

# The McCabe-Thiele method: a centenary tribute to an emblem of chemical engineering and engineering epistemology

Fabian Jirasek,<sup>†</sup> Johannes Lenhard,<sup>†,‡</sup> Jana Heiß,<sup>†</sup> and Hans Hasse<sup>\*,†</sup>

<sup>†</sup>*Laboratory of Engineering Thermodynamics (LTD), RPTU Kaiserslautern, Germany*

<sup>‡</sup>*Philosophy in Science and Engineering, RPTU Kaiserslautern, Germany*

E-mail: [hans.hasse@rptu.de](mailto:hans.hasse@rptu.de)

## Abstract

This essay is a tribute to the McCabe-Thiele method, which was first published 100 years ago in this journal. We argue that the McCabe-Thiele method has become an emblem of chemical engineering, as it establishes a balance that simultaneously enables *prediction* and *understanding*, making it also emblematic of engineering epistemology. The McCabe-Thiele method is analyzed here by going through its most important aspects: the historical background, the model assumptions and how these enable predictions and understanding, as well as the question of novelty. We argue that the remarkably long-lasting success of the McCabe-Thiele method is due to how it creates understanding through its graphical approach, and we identify elements of engineering epistemology that make it emblematic in a quite general sense. Finally, we briefly discuss the future role of the McCabe-Thiele method, considering that machine learning will change the way distillations are designed.

# Introduction

Ask any chemical engineer, young or old, from any country to name a method that could serve as an emblem for their field – and you are likely to get the same answer: the McCabe-Thiele method, a method for designing distillation processes. In 2025, we celebrate the 100th anniversary of this method – it was first presented at the 69th meeting of the American Chemical Society in Baltimore, April 6 - 10, 1925, and published in June that year in *Industrial and Engineering Chemistry*.<sup>1</sup>

Why is this method still held in such high esteem after 100 years? Why is it still taught to all chemical engineering students everywhere in the world? Why has this graphical method not gone the way of so many other constructions that disappeared with the rise of numerical methods and computers? Is there something to learn from its success beyond chemical engineering? We address these questions in the present essay. It is written to be understood without previous knowledge of distillation. However, we hope those familiar with the technical details of distillation will also find it interesting to look at a familiar topic from a new perspective. At the same time, addressing these questions leads us into the realm of epistemology. Also, in this regard, an analogous encouragement of our readers holds: The essay is written to be understood without previous knowledge of epistemology, but we hope that those who are familiar with the philosophical discussion will find it interesting to look at a familiar topic from a new perspective.

## The McCabe-Thiele method

### The stage and the actors

Chemical engineering has roots in the second half of the 19th century and was still a young and rapidly developing discipline in the 1920s. In particular, it has long been debated whether chemical engineering belongs under the umbrella of chemistry or should be established as

something institutionally separate. Two trailblazers for chemical engineering were William H. Walker, a student of German chemist Walter Nernst, and Arthur D. Little, the founder of an engineering consultancy. In a talk at MIT in 1905, they introduced unit operations as building blocks of all chemical processes (distillation is one of them) and elevated the unit operation to the rank of a scientific subject. Many consider this to be the birth of chemical engineering.<sup>2-4</sup> Walker established a new teaching program at MIT in 1907, and he was one of the central figures in founding the “American Institute of Chemical Engineering” (AIChE) in 1908 – against resistance from the American Chemical Society. (The early history of the American Institute of Chemical Engineers was dominated by two problems – establishing what exactly chemical engineering was and establishing the legitimacy of chemical engineering as an independent discipline<sup>3</sup>). Around 1920, the first general textbooks on chemical engineering were published: the “Principles of Chemical Engineering”<sup>5</sup> and the “Handbook of Chemical Engineering”.<sup>6</sup>

This is the point where our protagonists, Warren L. McCabe (1899 – 1982) and Ernest W. Thiele (1899 – 1982), enter the stage. Let us take a brief look at their biographies. The pictures in Figure 1 show them in their later years, long after they invented their famous method.



Warren L. McCabe  
1899 – 1982



Ernest W. Thiele  
1895 – 1993

Figure 1: The inventors of the McCabe-Thiele method in their later years. They were still in their twenties when they invented their famous method.

Ernest Thiele earned his Bachelor's degree from Loyola University in Chicago and then went to MIT for his Master's and PhD. He developed the McCabe-Thiele method while waiting for his dissertation in reaction engineering to be reviewed.<sup>7</sup> He is also well-known for the Thiele modulus, which measures the relationship between reaction and transport rates.<sup>8</sup> He retired as associate director of research after 35 years at Standard Oil (now BP; see Kohn<sup>7</sup> for a more comprehensive biographical appraisal).

