Data.* Wiley Series in Probability and Mathematical Statistics, New York, 1993

[28] Danielson, E., Golden, J. H., McFarland, E. W., Reaves, C. M., Weinberg, W. H., Wu, X. D.: *A Combinatorial Approach to the Discovery and Optimization of Luminescent Materials.* Nature, 389, 944-948, 1997

[29] Danielson, E., Devenney, M., Giaquinta, D. M., Golden, J. H., Haushalter, R. C., McFarland, E. W., Poojary, D. M., Reaves, C. M., Weinberg, W. H., Wu, X. D.: *A Rare-Earth Phosphor Containing One-Dimensional Chains Identified Through Combinatorial Methods.* Science, 279, 837-839, 1998

[30] Dar, Y. L.: *High-Throughput Experimentation: A Powerful Enabling Technology for the Chemicaly and Materials Industry.* Macromol. Rapid Commun., 25, 34-47, 2004

[31] Delmon, B., Froment, G. (Eds.): *8th International Symposium on Catalyst Deactivation*, Elsevier, New York, 1999

[32] Dittmeyer,R., Winnacker, K., Küchler, L.: *Chemische Technik*, Vol. 5, pp. 135, 5. Aufl., Wiley-VCH, Weinheim, 2005

[33] Drochner, A., Fehlings, M., König, D., Vogel, H.: *Kinetisches Modell der Partialoxidation von Propen zu Acrolein an Bi/Mo-Mischoxiden.* Chem.-Ing. Techn. 75, 864-876, 2003

[34] Duff, D. G., Ohrenberg, A., Voelkening, S., Boll, M.: *A Screening Workflow for Synthesis and Testing of 10,000 Heterogeneous Catalysts per Day - Lessons Learned.* Macromol. Rapid Commun., 25, 169-177, 2004

[35] Dyn, N. *Interpolation and Approximation by Radial and Related Functions.* in: Chui, C., Schumaker, L., Ward, J. (Eds.) Approximation Theory VI, pp. 211-234, San Diego, Calif. Academic Press, 1989

[36] Dyn, N., Levin, D., Ruppa, S.: *Numerical Procedures for Global Surface Fitting of Scattered Data by Radial Functions.* SIAM J. Sci. Statis. Comput., Vol. 7, pp. 639-659, 1986

[37] Ebelmen, M. *Sur les Combinaisons des Acides Borique et Silicique avec les Ethers.* Ann. Chimie Phys., 16, 129, 1846

[38] Ebner, J. R., Thompson, M. R.: *An Active Site Hypothesis for Well-Crystallized Vanadium Phosphorus Oxide Catalyst Systems.* Catal. Today, 16, 51, 1993

[39] Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G.: *Statistik.* Springer-Verlag, Berlin, 2004

[40] Farin, G.: *A Modified Clough-Tocher Interpolant.* Comput. Aided Design, No. 2, pp. 19-27, 1985

[41] Fayyad, U. M., Piatetski-Shapiro, G., Smyth, P.: *The KDD Process for Extracting Useful Knowledge from Volumes of Data.* Comm. of the ACM, 39(11), 27-34, 1996

[42] Fayyad, U. M., Grinstein, G. G., Wierse, A.: *Information Visualization in Data Mining and Knowledge Discovery.* Morgan Kaufman Publishers, San Francisco, 2002

[43] Fehlings, M., König, D., Vogel, H.: *Explosionsgrenzen von Propen/Sauerstoff/Alkan-Mischungen und ihre Reaktivität an Bi/Mo-Mischoxiden.* Chem. Technik, 50, 241-245, 1998

[44] Fehlings, M., König, D., Vogel, H.: *Partialoxidation von Propen an Bi/Mo-Mischoxiden* Chem. Technik, 52, 90-95, 2000

[45] Franke, R., Nielson, G. M.: *Scattered Data Interpolation and Applications: A Tutorial and Survey.* in: Geometric Modelling: Methods and Their Application, Hagen, H., Roller, D.: (Eds.), pp. 131-160, Berlin: Springer-Verlag, 1991

[46] Franke, R., Nielson, G. M.: *Scattered Data Interpolation of Large Sets of Scattered Data.* Int. J. Numerical Methods in Eng., Vol. 15, pp. 1691-1704, 1980

[47] Frantzen, A., Sanders, D., Scheidtman, J., Simon, U., Maier, W. F.: *A Flexible Database for Combinatorial and High-Throughput Materials Science.* QSAR & Comb. Sci., 24, 22-28, 2005

[48] Frenzer, G., Maier, W. F.: *Amorphous Porous Mixed Oxides: Sol-Gel Ways to a Highly Versatile Class of Materials and Catalysts.* Annu. Rev. Mater. Res., 36, 281-331, 2006

[49] Fujikawa, N., Wakui, K., Tomita, K., Ooue, N., Ueda, W.: *Selective Oxidation of Propane Over Nickel Molybdate Modified with Telluromolybdate.* Catal. Today, 71, 83-88, 2001

[50] Gabrosek, J. G.: *The Effect of Locational Uncertainty in Geostatistics.* PhD-Thesis, Iowa State University, Ames, Iowa, 1999

[51] Glaeser, L. C., Brazdil, J. F., Hazle, M. A., Mehicic, M., Grasselli, R. K.: *Identification of Active Oxide Ions in a Bismuth Molybdate Selective Oxidation Catalyst.* J. Chem. Soc., Faraday Trans. 1, 81, 2903-2912, 1985

[52] Glassner, A. *Principles of Digital Image Synthesis.* San Francisco, Calif.: Morgan Kaufmann, 1995

[53] Gonzalez, R. D., Lopez, T., Gomez, R.: *Sol-Gel Preparation of Supported Metal Catalysts.* Catal. Today, 35, 293, 1997

[54] Grasselli, R. K., Burrington, J. D.: *Selective Oxidation and Ammoxidation of Propylene by Heterogeneous Catalysis.* Adv. Catal., 30, 133, 1981

[55] Grinstein, G. G., Hoffman, P., Laskowski, S. J., Pickett, R. M.: *Benchmark Development for the Evaluation of Visualization for Data Mining.* in: Fayyad, U., Grinstein, G. G., Wierse, A. (Eds.), Information Visualization in Data Mining and Knowledge Discovery, pp. 129-176, Morgan Kaufman Publishers Inc., San Francisco, 2002

[56] Grinstein, G. G., Trutschl, M., Cvek, U.: *High-Dimensional Visualizations.* Proceedings of the Visual Data Mining Workshop KDD, San Francisco, 2001

[57] Guglielmi, M., Carturan, G.: *Precursors for Sol-Gel Preparations.* J. Non-Cryst. Solids, 100, 16, 1988

[58] Haas-Santo, K., Fichtner, M., Schuber, K.: *Preparation of Microstructure Compatible Porous Supports by Sol-Gel Synthesis for Catalyst Coatings.* Appl. Catal. A: General, 220, 79, 2001

[59] Hahndorf, I., Buyevskaya, O., Langpape, M., Grubert, G., Kolf, S., Guillon, E., Baerns, M.: *Experimental Equipment for High-Throughput Synthesis and Testing of Catalytic Materials.* Chemical Engineering Journal, 89, 119-125, 2002

[60] Hamprecht, F. A., Agrell, E.: *Exploring a Space of Materials: Spatial Sampling Design and Subset Selection.* in: Cawse, J. N (Edt.), Experimental Design for Combinatorial and High-Throughput Materials Development, John Wiley & Sons, New Jersey, 2003

[61] Hanak, J. J.: *The 'Multiple-Sample Concept' in Materials Research: Synthesis, Compositional Analysis and Testing of Entire Multicomponent Systems.* Journal of Materials Science, 5, 964, 1970

[62] Härdle, W. *Applied Multivariate Statistical Analysis.* Springer, 2003

[63] Hardy, R.: *Multiquadratic Equations of Topography and Other Irregular Surfaces.* J. Geophysical Research, Vol. 76, No. 8, pp. 1905-1915, 1971

[64] Haslett, J., Bradley, R., Craig, P. S., Wills, G., Unwin, A. R.: *Dynamic Graphics for Exploring Spatial Data, with Application to Locating Global and Local Anomalies.* The American Statistican, 45, 234-242, 1991

[65] Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning.* Springer-Verlag, New York, 2001

[66] Hayashi, H., Ishizaka, A., Haemori, M., Koinuma, H.: *Bright Blue Phosphors in ZnO-WO_3 Binary System Discovered Through Combinatorial Methodology.* Appl. Phys. Lett., 82, 1365, 2003

[67] Hench, L. L., West, J. L.: *The Sol-Gel Process.* Chem. Rev., 90, 33, 1990

[68] Hoffman, P., Grinstein, G. G.: *A Survey of Visualization for High-Dimensional Data Mining.* in: Fayyad, U., Grinstein, G. G., Wierse, A. (Eds.), Information Visualization in Data Mining and Knowledge Discovery, pp. 47-82, Morgan Kaufman Publishers Inc., San Francisco, 2002

[69] Hoffman, P., Grinstein, G.: *Dimensional Anchors: A Graphic Primitive for Multidimensional Multivariate Information Visualizations.* presented at NPIV'99 (Workshop on New Paradigms in Information Visualization amd Manipulation), 1999

[70] Hoffmann, C., Wolf, A., Schüth, F.: *Parallel Synthesis and Testing of Catalysts under Nearly Conventional Testing Conditions.* Angew. Chem., 111, 2971, 1999; Angew. Chem. Int. Ed., 38, 2800, 1999

[71] Holzwarth, A., Denton, P., Zanthoff, H., Mirodatos, C.: *Combinatorial Approaches to Heterogeneous Catalysis: Strategies and Perspectives for Academic Research.* Catalysis Today, 67, 309-318, 2001

[72] Holzwarth, A., Maier, W. F.: *Catalytic Phenomena in Combinatorial Libraries of Heterogeneous Catalysts.* Platimum Metals Review, 44(1), 16-21, 2000

[73] Holzwarth, A., Schmidt, H.-W., Maier, W. F.: *Detection of Catalytic Activity in Combinatorial Libraries of Heterogeneous Catalysts by IR Thermography.* Angew. Chem., 110, 2788-2791, 1998; Angew. Chem. Int. Ed., 37, 2644-2647, 1998

[74] Hoschek, J., Lasser, D.: *Computer Aided Geometric Design.* Wellesley, Mass.: A.K. Peters, Ltd., 1993

[75] Hsu, W. M., Hughes, J. F., Kaufman, H.: *Direct Manipulation of Free-Form Deformations.* Computer Graphics (Proc. SIGGRAPH '92), Vol. 26, No. 2, pp. 177-184, 1992

[76] Inselberg, A.: *n-dimensional Graphics, Part I - Lines and Hyperplanes.* Technical Report G320-2711, IBM Los Angeles Scientific Center, 1981

[77] Inselberg, A.: *The Plane with Parallel Coordinates.* The Visual Computer, 1: pp. 69-91, 1985

[78] Inselberg, A, Dimsdale, B.: *Parallel Coordinates: A Tool for Visualization Multivariate Relations.* Chapter 9, pp. 199-233, Plenum Publishing Corporation, New York 1991

[79] Isaaks, E. H., Srivastava, R. M.: *An Introduction to Applied Geostatistics.* Oxford University Press, 1989

[80] Jandeleit, B., Weinberg, H. W.: *Putting Catalysis on the Fast Track.* Chem. Ind., No. 19, 795-798, 1998

[81] Jandeleit, B., Turner, H. W., Uno, T., van Beek, J. A. M., Weinberg, W. H.: *Combinatorial Methods in Catalysis.* Science, Vol. 2, No. 2, 1998

[82] Jandeleit, B., Schaefer, D. J., Powers, T. S., Turner, H. W., Weinberg, W. H.: *Kombinatorische Materialforschung und Katalyse.* Angew. Chemie, 111, 2648-2689, 1999

[83] Johann, T., Brenner, A., Schwickardi, M., Busch, O., Marlow, F., Schunk, S., Schüth, F.: *Parallele, photoakustische Echtzeit-Detektion von Reaktionsprodukten aus Katalysatorbibliotheken.* Angew. Chem., 114, 3096, 2002

[84] Joliffe, I. T.: *Principle Component Analysis.* Springer Verlag, New York, 2nd edition, 2000

[85] Katada, N., Niwa, M.: *Silica Monolayer Solid-Acid Catalyst Prepared by CVD.* Chem. Vapor Disposition, 2, 125, 1996

[86] Kim, D. K.: *Combinatorial and Conventional Development of New Catalysts for the CO_2 Reforming of Methane* PhD-Thesis, Saarland University, 2006

[87] Kim, D. K., Maier, W. F.: *Combinatorial Discovery of New Autoreduction Catalysts for the CO_2 Reforming of Methane.* J. Catal., 238, 142-152, 2006

[88] Kim, Y. J., Gao, K., Chambers, G. A.: *Selective Growth and Characterization of Pure, Epitaxial α-Fe_2O_3 (0001) and Fe_3O_4 (001) Films by Plasma-Assisted Molecular Beam Epitaxy.* Surf. Sci., 371, 358, 1997

[89] Kirsten, G., Maier, W. F.: *Strategies for the Discovery of New Catalysts with Combinatorial Chemistry.* Appl. Surf. Sci., 223, 87-101, 2004

[90] Kizling, M. B., Järas, S. G.: *A Review of the Use of Plasma Techniques in Catalyst Preparation and Catalytic Reactions.* Appl. Catal. A: General, 147, 1, 1996

[91] Klanner, C., Farrusseng, D., Baumes, L., Mirodatos, C.: *How to Design Diverse Libraries of Solid Catalysts.* QSAR & Combinatorial Science, 22, 729-736, 2003

[92] Klanner, C., Farrusseng, D., Baumes, L., Lengliz, M., Mirodatos, C., Schüth, F.: *The Development of Descriptors for Solids: Teaching "Catalytic Intuition" to a Computer.* Angew. Chem. Int. Ed., 43, 5347-5349, 2004

[93] Klein, S., Thorimbert, S., Maier, W. F.: *Amorphous Microporous Titania-Silica Mixed Oxides: Preparation, Characterization, and Catalytic Redox Properties.* J. Catal., 163, 476-488, 1996

[94] Klein, J., Lehmann, K. W., Schmidt, H.-W., Maier, W. F.: *Kombinatorische Materialbibliothek im µg-Maßstab am Beispiel der Hydrothermalsynthese.* Angew. Chem. 110, 3557-3561, 1998; Angew. Chem. Int. Ed., 37, 3369-3372, 1998

[95] Klein, J., Maier, W. F.: *Thermal Stability of Sol-Gel Derived Porous AM-Al_xZr Mixed Oxides.* Chem. Mater., 11, 2584, 1999

[96] Klein, J., Lettmann, C., Maier, W. F.: *Thermally Stable Silica-Based Amorphous Porous Mixed Oxides Prepared by Sol-Gel Procedures.* J. Non-Cryst. Solids, 282, 203, 2001

[97] Lebl, M.: *Parallel Personal Comments on "Classical" Papers in Combinatorial Chemistry.* J. Comb. Chem. 1, 3-24, 1999

[98] Lee, S., Chwa, K.-Y., Shin, S. Y., Wolberg, G.: *Image Metamorphosis Using Snakes and Free-Form Deformations.* Computer Graphics (Proc. SIGGRAPH'95), pp. 439-448, 1995

[99] Lee, S., Wolberg, G., Chwa, K.-Y., Shin, S.Y.: *Image Metamorphosis with Scattered Feature Constraints.* IEEE Trans. Visualization and Computer Graphics, Vol. 2, No. 4, pp. 337-354, 1996

[100] Lee, S., Wolberg, G., Shin, S. Y.: *Scattered Data Interpolation with Multilevel B-Splines.* IEEE Transactions on Visualization and Computer Graphics, Vol. 3, No. 3, 228-244, 1997

[101] Li, W. Y., Lettmann, C., Maier, W. F.: *Amorphous Porous M_xTi Mixed Oxides as Catalysts for the Oxidative Dehydrogenation of Ethylbenzene.* Catal. Lett., 69, 181, 2000

