

Model-driven catalyst design for ethanol dehydration into ethylene

Diogo Rafael Teixeira de Ascensão
diogo.ascensao@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2022

Abstract

Olefins, mainly produced by steam cracking, are playing a bigger part in the chemical industry, year by year. However, with the rise in environmental awareness and the new environmental policies, the search for greener production processes is very important. Dehydration of ethanol into ethylene, the process of interest in this thesis appears as one of the main alternatives to this problem. However, the ideal catalyst for this is yet to be found. Hence, a methodology previously developed, named "Model-driven catalyst design", was applied in this work to identify the most suitable catalyst properties to maximize ethylene yield. This methodology integrates the kinetics of the process, by using identified catalyst descriptors, namely chemisorption enthalpies, sticking coefficients and active site density, with statistical tools. It has as main objective to establish relationships between catalyst properties and the observed performance. Experimental data reported in two different studies obtained over zeolite-based catalysts modified through dealumination and metal doping processes, with phosphorous and lanthanum, were used in this work to correlate catalyst physical properties with their catalytic activity. Obtained relationships show that lower chemisorption enthalpies of ethylene are associated with good performances while higher values are related to worst performance selectivity-wise. This appeared associated with catalysts doped with phosphorous, which leads to modifications in the catalyst's porous structure and in the acidity of the catalyst. Yet, this constitutes preliminary conclusions and broader datasets should be studied in the future, so getting to an optimal catalyst composition can be an easier process.

Keywords: Catalyst descriptor, Kinetic modelling, Ethylene, Dehydration, Catalyst design, Virtual Catalyst

1. Introduction

With a claimed production capacity of 200 million tons by 2024 and a predicted increase of almost 40% for the decade between 2016 and 2026, ethylene is one of the chemicals with the largest production in the world, being not only an important product but a very important raw material, or intermediate, for the production of several products, such as acetic acid and styrene [1]. The conventional processes for the production of ethylene rely majorly on fossil fuels, since the main production route for it is steam cracking. However, due to the rise of environmental concerns and the demand for this fuel, finding new ways to produce this kind of product is a priority. Among many processes like methanol to olefins, oxidative coupling of methane or Fischer-Tropsch synthesis, ethanol dehydration to ethylene is reported to be one of the best alternatives [2], especially with the rise of bioethanol production, and the ease of how this can be produced from biomass. However, for this process to be viable more insights and search needs to be done, especially regarding the operating condi-

tions and catalyst. There are three types of catalyst development techniques, statistics-driven design, performance-driven design and information-driven design. There is also an emerging approach related to information-driven design, entitled "Model-driven catalyst design", which is based on the concept of catalyst descriptors [3].

The objective of the present work was to apply the model-driven catalyst design to identify the properties of zeolite-based catalysts that largely affect ethanol conversion and ethylene selectivity and therefore provide guidelines for the catalyst synthesis. To this end, a kinetic model based on the reaction mechanism and catalyst descriptors was developed considering an ideal plug flow reactor. Various virtual catalysts, defined by a combination of the descriptor values (which will be defined later), were formulated and tested *in silico* with the developed kinetic model. The performance of virtual catalysts was compared with the one of the real catalysts and the match was statistically evaluated to establish relationships between the physical properties and observed performance, aiming

to find a catalyst composition that would provide, maximal yield and selectivity of ethylene.

2. Ethylene industrial production

Steam cracking is a well-established technology that is the leader among the technologies for olefin production. It consists of a petrochemical process that breaks down saturated hydrocarbons into smaller molecules. However, due to the rise of environmental concerns and the implementation of new governmental policies associated with it, the search for new processes is being given more importance as time goes by [2, 4, 5].

2.1. Ethanol dehydration into ethylene

Catalytic dehydration of ethanol into ethylene appears as one of the main alternatives for ethylene's main production process. It is a process that has been known for a long time but has never been exploited for large-scale ethylene synthesis, mainly due to the cost and raw material availability. However, this technique must be reassessed, due to the big increase in bioethanol production, especially in the American and Asian continents. This could solve the environmental problems related to the conventional steam cracking route and lead to a disruption of ethylene's value chain reliance on crude oil price fluctuations [2, 4].

2.2. Previously used catalysts

Ethanol dehydration can be considered an acid-catalyzed reaction for which many catalysts have already been tried: phosphoric acid, alumina-based, heteropolyacid, and zeolites. The first three mentioned types of catalysts are no longer used either because of low stability related to easy deactivation by coke deposition or the high cost associated with the need for loading in supports, for example [6]. The last ones, zeolites, are considered to be the most promising catalysts in this process. This is due to their porous structure that can be easily modified to enhance selectivity to the wanted products. As with everything, these still have some disadvantages regarding, similarly to the previous ones, less stability due to coke deposition [6]. Hence finding the right catalyst for the studied process is still a task that needs to be done. Therefore making use of new catalyst development techniques that aid the experimental work developed is more important than never.

2.3. Mechanism

Many studies have been made on the reaction mechanism over different kinds of catalysts. For zeolites, which are the main focus of this work, the mechanism is consensual and it considers three different reaction paths. The first one consists of the ethanol dehydration to water and ethylene and

is thermodynamically endothermic while the second is a bi-molecular step where ethanol is dehydrated into water and diethyl ether and is an exothermic reaction. The third path embodies a diethyl ether conversion to ethylene and ethanol, as reported in previously conducted density-functional theory studies by K. Alexopoulos *et al.*[7]. Finally, their performances will be compared with the ones from real catalysts to establish relationships between the physical properties of these and the observed performance, to try to find an optimal catalyst for the process.