Warren McCabe graduated in chemical engineering at the University of Michigan (BSc 1923, MSc 1924) and then went to MIT, where he met Thiele. He became the co-inventor of this seminal method having hardly started his doctoral studies (PhD in 1928). While Thiele's career was firmly based in industrial research (even though he became a professor at Notre Dame after his retirement), McCabe held professorships at numerous prestigious universities, including MIT. He is also famous for his monograph with Smith "Unit Operations of Chemical Engineering",<sup>9</sup> and held leading positions in AIChE), including the presidency.<sup>10</sup>

While both had distinguished careers, neither Thiele nor McCabe had completed their doctorates when they invented the McCabe-Thiele method, which for both of them was different from the subject of their doctoral theses.

Interestingly, the McCabe-Thiele method was first presented at a meeting of the American Chemical Society (ACS) (April 6 - 10, 1925) and not at an AIChE meeting, which might have been a more natural choice. (The AIChE was already founded in 1908 and held annual meetings too). The reason for this may have been that in the 1920s, presentations at the AIChE meetings required publishing a corresponding paper in proceedings (Bulletin) - which, however, appeared regularly only with a long delay. It was not until 1944 that AIChE responded and separated presentation and publication.<sup>11</sup> The monthly publications of ACS plausibly persuaded McCabe and Thiele. However, we emphasize that this is a speculation.

## Design and understanding

Distillation is the most important technique for separating mixtures and is widely used in the chemical industry. The McCabe-Thiele method is often introduced as a method for designing distillation columns - starting with the title McCabe and Thiele have chosen for their famous paper:<sup>1</sup> “*Graphical Design of Fractionating Columns*”. However, after half a century, during which chemical engineers used the graphical method of McCabe and Thiele to design distillation columns, they switched to numerical methods when computers became accessible in the 1970s. The fact that the McCabe-Thiele method is still considered to be important enough to be taught to all chemical engineering students worldwide to date shows that it must be more than a design method: It gives those who master it insights into distillation – it creates a particular kind of understanding. The McCabe-Thiele method as a design method has limitations, the most important one is that it works only for binary mixtures, whereas in most industrial distillations, mixtures with more than two components have to be separated. On the other hand, focusing on a simple case, here binary systems, can be helpful when it comes to creating a basic understanding of the working principles of processes.

Thus, the McCabe-Thiele method is valuable in different respects: it contributes to *what* engineers know and *how* they know it. It makes *predictions* for design *and* provides *understanding*. Both aspects are interrelated and based on insightful process modeling. How modeling, prediction, and understanding are intertwined is a seminal characteristic of engineering epistemology. Our paper contributes to an evolving discussion.<sup>12,13</sup>

## The topic is complex

Developing an understanding of a complex process like distillation and methods for their design is far from being trivial.<sup>14–18</sup> The physical processes in distillation columns are highly intricate: a rising gas phase and a liquid phase that flows down are brought into close contact using so-called column internals. Generations of engineers have worked on developing

and optimizing column internals, and a wide variety of internals of different types and performance is offered by competing suppliers, the two main types being trays and packings (structured or random). The counter-current flow of the gas and the liquid in the column is created by evaporating a part of the bottom product, condensing a part of the top product, and feeding these streams back to the column. The condensate stream fed back to the column is called reflux, and the reflux ratio  $R$  (the mass flow of the reflux divided by the mass of the top product) is an important operating parameter of the distillation column. High reflux ratios are good for the separation. The internal streams are not only influenced by the reflux but also by the feed that is added at a certain position to the column. Of course, the properties of the mixture to be separated are of paramount importance, particularly the (relative) volatility of its components, i.e., their tendency to accumulate in the gas phase. Another important parameter is the column pressure. All the aspects mentioned above are closely entwined, and their interactions need to be taken into account for a sound design and an adequate understanding.

Even today, despite all advances in simulations of fluid flow and fluid property modeling, it is still not possible to design a distillation apparatus based on a realistic representation of the fluid dynamics in the column. Hence, simpler models have to be used, representing compromises between conflicting goals, such as scope, accuracy, epistemic value, and, last but not least, feasibility. Finding such a compromise was precisely the challenge McCabe and Thiele have tackled.

## **A short description**

The McCabe-Thiele method is summarized so well in the abstract of the original paper<sup>1</sup> that we cite it here in full length:

*“A rapid and accurate graphical method is described for computing the theoretical number of plates required in a column for performing a given separation of a binary mixture. The method consists essentially in drawing, on the same rectan-*

*gular plot, the equilibrium curve for vapor and liquid compositions and straight lines representing the equations for enrichment from plate to plate, and passing from one to the other in a series of steps. It is shown that the effect of varying feed, product, waste, and overflow, as well as the effect of varying temperature of feed, may be represented by very simple geometrical constructions. The application of the method to some more complicated problems is discussed.”*

Figure 2 illustrates the McCabe-Thiele construction. We will not explain here how exactly the construction works, this is nicely done in the original paper of McCabe and Thiele<sup>1</sup> and in basically any textbook on fluid separation processes written after 1925, see Refs.<sup>18-20</sup> for some examples.