[102] Lippmaa, M., Koida, T., Minami, H., Jin, Z. W., Kawasaki, M., Koinuma, H.: *Design of Compact Pulsed Laser Deposition Chambers for the Growth of Combinatorial Oxide Thin Film Libraries.* Appl. Surf. Sci., 189, 205, 2002

[103] Maier, W. F., Bohnen, F. M., Heilmann, J., Klein, S., Ko, H.-C., Mark, M. F., Thorimbert, S., Tilgner, I.-C., Wiedorn, M.: *Sol-Gel Methods for the Production of Novel Catalytic Materials.* in: Harrod, J. F., Laine, R. M. (eds.), Application of Organometallic Chemistry in the Preparation and Processing of Advanced Materials , 27-46, Kluwer Academic Publishers, Dordrecht, 1995

[104] Maier, W. F.: *Kombinatorische Chemie - Herausforderung und Chance für die Entwicklung neuer Katalysatoren und Materialien.* Angew. Chemie, 111, Nr. 9, 1294-1296, 1999

[105] Maier, W. F.: *Sol-Gel Catalysts.* in: Cornils, B., Herrmann, W. A., Schlögl, R., Wong, C. H. (eds.): Catalysis from A to Z, Wiley-VCH, 534, 2001

[106] Maier, W. F., Stöwe, K., Sieg, S. C.: *Kombinatorische und Hochdurchsatz-Techniken in der Materialforschung.* Angew. Chemie, 119, 32, 6122-6179, 2007; Angew. Chem. Int. Ed., 46, 32, 6016-6067, 2007

[107] Marengo, S., Patrono, P., Comotti, P., Galli, G., et al.: *Propane Partial Oxidationover M^{3+}-substituted Vanadyl Phosphates Dispersed on Titania and Silica.* Appl. Catal. A: General, 230, 219-231, 2002

[108] Matérn, B.: *Spatial Variation.* Second Edition Springer-Verlag, New York, 1986

[109] Matheron, G.: *Traité de Géostatistique Appliquée, Tome I; Tome II: Le krigeage.* In: Mémoires du Bureau de Recherches Géologiques et Minières, No. 14 (1962), Editions Technip, Paris, II: Mémoires du Bureau de Recherches Géologiques et Minières, No. 24 (1963), Editions B.R.G.M., Paris

[110] Matheron, G.: *Les Variables Régionalisées et leur Estimation: Une Application de la Théorie des Fonctions Aléatoires aux Sciences de la Nature.* Masson, Paris, 1965

[111] Matheron, G.: *La Théorie des Variables Régionalisées et ses Applications.* Cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fasc. 5, Ecole des Mines de Paris. Translation (1971): *The Theory of Regionalized Variables and Its Applications.*

[112] Matheron, G.: *Estimating and Choosing.* Springer-Verlag, Berlin, 1989

[113] Matsumoto, Y., Murakami, M. et al.: *Structual Control and Combinatorial Doping of Titanium Dioxide Thin Films by Laser Molecular Beam Epitaxy.* Appl. Surf. Sci., 189, 344, 2002

[114] Maxwell, I. E.: *Combinatorial Chemistry - Connecting with Catalysis.* Nature, Vol. 394, 325-326, 1998

[115] Meurig, J.T.: *Design, Synthesis, and In Situ Characterization of New Solid Catalysts.* Angew. Chem. Int. Ed., 38, 3588, 1999

[116] Miertus, S., Fassina, G., Seneci, P. F.: *Concepts of Combinatorial Chemistry and Combinatorial Technologies.* Chem. Listy, 94, 1104-1110, 2000

[117] Mittasch, A.: *Early Studies of Multicomponent Catalysts.* Adv. Catal., 2, 81, 1950

[118] Mittasch, A.: *Geschichte der Ammoniaksynthese.* Verlag Chemie, Weinheim, 1951

[119] Miyao, T., Shishikura, I., Matsuoka, M., Nagai, M.: *CVD Synthesis of Alumina-Supported Molybdenum Carbide Catalyst.* Chemical Letters, 121, 561, 1996

[120] Moates, F. C., Somani, M., Annamalai, M., Richardson, J. T., Luss, D., Willson, R. C.: *Infrared Thermographic Screening of Combinatorial Libraries of Heterogeneous Catalysts.* Ind. Eng. Chem. Res., 35, 4801, 1996

[121] Mohebi, M.M., Evans, J. R. G.: *A Drop-On-Demand Ink-Jet Printer for Combinatorial Libraries and Functionally Graded Ceramics.* J. Comb. Chem., 4, 267, 2002

[122] Moro-Oka, Y., Ueda, W.: *Multicomponent Bismuth Molybdate Catalyst: A Highly Functionalized Catalyst System for the Selective Oxidation of Olefin.* Adv. Catal., 40, 233, 1994

[123] Murrell, L. L.: *Sols and Mixtures of Sols as Precursors of Unique Oxides.* Catal. Today, 35, 225, 1997

[124] Oellien, F.: *Algorithmen und Applikationen zur interaktiven Visualisierung und Analyse chemiespezifischer Datensätze.* PhD-Thesis, Friedrich-Alexander University Erlangen-Nürnberg, Department of Natural Sciences, 2003

[125] Ohnishi, T.,Komiyama, D., Koida, T., Ohashi, S., Stanter, C., Koinuma, H., Ohtomo, A., Lippmaa, M., Nakayawa, N., Kawasaki, M., Kikuchi, T., Omote, K.: *Parallel Integration and Characterization*

of Nanoscaled Epitaxial Lattices by Concurrent Molecular Layer Epitaxy and Diffractometry. Appl. Phys. Lett., 79, 536, 2001

[126] Orschel, M., Klein, J., Schmidt, H.-W., Maier, W. F.: *Erkennung der Selektivität von Oxidationsreaktionen auf Katalysatorbibliotheken durch ortsaufgelöste Massenspektrometrie.* Angew. Chemie, 111, Nr. 18, 2961, 1999; Angew. Chem. Int. Ed. 38, 2791, 1999

[127] Paul, J. S., Urschey, J., Jacobs, P. A., Maier, W. F., Verpoort, F.: *Combinatorial Screening and Conventional Testing of Antimony-Rich Selective Oxidation Catalysts.* Journal of Catalysis, Vol. 220, 136-145, 2003

[128] Pescarmona, P. P., van der Waal, J. C., Maxwell, I. E., Maschmeyer, T.: *Combinatorial Chemistry, High-Speed Screening and Catalysis.* Catalysis Letters, 63, 1-11, 1999

[129] Plunkett, M. J., Ellman, J. A.: *Combinatorial Chemistry and New Drugs.* Scientific American 276(4), 68-73, 1997

[130] Poncelet, G., Martens, J., Delmon, B., Jacobs, P. A., Grange, P. (eds.): *Preparation of Catalysts VI.*, Elsevier, Amsterdam, 1995

[131] Redlingshöfer, H., Kröcher, O., Böck, W., Huthmacher, K., Emig, G.: *Catalytic Wall Reactor as a Tool for Isothermal Investigations in the Heterogeneously Catalyzed Oxidation of Propene to Acrolein.* Ind. Eng. Chem. Res., 41, 1445, 2002

[132] Redlingshöfer, H., Fischer, A., Weckbecker, C., Huthmacher, K., Emig, G.: *Kinetic Modeling of the Heterogeneous Catalyzed Oxidation of Propene to Acrolein in a Catalytic Wall Reactor.* Ind. Eng. Chem. Res., 42, 5482, 2003

[133] Reetz, M. T., Becker, M. H., Kuhling, K. M., Holzwarth, A.: *Time-Resolved IR-Thermographic Detection and Screening of Enatioselectivity in Catalytic Reactions.* Angw. Chem., 110, 2792, 1998; Angew. Chem. Int. Ed., 37, 2647, 1998

[134] Renka, R. J.: *Algorithm 752; SRFPACK: Software for Scattered Data Fitting with a Constrained Surface Tension.* ACM Transactions on Mathematical Software, Vol. 22, No. 1, pp. 9-17, 1996

[135] Richardson, J. T.: *Principles of Catalyst Development.* Plenum, New York, 1989

[136] Riedel, E. *Anorganische Chemie.* Walter de Gruyter, Berlin, New York, 3.Auflage, 1994

[137] Rodemerck, U., Ignaszewski, Lukas, M., Claus, P.: *Parallel Synthesis and Fast Catalytic Testing of Catalyst Libraries for Oxidation Reactions.* Chem. Eng. Technol., 23, 413, 2000

[138] Rodemerck, U., Wolf, D., Buyevskaya, O. V., Claus, P., Senkan, S., Baerns, M.: *High-Throughput Synthesis and Screening of Catalytic Materials: Case Study on the Search for a Low-Temperature Catalyst for the Oxidation of Low-Concentration Propane.* Chem. Eng. J., 82, 3-11, 2001

[139] Rodemerck, U., Baerns, M., Holena, M., Wolf, D.: *Application of a Genetic Algorithm and a Neural Network for the Discovery and Optimization of New Solid Catalytic Meatrials.* Applied Surface Science, 223, 168, 2004

[140] Ruprecht, D., Muller, H.: *Image Warping with Scattered Data Interpolation.* IEEE Computer Graphics and Applications, Vol. 15, No. 2, pp. 37-43, 1995

[141] Saalfrank, J. W.: *Kombinatorische und konventionelle Entwicklung von CO-Oxidationskatalysatoren für Atemschutzgeräte und Niedertemperatur-Brennstoffzellen.* PhD-Thesis, Saarland University, 2003

[142] Saalfrank, J., Maier, W. F.: *Edelmetallfreie Katalysatoren für die CO-Oxidation bei Raumtemperatur durch gezielte Evolution.* Angw. Chemie, 116, 2062-2066, 2004; Angew. Chem. Int. Ed. 43, 2028-2031, 2004

[143] Sada, K., Yoshikawa, K., Miyata, M. *New Ammonium Carboxylate Host Compounds Screened by Combinatorial Chemistry.* J. Chem. Soc. Chem. Comm., 1763, 1998

[144] Satterfield, C. N.: *Heterogeneous Catalysis in Practice.* 2nd. ed. McGraw Hill, New York, 1991

[145] Scheidtmann, J., Weiss, P.-A., Maier, W. F.: *Hunting for Better Catalysts and Materials Using Combinatorial Chemistry and High Throughput Technology.* Applied Catalysis A: General, 222, 79-89, 2001

[146] Scheidtmann, J., Saalfrank, J. W., Maier, W. F.: *Plattenbau - Automated Synthesis of Catalysts and Materials Libraries.* in: Anpo, M., Onaka, M., Yamashita, H.: (Eds.) Studies in Surface Science and Catalysis 145, Elsevier, Tokyo, 13-21, 2003

[147] Scheidtmann, J., Klär, D., Saalfrank, J. W., Simon, U., Maier, W. F.: *Quantitative Composition Activity Relationships (QCAR) of Co-Ni-Mn-Mixed Oxide and M_1-M_2-Mixed Oxide Catalysts.* QSAR & Combinatorial Science, 24, 203-210, 2004

[148] Schlögl, R.: *Kombinatorische Chemie in der heterogenen Katalyse: ein neuer wissenscahftlicher Ansatz oder "des Kaisers neue Kleider"?.* Angew. Chem. 110, 2467-2470, 1998 Angew. Chem. Int. Ed., 37, 2333-2336, 1998

[149] Schmidt, T.: *Maßgeschneiderte Katalysatoren für die Verbrennung flüchtiger organischer Kohlenwasserstoffe in Abluftströmen.* PhD-Thesis, Saarland University, 2006

[150] Schüth, F., Busch, O., Hoffmann, C., Johann, T., Kiener, C., Demuth, D., Klein, J., Schunk, S., Strehlau, W., Zech, T. *High-Throughput Experimentation in Oxidation Catalysis.* Top. Catal., 21, 55, 2002

[151] Schumaker, L.: *Fitting Surfaces to Scattered Data.* in: Chui, C., Schumaker, L., Lorentz, G.: (Eds.) Approximation Theory II, pp. 203-268, Wiley, New York, 1976

[152] Schwarz, J. A. Contescu, C., Contescu, A.: *Methods for Preparation of Catalytic Materials.* Chem. Rev., 95, 477, 1995

[153] Senkan, S.: *High-Throughput Screening of Solid-State Catalyst Libraries.* Nature, 394, 350, 1998

[154] Senkan, S., Ozturk, S.: *Discovery and Optimization of Heterogeneous Catalysts by Using Combinatorial Chemistry.* Angew. Chem., 111, 867, 1999; Angew. Chem. Int. Ed., 38, 791, 1999

[155] Senkan, S., Krantz, K., Ozturk, S., Zengin, V., Onal, I.: *High-Throughput Testing of Heterogeneous Catalyst Libraries Using Array Microreactors and Mass Spectrometry.* Angew. Chem., 111, 2965, 1999; Angew. Chem. Int. Ed., 38, 2794, 1999

[156] Senkan, S.: *Combinatorial Heterogeneous Catalysis - A New Path in an Old Field.* Angew. Chem. Int. Ed., 40, 312-329, 2001; Angew. Chem., 113, 322-341, 2001

[157] Service, R. F.: *Combinatorial Chemistry: High Speed Materials Design.*, Science, 277 (5325), 474-475, 1997

[158] Service, R. F.: *Combinatorial Chemistry Hits the Drug Market.*, Science, 272 (5266), 1266-1268, 1996

[159] Shepard, D.: *A Two Dimensional Interpolation Function for Irregularly Spaced Data.* Proceedings of the 23rd ACM National Conference, pp. 517-524, ACM, New York, 1968

[160] Shimizu, K. D., Snapper, M. L., Hoveyda, A. H.: *High-Throughput Strategies for the Discovery of Catalysts.* Chem. Eur. J., 4, No. 10, 1885-1889, 1998

[161] Shneiderman, B.: *The Eyes Have It: A Task by Data-type Taxonomy for Information Visualization.* In: Proceedings of Visual Languages, IEEE Computer Science Press, Los Alamitos, CA, USA, 336-343, 1996

[162] Sieg, S. C., Stutz, B., Schmidt, T., Hamprecht, F., Maier, W. F.: *A QCAR-Approach to Materials Modeling.* J. Mol. Model., 12, 611-619, 2006

[163] Sieg, S. C., Suh, C., Schmidt, T., Stukowski, M., Rajan, K., Maier, W. F.: *Principal Component Analysis of Catalytic Functions in the Composition Space of Heterogeneous Catalysts.* QSAR & Comb. Sci., 26, 4, 528-535, 2007

[164] Snively, C. M., Oskarsdottir, G., Lauterbach, J.: *Chemically Sensitive Parallel Analysis of Combinatorial Catalyst Libraries.* Catal. Today, 67, 357, 2001

[165] Soukup, T., Davidson, I.: *Visual Data Mining - Techniques and Tools for Data Visualization and Mining.* Wiley Publishing Inc., New York, 2002

[166] Spinolo, G., Ardizzone, S., Trasatti, S. J.: *Surface Characterization of Co_3O_4 Electrodes Prepared by the Sol-Gel Method.* Electroan. Chem., 423, 49, 1997

[167] Stein, M. L.: *Interpolation of Spatial Data - Some Theory for Kriging.* Springer Series in Statistics, New York, 1999

[168] Stern, D. L., Grasselli, R. K.: *Propane Oxydehydrogenation over Molybdate-Based Catalysts.* J. Catal., 167, 550-569, 1997

[169] Strang, G.: *Introduction to Applied Mathematics.* Wellesley-Cambridge Press, Wellesley, 1986

[170] Stutz, B.: *Suchstrategien für Optimale Katalysatoren in der Kombinatorischen Chemie.* Diploma Thesis, Saarland University, Department of Informatics, 2004

[171] Su, H., Yeung, E. S.: *High-Throughput Screening of Heterogeneous Catalysts by Laser-Induced Fluorescence Imaging.* J. Am. Chem. Soc., 122, 7422, 2000