3. Methodology

Figure 1 shows a schematical representation of the methodology proposed in Ref.[3] and followed in this work. It is based on three different concepts. The first one is the concept of virtual catalyst which is a computational representation of materials that are defined as a vector of m catalyst descriptors ($D_{i,j}$). The number of virtual catalysts to be tested is defined as n . The second concept is one of the real catalysts that are materials with known composition and structure and have already been studied, previously, hence their performance is already available and quantified. The third concept on which the methodology relies is the concept of catalyst descriptor. This works as a bridge between virtual and real catalysts being defined as a factor that has an impact on the reaction behaviour and kinetics varying from catalyst to catalyst [3].

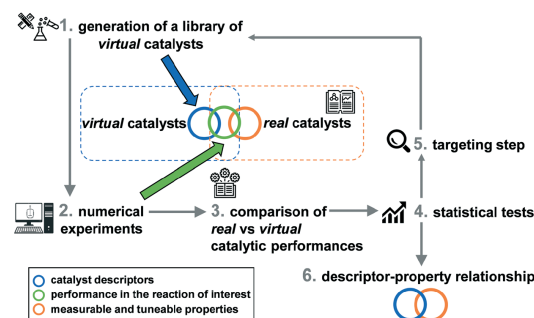


Figure 1: Summary of the followed methodology for the model-driven catalyst design technique [3].

This methodology comprises 6 different steps that will be now briefly explained.

Step 1. Virtual catalysts generation.

This first step has as its main objective to generate the library of virtual catalysts. These libraries need to be statistically representative and meaningful and for that, a Design of Experiments technique named Fast flexible space-filling design was used. The n virtual catalysts obtained from the implementation of this technique are constituted by a set of m descriptors, leading to a $n \times m$ matrix, where each row represents a vector of descriptor

combinations, and consequently, a virtual catalyst [3].

Step 2. Numerical experiments.

In this step, the previously generated virtual catalysts, are screened in a numerical set-up. The numerical set-up in this context represents a microkinetic model, where the kinetics is correlated to the selected descriptors, coupled with the adequate reactor model. With this numerical set-up, the behaviour of all virtual catalysts, that compose the discovery library, can be evaluated, and parameters that can characterize catalysts' performance like conversion or selectivity can be calculated [3]. So evaluated performance of virtual catalysts should be further compared to the performance of real catalysts.

Step 3. Comparison between the performances of virtual and real catalysts.

Once the performance of the virtual catalysts is studied in the numerical experiments step, it should be compared with the real performance obtained from real catalysts. To achieve this, clustering algorithms like the *k-Means* clustering technique can be used. A cluster consists of data points, simulated and/or experimental, that have been grouped due to particular commonalities [8] and can be performed through the use of specific software like *Orange3*, for example. This consists of a very intuitive tool that can assess some problems related to *k-Means* clustering, namely the choice of a correct number of clusters for a specific dataset and the high influence on the obtained results of the initialization step of the method [3].

Step 4. Statistical analysis of the descriptor distributions.

The descriptors present different probability distributions in each one of the clusters obtained before. By taking this concept into account, through comparing these distributions within all the clusters, we can get to the more relevant descriptors, which allows us to differentiate from one to the other and start to relate these with some real properties already [3]. To test the descriptor distributions some specific statistical tests need to be applied. In this case, tests like the ANOVA or the t-student test, which are the most commonly used ones, should not be applied, as these are indicated for data that follows a normal distribution type and catalyst descriptors are likely to be non-normally distributed within the different clusters. Thus the non-parametric alternatives of the mentioned tests, like the Kruskal-Wallis and the Mann-Whitney test, should be used [3]. The assessment parameter that will be used in the application of this statistical test to evaluate if a descriptor is a discriminating one or not will be the p-value being 0.05 the value

considered as the threshold for the identification of a descriptor as relevant or not [3].

Step 5. Building targeted libraries of catalysts.

To accomplish the methodology's objective, and to be able to establish relationships between catalyst descriptors and physical properties, the number of virtual catalysts close to real catalysts should be high enough. For this, the number of virtual catalysts in the surrounding of real catalysts should be increased across iterations, which is done by generating targeted libraries from the discovered library used in the first iteration. Consequently, targeting means reducing the design space iteratively through the diminishing of discriminating descriptor ranges with the main objective of getting virtual catalysts closer to the performance of real catalysts [3]. Concerning this narrowing down of the value ranges, standard deviations of around $\pm\sqrt{2}$ times the standard deviations of the descriptor values for the studied cluster can be applied to the mean of the cluster. According to Chebyshev inequality, this formula assures that at least 50% of the virtual catalysts are retained on the new cluster that is being generated [3].

Step 6. Developing descriptor-property relationships from real catalyst properties.

This step aims at establishing a relationship between a qualitative property, which is measurable and tuneable property of the studied catalyst and the descriptor that reflects this property in the kinetic model (i.e. reflects on the kinetics).

4. Catalyst design

4.1. Generation of the virtual catalyst library

The discovery library initially generated needs to be built with a considerable amount of virtual catalysts so several different performances can be obtained for comparison. A virtual catalyst is defined by selecting a value for each descriptor from a priori-defined descriptor ranges based on Design of experiments and fast-flexible filling design methods. As mentioned earlier, a virtual catalyst is a combination of identified descriptors. The descriptors chosen in this work were: 1) the sticking coefficients of ethanol, 2) the sticking coefficient of ethylene, 3) the chemisorption enthalpies of ethanol, 4) the chemisorption enthalpy of ethylene, 5) the chemisorption enthalpy of DEE, and 6) the density of active sites of the used catalyst. The used ranges for these parameters were estimated based on a literature survey and their values are presented in Table 1. In the case of sticking coefficients, due to a lack of data in the literature, their initial range was considered to be between 0 and 1 since they represent probabilities.