The graphical construction of McCabe and Thiele is easy to carry out – everything is done in a single diagram. In that diagram, the gas phase concentration of the light boiling component  $y_1$  is plotted over its liquid phase concentration  $x_1$ . We use mole fractions here for simplicity; other options are feasible. The basic elements in that diagram are the equilibrium line, the operating lines for the rectifying section and the stripping section, and the feed line. The equilibrium line EL describes the vapor-liquid equilibrium at the pressure of the column, i.e., the relevant fluid properties of the mixture. The operating lines represent the material balances in the two column sections above and below the feed and the feed line that captures the influence of the composition and the state of the feed. The zig-zag construction yields the number of theoretical stages  $N$  below and above the feed required for a given separation, as specified by the compositions of the top and bottom product,  $x_1^{\text{top}}$  and  $x_1^{\text{bottom}}$ . The influence of the reflux ratio  $R$  is also captured, as it influences the slope of the operating lines. The number  $N$  of theoretical stages is the basic result of the McCabe-Thiele construction.  $N$  is central to the column design as it is directly related to the height of the column.

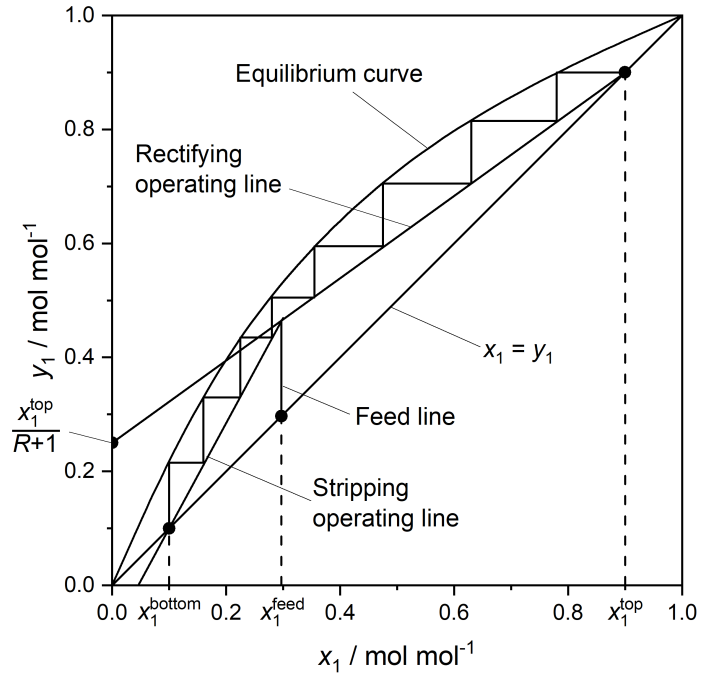


Figure 2: McCabe-Thiele construction with its main elements: the equilibrium line (for the operating pressure of the column), the operating lines for the rectifying section and the stripping section, and the feed line (here: for the case of a liquid boiling feed). The diagram shows the concentration  $y_1$  of the light boiling component in the gas phase as a function of its concentration in the liquid phase  $x_1$  (we use mole fractions here, for simplicity). Also, the diagonal  $x_1 = y_1$  is indicated. The composition of the bottom product, the top product, and the feed are specified by  $x_1^{\text{bottom}}$ ,  $x_1^{\text{top}}$ , and  $x_1^{\text{feed}}$ , respectively. The construction yields the number of theoretical stages, which is  $N = 8$  here. The feed is located on stage 4 (counted from the bottom).

## What was new?

The basis of the McCabe-Thiele method is the equilibrium-stage model. Assuming equilibrium conditions was decisive for being able to build a quantitative, mathematized model of distillation that remained tractable. Note that in distillation, at no point equilibrium is achieved - if it were, separation would come to an end. Realizing that such a process can be efficiently modeled using the equilibrium assumption is a significant creative step. McCabe and Thiele give credit for introducing the equilibrium-stage model to Sorel.<sup>21</sup> In his series of papers on distillation,<sup>22,23</sup> Sorel describes a rational method for the design of distillation columns that is based on mass balances, the energy balance, and the equilibrium condition. In fact, Sorel only implicitly mentions the equilibrium condition. Hausbrand<sup>24</sup> 1893 more clearly sets out a similar approach. By the way, both dealt with alcohol-water systems – the only mixtures for which equilibrium data were then available.<sup>2</sup> For a wider discussion of rationality and mathematical tools in the context of 19th-century engineering, see Johnson and Lenhard.<sup>13</sup>

McCabe and Thiele add an important twist to these idealizations. Rather than considering the energy balance, they use the constant molar overflow assumption, i.e., constant flow rates of liquid and vapor throughout the column, leading to straight operating lines, see Figure 2. We do not go into the technical details of this assumption and discuss only some general aspects: Firstly, using the energy balance as a starting point creates practical problems, as very often, only little is known about the caloric properties of the studied mixtures. The caloric properties must then be estimated, and the uncertainty introduced by this may be of the same order as the errors resulting from the simplifying assumption used by McCabe and Thiele. It should also be noted that the constant molar overflow assumption was not new in 1925; it had been used previously by, e.g., Lewis.<sup>25</sup> The novelty of the McCabe-Thiele method was, hence, basically the judicious combination of elements that had practically all been on the table before. The result was nevertheless ground-breaking: a clear-cut, simple, and practically useful method that strongly supports the understanding of distillation.