[172] Sun, X. D., Wang, K.-A., Yoo, Y., Wallace-Freedman, Gao, C., Xiang, X. D., Schultz, P. G.: *Solution-Phase Synthesis of Luminescent Materials Libraries.* Adv. Mater., 9, 1046-1049, 1997

[173] Tilgner, I.C., Fischer, P., Bohnen, F. M., Rehage, H., Maier, W. F.: *Effect of Acidic, Basic and Fluoride-Catalyzed Sol-Gel Transitions on the Preparation of Sub-Nanostructured Silica.* Microp. Mat., 5, 77, 1995

[174] Urschey, J., Weiss, P.-A., Scheidtmann, J., Richter, R., Maier, W. F.: *A Low Cost Reactor for High-Throughput Activity Screening of Heterogeneous Catalysts by Mass Spectrometry.* Solid State Science, 5, 909-916, 2003

[175] Urschey, J., Kühnle, A., Maier, W. F.: *Combinatorial and Conventional Development of Novel Dehydrogenation Catalysts.* Appl. Cat. A: General, 252, 91-106, 2003

[176] van Dover, R. B., Schneemeyer, L. F.: *The Codeposited Composition Spread Approach to High-Throughput Discovery/Exploration of Inorganic Materials.* Macromolecular Rapid Communications, 25, 150-157, 2004

[177] Vapnik, V.: *Statistical Learning Theory.* Wiley, New York, 1999

[178] Vissokov, G. P., Pirgov, P.S.: *Experimental Studies on the Plasma-Chemical Synthesis of a Catalyst for Natural Gas Reforming.* Appl. Catal. A: General, 168, 229, 1998

[179] Wachs, I.E.: *Recent Conceptual Advances in the Catalysis Science of Mixed Metal Oxide Catalytic Materials.* Catal. Today, 100, 79, 2005

[180] Wackernagel, H.: *Multivariate Geostatistics.* Springer-Verlag, Berlin, Heidelberg, 2003

[181] Wang, J., Yoo, Y., Gao, C., Takeuchi, I., Sun, X., Chang, H., Xiang, X.-D., Schultz, P. G.: *Identification of a Blue Photoluminescent Composite Material from a Combinatorial Library.* Science, 279, 837, 1998

[182] Ward, D. A., Ko, E. I.: *Preparing Catalytic Materials by the Sol-Gel Method.* Ind. Eng. Chem. Res. 34, 421, 1995

[183] Ware, C.: *Information Visualization - Perception for Design.* Morgan Kaufman Publishers, San Francisco, 1999

[184] Weinberg W. H., Jandeleit, B., Self, K., Turner, H. W.: *Combinatorial Methods in Homogeneous and Heterogeneous Catalysis.* Curr. Opin. Solid State Mat. Sci., 3, 104-110, 1998

[185] Weiss, P.-A., Saalfrank, J. W., Scheidtmann, J., Schmidt, H.-W., Maier, W. F.: in: Potyrailo, R. A., Amis, E. J. (Eds.) High-Throughput Analysis: *A Tool for Combinatorial Materials Science.* Kluwer Academic/Plenum, New York, 125, 2003

[186] Weiss, P.-A., Saalfrank, J. W., Scheidtmann, J., Schmidt, H.-W., Maier, W. F.: *High-Throughput Gas Chromatography and Mass-Spectrometry for Heterogenenous Catalysis: Screening of Catalytic Activities and Selectivities.* ChemInform, 36, 21, 2005

[187] Weissermel, K., Arpe, H.-J.: *Industrial Organic Chemistry.* Wiley-VCH, 4. Aufl., pp. 291, 2002

[188] Wilson, E. K.: Chem. Eng. News, 75(49), 24-25, 1997

[189] Wippert, C.: *Neue Allyl- und Oxoverbindungen des Molybdäns: Synthese, Struktur und Eigenschaften.* PhD-Thesis, University of Heidelberg, 2004

[190] Witten, I. H., Frank, E.: *Data Mining - Practical Machine Learning Tools and Techniques with Java Implementations.* Morgan Kaufman, San Francisco, 1999

[191] Wolf, D., Buyevskaya, O., Baerns, M.: *An Evolutionary Approach in the Combinatorial Selection and Optimization of Catalytic Materials.* Appl. Cat. A: General, 200(1-2), 63-77, 2000

[192] Wolf, D.: *An Evolutionary Strategy for the Design and Evaluation of High-Throughput Experiments.* in: Derouane, E. G., Parmon, V., Lemos, F., Ribeiro, F. R. (Eds.), Principles and Methods for Accelerated Catalyst Design and Testing, Kluwer Academic Publishers, Dordrecht, pp. 125-133, 2002

[193] Wolter, T., Maier, W. F.: *Combinatorial Search for Low-Temperature Combustion Catalysts.* in: Potyrailo, R. A., Karim, A., Wung, Q., Chikyow, T. (Eds.), MRS-Symposium Proceedings, Vol. 804, 283-293, 2004

[194] Wong, P. C., Bergeron, R. D.: *30 Years of Multidimensional Multivariate Visualization.* In: Proceedings of the Workshop of Scientific Visualization, IEEE Computer Society Press, Los Alamitos, CA, USA, 1995

[195] Xiang, X. D., Sun, X., Briceño, G., Lou, Y., Wang, K.-A., Chang, H., Wallace-Freedman, W. G., Chen, S. W., Schultz, P. G.: *A Combinatorial Approach to Materials Discovery.* Science, 268, 1738, 1995

[196] Yamamotoa, Y., Takahashib, R., Matsumotoc, Y., Chikyowa, T., Koinumaa, H.: *Mathematical Design of Linear Action Masks for Binary and Ternary Composition Spread Film Libraries.* Applied Surface Science, 223, 9-13, 2004

[197] Zanthoff, H. W., Lahmer, M., Baerns, M., Klemm, E. et al.: *Enhanced Product Selectivity in Partial Oxidation of Propane on Multicomponent Oxide Catalysts by Masking of Total Oxidation Sites.* J. Catal., 172, 203-210, 1997

List of Figures

List of Tables

Part IV

Appendix

A List of Abbreviations

°C	degree Centigrade
h	hour(s)
min	minutes
l	liter
ml	milliliter
μl	microliter
nl	nanoliter
Vol-%	volume percent
mbar	millibar
DOE	Design of Experiment
GA	Genetic Algorithm
ANN	Artificial Neural Network
GUI	Graphical User Interface
PCA	Principal Component Analysis
QSAR	Quantitative Structure Activity Relationship
QCAR	Quantitative Composition Activity Relationship
KDD	Knowledge Discovery in Databases
PVD	Physical Vapor Deposition
HT	High-Throughput
DM	Data Mining
IR	Infrared
TMOS	Tetramethoxysilane
TEOS	Tetraethoxysilane
LIFI	Laser Induced Fluorescence Imaging
REMPI	Resonance-Enhanced Multiphoton Ionization
GHSV	Gas Hourly Space Velocity
HTAD	High Temperature Aerosol Decomposition
BLUE	Best Linear Unbiased Estimator
GIS	Geographic Information System
RF	Random Function
OK	Ordinary Kriging
SK	Simple Kriging
MBA	Multilevel B-Spline Algorithm
VoF	Value of Fit
PA	Propionic Acid
PS	Pure Solvent
CA	Complexing Agent
GC	Gas Chromatography

B Used Chemicals and Equipment

B.1 Metal Precursors and Other Chemicals

Element Precursor	Producer
$CoCO_3$	Alfa Aesar
MnO	Aldrich
$Ni(OH)_2$	Alfa Aesar
$Cr(NO_3)_3 \cdot 9\ H_2O$	Aldrich
H_6TeO_6	Aldrich

Table B.1: Used Metal Precursors

Chemical	Producer
propionic acid	Acros Organics
methanol	ZChLUdS [1]
4-hydroxy-4-methyl-2-pentanone	Alfa Aesar

Table B.2: Other Chemicals.

B.2 Used Equipment and Software

Description	Type	Producer
software	Plattenbau	Jens Scheidtmann
software	Matlab	The MathWorks
pipetting robot	Lissy	Zinsser Analytic
robot control software	Zinsser REDI v. 5.3.0	Zinsser Analytic
vibrating plate	Titramax 100	Heidolph Instruments
oven	CWF 1100	Carbolite
oven software controller	S27	Nabertherm
mass flow controller	F-201C & F-200D	Bronkhorst Hi-Tec
temperature controller	dTron 16.1	Jumo
gas chromatograph	CP-4900	Varian
software GC control	CP-Maître Elite	Varian

Table B.3: Used Equipment and Software.

[1] Zentrales Chemikalienlager der Universität des Saarlandes

C Compositions of Tested Catalysts

C.1 Pentanary Composition Spread of Cr, Mn, Co, Te and Ni

Sample No.	Cr	Mn	Co	Te	Ni
1	0.1	0.6	0.2	0.0	0.1
2	0.1	0.3	0.6	0.0	0.0
3	0.1	0.3	0.0	0.0	0.6
4	0.0	0.7	0.2	0.1	0.0
5	0.2	0.1	0.2	0.0	0.5
6	0.1	0.4	0.1	0.1	0.3
7	0.4	0.0	0.6	0.0	0.0
8	0.1	0.0	0.1	0.3	0.5
9	0.2	0.5	0.2	0.0	0.1
10	0.1	0.2	0.0	0.5	0.2
11	0.0	0.2	0.3	0.0	0.5
12	0.2	0.1	0.0	0.0	0.7
13	0.0	0.0	0.3	0.4	0.3
14	0.2	0.2	0.0	0.3	0.3
15	0.1	0.0	0.6	0.1	0.2
16	0.3	0.2	0.0	0.4	0.1
17	0.2	0.0	0.0	0.6	0.2
18	0.3	0.0	0.4	0.1	0.2
19	0.3	0.5	0.1	0.1	0.0
20	0.0	0.0	0.1	0.4	0.5
21	0.0	0.0	0.3	0.3	0.4
22	0.3	0.1	0.4	0.0	0.2
23	0.4	0.1	0.0	0.0	0.5
24	0.3	0.0	0.2	0.1	0.4
25	0.1	0.4	0.1	0.2	0.2
26	0.1	0.7	0.1	0.1	0.0
27	0.0	0.1	0.6	0.0	0.3
28	0.1	0.0	0.5	0.1	0.3
29	0.7	0.1	0.0	0.1	0.1
30	0.6	0.0	0.3	0.0	0.1
31	0.0	0.2	0.1	0.6	0.1
32	0.2	0.6	0.1	0.1	0.0
33	0.0	0.3	0.0	0.7	0.0
34	0.2	0.7	0.0	0.0	0.1
35	0.0	0.7	0.1	0.1	0.1
36	0.8	0.0	0.1	0.1	0.0
37	0.2	0.0	0.0	0.7	0.1
38	0.1	0.5	0.0	0.1	0.3
39	0.0	0.3	0.7	0.0	0.0
40	0.1	0.2	0.1	0.4	0.2
41	0.4	0.5	0.1	0.0	0.0
42	0.1	0.1	0.5	0.0	0.3
43	0.5	0.1	0.0	0.1	0.3
44	0.3	0.0	0.2	0.0	0.5
45	0.4	0.2	0.2	0.1	0.1
46	0.5	0.3	0.1	0.0	0.1
47	0.2	0.4	0.2	0.1	0.1
48	0.2	0.0	0.0	0.4	0.4
49	0.0	0.3	0.0	0.6	0.1
50	0.3	0.0	0.0	0.3	0.4
51	0.6	0.1	0.1	0.2	0.0
52	0.2	0.5	0.0	0.1	0.2
53	0.4	0.2	0.3	0.1	0.0
54	0.1	0.4	0.1	0.4	0.0
55	0.4	0.0	0.4	0.2	0.0
56	0.0	0.5	0.3	0.2	0.0
57	0.0	0.4	0.0	0.1	0.5
58	0.2	0.0	0.3	0.5	0.0
59	0.6	0.2	0.1	0.0	0.1
60	0.0	0.3	0.1	0.4	0.2
61	0.0	0.5	0.1	0.2	0.2
62	0.1	0.5	0.0	0.0	0.4
63	0.1	0.3	0.4	0.2	0.0
64	0.0	0.4	0.1	0.1	0.4
65	0.0	0.0	0.5	0.1	0.4
66	0.1	0.2	0.1	0.0	0.6
67	0.5	0.0	0.4	0.1	0.0
68	0.0	0.1	0.5	0.1	0.3
69	0.1	0.0	0.1	0.2	0.6
70	0.0	0.3	0.5	0.0	0.2
71	0.0	0.1	0.0	0.9	0.0
72	0.5	0.0	0.3	0.2	0.0
73	0.1	0.2	0.6	0.0	0.1
74	0.1	0.1	0.0	0.1	0.7
75	0.1	0.4	0.2	0.1	0.2
76	0.3	0.1	0.5	0.0	0.1
77	0.1	0.5	0.0	0.4	0.0
78	0.0	0.2	0.7	0.0	0.1
79	0.1	0.6	0.1	0.0	0.2
80	0.2	0.1	0.3	0.3	0.1
81	0.3	0.0	0.0	0.6	0.1
82	0.2	0.1	0.0	0.3	0.4
83	0.1	0.1	0.0	0.5	0.3
84	0.0	0.3	0.5	0.2	0.0
85	0.4	0.0	0.1	0.0	0.5
86	0.0	0.3	0.1	0.2	0.4
87	0.2	0.1	0.5	0.2	0.0
88	0.7	0.2	0.1	0.0	0.0
89	0.3	0.3	0.1	0.1	0.2
90	0.1	0.0	0.1	0.4	0.4
91	0.0	0.5	0.1	0.1	0.3
92	0.2	0.0	0.1	0.1	0.6
93	0.5	0.1	0.0	0.0	0.4
94	0.5	0.1	0.1	0.3	0.0
95	0.4	0.2	0.4	0.0	0.0
96	0.3	0.0	0.1	0.0	0.6
97	0.0	0.6	0.3	0.1	0.0
98	0.1	0.1	0.0	0.6	0.2
99	0.0	0.2	0.8	0.0	0.0
100	0.1	0.5	0.3	0.1	0.0
101	0.2	0.6	0.2	0.0	0.0
102	0.3	0.4	0.0	0.2	0.1