Table 1: Value ranges used for the catalyst descriptors in the discovery library.[9, 10, 11, 12, 13, 14, 15]

Descriptor	Value
$\Delta H_{\text{chemisorption,Ethanol}}$ (kJ/mol)	35-135
$\Delta H_{\text{chemisorption,Ethylene}}$ (kJ/mol)	35-165
$\Delta H_{\text{chemisorption,DEE}}$ (kJ/mol)	50-90
$S_{0,\text{Ethanol}}$	0-1
$S_{0,\text{Ethylene}}$	0-1
$\rho_{\text{active sites}}$	$2 \times 10^{-7} - 2 \times 10^{-6}$

4.2. Kinetic model formulation

Due to numerical problems that occurred with the simulations due to the high stiffness of the studied process, associated, mainly, with the kinetic constants of the system when modelling in terms of concentrations, the developed kinetic model was an adaptation of a previously built code, following the mechanism stated on K. Alexopoulos *et al.*[7]. The model was developed by making use of the following equations. Where F is the flow of the components in mol/s, W is the mass of catalyst in kg, θ^* , θ_0^* , and θ_i are the amount of active sites available for adsorption at a given time, the initial amount of active sites and the surface coverage by component i , respectively.

$$\frac{dF_i}{dW_{\text{cat}}} = R_i \quad (1)$$

$$R_i = C_t \times \left(\sum_1^N r_{i,\text{ads}} - \sum_1^N r_{i,\text{des}} \pm \sum_1^N r_{i,\text{reaction}} \right) \quad (2)$$

Where $F_i = F_{i,0}$ at $W = 0$.

$$\frac{d\theta_k}{dt} = 0 = \sum_1^N r_{k,\text{ads}} - \sum_1^N r_{k,\text{des}} \pm \sum_1^N r_{k,\text{reaction}} \quad (3)$$

$$\theta^* = \theta_0^* - \sum_{i=1}^N \theta_i \quad (4)$$

To avoid dealing with numeric problems as in other previous approaches, a scaling factor was used for the adsorption and desorption rates. The scaling factor used was 10^3 to follow what had already been done in the previous modelling of this process [16]. This model was solved in Python by making use of the *numpy* and *scipy* packages.

The used parameters to assess the catalyst's performance were the conversion of ethanol and the selectivity to ethylene and diethyl ether.

4.3. Descriptor Implementation

The catalyst descriptors will be implemented through the rate coefficients of the reaction. The

adsorption and desorption steps of the gaseous species C_2H_4 , DEE, and C_2H_5OH were the chosen ones for descriptor implementation. These steps were assumed to be non-activated ones so the reaction rates can be determined by using catalyst descriptors. This way, the adsorption steps rate constants were obtained through Equation 5 as in [17], where k represents the rate constant of the reaction, $S_{0,i}$ is the initial sticking probability, σ the density of active sites in mol/m^2 , n is the order of the reaction, R is the universal gas constant, T is the temperature at which the process is being studied in Kelvin and, lastly, M is the molar mass of the adsorbed component in kg/mol .

$$k = \frac{S_{0,i}}{\sigma^n} \sqrt{\frac{RT}{2\pi M}} \quad (5)$$

This equation was then converted to pressure units so it could be used in the model by making use of the ideal gas law [18].

Regarding the desorption steps, descriptors were implemented on this by making use of the Arrhenius equation, presented in Equation 6 [17], where, similarly to Equation 5 the R and T represent the universal gas constant in J/mol.K and the process temperature in K , respectively. A is the pre-exponential factor, that was assumed to be placed in the interval from 10^{13} s^{-1} to 10^{16} s^{-1} as proposed by James A. Dumesic *et al.* [19]. Lastly, E_a represents the activation energy of the reaction step in J/mol . The activation energies of these desorption steps were considered equal to the respective chemisorption enthalpies of the different species. This is based on the, previously mentioned, assumption that these steps are non-activated.

$$k = A \times \exp\left(\frac{-E_a}{RT}\right) \quad (6)$$

4.4. Case Study 1

The experimental dataset studied by N.Zhan *et al.*[12] was selected in this work to perform the methodology. In their work, a set of 6 different HZSM-5 zeolites, pure and doped with lanthanum and phosphorous, were tested under different operational conditions. The used percentages of each element were 0.25%, 0.5% and 1% for lanthanum and 2% for phosphorous. Upon the generation of the discovery library, all the virtual catalysts were tested *in silico* (by the mean of a microkinetic model) and the obtained performances have been displayed in Figure 2, which shows ethylene selectivity as a function of ethanol conversion. The real catalysts performance is also presented in Figure 2 for an easier evaluation of virtual ones in terms of matching the real catalyst behaviour.

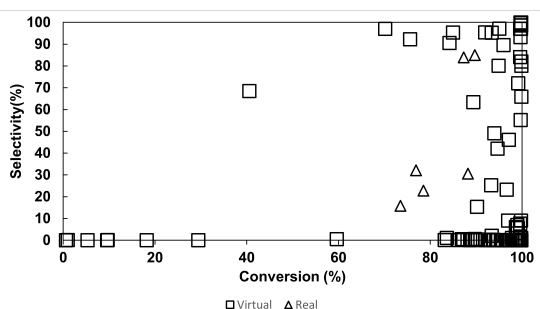


Figure 2: Obtained performances for the simulated virtual catalysts and the respective real catalysts from dataset 1. Reaction conditions: $T=493\text{K}$, $W=0.5\text{g}$ of catalyst, atmospheric pressure, $\text{WHSV} = 2.0\text{h}^{-1}$.