We also note that more accurate graphical methods were available before 1925, namely the method of Ponchon-Savarit,<sup>26,27</sup> which, in contrast to the McCabe-Thiele method, accounts for the energy balance, however, at the expense of introducing a second diagram, in which the caloric properties are represented, which makes it more tedious to apply. While, in the end, a slightly more accurate result may be obtained using the Ponchon-Savarit construction compared to the McCabe-Thiele construction, the ease of the graphical interpretation is lost and, thereby, a major epistemic value.

In the introduction of their original paper,<sup>1</sup> McCabe and Thiele nicely summarize and reflect upon previous efforts to develop useful approaches for working with the equilibrium-stage model:

*“Sorel’s method, though sound, is tedious when applied analytically. Many writers have simplified or modified Sorel’s method, among them being Lewis, Leslie, Peters, Van Nys, and Gay, who have all used analytical methods... Graphical methods have been presented by Ponchon, Rodebush, and Savarit. However, it is felt that the following method is simpler, exhibits its results in a plainer manner than any method, analytical or graphical, so far proposed, and is accurate enough for all practical use.”*

This is an example of a clear and modest statement. At the same time, it contains important features of engineering epistemology: The benefits of a method may be meager if it is too cumbersome to use. And: while idealization is needed, it must be oriented at a result that is “accurate enough”. The claim is that the new method is simple and plain and still practically useful - nothing more and nothing less. The authors also justify the simplifications they introduce very clearly. Speaking about the adjustable parameters plate efficiency  $\eta$  and number of theoretical stages per meter NTSM, McCabe and Thiele state:<sup>1</sup>

*“Since they [i.e. the parameters  $\eta$  and NTSM] have not been accurately determined for all types of columns or for all mixtures, there is little practical value in*

*determining [the number of] theoretical plates [stages  $N$ ] with great precision.”*

In other words: the justification of their simplifying assumptions is rooted in practice. A more detailed model might give a somewhat better result for the number of theoretical stages  $N$ , but this is irrelevant because the uncertainties in parameter adjustment to empirical data would override the gain in precision. This type of justification via relevance in practice is typical for the complexities of engineering epistemology.

## **What made it so successful and why it still is?**

Why did this construction gain such a prominent role? We cannot single out a single cause; instead, we identify factors contributing to its success. Firstly, despite its simplicity, it accounts for the essential features of distillation. We go a step further and argue that the McCabe-Thiele construction has attained such a status that it is commonly accepted that the features it accounts for are the essential ones, whereas all others are secondary. This does not mean that they are not important, but rather that they can be accounted for in subsequent steps. In particular, this holds for all fluid dynamic considerations and everything related to caloric properties. In other words, the method has shaped the picture chemical engineers have of distillation.

Secondly, the output of the method, the number of theoretical stages  $N$ , leaves room for interpretation. It is so flexible that it can be mapped to any type of column internals. However, this requires using empirical parameters: for tray columns, the tray efficiency  $\eta$  is used, and for packed columns, the number of theoretical stages per meter NTSM. The numbers for  $\eta$  or NTSM must be determined for each studied internal in test measurements (with test mixtures), and manufacturers of internals supply recommended values for their products. They may depend on operating conditions, such as the gas load (volume flow of the gas divided by the cross-sectional area of the column, i.e., a velocity), and sometimes different values are used for different classes of mixtures. Even though there is no guarantee that the recommended values (obtained from the studies of test systems) will hold for a

given application, experience shows that mapping the number of theoretical stages to the column design is sufficiently robust. Thus, the adjustable parameters  $\eta$  and NTSM mediate between the method and the concrete case. For a general discussion of the role of adjustable parameters in physical modeling, we refer the reader to Hasse and Lenhard.<sup>28</sup>

Thirdly, the method allows to study and illustrate the influence of the main design parameters on the distillation process, e.g., the product specifications or the reflux ratio. Furthermore, the method covers important limiting cases. Here are some examples: it is *immediately clear* from Figure 2 that if a product specification approaches a pure component, the number of stages  $N$  has to go to infinity. Furthermore, if the reflux ratio  $R$  goes to infinity (e.g., practically nothing is withdrawn from the column), the number of stages becomes minimal  $N = N_{\min}$  (and the operating lines coincide with the diagonal  $x_1 = y_1$ ). On the other hand, there is a minimal value of the reflux ratio  $R_{\min}$  that is required to solve a given separation task.

Furthermore, and fourthly, the influence of the fluid properties can be easily assessed. The closer the equilibrium line comes to the diagonal  $x_1 = y_1$ , the more difficult the separation gets (the more stages will be needed) – and if the equilibrium line intersects the diagonal (i.e., for azeotropic mixtures), no separation beyond the point of the intersection (the azeotropic point) is possible.