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
103	0.1	0.5	0.3	0.0	0.1
104	0.0	0.0	0.3	0.6	0.1
105	0.3	0.2	0.0	0.5	0.0
106	0.0	0.4	0.1	0.2	0.3
107	0.5	0.0	0.4	0.0	0.1
108	0.4	0.0	0.0	0.5	0.1
109	0.0	0.1	0.0	0.4	0.5
110	0.2	0.5	0.0	0.2	0.1
111	0.1	0.2	0.2	0.1	0.4
112	0.4	0.0	0.3	0.3	0.0
113	0.2	0.0	0.0	0.1	0.7
114	0.0	0.0	0.3	0.2	0.5
115	0.1	0.3	0.1	0.5	0.0
116	0.2	0.2	0.0	0.5	0.1
117	0.1	0.0	0.1	0.8	0.0
118	0.6	0.2	0.0	0.2	0.0
119	0.3	0.0	0.3	0.1	0.3
120	0.2	0.5	0.0	0.0	0.3
121	0.1	0.1	0.2	0.3	0.3
122	0.2	0.3	0.1	0.3	0.1
123	0.1	0.1	0.4	0.1	0.3
124	0.1	0.6	0.3	0.0	0.0
125	0.1	0.3	0.3	0.2	0.1
126	0.1	0.1	0.0	0.8	0.0
127	0.9	0.1	0.0	0.0	0.0
128	0.4	0.1	0.4	0.1	0.0
129	0.0	0.1	0.4	0.1	0.4
130	0.1	0.2	0.2	0.0	0.5
131	0.2	0.0	0.2	0.0	0.6
132	0.2	0.0	0.1	0.3	0.4
133	0.0	0.0	0.2	0.1	0.7
134	0.2	0.1	0.2	0.1	0.4
135	0.1	0.3	0.1	0.2	0.3
136	0.1	0.3	0.1	0.4	0.1
137	0.0	0.3	0.2	0.0	0.5
138	0.2	0.0	0.2	0.6	0.0
139	0.0	0.6	0.0	0.0	0.4
140	0.2	0.3	0.0	0.0	0.5
141	0.9	0.0	0.1	0.0	0.0
142	0.2	0.1	0.1	0.0	0.6
143	0.1	0.6	0.2	0.1	0.0
144	0.0	0.0	0.1	0.5	0.4
145	0.2	0.0	0.2	0.5	0.1
146	0.0	0.1	0.0	0.0	0.9
147	0.1	0.2	0.0	0.2	0.5
148	0.3	0.4	0.0	0.3	0.0
149	0.5	0.0	0.1	0.0	0.4
150	0.0	0.0	0.9	0.0	0.1
151	0.2	0.2	0.0	0.2	0.4
152	0.1	0.2	0.2	0.2	0.3
153	0.3	0.2	0.3	0.2	0.0
154	0.3	0.0	0.0	0.4	0.3
155	0.0	0.4	0.2	0.0	0.4
156	0.0	0.2	0.3	0.5	0.0
157	0.0	0.1	0.4	0.4	0.1
158	0.3	0.4	0.3	0.0	0.0
159	0.4	0.1	0.1	0.3	0.1
160	0.0	0.7	0.0	0.2	0.1
161	0.1	0.4	0.2	0.3	0.0
162	0.2	0.0	0.1	0.6	0.1
163	0.1	0.3	0.0	0.1	0.5
164	0.3	0.0	0.1	0.5	0.1
165	0.0	0.0	0.0	0.8	0.2
166	0.2	0.1	0.2	0.5	0.0
167	0.7	0.1	0.2	0.0	0.0
168	0.6	0.0	0.3	0.1	0.0
169	0.0	0.3	0.3	0.0	0.4
170	0.4	0.0	0.2	0.3	0.1
171	0.3	0.1	0.3	0.1	0.2
172	0.0	0.0	0.5	0.3	0.2
173	0.0	0.3	0.3	0.2	0.2
174	0.0	0.3	0.4	0.2	0.1
175	0.1	0.2	0.3	0.2	0.2
176	0.0	0.1	0.1	0.8	0.0
177	0.5	0.0	0.0	0.5	0.0
178	0.1	0.2	0.1	0.6	0.0
179	0.3	0.7	0.0	0.0	0.0
180	0.2	0.0	0.3	0.1	0.4
181	0.0	0.1	0.4	0.3	0.2
182	0.5	0.2	0.1	0.0	0.2
183	0.1	0.2	0.2	0.4	0.1
184	0.0	0.6	0.0	0.4	0.0
185	0.2	0.0	0.5	0.3	0.0
186	0.0	0.0	0.7	0.0	0.3
187	0.2	0.2	0.3	0.2	0.1
188	0.0	0.3	0.0	0.0	0.7
189	0.2	0.0	0.3	0.4	0.1
190	0.8	0.0	0.0	0.0	0.2
191	0.3	0.1	0.0	0.1	0.5
192	0.4	0.3	0.1	0.0	0.2
193	0.1	0.1	0.8	0.0	0.0
194	0.1	0.0	0.4	0.5	0.0
195	0.3	0.0	0.2	0.4	0.1
196	0.6	0.1	0.0	0.2	0.1
197	0.1	0.0	0.0	0.7	0.2
198	0.0	0.0	0.3	0.5	0.2
199	0.5	0.0	0.0	0.1	0.4
200	0.2	0.1	0.2	0.3	0.2
201	0.2	0.2	0.1	0.2	0.3
202	0.0	0.5	0.0	0.2	0.3
203	0.2	0.0	0.0	0.5	0.3
204	0.3	0.3	0.0	0.4	0.0
205	0.0	0.0	0.3	0.7	0.0
206	0.3	0.0	0.6	0.1	0.0
207	0.7	0.0	0.2	0.1	0.0
208	0.3	0.3	0.3	0.0	0.1
209	0.2	0.3	0.1	0.2	0.2
210	0.0	0.0	0.6	0.4	0.0
211	0.0	0.1	0.5	0.3	0.1
212	0.2	0.6	0.0	0.0	0.2
213	0.2	0.1	0.4	0.0	0.3
214	0.3	0.0	0.0	0.1	0.6
215	0.2	0.2	0.4	0.1	0.1
216	0.3	0.1	0.1	0.2	0.3
217	0.1	0.5	0.2	0.2	0.0
218	0.1	0.0	0.0	0.2	0.7
219	0.0	0.2	0.0	0.4	0.4
220	0.2	0.1	0.6	0.1	0.0
221	0.0	0.2	0.5	0.2	0.1
222	0.1	0.1	0.0	0.0	0.8
223	0.1	0.1	0.0	0.4	0.4
224	0.2	0.1	0.1	0.4	0.2
225	0.0	0.8	0.0	0.0	0.2
226	0.1	0.0	0.2	0.3	0.4
227	0.4	0.1	0.4	0.0	0.1
228	0.6	0.0	0.0	0.2	0.2
229	0.1	0.8	0.0	0.0	0.1
230	0.2	0.1	0.1	0.3	0.3
231	0.3	0.0	0.1	0.3	0.3
232	0.0	0.1	0.1	0.7	0.1

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
233	0.0	0.5	0.4	0.1	0.0
234	0.0	0.4	0.1	0.0	0.5
235	0.5	0.2	0.0	0.0	0.3
236	0.1	0.3	0.2	0.4	0.0
237	0.0	0.5	0.0	0.3	0.2
238	0.5	0.3	0.0	0.2	0.0
239	0.2	0.2	0.3	0.3	0.0
240	0.1	0.5	0.1	0.0	0.3
241	0.3	0.1	0.2	0.0	0.4
242	0.1	0.1	0.1	0.1	0.6
243	0.0	0.6	0.1	0.2	0.1
244	0.2	0.2	0.2	0.4	0.0
245	0.1	0.2	0.2	0.3	0.2
246	0.0	0.1	0.3	0.6	0.0
247	0.0	0.2	0.6	0.2	0.0
248	0.6	0.1	0.0	0.0	0.3
249	0.2	0.5	0.1	0.1	0.1
250	1.0	0.0	0.0	0.0	0.0
251	0.4	0.2	0.2	0.0	0.2
252	0.2	0.2	0.3	0.1	0.2
253	0.0	0.4	0.4	0.2	0.0
254	0.3	0.0	0.2	0.3	0.2
255	0.1	0.3	0.5	0.1	0.0
256	0.0	0.3	0.0	0.3	0.4
257	0.0	0.5	0.3	0.0	0.2
258	0.2	0.2	0.0	0.1	0.5
259	0.1	0.7	0.0	0.1	0.1
260	0.5	0.1	0.4	0.0	0.0
261	0.1	0.2	0.5	0.0	0.2
262	0.2	0.3	0.3	0.0	0.2
263	0.2	0.1	0.4	0.1	0.2
264	0.0	0.1	0.2	0.1	0.6
265	0.0	0.1	0.0	0.2	0.7
266	0.3	0.3	0.2	0.0	0.2
267	0.0	0.0	0.0	0.0	1.0
268	0.4	0.0	0.5	0.0	0.1
269	0.5	0.0	0.5	0.0	0.0
270	0.1	0.1	0.4	0.2	0.2
271	0.4	0.6	0.0	0.0	0.0
272	0.0	0.1	0.6	0.3	0.0
273	0.3	0.1	0.2	0.2	0.2
274	0.4	0.3	0.0	0.3	0.0
275	0.1	0.1	0.2	0.6	0.0
276	0.3	0.2	0.2	0.1	0.2
277	0.1	0.0	0.2	0.7	0.0
278	0.1	0.0	0.7	0.1	0.1
279	0.0	0.2	0.1	0.7	0.0
280	0.1	0.1	0.3	0.4	0.1
281	0.3	0.2	0.2	0.3	0.0
282	0.1	0.2	0.3	0.4	0.0
283	0.0	0.1	0.4	0.5	0.0
284	0.1	0.1	0.2	0.5	0.1
285	0.2	0.2	0.3	0.0	0.3
286	0.0	0.2	0.3	0.3	0.2
287	0.2	0.0	0.2	0.1	0.5
288	0.4	0.3	0.2	0.0	0.1
289	0.7	0.0	0.1	0.2	0.0
290	0.1	0.0	0.2	0.6	0.1
291	0.4	0.1	0.2	0.3	0.0
292	0.2	0.0	0.0	0.2	0.6
293	0.7	0.0	0.2	0.0	0.1
294	0.0	0.4	0.1	0.3	0.2
295	0.2	0.3	0.3	0.1	0.1
296	0.0	0.0	0.1	0.9	0.0
297	0.4	0.0	0.0	0.1	0.5
298	0.2	0.4	0.1	0.3	0.0
299	0.1	0.2	0.4	0.3	0.0
300	0.1	0.0	0.3	0.1	0.5
301	0.2	0.3	0.2	0.2	0.1
302	0.0	0.1	0.2	0.0	0.7
303	0.3	0.2	0.2	0.2	0.1
304	0.7	0.1	0.1	0.0	0.1
305	0.0	0.0	0.5	0.4	0.1
306	0.0	0.2	0.2	0.4	0.2
307	0.2	0.3	0.5	0.0	0.0
308	0.1	0.3	0.3	0.1	0.2
309	0.1	0.1	0.2	0.1	0.5
310	0.4	0.2	0.1	0.3	0.0
311	0.1	0.4	0.1	0.0	0.4
312	0.1	0.1	0.3	0.3	0.2
313	0.1	0.0	0.7	0.2	0.0
314	0.1	0.7	0.0	0.2	0.0
315	0.0	0.1	0.2	0.2	0.5
316	0.2	0.2	0.2	0.1	0.3
317	0.0	0.4	0.0	0.4	0.2
318	0.1	0.4	0.0	0.2	0.3
319	0.3	0.2	0.0	0.2	0.3
320	0.1	0.0	0.3	0.3	0.3
321	0.1	0.2	0.3	0.1	0.3
322	0.1	0.2	0.6	0.1	0.0
323	0.1	0.0	0.4	0.0	0.5
324	0.0	0.1	0.1	0.6	0.2
325	0.8	0.2	0.0	0.0	0.0
326	0.3	0.0	0.4	0.2	0.1
327	0.1	0.3	0.1	0.1	0.4
328	0.1	0.1	0.5	0.1	0.2
329	0.4	0.1	0.0	0.4	0.1
330	0.0	0.0	0.7	0.2	0.1
331	0.2	0.0	0.6	0.0	0.2
332	0.0	0.2	0.2	0.2	0.4
333	0.0	0.5	0.0	0.1	0.4
334	0.0	0.1	0.3	0.4	0.2
335	0.0	0.5	0.0	0.0	0.5
336	0.2	0.7	0.1	0.0	0.0
337	0.1	0.2	0.1	0.5	0.1
338	0.6	0.1	0.0	0.1	0.2
339	0.3	0.5	0.2	0.0	0.0
340	0.0	0.4	0.4	0.1	0.1
341	0.7	0.1	0.1	0.1	0.0
342	0.3	0.1	0.3	0.0	0.3
343	0.4	0.2	0.0	0.2	0.2
344	0.0	0.1	0.3	0.0	0.6
345	0.0	0.4	0.3	0.3	0.0
346	0.0	0.2	0.0	0.6	0.2
347	0.4	0.0	0.4	0.1	0.1
348	0.1	0.0	0.0	0.9	0.0
349	0.7	0.2	0.0	0.1	0.0
350	0.2	0.4	0.0	0.0	0.4
351	0.1	0.2	0.1	0.2	0.4
352	0.3	0.5	0.0	0.2	0.0
353	0.0	0.2	0.3	0.4	0.1
354	0.4	0.1	0.2	0.1	0.2
355	0.1	0.0	0.6	0.3	0.0
356	0.3	0.0	0.3	0.2	0.2
357	0.1	0.4	0.0	0.4	0.1
358	0.2	0.1	0.3	0.1	0.3
359	0.5	0.0	0.0	0.0	0.5
360	0.3	0.0	0.4	0.3	0.0
361	0.0	0.0	0.1	0.8	0.1
362	0.2	0.0	0.4	0.2	0.2

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
363	0.0	0.2	0.0	0.2	0.6
364	0.3	0.0	0.0	0.2	0.5
365	0.3	0.2	0.0	0.3	0.2
366	0.0	0.2	0.3	0.1	0.4
367	0.1	0.1	0.0	0.7	0.1
368	0.1	0.2	0.0	0.3	0.4
369	0.2	0.4	0.3	0.1	0.0
370	0.4	0.0	0.1	0.5	0.0
371	0.0	0.5	0.0	0.5	0.0
372	0.0	0.0	0.0	0.4	0.6
373	0.4	0.5	0.0	0.1	0.0
374	0.0	0.8	0.1	0.1	0.0
375	0.2	0.3	0.0	0.3	0.2
376	0.0	0.2	0.1	0.5	0.2
377	0.0	0.5	0.3	0.1	0.1
378	0.0	0.4	0.2	0.4	0.0
379	0.1	0.2	0.4	0.1	0.2
380	0.5	0.1	0.1	0.2	0.1
381	0.2	0.2	0.1	0.4	0.1
382	0.0	0.6	0.1	0.0	0.3
383	0.5	0.3	0.2	0.0	0.0
384	0.2	0.0	0.2	0.3	0.3
385	0.0	0.4	0.4	0.0	0.2
386	0.3	0.1	0.4	0.1	0.1
387	0.4	0.1	0.3	0.1	0.1
388	0.0	0.0	0.3	0.1	0.6
389	0.0	0.0	0.4	0.3	0.3
390	0.3	0.1	0.2	0.3	0.1
391	0.2	0.5	0.1	0.0	0.2
392	0.1	0.2	0.4	0.0	0.3
393	0.0	0.3	0.2	0.1	0.4
394	0.0	0.4	0.5	0.1	0.0
395	0.4	0.1	0.3	0.2	0.0
396	0.2	0.4	0.1	0.2	0.1
397	0.0	0.1	0.2	0.3	0.4
398	0.2	0.0	0.4	0.3	0.1
399	0.0	0.0	0.4	0.4	0.2
400	0.2	0.3	0.0	0.4	0.1
401	0.0	0.1	0.1	0.4	0.4
402	0.0	0.0	0.6	0.1	0.3
403	0.1	0.0	0.1	0.5	0.3
404	0.1	0.2	0.0	0.0	0.7
405	0.0	0.0	0.9	0.1	0.0
406	0.6	0.0	0.0	0.0	0.4
407	0.3	0.1	0.1	0.3	0.2
408	0.1	0.5	0.0	0.3	0.1
409	0.5	0.1	0.2	0.1	0.1
410	0.1	0.1	0.2	0.0	0.6
411	0.2	0.0	0.2	0.4	0.2
412	0.3	0.2	0.1	0.4	0.0
413	0.4	0.1	0.2	0.2	0.1
414	0.1	0.2	0.1	0.1	0.5
415	0.0	0.0	0.1	0.2	0.7
416	0.0	0.7	0.0	0.3	0.0
417	0.0	0.4	0.1	0.4	0.1
418	0.3	0.1	0.0	0.3	0.3
419	0.0	0.6	0.0	0.2	0.2
420	0.5	0.3	0.0	0.1	0.1
421	0.2	0.1	0.0	0.5	0.2
422	0.0	0.1	0.7	0.1	0.1
423	0.1	0.2	0.0	0.6	0.1
424	0.0	0.1	0.4	0.0	0.5
425	0.0	0.4	0.3	0.2	0.1
426	0.0	0.2	0.0	0.1	0.7
427	0.2	0.2	0.1	0.5	0.0
428	0.0	0.3	0.1	0.1	0.5
429	0.0	0.4	0.0	0.0	0.6
430	0.4	0.1	0.1	0.1	0.3
431	0.3	0.0	0.7	0.0	0.0
432	0.1	0.0	0.6	0.2	0.1
433	0.5	0.2	0.3	0.0	0.0
434	0.2	0.1	0.5	0.0	0.2
435	0.4	0.1	0.1	0.0	0.4
436	0.1	0.1	0.1	0.6	0.1
437	0.0	0.2	0.4	0.1	0.3
438	0.0	0.1	0.2	0.5	0.2
439	0.0	0.2	0.3	0.2	0.3
440	0.1	0.1	0.3	0.2	0.3
441	0.3	0.2	0.1	0.3	0.1
442	0.0	0.1	0.1	0.2	0.6
443	0.4	0.2	0.0	0.1	0.3
444	0.0	0.5	0.1	0.3	0.1
445	0.3	0.2	0.4	0.1	0.0
446	0.0	0.0	0.2	0.0	0.8
447	0.1	0.3	0.1	0.0	0.5
448	0.4	0.3	0.0	0.2	0.1
449	0.0	0.3	0.2	0.4	0.1
450	0.3	0.0	0.0	0.5	0.2
451	0.5	0.3	0.0	0.0	0.2
452	0.1	0.0	0.2	0.2	0.5
453	0.0	0.0	0.2	0.5	0.3
454	0.0	0.1	0.7	0.2	0.0
455	0.3	0.0	0.5	0.1	0.1
456	0.0	0.4	0.0	0.3	0.3
457	0.3	0.0	0.0	0.0	0.7
458	0.0	0.1	0.4	0.2	0.3
459	0.1	0.5	0.1	0.3	0.0
460	0.4	0.0	0.3	0.2	0.1
461	0.4	0.4	0.0	0.2	0.0
462	0.1	0.1	0.7	0.0	0.1
463	0.1	0.0	0.3	0.2	0.4
464	0.8	0.0	0.0	0.2	0.0
465	0.0	0.1	0.2	0.7	0.0
466	0.4	0.4	0.1	0.1	0.0
467	0.2	0.1	0.1	0.6	0.0
468	0.0	0.2	0.0	0.8	0.0
469	0.0	0.2	0.6	0.0	0.2
470	0.1	0.1	0.1	0.0	0.7
471	0.2	0.3	0.0	0.2	0.3
472	0.1	0.1	0.4	0.4	0.0
473	0.0	0.2	0.7	0.1	0.0
474	0.1	0.2	0.0	0.7	0.0
475	0.2	0.4	0.0	0.3	0.1
476	0.2	0.4	0.1	0.1	0.2
477	0.1	0.6	0.1	0.2	0.0
478	0.2	0.1	0.1	0.1	0.5
479	0.0	0.0	0.7	0.1	0.2
480	0.2	0.0	0.6	0.2	0.0
481	0.1	0.0	0.0	0.5	0.4
482	0.0	0.1	0.5	0.2	0.2
483	0.2	0.3	0.4	0.0	0.1
484	0.4	0.1	0.2	0.0	0.3
485	0.1	0.0	0.3	0.6	0.0
486	0.3	0.2	0.3	0.0	0.2
487	0.0	0.3	0.2	0.2	0.3
488	0.2	0.0	0.1	0.7	0.0
489	0.1	0.3	0.0	0.4	0.2
490	0.0	0.2	0.2	0.5	0.1
491	0.2	0.0	0.2	0.2	0.4
492	0.8	0.0	0.0	0.1	0.1