Step 3- Comparison between the performances of virtual and real catalysts

By making use of the *Orange3* software to statistically evaluate the match between virtual and real catalysts, three different clusters were identified for all the performances obtained with virtual and real catalysts as the ideal number of division groups, as shown in Figure 3. The first cluster, coloured in blue, corresponds to very low conversion and selectivity. The second one, coloured in orange, corresponds to high conversions associated but lower selectivities. Lastly, the third cluster in yellow presents high selectivities and conversions. Given that only clusters, number two and three have both virtual and real catalysts these two are the only relevant clusters that will be statistically studied in further steps.

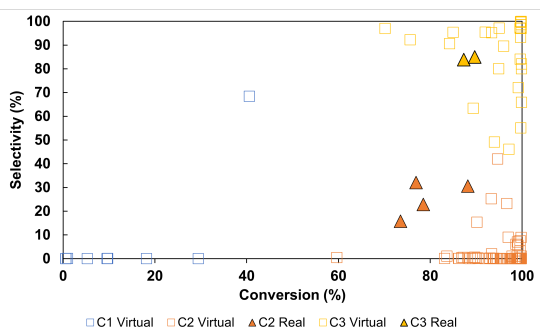


Figure 3: Output of the clustering step performed in *Orange3* for case study 1.

Step 4- Statistical analysis of the descriptor distributions

After applying the statistical tests described earlier, the relevance of used descriptors for observed catalyst behaviour has been assessed by using the p-value of the statistical tests as a quantitative criterion. Accordingly, only the chemisorption enthalpy of ethylene satisfied this criterion and was thus identified as the most influential descriptor for the obtained results. Figure 4 displays a comparison between the distributions of this descriptor in box plots, where the central line represents the median of the descriptor distribution in the cluster, the

cross the mean of the descriptor values in each cluster, and the edges of the box represent the 25th percentile at the bottom and 75th percentile at the top. The lines that stand outside the box are the respective standard deviations of each one of the clusters. These box plots work as a very good tool when several groups of data are being studied simultaneously [20]. Because, not only this representation can already help establish relationships since it is already possible to see a trend in terms of cluster 2 presenting higher values of chemisorption enthalpy, while cluster 3 presents very lower chemisorption enthalpy values, but also it proves that the descriptors are not normally distributed, otherwise the median would be centred inside the box and the rest of the diagram would be symmetric, proving what was said in above, regarding the application of non-parametric statistical tests.

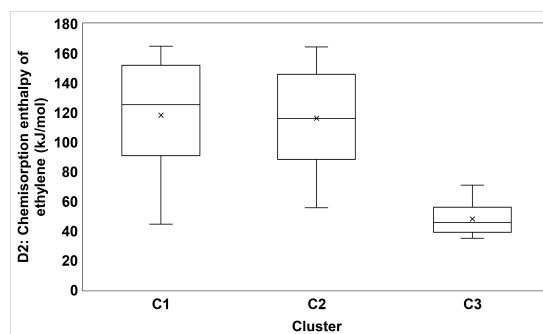


Figure 4: Comparison of the distributions of the relevant descriptor between clusters for the discovery library of case study 1.

Step 5-Building targeted libraries of virtual catalyst

Considering all the results obtained, a targeted library of virtual catalysts for each cluster will be created by applying the Chebyshev inequality to the values taken into account for the chemisorption enthalpy of ethylene within each cluster. Figures 5 and 6 show the results obtained upon applying the targeting process to clusters 2 and 3, respectively. As expected, the shrinking of the descriptor ranges leads to a less broad range of performances, since the dots that represent virtual catalysts are closer to each other, but with a higher number of virtual catalysts per cluster, approximating the distance between virtual and real catalysts performances. This was accomplished without an increase in the total number of virtual catalysts, which proves the targeting step was successful. The trend obtained after the first iteration was retained after performing this step. Namely, cluster 2 is still characterized by catalysts with high conversion but lower selectivities, while cluster 3 contains catalysts with high conversions and high selectivities.

Figure 7 represents the distribution of the identified relevant descriptor, explained previously, for

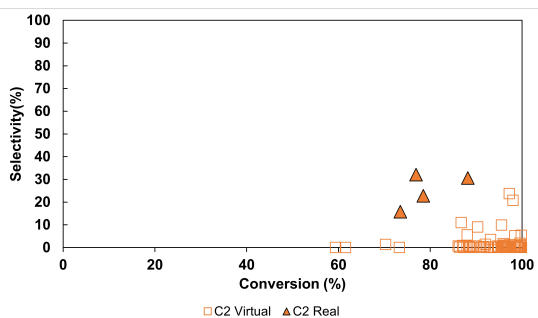


Figure 5: Result of the targeting step (step 5) for cluster 2 of case study 1.

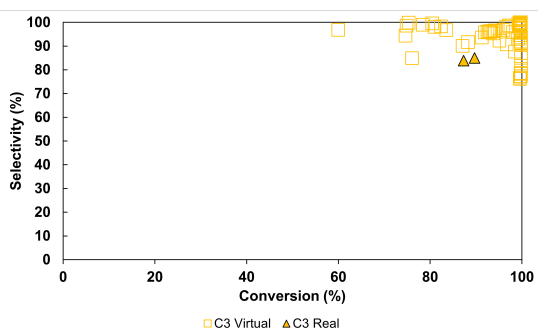


Figure 6: Result of the targeting step (step 5) for cluster 3 of case study 1.

the discovery library, and both the targeted libraries for clusters 2 and 3, respectively.

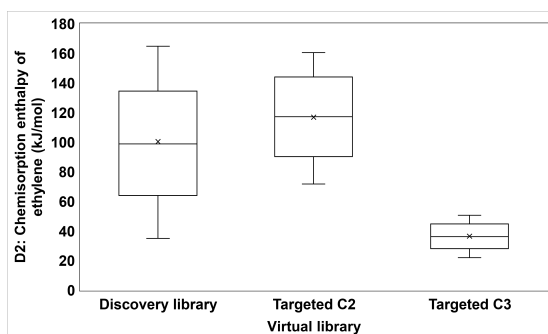


Figure 7: Comparison of the distribution of the relevant descriptor in the discovering and targeted libraries of virtual catalysts for case study 1.