The strength of the McCabe-Thiele method and the key to its success is that all these relations (and many others we have not discussed here for brevity) can be grasped through the graphical representation. If we accept this graphical representation as an adequate model, the relationships and connections just mentioned arise from the properties of the graphic. We argue that the ideas of *understanding* and *immediacy* are closely related to the fact that they are grasped from the image. This is why the McCabe-Thiele diagram quickly became and still is the workhorse for explaining distillation - which is quite remarkable, as the McCabe-Thiele construction is based on a model of distillation with many simplifying assumptions. Based on the foundations laid by McCabe and Thiele, many extensions of their

method to more complex distillation processes have been developed.<sup>18</sup>

The McCabe-Thiele method is a nice example of pragmatic understanding in engineering. As a side note, we mention here that, despite the significance of pragmatic understanding for engineering epistemology, the corresponding philosophical discussion is highly underdeveloped, for exceptions see Ferguson<sup>29</sup> and Lenhard.<sup>30</sup>

## Engineering epistemology

The McCabe-Thiele method is an emblem of chemical engineering. Significant contributing elements can also be discussed in more general terms, making the method instructive for engineering epistemology in general. This section briefly discusses these elements and puts them into a philosophical context.

### Idealization

Practically all models in science and engineering are based on idealizations. One can distinguish different types of idealization depending on how the idea or the ideal is conceived. We want to focus on one particular distinction, which we call *objective* versus *pragmatic*. Objective idealization refers to the structure of the world. According to this point of view, every object in the world has many properties, and idealization in model building means constructing an image that leaves out irrelevant properties and retains only the relevant ones. Objective idealization is regularly thought to steer models away from noisy complexities and bring them closer to the true fundamentals. The philosophical point is that this kind of idealization still maintains a connection to the truth since it deals with properties that the object actually has. Objective idealization, in this sense, resembles the process of abstraction according to Aristotle.

In an influential paper, McMullin<sup>31</sup> contrasted “Aristotelian idealization” with a “Galilean idealization”. The latter introduces deliberative distortions, like point masses moving with-

out friction or omniscient agents assigning prices on a market. Using Galileo as a source of inspiration for the name is apt because Galileo mathematized movements in physics. Understandably, he had to invent models for this purpose that were tractable with the mathematical tools available to him, i.e., geometry. The most famous example is his law of falling bodies, which he was able to formulate because he idealized away friction. Thus, the model works with properties that are different from those of the target the model should capture. Truth is not preserved, deliberately so. The philosophical point is that modeling is after a balance of representation and tractability. The model is not created as an image but rather as a means to do something, like predicting the movement of a falling body. Therefore, we call this type of idealization pragmatic. To be relevant, the model must be tractable. Galileo's ingenious achievement was finding a mathematization to work with and to demonstrate something. The question of how well the model actually fit was of secondary importance. However, this question is of primary importance for the process of designing a distillation column. The McCabe-Thiele method uses pragmatic idealizations (equilibrium assumption, constant molar overflow). The point is to find the right dosage for idealization, thus building a model that is tractable and accurate enough at the same time.

## Representation and prediction

Representation belongs to the basic rationale of modeling. Typically, it is closely intertwined with prediction. A classic formulation is the one by German physicist Heinrich Hertz:<sup>32</sup>

*“The most direct, and in a sense the most important problem which our conscious knowledge of nature should enable us to solve is the anticipation of future events, so that we may arrange our present affairs in accordance with such anticipation. As a basis for the solution of this problem we always make use of our knowledge of events which have already occurred, obtained by chance observation or by pre-arranged experiment. In endeavouring thus to draw inferences as to the future from the past, we always adopt the following process. We form for ourselves im-*

*ages or symbols of external objects; and the form which we give them is such that the necessary consequents of the images in thought are always the images of the necessary consequents in nature of the things pictured.”*

Hertz says: the logic in the model needs to represent the logic in the studied object in a way that is good enough to enable predictions. However, it is not at all clear how well and how accurately a model has to represent in order to achieve this. The McCabe-Thiele method is a prime example of a strong idealization that yields, nevertheless, good predictions. It is also an example of an important related point: successful tractable models set standards for what “good enough” means.

Considerations about what idealizations achieve have led to a controversy about the sense in which a model represents. If a model presents an idealized image of its target, does that mean there is a stepwise procedure of idealization that leads from the real world to the model? Following this procedure, would it be possible to go back from the model to the world through a step-by-step de-idealization? If this back-and-forth would work, then the simplifications and idealizations would generally be provisional in the sense that they can be eliminated later. This standpoint has, e.g., been taken by McMullin,<sup>31</sup> who argued that Galilean idealizations were distortive idealizations but also controlled idealizations in the sense that one could get rid of the distortions by de-idealizing, successively removing distorting assumptions. However, philosophers like Batterman and Rice<sup>33,34</sup> discuss distortive idealizations that cannot be removed without ruining the model altogether (on that discussion, see the encyclopedia entry on modeling by Frigg and Hartmann<sup>35</sup> for an overview).