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
493	0.0	0.1	0.1	0.3	0.5
494	0.3	0.3	0.0	0.3	0.1
495	0.2	0.3	0.0	0.5	0.0
496	0.1	0.0	0.0	0.1	0.8
497	0.3	0.1	0.3	0.3	0.0
498	0.5	0.4	0.0	0.0	0.1
499	0.7	0.3	0.0	0.0	0.0
500	0.4	0.3	0.0	0.1	0.2
501	0.1	0.0	0.2	0.4	0.3
502	0.1	0.3	0.2	0.3	0.1
503	0.3	0.2	0.3	0.1	0.1
504	0.6	0.1	0.1	0.1	0.1
505	0.2	0.0	0.5	0.2	0.1
506	0.1	0.1	0.6	0.1	0.1
507	0.0	0.4	0.2	0.1	0.3
508	0.1	0.0	0.3	0.4	0.2
509	0.0	0.4	0.3	0.1	0.2
510	0.1	0.4	0.4	0.0	0.1
511	0.4	0.2	0.1	0.0	0.3
512	0.0	0.0	0.1	0.6	0.3
513	0.4	0.5	0.0	0.0	0.1
514	0.0	0.2	0.0	0.3	0.5
515	0.1	0.3	0.2	0.1	0.3
516	0.1	0.4	0.3	0.1	0.1
517	0.0	0.0	0.5	0.5	0.0
518	0.5	0.0	0.0	0.2	0.3
519	0.5	0.1	0.0	0.3	0.1
520	0.0	0.0	0.0	0.2	0.8
521	0.8	0.0	0.1	0.0	0.1
522	0.1	0.2	0.0	0.1	0.6
523	0.4	0.1	0.5	0.0	0.0
524	0.2	0.4	0.0	0.4	0.0
525	0.3	0.0	0.1	0.2	0.4
526	0.4	0.1	0.0	0.1	0.4
527	0.1	0.0	0.8	0.1	0.0
528	0.1	0.0	0.1	0.7	0.1
529	0.0	0.0	0.2	0.2	0.6
530	0.3	0.1	0.1	0.0	0.5
531	0.1	0.1	0.4	0.3	0.1
532	0.4	0.0	0.1	0.3	0.2
533	0.3	0.4	0.2	0.0	0.1
534	0.2	0.1	0.3	0.4	0.0
535	0.2	0.1	0.6	0.0	0.1
536	0.1	0.4	0.0	0.1	0.4
537	0.0	0.5	0.2	0.3	0.0
538	0.0	0.3	0.0	0.1	0.6
539	0.0	0.1	0.5	0.0	0.4
540	0.0	0.5	0.2	0.1	0.2
541	0.5	0.0	0.1	0.1	0.3
542	0.3	0.5	0.0	0.0	0.2
543	0.2	0.2	0.2	0.0	0.4
544	0.0	0.5	0.1	0.4	0.0
545	0.3	0.1	0.0	0.6	0.0
546	0.0	0.0	1.0	0.0	0.0
547	0.0	0.7	0.1	0.2	0.0
548	0.2	0.2	0.1	0.1	0.4
549	0.0	0.5	0.2	0.2	0.1
550	0.2	0.0	0.7	0.0	0.1
551	0.4	0.0	0.4	0.0	0.2
552	0.1	0.0	0.1	0.1	0.7
553	0.0	0.2	0.4	0.3	0.1
554	0.4	0.2	0.0	0.4	0.0
555	0.0	0.4	0.1	0.5	0.0
556	0.2	0.1	0.0	0.2	0.5
557	0.2	0.2	0.5	0.1	0.0
558	0.3	0.1	0.3	0.2	0.1
559	0.5	0.1	0.1	0.0	0.3
560	0.7	0.0	0.1	0.0	0.2
561	0.3	0.3	0.0	0.0	0.4
562	0.1	0.3	0.4	0.1	0.1
563	0.3	0.0	0.2	0.5	0.0
564	0.1	0.0	0.1	0.6	0.2
565	0.3	0.0	0.6	0.0	0.1
566	0.0	0.1	0.7	0.0	0.2
567	0.0	0.3	0.1	0.6	0.0
568	0.5	0.1	0.3	0.0	0.1
569	0.3	0.0	0.3	0.3	0.1
570	0.0	0.4	0.6	0.0	0.0
571	0.2	0.7	0.0	0.1	0.0
572	0.2	0.2	0.5	0.0	0.1
573	0.0	0.5	0.1	0.0	0.4
574	0.1	0.1	0.1	0.3	0.4
575	0.2	0.1	0.0	0.4	0.3
576	0.6	0.0	0.1	0.1	0.2
577	0.0	0.0	0.4	0.2	0.4
578	0.1	0.0	0.2	0.5	0.2
579	0.1	0.7	0.2	0.0	0.0
580	0.1	0.2	0.5	0.2	0.0
581	0.6	0.0	0.0	0.4	0.0
582	0.4	0.0	0.2	0.4	0.0
583	0.3	0.3	0.0	0.2	0.2
584	0.2	0.0	0.0	0.0	0.8
585	0.2	0.0	0.6	0.1	0.1
586	0.3	0.3	0.1	0.0	0.3
587	0.1	0.3	0.3	0.0	0.3
588	0.4	0.1	0.0	0.3	0.2
589	0.2	0.3	0.1	0.0	0.4
590	0.3	0.1	0.4	0.2	0.0
591	0.4	0.0	0.2	0.2	0.2
592	0.3	0.4	0.1	0.1	0.1
593	0.0	0.6	0.2	0.1	0.1
594	0.2	0.0	0.1	0.0	0.7
595	0.5	0.1	0.1	0.1	0.2
596	0.3	0.2	0.5	0.0	0.0
597	0.1	0.3	0.0	0.3	0.3
598	0.4	0.2	0.2	0.2	0.0
599	0.1	0.6	0.0	0.0	0.3
600	0.0	0.2	0.4	0.4	0.0
601	0.6	0.0	0.2	0.0	0.2
602	0.5	0.2	0.2	0.0	0.1
603	0.0	0.3	0.5	0.1	0.1
604	0.1	0.0	0.4	0.2	0.3
605	0.2	0.2	0.0	0.4	0.2
606	0.2	0.0	0.4	0.0	0.4
607	0.3	0.4	0.2	0.1	0.0
608	0.1	0.5	0.2	0.1	0.1
609	0.3	0.6	0.0	0.0	0.1
610	0.6	0.1	0.3	0.0	0.0
611	0.8	0.1	0.1	0.0	0.0
612	0.1	0.1	0.3	0.1	0.4
613	0.2	0.1	0.1	0.2	0.4
614	0.1	0.0	0.5	0.3	0.1
615	0.4	0.1	0.1	0.4	0.0
616	0.3	0.1	0.0	0.2	0.4
617	0.0	0.8	0.1	0.0	0.1
618	0.0	0.7	0.2	0.0	0.1
619	0.1	0.0	0.4	0.1	0.4
620	0.0	0.2	0.2	0.0	0.6
621	0.1	0.3	0.0	0.6	0.0
622	0.0	0.4	0.0	0.6	0.0

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
623	0.0	0.3	0.6	0.0	0.1
624	0.2	0.0	0.7	0.1	0.0
625	0.1	0.0	0.3	0.0	0.6
626	0.0	0.3	0.3	0.3	0.1
627	0.1	0.3	0.2	0.2	0.2
628	0.2	0.2	0.6	0.0	0.0
629	0.3	0.1	0.2	0.1	0.3
630	0.2	0.0	0.3	0.0	0.5
631	0.0	0.2	0.5	0.1	0.2
632	0.3	0.1	0.1	0.4	0.1
633	0.2	0.4	0.4	0.0	0.0
634	0.0	0.2	0.0	0.7	0.1
635	0.1	0.1	0.5	0.2	0.1
636	0.0	0.6	0.4	0.0	0.0
637	0.2	0.2	0.1	0.0	0.5
638	0.0	0.3	0.0	0.4	0.3
639	0.1	0.3	0.5	0.0	0.1
640	0.3	0.3	0.1	0.3	0.0
641	0.0	0.0	0.8	0.1	0.1
642	0.6	0.0	0.0	0.1	0.3
643	0.0	0.2	0.2	0.1	0.5
644	0.0	0.2	0.1	0.3	0.4
645	0.0	0.5	0.5	0.0	0.0
646	0.0	0.6	0.2	0.2	0.0
647	0.1	0.4	0.5	0.0	0.0
648	0.2	0.3	0.0	0.1	0.4
649	0.0	0.0	0.2	0.8	0.0
650	0.0	0.1	0.0	0.1	0.8
651	0.0	0.0	0.2	0.3	0.5
652	0.4	0.1	0.0	0.2	0.3
653	0.0	0.0	0.1	0.0	0.9
654	0.0	0.3	0.4	0.0	0.3
655	0.5	0.5	0.0	0.0	0.0
656	0.3	0.3	0.2	0.1	0.1
657	0.5	0.0	0.2	0.2	0.1
658	0.3	0.5	0.0	0.1	0.1
659	0.0	0.8	0.0	0.2	0.0
660	0.1	0.0	0.2	0.1	0.6
661	0.4	0.4	0.0	0.0	0.2
662	0.0	0.4	0.2	0.3	0.1
663	0.1	0.0	0.0	0.4	0.5
664	0.1	0.1	0.1	0.5	0.2
665	0.0	0.3	0.1	0.0	0.6
666	0.4	0.0	0.0	0.2	0.4
667	0.0	0.9	0.0	0.1	0.0
668	0.4	0.0	0.1	0.2	0.3
669	0.2	0.1	0.2	0.2	0.3
670	0.0	0.9	0.0	0.0	0.1
671	0.1	0.0	0.5	0.0	0.4
672	0.2	0.1	0.2	0.4	0.1
673	0.0	0.1	0.1	0.0	0.8
674	0.1	0.4	0.0	0.5	0.0
675	0.2	0.8	0.0	0.0	0.0
676	0.0	0.0	0.4	0.6	0.0
677	0.6	0.1	0.2	0.1	0.0
678	0.5	0.0	0.1	0.3	0.1
679	0.0	0.3	0.2	0.3	0.2
680	0.1	0.1	0.7	0.1	0.0
681	0.1	0.1	0.1	0.4	0.3
682	0.2	0.0	0.3	0.2	0.3
683	0.2	0.4	0.2	0.0	0.2
684	0.5	0.2	0.0	0.3	0.0
685	0.1	0.0	0.4	0.3	0.2
686	0.1	0.4	0.0	0.3	0.2
687	0.0	0.4	0.3	0.0	0.3
688	0.4	0.0	0.0	0.0	0.6
689	0.4	0.2	0.3	0.0	0.1
690	0.1	0.2	0.3	0.0	0.4
691	0.4	0.0	0.0	0.3	0.3
692	0.5	0.2	0.1	0.1	0.1
693	0.2	0.2	0.4	0.2	0.0
694	0.1	0.5	0.1	0.1	0.2
695	0.1	0.1	0.6	0.2	0.0
696	0.4	0.0	0.2	0.1	0.3
697	0.0	0.2	0.4	0.0	0.4
698	0.5	0.3	0.1	0.1	0.0
699	0.9	0.0	0.0	0.1	0.0
700	0.2	0.2	0.0	0.0	0.6
701	0.0	0.0	0.0	0.3	0.7
702	0.3	0.1	0.1	0.1	0.4
703	0.0	0.0	0.2	0.6	0.2
704	0.2	0.4	0.3	0.0	0.1
705	0.2	0.1	0.7	0.0	0.0
706	0.0	0.0	0.4	0.0	0.6
707	0.4	0.3	0.2	0.1	0.0
708	0.4	0.0	0.3	0.0	0.3
709	0.7	0.0	0.0	0.2	0.1
710	0.9	0.0	0.0	0.0	0.1
711	0.1	0.2	0.2	0.5	0.0
712	0.2	0.5	0.3	0.0	0.0
713	0.0	0.4	0.0	0.5	0.1
714	0.1	0.1	0.4	0.0	0.4
715	0.0	0.3	0.4	0.1	0.2
716	0.0	0.9	0.1	0.0	0.0
717	0.1	0.7	0.1	0.0	0.1
718	0.6	0.3	0.0	0.0	0.1
719	0.2	0.1	0.4	0.2	0.1
720	0.0	0.1	0.0	0.5	0.4
721	0.0	0.0	0.1	0.1	0.8
722	0.3	0.4	0.0	0.1	0.2
723	0.0	0.2	0.0	0.0	0.8
724	0.0	0.0	0.2	0.4	0.4
725	0.6	0.2	0.0	0.1	0.1
726	0.1	0.2	0.1	0.3	0.3
727	0.2	0.0	0.4	0.4	0.0
728	0.3	0.0	0.1	0.1	0.5
729	0.2	0.2	0.1	0.3	0.2
730	0.6	0.0	0.1	0.2	0.1
731	0.4	0.3	0.3	0.0	0.0
732	0.0	1.0	0.0	0.0	0.0
733	0.1	0.0	0.2	0.0	0.7
734	0.4	0.0	0.1	0.4	0.1
735	0.7	0.1	0.0	0.2	0.0
736	0.0	0.0	0.3	0.0	0.7
737	0.2	0.3	0.3	0.2	0.0
738	0.2	0.2	0.0	0.6	0.0
739	0.1	0.1	0.3	0.0	0.5
740	0.6	0.3	0.0	0.1	0.0
741	0.0	0.1	0.2	0.6	0.1
742	0.3	0.2	0.1	0.2	0.2
743	0.1	0.0	0.9	0.0	0.0
744	0.1	0.3	0.3	0.3	0.0
745	0.1	0.3	0.0	0.2	0.4
746	0.0	0.1	0.3	0.1	0.5
747	0.0	0.3	0.3	0.1	0.3
748	0.2	0.1	0.4	0.3	0.0
749	0.7	0.0	0.3	0.0	0.0
750	0.8	0.1	0.0	0.0	0.1
751	0.1	0.9	0.0	0.0	0.0
752	0.1	0.3	0.2	0.0	0.4