Step 6 - Developing descriptor-property relationships from real catalysts properties

To relate catalyst compositions to these observations presented in Figure 7, a closer look at the catalyst characterization needs to be done. Hence by looking at the characterization done in [12], it is possible to see that, as stated before, 6 different catalysts were studied being the BET surface one of the most pertinent studied characteristics for the studied dataset, given the fact that it is a relevant property for the chemisorption enthalpy as well. The HZSM-5 catalyst that was modified by the addition of only phosphorous presents a very big performance drop, when compared to the unmodified HZSM-5 zeolite, especially in terms of se-

lectivity. This can be due to a high reduction in the BET surface area of the catalyst, also to lower pore volume and width and, lastly, to a decrease in the total acidity amount of the catalyst as reported by N.Zhan *et al.*[12]. Considering that the density of active sites was not identified as a relevant descriptor, this change in the total acidity could not be a major influence on the observed results, however, the p-value for the density of active sites was very close to the one taken into account as a threshold, so the acidity can end up being an impactful factor. This decrease in the total acidity amount is reported to have led to a major decrease in the quantity of stronger acid sites diminishing the acidic strength of the catalyst and consequently hurting its performance given the fact that this process is usually catalyzed by strongly acidic zeolites [21]. On the other hand modification with lanthanum has enhanced catalyst performance at the studied operational conditions. This is probably associated with the increase in BET surface area and micropore volume. Moreover, it is reported that adding lanthanum to HZSM-5 can improve the stability of the $[AlO_4]^-$ anion, which can be also one of the reasons why the HZSM-5 zeolite modified with just lanthanum is the one that presents a better performance out of all of the catalysts present in the dataset[22]. By analysing Figure 7 it is possible to observe that the targeted library for cluster 2, which is associated with lower selectivities for ethylene, reveals to have higher values for the ethylene chemisorption enthalpy when compared to the targeted library of cluster 3, related to a high selectivity for this compound that presents lower values for the ethylene chemisorption enthalpy. So, taking into account these considerations on catalyst characteristics and the shown relations in Figure 7, one can conclude that higher values of the chemisorption enthalpy of ethylene are directly correlated to the presence of phosphorous, lower amounts of acidic strength of the catalyst and BET surface area, and consequently, to worst catalytic performance.

4.5. Case Study 2

Another set of experimental data was taken from [23], where zeolites have been dealuminated apart from being only doped with metals, as it was the case in Case study 1. In this study, six different ZSM-5 zeolites were prepared, characterized, and tested, differing in the doping metal or the degree of dealumination. Having performed the numerical experiments step by making use of the developed microkinetic model, the performances obtained for the simulated virtual catalysts are shown in Figure 8.

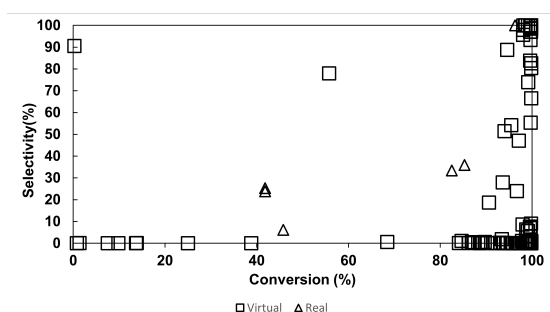


Figure 8: Obtained performances for the simulated virtual catalysts and the respective real catalysts from the dataset 2. Reaction conditions: $T=493K$, $W=0.4g$ of catalyst, atmospheric pressure, $WHSV = 1.5h^{-1}$.

Step 3- Comparison between the performances of virtual and real catalysts

Three different clusters were also identified, in this case, by applying the silhouette method in the software *Orange3*. All three clusters contain virtual and real catalysts. In analogy to the previous case study, these three clusters are marked in blue, orange and yellow corresponding, respectively, to lower conversions and selectivities, with high conversion but low selectivity and high conversion and selectivities.

Figure 9 visualizes the results of the clustering step after this first iteration of the methodology.

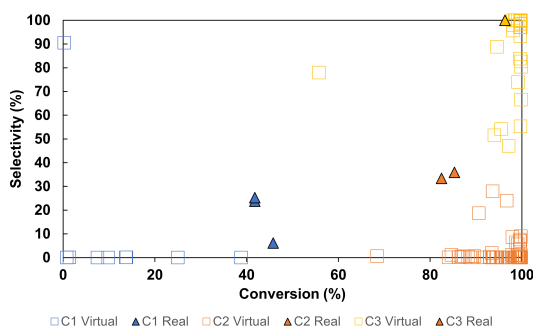


Figure 9: Output of the clustering step performed in *Orange3* for the 2nd dataset.

Step 4- Statistical analysis of the descriptor distributions

Having applied the statistical tests, which this time was the Kruskal Wallis test, due to having more than two clusters with real catalysts, it was possible to identify the discriminating descriptors for this case.

The same threshold of 0.05 was applied as the boundary for the p-value to filter more from less influential descriptors, which were the chemisorption enthalpies of ethanol, ethylene and diethyl ether. Figures 10, 11 and 12 show the comparison of the obtained distributions for the three relevant descriptors in the three clusters. Cluster 1 which presents very low values for ethanol chemisorption enthalpy, also cluster 3 presents very low values for

ethylene chemisorption enthalpy, similarly to Case study 1.