The case of McCabe-Thiele shows that a nuanced picture is desirable. On the one hand, one can argue that the McCabe-Thiele method stands in a line between a de-idealized model, in which the constant molar overflow assumption is replaced by the energy balance (as it is done in state-of-the-art process simulation) and more idealized models that, e.g., additionally assume that the phase equilibrium can be described by a constant relative volatility, as it is done in the formulae of Fenske and Underwood for calculating the minimal reflux ratio

$R_{\min}$  and the minimal number of theoretical stages  $N_{\min}$ .<sup>36,37</sup> On the other hand, it is not fruitful to try to apply the concept of de-idealization to the equilibrium-stage model itself. One could formally argue that an equilibrium stage can be de-idealized by explicitly taking mass transfer kinetics into account, as it is done in the so-called rate-based models,<sup>38</sup> but this disregards that these model classes fundamentally differ in how they are applied: if a rate-based model were just a de-idealized version of the equilibrium-stage model, one would assume that the stage efficiency or the NTSM value would be calculated from the rate-based model and the result would then be plugged into the equilibrium-stage model - which is not how it is done. The two model classes are separate entities with different inputs and outputs and only some potential overlaps in the fluid property model. Here, the concept of de-idealization of McMullin does not apply. Still, one can construct models that are more “realistic” as this is often called. When models are made more realistic, further assumptions and conditions are added, typically with the help of parameterizations (as is done in the rate-based models). These then relate to aspects that were previously not represented explicitly. (Note that the additional details described by rate-based models are implicitly included in the equilibrium-stage model by using  $\eta$  or NTSM as adjustable parameters.) However, it is questionable to interpret the rate-based model as a more realistic image of reality than the equilibrium-stage model; it is rather a different type of idealization containing more detail and more parameters that need to be adjusted.

This leads us to the predominant role of parameters in engineering models, see Vincenti,<sup>12</sup> who has pointed out that experimentation and parameter adjustment are pivotal in engineering epistemology. Modeling essentially involves a feedback loop in which (ideal and, at the same time, false) assumptions are explored and in which parameterizations are set in such a way that the model behavior becomes acceptable. At the *end* of this feedback process, Hertz’s description is correct again. However, it is essential to include the phases of exploration and experimentation that take place during the modeling process. Thus, we observe a tight integration of parameterization, adjustment, and predictive capacity.<sup>28</sup>

## Distillation process design today

Distillation design has been based on numerical process simulation for about half a century now. Nobody uses the McCabe-Thiele method for design anymore, but it is still in all curricula because of its epistemic value.

Most of the simulations of distillation processes, both in industry and academia, are still carried out based on the equilibrium-stage model. We are looking forward to celebrating the 150th anniversary of the equilibrium-stage model, and its inventor Ernest Sorel, in 2038! As stated above for the McCabe-Thiele method, the key to the success of this approach is the simple link of the theoretical results to practice using the stage efficiency  $\eta$  or NTSM, which can be carried out sequentially after the numerical design. The balance between rigorous theory and pragmatic adaption established in this approach is hard to beat.

Unsurprisingly, numerical simulations based on the equilibrium-stage concept have been used in many sophisticated applications in academia and industry, often requiring extensions of the basic concepts. We mention here only simulations of reactive distillations and dividing wall columns as examples.<sup>39–42</sup> Much work has also gone into the numerical solution of the equations of the equilibrium-stage model.<sup>43,44</sup> and the development of advanced methods for carrying out optimizations based on these models, which are central to process design.<sup>45,46</sup> There is also an important body of work in which the equilibrium-stage model is used in conceptual distillation process design, see, e.g., Refs.<sup>47–51</sup> Among these are the highly developed and very successful methods for the analysis of the feasibility of distillation of ternary mixtures based on distillation line diagrams.<sup>48,52–56</sup> They follow a similar basic approach as it was put into practice by McCabe and Thiele: in both approaches, the energy balance is not considered (although this is done differently), which greatly simplifies the analysis and allows to focus on the essentials and a graphical representation gives clout to the results. However, it has to be noted that many of these advanced methods for distillation process design are still academic because of the important hurdles that have to be overcome before their benefits can be reaped in practice. On the other hand, there are also encouraging

counter-examples in which this has succeeded, see, e.g., Refs. <sup>57-59</sup>

There is also a rich body of work in the literature dealing with the development of models that provide a more detailed picture of distillation than the equilibrium-stage model, in particular, the so-called rate-based models,<sup>60-62</sup> in which the transfer processes in the column are explicitly modeled - and not lumped into  $\eta$  or NTSM as in the equilibrium-stage model. Unfortunately, this involves using other crude simplifying assumptions (as those in the film theory of mass transfer<sup>63</sup>) and introducing other adjustable parameters in the model (such as the film thickness). These models have merits in certain applications, e.g., in adsorption process design, but the benefits often do not outweigh the higher effort for the parametrization, the higher challenges of the numerical solution, and the higher opacity of the complex models in distillation process design. Many of the drawbacks of rate-based models could, in principle, be overcome by CFD models giving a realistic description of the fluid mechanics in distillation columns.<sup>64</sup> However, due to the extreme complexity and size of the problem, this seems unfeasible for practical design for the foreseeable future.