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
753	0.0	0.2	0.5	0.0	0.3
754	0.1	0.6	0.0	0.2	0.1
755	0.0	0.6	0.1	0.3	0.0
756	0.3	0.3	0.2	0.2	0.0
757	0.1	0.5	0.4	0.0	0.0
758	0.8	0.0	0.2	0.0	0.0
759	0.1	0.2	0.0	0.4	0.3
760	0.0	0.2	0.5	0.3	0.0
761	0.0	0.1	0.5	0.4	0.0
762	0.3	0.5	0.1	0.0	0.1
763	0.1	0.4	0.1	0.3	0.1
764	0.0	0.0	0.1	0.7	0.2
765	0.0	0.3	0.1	0.3	0.3
766	0.1	0.0	0.5	0.4	0.0
767	0.0	0.0	0.4	0.5	0.1
768	0.1	0.4	0.0	0.0	0.5
769	0.3	0.2	0.0	0.1	0.4
770	0.7	0.0	0.0	0.1	0.2
771	0.3	0.3	0.1	0.2	0.1
772	0.3	0.0	0.4	0.0	0.3
773	0.5	0.2	0.2	0.1	0.0
774	0.1	0.1	0.0	0.3	0.5
775	0.0	0.2	0.1	0.4	0.3
776	0.3	0.2	0.2	0.0	0.3
777	0.0	0.6	0.0	0.3	0.1
778	0.0	0.2	0.1	0.1	0.6
779	0.2	0.3	0.2	0.0	0.3
780	0.3	0.2	0.4	0.0	0.1
781	0.0	0.4	0.2	0.2	0.2
782	0.6	0.2	0.1	0.1	0.0
783	0.0	0.6	0.0	0.1	0.3
784	0.4	0.4	0.2	0.0	0.0
785	0.4	0.4	0.0	0.1	0.1
786	0.2	0.0	0.0	0.8	0.0
787	0.1	0.1	0.6	0.0	0.2
788	0.1	0.0	0.3	0.5	0.1
789	0.7	0.0	0.0	0.0	0.3
790	0.0	0.2	0.4	0.2	0.2
791	0.1	0.1	0.2	0.4	0.2
792	0.3	0.2	0.1	0.1	0.3
793	0.0	0.6	0.1	0.1	0.2
794	0.6	0.1	0.0	0.3	0.0
795	0.4	0.0	0.0	0.4	0.2
796	0.1	0.4	0.3	0.0	0.2
797	0.0	0.8	0.0	0.1	0.1
798	0.6	0.0	0.2	0.1	0.1
799	0.2	0.1	0.5	0.1	0.1
800	0.3	0.0	0.1	0.6	0.0
801	0.2	0.1	0.1	0.5	0.1
802	0.0	0.1	0.8	0.0	0.1
803	0.4	0.4	0.1	0.0	0.1
804	0.0	0.7	0.0	0.0	0.3
805	0.5	0.4	0.1	0.0	0.0
806	0.2	0.0	0.8	0.0	0.0
807	0.0	0.5	0.2	0.0	0.3
808	0.2	0.0	0.3	0.3	0.2
809	0.4	0.2	0.1	0.1	0.2
810	0.5	0.1	0.2	0.0	0.2
811	0.0	0.6	0.3	0.0	0.1
812	0.0	0.2	0.2	0.3	0.3
813	0.5	0.0	0.0	0.4	0.1
814	0.0	0.1	0.0	0.8	0.1
815	0.0	0.0	0.2	0.7	0.1
816	0.0	0.0	0.5	0.0	0.5
817	0.5	0.0	0.2	0.1	0.2
818	0.6	0.2	0.2	0.0	0.0
819	0.1	0.5	0.0	0.2	0.2
820	0.0	0.2	0.6	0.1	0.1
821	0.2	0.4	0.2	0.2	0.0
822	0.3	0.1	0.0	0.0	0.6
823	0.1	0.0	0.6	0.0	0.3
824	0.3	0.0	0.3	0.4	0.0
825	0.0	0.0	0.0	1.0	0.0
826	0.4	0.2	0.0	0.0	0.4
827	0.2	0.1	0.0	0.7	0.0
828	0.1	0.5	0.1	0.2	0.1
829	0.0	0.8	0.2	0.0	0.0
830	0.4	0.2	0.0	0.3	0.1
831	0.6	0.0	0.2	0.2	0.0
832	0.0	0.4	0.5	0.0	0.1
833	0.3	0.1	0.6	0.0	0.0
834	0.3	0.0	0.1	0.4	0.2
835	0.6	0.3	0.1	0.0	0.0
836	0.2	0.5	0.2	0.1	0.0
837	0.2	0.0	0.5	0.0	0.3
838	0.6	0.1	0.2	0.0	0.1
839	0.5	0.0	0.3	0.0	0.2
840	0.0	0.0	0.5	0.2	0.3
841	0.2	0.5	0.1	0.2	0.0
842	0.4	0.0	0.5	0.1	0.0
843	0.5	0.4	0.0	0.1	0.0
844	0.2	0.0	0.5	0.1	0.2
845	0.0	0.0	0.0	0.1	0.9
846	0.0	0.1	0.2	0.4	0.3
847	0.8	0.1	0.0	0.1	0.0
848	0.0	0.1	0.1	0.5	0.3
849	0.3	0.3	0.0	0.1	0.3
850	0.4	0.3	0.1	0.2	0.0
851	0.1	0.1	0.3	0.5	0.0
852	0.4	0.1	0.1	0.2	0.2
853	0.0	0.1	0.1	0.1	0.7
854	0.4	0.1	0.3	0.0	0.2
855	0.5	0.2	0.1	0.2	0.0
856	0.1	0.3	0.1	0.3	0.2
857	0.0	0.2	0.1	0.2	0.5
858	0.3	0.3	0.4	0.0	0.0
859	0.1	0.1	0.5	0.3	0.0
860	0.0	0.1	0.8	0.1	0.0
861	0.3	0.1	0.2	0.4	0.0
862	0.5	0.1	0.0	0.4	0.0
863	0.2	0.5	0.0	0.3	0.0
864	0.2	0.2	0.2	0.3	0.1
865	0.0	0.0	0.6	0.2	0.2
866	0.0	0.1	0.0	0.7	0.2
867	0.0	0.1	0.3	0.2	0.4
868	0.7	0.1	0.0	0.0	0.2
869	0.2	0.0	0.0	0.3	0.5
870	0.1	0.4	0.3	0.2	0.0
871	0.0	0.1	0.0	0.3	0.6
872	0.5	0.0	0.2	0.3	0.0
873	0.0	0.1	0.3	0.3	0.3
874	0.2	0.4	0.0	0.1	0.3
875	0.6	0.2	0.0	0.0	0.2
876	0.4	0.0	0.3	0.1	0.2
877	0.1	0.1	0.1	0.2	0.5
878	0.0	0.0	0.8	0.2	0.0
879	0.6	0.4	0.0	0.0	0.0
880	0.0	0.4	0.0	0.2	0.4
881	0.0	0.7	0.0	0.1	0.2
882	0.1	0.0	0.1	0.0	0.8

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
883	0.3	0.6	0.0	0.1	0.0
884	0.2	0.3	0.1	0.1	0.3
885	0.5	0.0	0.3	0.1	0.1
886	0.1	0.2	0.4	0.2	0.1
887	0.7	0.0	0.0	0.3	0.0
888	0.0	0.0	0.0	0.9	0.1
889	0.0	0.1	0.9	0.0	0.0
890	0.6	0.0	0.0	0.3	0.1
891	0.5	0.0	0.1	0.2	0.2
892	0.1	0.0	0.5	0.2	0.2
893	0.0	0.2	0.0	0.5	0.3
894	0.5	0.0	0.2	0.0	0.3
895	0.4	0.0	0.2	0.0	0.4
896	0.3	0.0	0.0	0.7	0.0
897	0.2	0.4	0.1	0.0	0.3
898	0.0	0.0	0.0	0.5	0.5
899	0.1	0.0	0.8	0.0	0.1
900	0.3	0.2	0.0	0.0	0.5
901	0.1	0.6	0.0	0.3	0.0
902	0.0	0.3	0.4	0.3	0.0
903	0.0	0.7	0.3	0.0	0.0
904	0.7	0.2	0.0	0.0	0.1
905	0.0	0.6	0.2	0.0	0.2
906	0.3	0.6	0.1	0.0	0.0
907	0.2	0.2	0.4	0.0	0.2
908	0.5	0.1	0.0	0.2	0.2
909	0.3	0.0	0.5	0.0	0.2
910	0.1	0.0	0.0	0.3	0.6
911	0.0	0.0	0.0	0.7	0.3
912	0.0	0.3	0.3	0.4	0.0
913	0.5	0.1	0.3	0.1	0.0
914	0.4	0.0	0.1	0.1	0.4
915	0.3	0.2	0.1	0.0	0.4
916	0.1	0.2	0.5	0.1	0.1
917	0.0	0.0	0.6	0.3	0.1
918	0.2	0.1	0.0	0.1	0.6
919	0.1	0.1	0.0	0.2	0.6
920	0.0	0.0	0.6	0.0	0.4
921	0.4	0.1	0.0	0.5	0.0
922	0.1	0.0	0.7	0.0	0.2
923	0.0	0.5	0.4	0.0	0.1
924	0.1	0.0	0.0	0.8	0.1
925	0.0	0.1	0.0	0.6	0.3
926	0.0	0.0	0.7	0.3	0.0
927	0.1	0.6	0.1	0.1	0.1
928	0.2	0.0	0.1	0.4	0.3
929	0.0	0.1	0.6	0.1	0.2
930	0.0	0.0	0.0	0.6	0.4
931	0.6	0.0	0.4	0.0	0.0
932	0.1	0.0	0.0	0.6	0.3
933	0.4	0.3	0.1	0.1	0.1
934	0.6	0.0	0.1	0.0	0.3
935	0.2	0.1	0.0	0.6	0.1
936	0.0	0.1	0.3	0.5	0.1
937	0.1	0.3	0.4	0.0	0.2
938	0.2	0.3	0.2	0.3	0.0
939	0.2	0.1	0.3	0.2	0.2
940	0.2	0.6	0.1	0.0	0.1
941	0.0	0.5	0.0	0.4	0.1
942	0.1	0.6	0.0	0.1	0.2
943	0.5	0.0	0.0	0.3	0.2
944	0.0	0.7	0.1	0.0	0.2
945	0.3	0.1	0.1	0.5	0.0
946	0.2	0.0	0.1	0.2	0.5
947	0.1	0.2	0.3	0.3	0.1
948	0.4	0.2	0.1	0.2	0.1
949	0.5	0.2	0.0	0.1	0.2
950	0.0	0.3	0.0	0.5	0.2
951	0.2	0.0	0.4	0.1	0.3
952	0.3	0.1	0.0	0.5	0.1
953	0.2	0.3	0.4	0.1	0.0
954	0.0	0.3	0.6	0.1	0.0
955	0.1	0.1	0.2	0.2	0.4
956	0.3	0.1	0.0	0.4	0.2
957	0.1	0.2	0.7	0.0	0.0
958	0.1	0.4	0.2	0.2	0.1
959	0.0	0.0	0.8	0.0	0.2
960	0.6	0.0	0.1	0.3	0.0
961	0.3	0.0	0.3	0.0	0.4
962	0.4	0.3	0.0	0.0	0.3
963	0.0	0.1	0.6	0.2	0.1
964	0.0	0.2	0.2	0.6	0.0
965	0.1	0.8	0.1	0.0	0.0
966	0.3	0.4	0.1	0.2	0.0
967	0.6	0.1	0.1	0.0	0.2
968	0.2	0.6	0.0	0.1	0.1
969	0.4	0.0	0.0	0.6	0.0
970	0.1	0.4	0.4	0.1	0.0
971	0.5	0.0	0.1	0.4	0.0
972	0.5	0.1	0.2	0.2	0.0
973	0.0	0.3	0.1	0.5	0.1
974	0.1	0.7	0.0	0.0	0.2
975	0.2	0.3	0.2	0.1	0.2
976	0.5	0.2	0.0	0.2	0.1
977	0.7	0.0	0.1	0.1	0.1
978	0.3	0.4	0.1	0.0	0.2
979	0.0	0.3	0.0	0.2	0.5
980	0.3	0.1	0.5	0.1	0.0
981	0.1	0.4	0.2	0.0	0.3
982	0.1	0.8	0.0	0.1	0.0
983	0.1	0.3	0.0	0.5	0.1
984	0.1	0.1	0.1	0.7	0.0
985	0.2	0.3	0.1	0.4	0.0
986	0.1	0.0	0.4	0.4	0.1
987	0.0	0.0	0.4	0.1	0.5
988	0.2	0.2	0.2	0.2	0.2
989	0.0	0.2	0.1	0.0	0.7
990	0.1	0.5	0.2	0.0	0.2
991	0.3	0.4	0.0	0.0	0.3
992	0.2	0.0	0.1	0.5	0.2
993	0.3	0.0	0.2	0.2	0.3
994	0.3	0.3	0.3	0.1	0.0
995	0.3	0.0	0.5	0.2	0.0
996	0.1	0.0	0.0	0.0	0.9
997	0.0	0.0	0.1	0.3	0.6
998	0.2	0.1	0.3	0.0	0.4
999	0.2	0.4	0.0	0.2	0.2
1000	0.2	0.6	0.0	0.2	0.0
1001	0.0	0.3	0.2	0.5	0.0

Table C.1: First Generation of Catalysts and their Compositions.