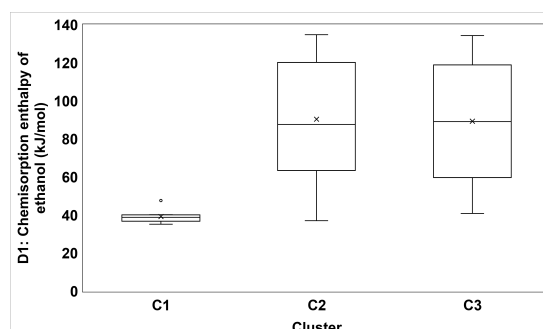


Figure 10: Comparison of the distributions of the chemisorption enthalpy of ethanol between clusters for the discovery library of case study 2.

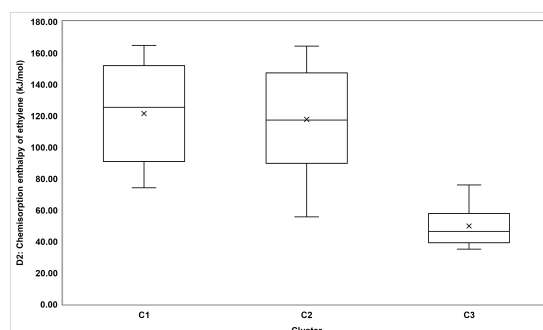


Figure 11: Comparison of the distributions of the chemisorption enthalpy of ethylene between clusters for the discovery library of case study 2.

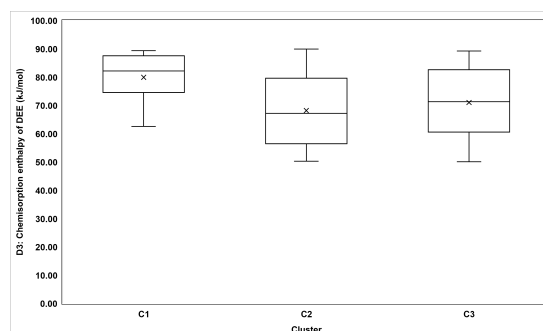


Figure 12: Comparison of the distributions of the chemisorption enthalpy of diethyl ether between clusters for the discovery library of case study 2.

Step 5-Building targeted libraries of virtual catalyst

In analogy to Case study 1, the Chebyshev inequality was applied to the clusters obtained in the previous step. In Figure 13 and Figure 14 the results obtained for the targeting process to clusters 1 and 2 are presented. By looking at the figures, by visual intuition, one could conclude that the targeting process has been successful for both of the clusters since the number of virtual catalysts in the cluster is way higher than that obtained for the discovery library. However, both of the clusters after

the targeting step do not contain real catalysts anymore. This could be because the real catalysts for both of these clusters were close to their boundaries, representing one of the major downsides of using a dataset with low quantities of real catalysts (<10 samples), as is the case.

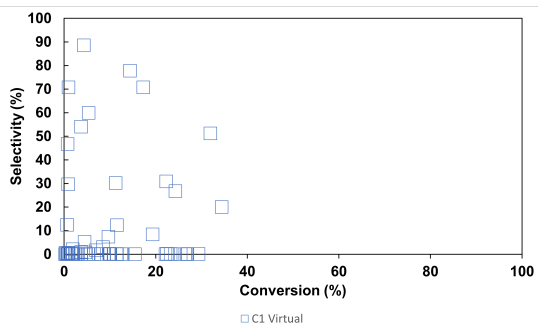


Figure 13: Result of the targeting step (step 5) for cluster 1 of case study 2.

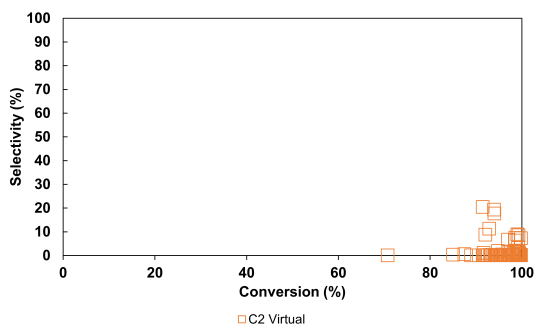


Figure 14: Result of the targeting step (step 5) for cluster 2 of case study 2.

Regarding cluster number three, one can say that the targeting step was developed with success. Not only the number of virtual catalysts was extended but also the real catalyst retained within the cluster with virtual catalyst performances being closer to the real one, as can be seen in Figure 15. So only this cluster can be considered relevant and will be the only one to be statistically studied from now on. Thus, the distributions of the descriptors on this cluster are presented as well from Figure 16 to Figure 18.

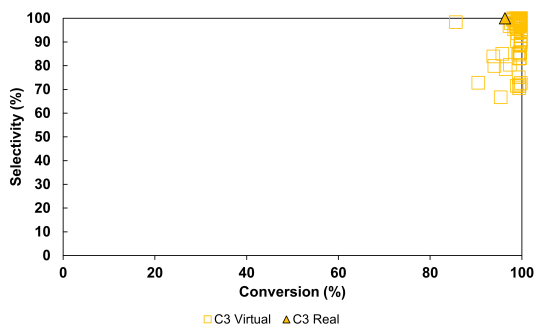


Figure 15: Result of the targeting step (step 5) for cluster 3 of case study 2.

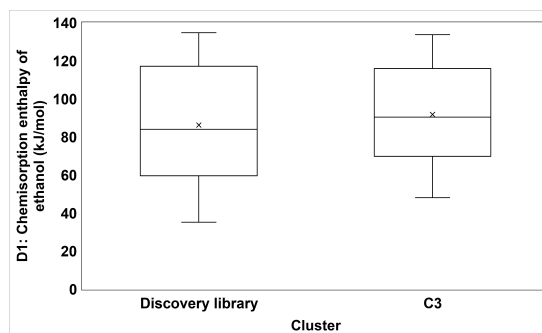


Figure 16: Comparison of the distributions of the chemisorption enthalpy of ethanol in the discovery library and target library of virtual catalysts for case study 2.