## Where do we go from here?

For the reasons outlined above, we believe that the equilibrium-stage model will remain the workhorse for the industrial design of distillation columns for quite some time to come. While the diverse routes of academic research in this field certainly allow for new work, the approaches outlined above can generally be described as mature. Important changes are coming from other quarters, some of which are imminent. These changes are being triggered by the rapid advances in machine learning (ML). Two issues stand out and will be briefly discussed in the following before we come to the role of the McCabe-Thiele method in all this.

## ML-based fluid property models

The first issue concerns the fluid property models - a central element of distillation design. The experimental data to parameterize the required fluid property models is often missing. The gaps are then typically filled using estimates from established group-contribution methods, but the uncertainties associated with this are a constant source of concern in distillation process design.<sup>65,66</sup> Measuring all necessary data at all relevant conditions is simply infeasible, which is why predicting fluid properties of unknown mixtures will remain important in the future. However, the methods available for this are about to change: the most important method for estimating activity coefficients in liquid mixtures (the central property for the calculation of vapor-liquid equilibria) has been UNIFAC.<sup>67,68</sup> This important method has been improved only incrementally for many years and regular updates are only available commercially. The application of matrix completion methods from ML<sup>69-76</sup> has recently changed this and enabled a complete overhaul of the parameters, which also included filling all gaps in the parameter tables (before, about 50 % of the entries were missing) and has also led to an important improvement of the predictions. The new model is called UNIFAC 2.0,<sup>77,78</sup> it is fully disclosed, and easy to implement in process simulators, as only the parameters need to be changed. This makes the method highly practical. However, UNIFAC 2.0 is only a starting point: a brand-new generation of ML-based fluid property models, which only need the structural formulae of the components as input, has very recently become available, such as GRAPPA for predicting pure-component vapor pressures<sup>79</sup> and HANNA for predicting activity coefficients in mixtures.<sup>80</sup> These models have unprecedented scope and accuracy. Furthermore, in contrast to widespread previous practice, the codes and parameterizations of these models have been fully disclosed and their application by users is actively supported.<sup>81</sup> These methods get their clout from the hybridization of ML methods with physical knowledge.<sup>82-84</sup> For example, HANNA<sup>80</sup> is a hard-constraint neural network that incorporates all relevant thermodynamic consistency criteria and ensures that the results strictly fulfill all relevant physical laws. The implementation of such models in process

simulation packages may require working with surrogate models to make the simulations technically feasible, but the route for integrating these new models in distillation process design is now open.

## Large language models in distillation process design

Large language models (LLMs) are extremely versatile tools currently developing at a breathtaking pace. The reasoning LLMs that have become available early in 2025, such as OpenAI's o3, can solve non-trivial engineering problems. In a recent test, we have given a regular exam in engineering thermodynamics, which is a high hurdle for most students, to OpenAI's o3 as a zero-shot prompt, i.e., without any additional input, and it solved all problems almost flawlessly - better than all students who took the exam.<sup>85</sup>

Encouraged by this, we have carried out preliminary tests to determine whether o3 also masters the McCabe-Thiele method. To this end, we have provided textbook-style problems on distillation design using the McCabe-Thiele method to o3. The quality of the answers depended on how the problem was formulated, but with suitable prompts, correct constructions in the McCabe-Thiele diagram were obtained. An example for the system (acetone + isopropanol) is shown in Figure 3. We only mention here that in these tests, we did not specify the fluid property data (the equilibrium line), so o3 had to search for the corresponding information, which it did successfully.

Does this mean that o3 understands the McCabe-Thiele method? And if so, following the widespread argument that those who understand the McCabe-Thiele method also understand distillation: does o3 understand distillation? These questions relate to the current and controversial discussion in the literature on the intellectual capabilities of LLMs.<sup>86-88</sup> LLMs successfully pass well-established tests designed to assess human understanding, which prompts a reconsideration of the very concept of understanding – both in the context of machines and of humans.

How will the rapid development of AI tools, namely LLMs, affect the practice of distil-