C.2 Second Generation of Catalyst Samples

Sample No.	Cr	Mn	Co	Te	Ni
1002	0	0.1	0.8	0.1	0
1003	0	0.05	0.85	0.1	0
1004	0	0.1	0.85	0.05	0
1005	0	0.05	0.8	0.15	0
1006	0	0.15	0.8	0.05	0
1007	0	0.2	0.75	0.05	0
1008	0	0.15	0.75	0.1	0
1009	0	0.1	0.75	0.15	0
1010	0	0.05	0.75	0.2	0
1011	0	0.05	0.7	0.25	0
1012	0	0.1	0.7	0.2	0
1013	0	0.15	0.7	0.15	0
1014	0	0.2	0.7	0.1	0
1015	0	0.25	0.7	0.05	0
1016	0	0.15	0.65	0.2	0
1017	0	0.2	0.65	0.15	0
1018	0	0.25	0.65	0.1	0
1019	0	0.3	0.65	0.05	0
1020	0	0.25	0.6	0.15	0
1021	0	0.3	0.6	0.1	0
1022	0	0.35	0.6	0.05	0
1023	0	0.2	0.55	0.25	0
1024	0	0.25	0.55	0.2	0
1025	0	0.3	0.55	0.15	0
1026	0	0.35	0.55	0.1	0
1027	0	0.15	0.5	0.35	0
1028	0	0.2	0.5	0.3	0
1029	0	0.25	0.5	0.25	0
1030	0	0.3	0.5	0.2	0
1031	0	0.35	0.5	0.15	0
1032	0	0.15	0.45	0.4	0
1033	0	0.2	0.45	0.35	0
1034	0	0.25	0.45	0.3	0
1035	0	0.3	0.45	0.25	0
1036	0	0.35	0.45	0.2	0
1037	0	0.15	0.4	0.45	0
1038	0	0.2	0.4	0.4	0
1039	0	0.25	0.4	0.35	0
1040	0	0.2	0.35	0.45	0
1041	0	0.25	0.35	0.4	0
1042	0	0.35	0.35	0.3	0
1043	0	0.4	0.35	0.25	0
1044	0	0.45	0.35	0.2	0
1045	0	0.5	0.35	0.15	0
1046	0	0.55	0.35	0.1	0
1047	0	0.6	0.35	0.05	0
1048	0	0.35	0.3	0.35	0
1049	0	0.4	0.3	0.3	0
1050	0	0.45	0.3	0.25	0
1051	0	0.5	0.3	0.2	0
1052	0	0.55	0.3	0.15	0
1053	0	0.6	0.3	0.1	0
1054	0	0.65	0.3	0.05	0
1055	0	0.35	0.25	0.4	0
1056	0	0.4	0.25	0.35	0
1057	0	0.45	0.25	0.3	0
1058	0	0.5	0.25	0.25	0
1059	0	0.55	0.25	0.2	0
1060	0	0.6	0.25	0.15	0
1061	0	0.65	0.25	0.1	0
1062	0	0.7	0.25	0.05	0
1063	0	0.75	0.25	0	0
1064	0	0.35	0.2	0.45	0
1065	0	0.4	0.2	0.4	0
1066	0	0.45	0.2	0.35	0
1067	0	0.65	0.2	0.15	0
1068	0	0.7	0.2	0.1	0
1069	0	0.75	0.2	0.05	0
1070	0	0.8	0.2	0	0
1071	0	0.25	0.15	0.6	0
1072	0	0.3	0.15	0.55	0
1073	0	0.35	0.15	0.5	0
1074	0	0.4	0.15	0.45	0
1075	0	0.45	0.15	0.4	0
1076	0	0.65	0.15	0.2	0
1077	0	0.7	0.15	0.15	0
1078	0	0.75	0.15	0.1	0
1079	0	0.8	0.15	0.05	0
1080	0	0.85	0.15	0	0
1081	0	0.25	0.1	0.65	0
1082	0	0.3	0.1	0.6	0
1083	0	0.35	0.1	0.55	0
1084	0	0.4	0.1	0.5	0
1085	0	0.45	0.1	0.45	0
1086	0	0.65	0.1	0.25	0
1087	0	0.7	0.1	0.2	0
1088	0	0.75	0.1	0.15	0
1089	0	0.8	0.1	0.1	0
1090	0	0.85	0.1	0.05	0
1091	0	0.3	0.05	0.65	0
1092	0	0.35	0.05	0.6	0
1093	0	0.4	0.05	0.55	0
1094	0	0.45	0.05	0.5	0
1095	0	0.7	0.05	0.25	0
1096	0	0.75	0.05	0.2	0
1097	0	0.8	0.05	0.15	0
1098	0	0.85	0.05	0.1	0
1099	0	0.9	0.05	0.05	0
1100	0	0.85	0	0.15	0
1101	0	0.9	0	0.1	0
1102	0	0.95	0	0.05	0
1103	0.95	0	0	0.05	0
1104	0.9	0	0	0.1	0
1105	0.9	0.05	0	0.05	0
1106	0.85	0	0	0.15	0
1107	0.85	0.05	0	0.1	0
1108	0.85	0.1	0	0.05	0
1109	0.8	0.05	0	0.15	0
1110	0.8	0.1	0	0.1	0
1111	0.8	0.15	0	0.05	0
1112	0.75	0.1	0	0.15	0
1113	0.75	0.15	0	0.1	0
1114	0.65	0.25	0	0.1	0
1115	0.65	0.3	0	0.05	0
1116	0.6	0.25	0	0.15	0
1117	0.6	0.3	0	0.1	0
1118	0.6	0.35	0	0.05	0
1119	0.55	0.3	0	0.15	0
1120	0.55	0.35	0	0.1	0
1121	0.55	0.4	0	0.05	0
1122	0.5	0.35	0	0.15	0
1123	0.5	0.4	0	0.1	0

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
1124	0.5	0.45	0	0.05	0
1125	0.45	0.35	0	0.2	0
1126	0.45	0.4	0	0.15	0
1127	0.45	0.45	0	0.1	0
1128	0.4	0.35	0	0.25	0
1129	0.4	0.4	0	0.2	0
1130	0.4	0.45	0	0.15	0
1131	0.35	0.4	0	0.25	0
1132	0.35	0.45	0	0.2	0
1133	0.25	0.25	0	0.5	0
1134	0.25	0.3	0	0.45	0
1135	0.25	0.35	0	0.4	0
1136	0.25	0.4	0	0.35	0
1137	0.25	0.45	0	0.3	0
1138	0.25	0.5	0	0.25	0
1139	0.2	0.25	0	0.55	0
1140	0.2	0.3	0	0.5	0
1141	0.2	0.35	0	0.45	0
1142	0.2	0.4	0	0.4	0
1143	0.2	0.45	0	0.35	0
1144	0.2	0.5	0	0.3	0
1145	0.2	0.55	0	0.25	0
1146	0.15	0.15	0	0.7	0
1147	0.15	0.2	0	0.65	0
1148	0.15	0.25	0	0.6	0
1149	0.15	0.3	0	0.55	0
1150	0.15	0.35	0	0.5	0
1151	0.15	0.4	0	0.45	0
1152	0.15	0.45	0	0.4	0
1153	0.15	0.5	0	0.35	0
1154	0.15	0.55	0	0.3	0
1155	0.15	0.75	0	0.1	0
1156	0.15	0.8	0	0.05	0
1157	0.1	0.15	0	0.75	0
1158	0.1	0.2	0	0.7	0
1159	0.1	0.25	0	0.65	0
1160	0.1	0.3	0	0.6	0
1161	0.1	0.35	0	0.55	0
1162	0.1	0.4	0	0.5	0
1163	0.1	0.45	0	0.45	0
1164	0.1	0.5	0	0.4	0
1165	0.1	0.55	0	0.35	0
1166	0.1	0.75	0	0.15	0
1167	0.1	0.8	0	0.1	0
1168	0.1	0.85	0	0.05	0
1169	0.05	0.2	0	0.75	0
1170	0.05	0.25	0	0.7	0
1171	0.05	0.3	0	0.65	0
1172	0.05	0.35	0	0.6	0
1173	0.05	0.4	0	0.55	0
1174	0.05	0.45	0	0.5	0
1175	0.05	0.5	0	0.45	0
1176	0.05	0.55	0	0.4	0
1177	0.05	0.8	0	0.15	0
1178	0.05	0.85	0	0.1	0
1179	0.05	0.9	0	0.05	0
1180	0.9	0	0	0.05	0.05
1181	0.85	0	0	0.05	0.1
1182	0.85	0	0	0.1	0.05
1183	0.8	0	0	0.05	0.15
1184	0.8	0	0	0.1	0.1
1185	0.8	0	0	0.15	0.05
1186	0.75	0	0	0.1	0.15
1187	0.75	0	0	0.15	0.1
1188	0.55	0	0	0.05	0.4
1189	0.55	0	0	0.1	0.35
1190	0.5	0	0	0.05	0.45
1191	0.5	0	0	0.1	0.4
1192	0.5	0	0	0.15	0.35
1193	0.45	0	0	0.05	0.5
1194	0.45	0	0	0.1	0.45
1195	0.45	0	0	0.15	0.4
1196	0.45	0	0	0.45	0.1
1197	0.45	0	0	0.5	0.05
1198	0.4	0	0	0.05	0.55
1199	0.4	0	0	0.1	0.5
1200	0.4	0	0	0.15	0.45
1201	0.4	0	0	0.45	0.15
1202	0.4	0	0	0.5	0.1
1203	0.4	0	0	0.55	0.05
1204	0.35	0	0	0.1	0.55
1205	0.35	0	0	0.15	0.5
1206	0.35	0	0	0.2	0.45
1207	0.35	0	0	0.25	0.4
1208	0.35	0	0	0.3	0.35
1209	0.35	0	0	0.45	0.2
1210	0.35	0	0	0.5	0.15
1211	0.35	0	0	0.55	0.1
1212	0.3	0	0	0.15	0.55
1213	0.3	0	0	0.2	0.5
1214	0.3	0	0	0.25	0.45
1215	0.3	0	0	0.3	0.4
1216	0.3	0	0	0.35	0.35
1217	0.3	0	0	0.45	0.25
1218	0.3	0	0	0.5	0.2
1219	0.3	0	0	0.55	0.15
1220	0.25	0	0	0.2	0.55
1221	0.25	0	0	0.25	0.5
1222	0.25	0	0	0.3	0.45
1223	0.25	0	0	0.35	0.4
1224	0.25	0	0	0.4	0.35
1225	0.25	0	0	0.45	0.3
1226	0.25	0	0	0.5	0.25
1227	0.25	0	0	0.55	0.2
1228	0.25	0	0	0.6	0.15
1229	0.2	0	0	0.35	0.45
1230	0.2	0	0	0.4	0.4
1231	0.2	0	0	0.45	0.35
1232	0.2	0	0	0.5	0.3
1233	0.2	0	0	0.55	0.25
1234	0.2	0	0	0.6	0.2
1235	0.2	0	0	0.65	0.15
1236	0.15	0	0	0.35	0.5
1237	0.15	0	0	0.4	0.45
1238	0.15	0	0	0.45	0.4
1239	0.15	0	0	0.5	0.35
1240	0.15	0	0	0.55	0.3
1241	0.15	0	0	0.6	0.25
1242	0.15	0	0	0.65	0.2
1243	0.1	0	0	0.35	0.55
1244	0.1	0	0	0.4	0.5
1245	0.1	0	0	0.45	0.45
1246	0.1	0	0	0.5	0.4
1247	0.1	0	0	0.55	0.35
1248	0.05	0	0	0.4	0.55
1249	0.05	0	0	0.45	0.5

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
1250	0.05	0	0	0.5	0.45
1251	0.05	0	0	0.55	0.4
1252	0	0	0.75	0.15	0.1
1253	0	0	0.75	0.2	0.05
1254	0	0	0.7	0.15	0.15
1255	0	0	0.7	0.2	0.1
1256	0	0	0.7	0.25	0.05
1257	0	0	0.7	0.3	0
1258	0	0	0.65	0.2	0.15
1259	0	0	0.65	0.25	0.1
1260	0	0	0.65	0.3	0.05
1261	0	0	0.65	0.35	0
1262	0	0	0.6	0.3	0.1
1263	0	0	0.6	0.35	0.05
1264	0	0	0.6	0.4	0
1265	0	0	0.55	0.4	0.05
1266	0	0	0.55	0.45	0
1267	0	0	0.45	0.05	0.5
1268	0	0	0.45	0.1	0.45
1269	0	0	0.4	0.05	0.55
1270	0	0	0.4	0.1	0.5
1271	0	0	0.4	0.15	0.45
1272	0	0	0.35	0.05	0.6
1273	0	0	0.35	0.1	0.55
1274	0	0	0.35	0.15	0.5
1275	0	0	0.3	0.05	0.65
1276	0	0	0.3	0.1	0.6
1277	0	0	0.3	0.15	0.55
1278	0	0	0.25	0.1	0.65
1279	0	0	0.25	0.15	0.6
1280	0	0	0.25	0.25	0.5
1281	0	0	0.25	0.3	0.45
1282	0	0	0.2	0.25	0.55
1283	0	0	0.2	0.3	0.5
1284	0	0	0.2	0.35	0.45
1285	0	0	0.15	0.3	0.55
1286	0	0	0.15	0.35	0.5
1287	0	0	0.15	0.4	0.45
1288	0	0	0.1	0.35	0.55
1289	0	0	0.1	0.4	0.5
1290	0	0	0.1	0.45	0.45
1291	0	0	0.05	0.4	0.55
1292	0	0	0.05	0.45	0.5
1293	0.9	0	0.05	0.05	0
1294	0.85	0	0.05	0.1	0
1295	0.85	0	0.1	0.05	0
1296	0.8	0	0.05	0.15	0
1297	0.8	0	0.1	0.1	0
1298	0.8	0	0.15	0.05	0
1299	0.75	0	0.1	0.15	0
1300	0.75	0	0.15	0.1	0
1301	0.65	0	0.15	0.2	0
1302	0.65	0	0.2	0.15	0
1303	0.6	0	0.15	0.25	0
1304	0.6	0	0.2	0.2	0
1305	0.6	0	0.25	0.15	0
1306	0.55	0	0.2	0.25	0
1307	0.55	0	0.25	0.2	0
1308	0.55	0	0.3	0.15	0
1309	0.5	0	0.25	0.25	0
1310	0.5	0	0.3	0.2	0
1311	0.5	0	0.35	0.15	0
1312	0.45	0	0.3	0.25	0
1313	0.45	0	0.35	0.2	0
1314	0.45	0	0.4	0.15	0
1315	0.4	0	0.35	0.25	0
1316	0.4	0	0.4	0.2	0
1317	0.4	0	0.45	0.15	0
1318	0.35	0	0.35	0.3	0
1319	0.35	0	0.4	0.25	0
1320	0.35	0	0.45	0.2	0
1321	0.3	0	0.35	0.35	0
1322	0.3	0	0.4	0.3	0
1323	0.3	0	0.45	0.25	0
1324	0.25	0	0.25	0.5	0
1325	0.25	0	0.3	0.45	0
1326	0.25	0	0.35	0.4	0
1327	0.25	0	0.4	0.35	0
1328	0.25	0	0.45	0.3	0
1329	0.25	0	0.5	0.25	0
1330	0.2	0	0.25	0.55	0
1331	0.2	0	0.3	0.5	0
1332	0.2	0	0.35	0.45	0
1333	0.2	0	0.4	0.4	0
1334	0.2	0	0.45	0.35	0
1335	0.2	0	0.5	0.3	0
1336	0.2	0	0.55	0.25	0
1337	0.15	0	0.3	0.55	0
1338	0.15	0	0.35	0.5	0
1339	0.15	0	0.4	0.45	0
1340	0.15	0	0.45	0.4	0
1341	0.15	0	0.5	0.35	0
1342	0.15	0	0.55	0.3	0
1343	0.1	0	0.45	0.45	0
1344	0.1	0	0.5	0.4	0
1345	0.1	0	0.55	0.35	0
1346	0.05	0	0.5	0.45	0
1347	0.05	0	0.55	0.4	0
1348	0.05	0	0.6	0.35	0
1349	0	0.2	0.5	0.1	0.2
1350	0	0.15	0.55	0.1	0.2
1351	0	0.25	0.45	0.1	0.2
1352	0	0.2	0.55	0.1	0.15
1353	0	0.15	0.6	0.1	0.15
1354	0	0.15	0.5	0.1	0.25
1355	0	0.2	0.45	0.1	0.25
1356	0	0.2	0.5	0.1	0.2
1357	0	0.15	0.55	0.1	0.2
1358	0	0.35	0.45	0.1	0.2
1359	0	0.25	0.3	0.15	0.3
1360	0	0.2	0.55	0.1	0.15
1361	0	0.15	0.6	0.1	0.15
1362	0	0.25	0.5	0.1	0.15
1363	0	0.3	0.45	0.1	0.15
1364	0	0.15	0.5	0.15	0.2
1365	0	0.2	0.45	0.15	0.2
1366	0	0.2	0.65	0.1	0.05
1367	0	0.25	0.6	0.1	0.05
1368	0	0.3	0.55	0.1	0.05
1369	0	0.35	0.5	0.1	0.05
1370	0	0.25	0.6	0.1	0.05
1371	0	0.3	0.45	0.1	0.15
1372	0	0.35	0.4	0.1	0.15
1373	0	0.4	0.35	0.1	0.15
1374	0	0.45	0.3	0.1	0.15
1375	0.1	0.2	0.05	0.4	0.25

continued on next page

Sample No.	Cr	Mn	Co	Te	Ni
1376	0.1	0.2	0.05	0.35	0.3
1377	0.1	0.1	0.2	0.4	0.2
1378	0.1	0.15	0.05	0.4	0.3
1379	0.1	0.1	0.3	0.4	0.1
1380	0.2	0.3	0.1	0.3	0.1
1381	0.1	0.2	0.2	0.4	0.1
1382	0.05	0.2	0.05	0.4	0.3
1383	0.2	0.1	0.3	0.3	0.1
1384	0.1	0.1	0.2	0.4	0.2
1385	0.1	0.1	0.3	0.4	0.1
1386	0.2	0.3	0.1	0.3	0.1
1387	0.1	0.2	0.2	0.4	0.1
1388	0.2	0.1	0.3	0.3	0.1
1389	0	0.3	0.3	0.1	0.3
1390	0	0.25	0.35	0.1	0.3
1391	0	0.35	0.25	0.1	0.3
1392	0	0.25	0.3	0.1	0.35
1393	0	0.35	0.2	0.1	0.35
1394	0	0.2	0.6	0.1	0.1
1395	0	0.25	0.55	0.1	0.1
1396	0	0.3	0.5	0.1	0.1
1397	0	0.1	0.7	0.1	0.1
1398	0	0.15	0.65	0.1	0.1
1399	0	0.4	0.4	0.1	0.1
1400	0	0.5	0.3	0.1	0.1

Table C.2: Second Generation of Catalysts and their Compositions.