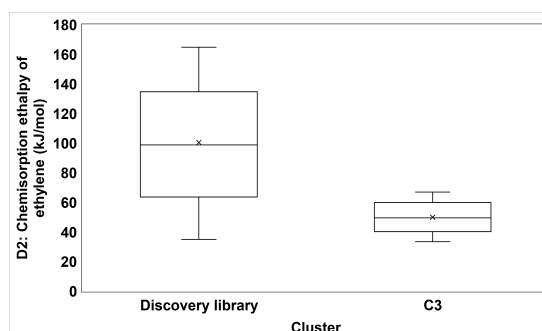


Figure 17: Comparison of the distributions of the chemisorption enthalpy of ethylene in the discovery library and target library of virtual catalysts for case study 2.

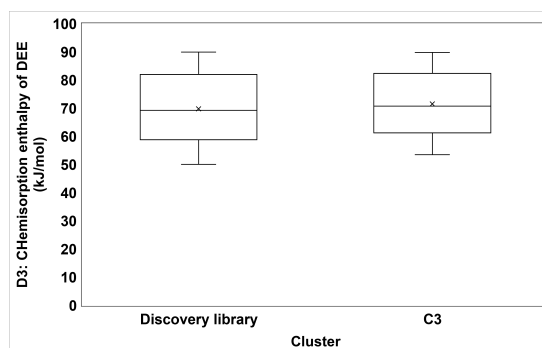


Figure 18: Comparison of the distributions of the chemisorption enthalpy of DEE in the discovery library and target library of virtual catalysts for case study 2.

Step 6 - Developing descriptor-property relationships from real catalysts properties

Firstly, through analysis of the previously presented boxplots, the hypothesis from case study 1, of lower ethylene chemisorption enthalpies being related to better catalyst performance is corroborated by the study of the second dataset as can be proven by Figure 18. Moreover, since both ZSM-5 doped with phosphorous and lanthanum present poor performance on this dataset, the presence of synergy between these two when catalysts are doped with both can be a possibility, due to the observed data from case study 1 as already previously reported [12]. Concerning the chemisorption

enthalpies of ethanol and DEE by taking a look at the distributions of these descriptors between the discovery library and the studied cluster, it is not possible to take any kind of conclusions, despite the fact they are statistically relevant. However, this reassures us that the chemisorption enthalpy of ethylene can be considered the most discriminating descriptor of the process in the simulated conditions. In the case of the third cluster, the only real catalyst present is a dealuminated ZSM-5 zeolite, being the one where 25% of oxalic acid was used during catalyst preparation. Since only this sample is present in the cluster and, on top of that, cluster 2 also has ZSM-5 zeolites that have been dealuminated but with higher percentages of oxalic acid, the assumption that a dealumination step during the catalyst synthesis process could influence the identified relevant descriptors and, in this case, the selectivity for ethylene of the process cannot be statistically proven. Thus, it is not possible to establish any kind of relationship between catalyst characteristics and the observed performance.

5. Conclusions

First of all, concerning the kinetic model development, this revealed to be hard to model due to numerical problems related to the very broad range of magnitude associated with the rate constants of the process. As a consequence, several different approaches needed to be considered and the usage of scaling factors was needed. Secondly, regarding the followed catalyst design technique, as it was already proven in previous works [3], the application of the methodology itself can be considered successful, representing an important tool in facing the biggest problems related to the kinetics-driven design of catalysts since it allows to create direct relations between the chemical properties and the results obtained on the statistical tests while being less expensive and time-consuming. Furthermore, even though it helps to have big sets of experimental data, it is proven by this work that it still works with smaller sets, being this a very promising feature to be taken into account for the future. However, some improvements still need to be made, especially in trying to use fewer tools so it is more integrated and efficient, as stated in [3]. When applied to ethanol dehydration to ethylene, for Case study 1, one can conclude that the chemisorption enthalpy of ethylene is the discriminating descriptor and that lower values for this are associated with better catalytic performance, while higher values for this are related to worse catalytic performance, especially in terms of selectivity. This is possibly connected to a reduction in the BET surface area and in the volume and width of the pores related to metal doping a catalyst, especially

with phosphorous in this case. On top of this, a big decrease in the total acid amount, especially in the acidic strength of the catalysts, which constitutes an important property for the studied process, is verified for the catalyst associated with the worst performances. This conclusion is based on three different clusters that were obtained, having two both virtual and real catalysts. On the other hand, it was possible to infer that lower values of the ethylene chemisorption enthalpy are associated with higher ethanol conversion and selectivity for ethylene. As for case study 2, the low number of real catalysts proved to be a setback for this study, leading to a non-successful targeting step and a lack of statistically relevant results that didn't allow taking conclusions. Thus, even though, it is possible to perform the methodology and reach conclusions and results with samples containing lower amounts of real catalysts, it can be concluded that a big dataset should always be used to make sure that the methodology is always successful. Overall, even though this constitutes a set of preliminary conclusions and more studies should be done to acquire more relationships and prove what was concluded until this point, the applicability of the methodology was successful with the chemisorption enthalpy of ethylene being discovered as the main factor influencing the performance of zeolites for the studies datasets. This appeared related to the acid strength of the catalyst and the surface area available for adsorption. Both of these properties can be manipulated by modifying catalyst structure mainly through metal doping processes and this should be taken into account for further studies, especially in design phases for real catalysts, so the ethanol dehydration process can overcome the issue associated with catalysts and reach its full potential in order to aid the environmental change that is needed nowadays.

References

- [1] H. Rahmanifard, N. Experience, S. Alpha, and K. Mascarenhas, "Economic impacts and market challenges for the methane to derivatives petrochemical sub-sector," March 2018.
- [2] I. Amghizar, L. A. Vandewalle, K. M. V. Geem, and G. B. Marin, "New trends in olefin production," *Engineering*, vol. 3, pp. 171–178, April 2017.
- [3] L. Pirro, P. S. Mendes, S. Paret, B. D. Vandegheuchte, G. B. Marin, and J. W. Thybaut, "Descriptor-property relationships in heterogeneous catalysis: Exploiting synergies between statistics and fundamental kinetic modelling," *Catalysis Science and Technology*, vol. 9, pp. 3109–3125, 2019.

- [4] H. Zimmermann and R. Walz, *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, April 2009, vol. 13.
- [5] H. Al-Megren and T. Xiao, *Petrochemical catalyst materials, processes, and emerging technologies*. IGI Global, 2016.
- [6] K. V. der Borght, "Omzetting van bio-ethanol naar koolwaterstoffen: van laboratorium tot industriële schaal," Ph.D. dissertation, Universiteit Ghent, 2015.
- [7] K. Alexopoulos, M. John, K. V. D. Borght, V. Galvita, M. F. Reyniers, and G. B. Marin, "Dft-based microkinetic modeling of ethanol dehydration in h-zsm-5," *Journal of Catalysis*, vol. 339, pp. 173–185, July 2016.
- [8] Understanding k-means clustering in machine learning — by dr.michael j.garbade — towards data science. Accessed on 10/05/2022. [Online]. Available: <https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1>
- [9] C. chau Chiu, G. N. Vayssilov, A. Genest, A. Borgna, and N. Rösch, "Predicting adsorption enthalpies on silicalite and hzsm-5: A benchmark study on dft strategies addressing dispersion interactions," *Journal of Computational Chemistry*, vol. 35, pp. 809–819, April 2014.
- [10] C. G. Pope, "Adsorption of methanol and related molecules on zeolite h-zsm-5 and silicalite," *Journal of the Chemical Society, Faraday Transactions*, vol. 89, pp. 1139–1141, January 1993.
- [11] J. Hunns, M. Arroyo, A. Lee, J. Escola, D. Serano, and K. Wilson, "Hierarchical mesoporous pd/zsm-5 for the selective catalytic hydrodeoxygenation of m-cresol to methylcyclohexane," *Catal. Sci. Technol.*, vol. 6, 12 2015.
- [12] N. Zhan, Y. Hu, H. Li, D. Yu, Y. Han, and H. Huang, "Lanthanum-phosphorous modified hzsm-5 catalysts in dehydration of ethanol to ethylene: A comparative analysis," *Catalysis Communications*, vol. 11, no. 7, pp. 633–637, 2010.
- [13] K. S. Abdulazizovna, Y. Y. Yusupboyevich, A. T. Dolimjonovich, and R. L. Sobirjonovna, "Adsorption energetics in zsm-5 zeolites," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 7, pp. 887–901, 2020.
- [14] Y.-H. Yeh, C. Rzepa, S. Rangarajan, and R. J. Gorte, "Influence of brønsted-acid and cation-exchange sites on ethene adsorption in zsm-5," *Microporous and Mesoporous Materials*, vol. 284, pp. 336–342, 2019.
- [15] M. M. Dubinin, G. U. Rakhmatkariev, and A. A. Isirikyan, "Heats of adsorption of methanol and ethanol on high-silicon ZSM-5 zeolite," *Bulletin of the Academy of Sciences of the USSR, Division of chemical science*, vol. 38, no. 11, pp. 2419–2421, Nov. 1989.
- [16] W. Buysse, "Ab initio microkinetic modelling of bio-alcohol dehydration in zeolites: Influence of carbon number and framework topology," Master's thesis, Universiteit Gent, 2016.
- [17] J. Sun, J. W. Thybaut, and G. B. Marin, "Microkinetics of methane oxidative coupling," *Catalysis Today*, vol. 137, no. 1, pp. 90–102, 2008, recent Developments in Combinatorial Catalysis Research and High-Throughput Technologies.
- [18] P. L. Houston, *Chemical kinetics and reaction dynamics*. Dover Publications, 2006, pp. 34–91.
- [19] J. A. Dumesic, D. F. Rudd, L. M. Aparicio, J. E. Rekoske, and A. A. Treviño, *The Microkinetics of Heterogeneous Catalysis*. American Chemical Society, 1993.
- [20] R. N. Forthofer, E. S. Lee, and M. Hernandez, "3 - descriptive methods," in *Biostatistics (Second Edition)*, second edition ed., R. N. Forthofer, E. S. Lee, and M. Hernandez, Eds. San Diego: Academic Press, 2007, pp. 21–69.
- [21] A. Styskalik, V. Vykoukal, L. Fusaro, C. Aprile, and D. P. Debecker, "Mildly acidic aluminosilicate catalysts for stable performance in ethanol dehydration," *Applied Catalysis B: Environmental*, vol. 271, p. 118926, 8 2020.
- [22] G. Yang, Y. Wang, D. Zhou, J. Zhuang, X. Liu, X. Han, and X. Bao, "On configuration of exchanged la³⁺ on zsm-5: A theoretical approach to the improvement in hydrothermal stability of la-modified zsm-5 zeolite," *The Journal of Chemical Physics*, vol. 119, no. 18, pp. 9765–9770, 2003.
- [23] C.-Y. Wu and H.-S. Wu, "Ethylene formation from ethanol dehydration using zsm-5 catalyst," *ACS Omega*, vol. 2, no. 8, pp. 4287–4296, 2017, PMID: 31457720.