D Synthesized Libraries with Distribution of Samples, References and Empty Wells

In the following the distribution of the catalyst samples on the slate libraries can be found. The wells are marked by the sample number or by E ('empty well') and H ('Hopcalite'). Only Libraries 1a-5a are shown since libraries 1b-5b have been identically filled. Libraries 6 and 7 contain the 400 refinement catalysts with 5% and 10 %-wise variations of composition.

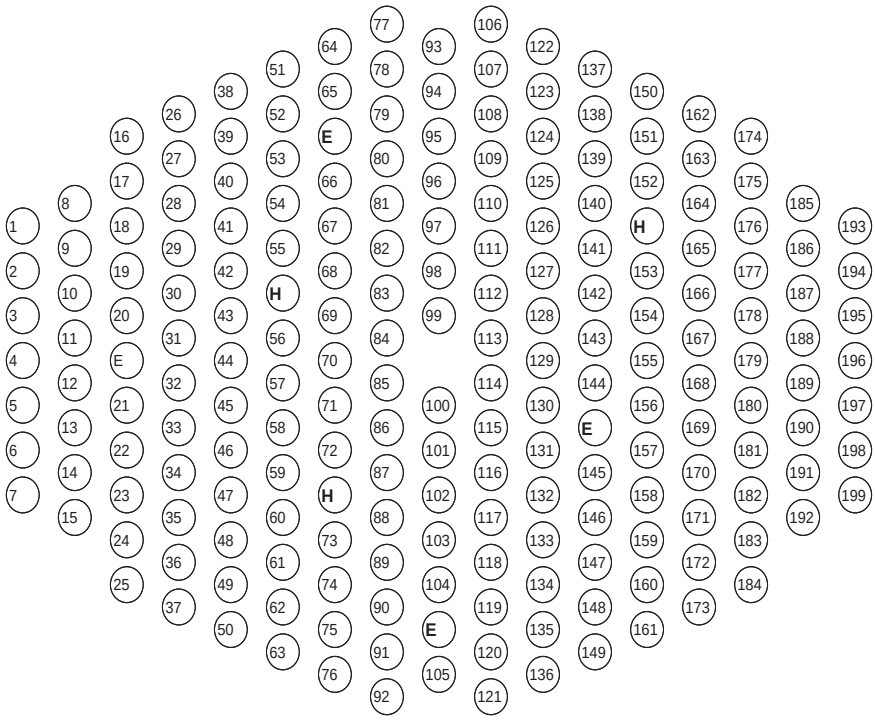


Figure D.1: Layout of Library 1a.

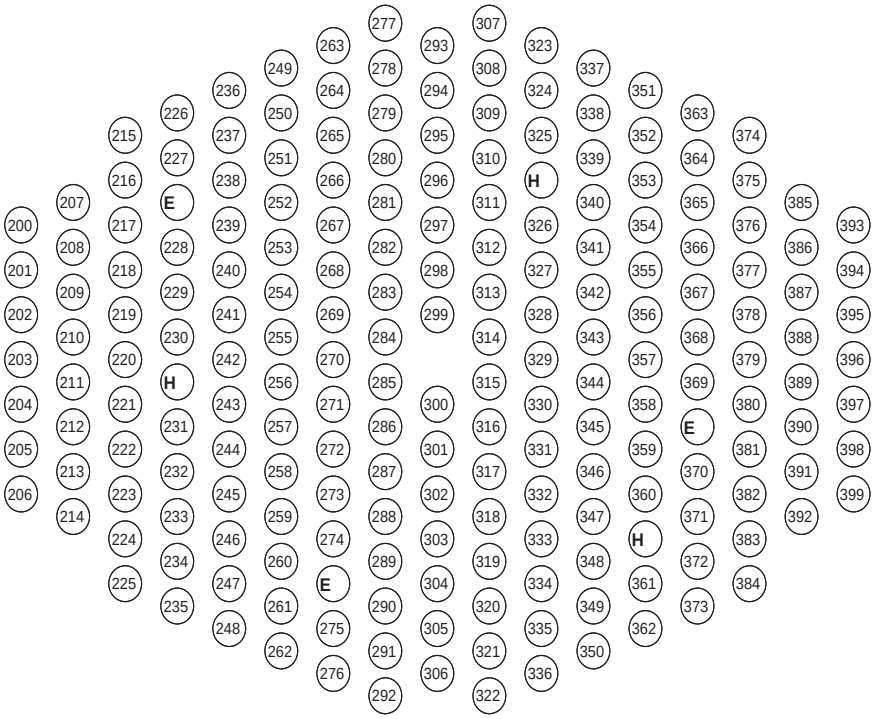


Figure D.2: Layout of Library 2a.

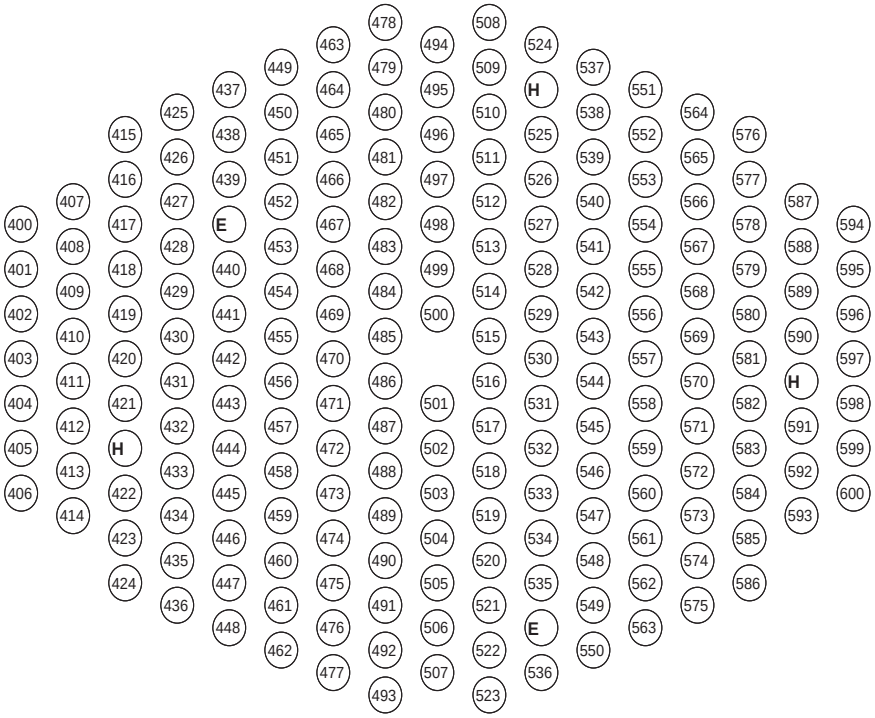


Figure D.3: Layout of Library 3a.

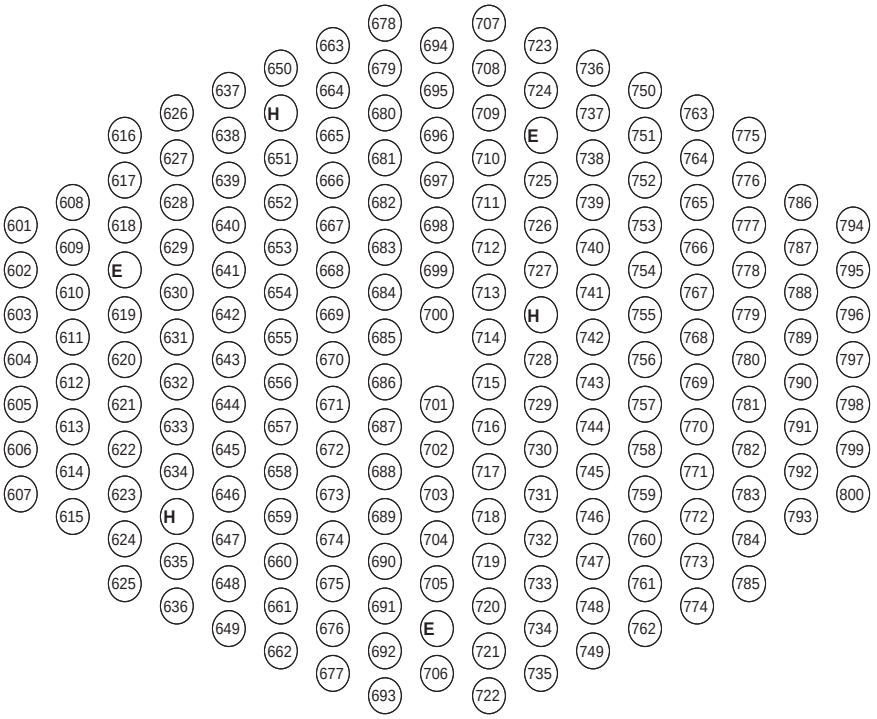


Figure D.4: Layout of Library 4a.

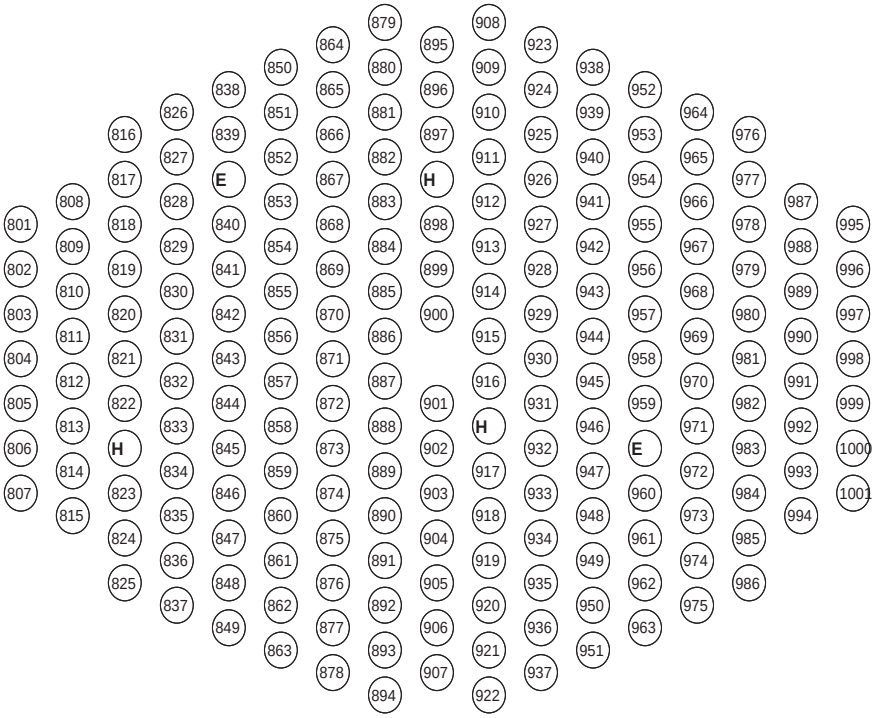


Figure D.5: Layout of Library 5a.

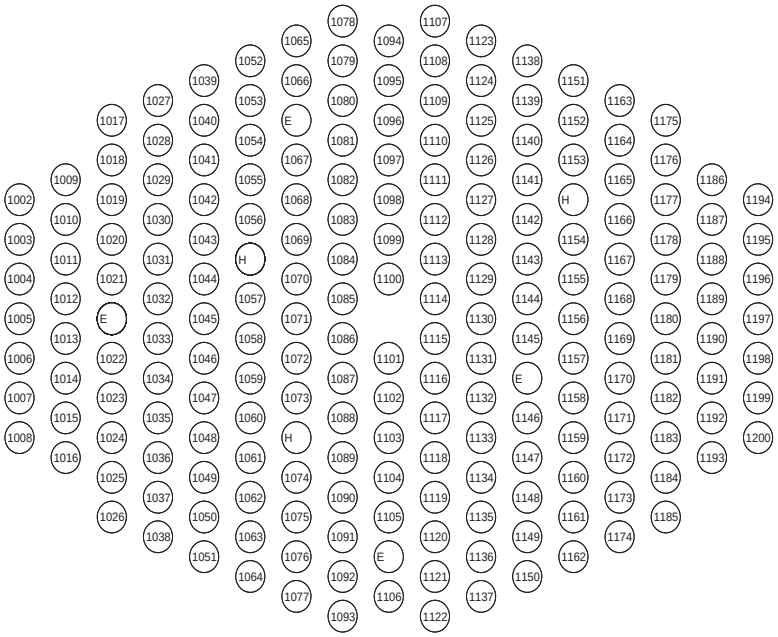


Figure D.6: Layout of Library 6.

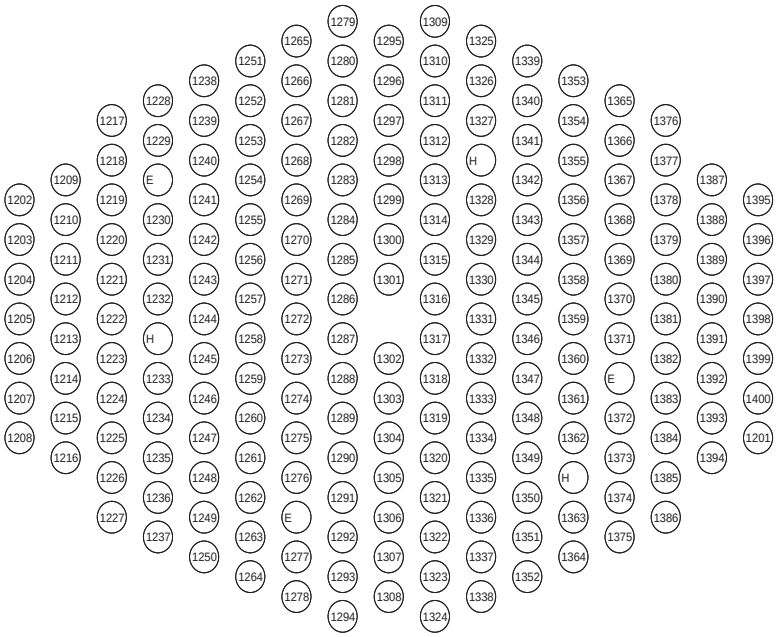


Figure D.7: Layout of Library 7.