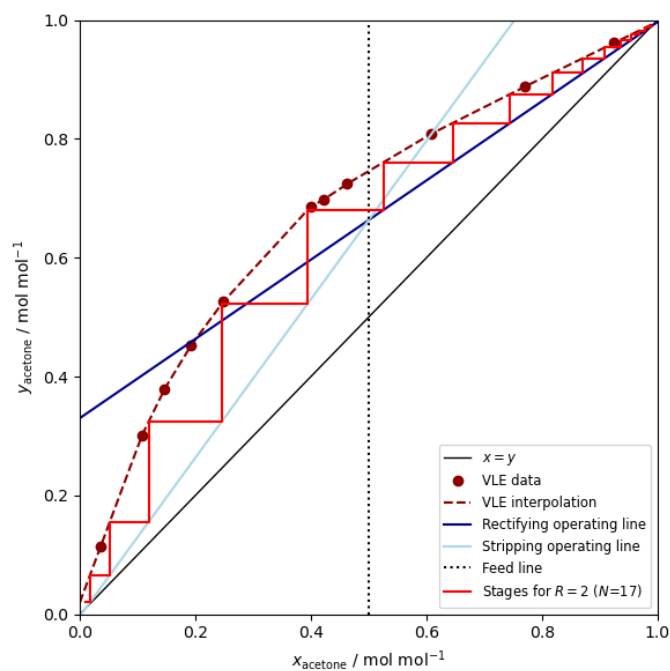


Figure 3: McCabe-Thiele plot obtained with OpenAI’s reasoning LLM o3 in June 2025, prompting a textbook-style problem. The compositions of the top product, the bottom product, and the feed (boiling liquid) were specified. The task was to determine the number of stages  $N$  for a reflux ratio of  $R = 2$ . The pressure was 1 bar. The LLM autonomously retrieved reasonable equilibrium data for the system (acetone + isopropanol) from the literature.

lation process design? We believe that standard tasks may be solved in the near future by successors of LLMs like o3, which will require LLMs to interact with process simulators and data banks (which is basically already feasible now). The key issue here is whether or not the results can be trusted. This question is not really new; it is, in principle, always on the table when decisions have to be taken based on information provided by others. We are experienced (although not always right) in assessing the capabilities of humans with whom we work together, but we are only starting to learn to do the same with ML systems. A tremendous problem here is the pace at which the capabilities of LLMs are currently evolving. What you learn today about a certain version can be outdated tomorrow when the next version is released. However, this rapid development may also open up many new opportunities in distillation process design. Many advanced methods developed in academia for solving tasks such as conceptual design or optimization are rarely used in industrial engineering. This is because their application generally requires special tools and intensive training, which only makes sense if the method is applied routinely rather than occasionally, as is often the case with specialized methods. LLMs could potentially take over such tasks in the future, thereby enhancing the design capabilities of companies.

All this has enormous consequences for the training of future engineers. Training operational capabilities will become less important as LLMs can take over many tasks. The engineers will have to be trained to co-operate efficiently with LLMs, prompt them in a suitable way, and assess their results. Will McCabe and Thiele continue to have a place in this rapidly changing world? We think that the answer is a resounding yes! The McCabe-Thiele method provides insights into the inner workings of distillation – and such insights are more important than ever!

## Conclusions

We have argued that the major achievement of McCabe and Thiele consisted in assembling a *package* that balances idealization, tractability, and usefulness. This makes the McCabe-Thiele method a hallmark of chemical engineering. We think that, moreover, such an artful combination is a hallmark of engineering epistemology.

This package contains, as we have analyzed, both prediction and understanding. The prediction part is a must because it is essential for design. McCabe and Thiele have based their method on an artful combination of assumptions to make a prediction workable. Such a package re-defines or re-adjusts fundamental conceptions, like what counts as useful, what is expected from a design method, or what sort of expertise an engineer should have. To a degree that even *defines* what distillation is, in the sense of what chemical engineers agree upon when they work on distillation. This alludes to historian-philosopher Thomas Kuhn,<sup>89</sup> who argued that scientific paradigms come as a package: they provide the framework for formulating problems and solving them. A new paradigm, in the words of Kuhn, constitutes a new “world view”. It is this view of distillation that is taught to the students when introducing distillation using the McCabe-Thiele construction.

Essential for the enduring popularity of the McCabe-Thiele method is that understanding is also part of the package. They give a fine example of how graphical or visual understanding can be utilized in chemical engineering. The understanding part is also why the McCabe-Thiele “package” did not become obsolete when computer methods for distillation design were introduced in the 1970s.

The immediate success of using computers for distillation design was largely due to the fact that the (numerical) solution of the equations of the equilibrium-stage model was now possible also for systems with any number of components. Arguably, the strongly increasing computational capacities enlarged the realm of tractability in a way that made simulation modeling a new type of mathematical modeling.<sup>30</sup> Caution is advised, however. The downside of the numerical solution is that it creates basically no insights. If no numerical solution

is found, it is notoriously difficult to decide whether this is the case because no solution exists or only numerical problems occurred. Insights can still be gained by a judicious representation of the simulation results, interestingly involving generally graphical representations. Computer methods have given rise to a new “culture of prediction”.<sup>13</sup> The recent advent of Machine Learning seems to further boost predictive capacities at the cost of understanding. Will this tear apart prediction and understanding? We think that it will make methods that intertwine the two, such as the McCabe-Thiele method, even more valuable in the future.

## Acknowledgement

H.H. is grateful for the inspiring discussions on distillation with his colleagues Heinz Auer, Bernd Bessling, Sergej Blagov from BASF Ludwigshafen, and Andrzej Gorak from TU Dortmund. J.L. thanks historians of science and technology Ernst Homburg (U Maastricht) and Carsten Reinhardt (U Bielefeld) for valuable hints as well as the German Research Foundation (DFG) for funding his research (grant no. LE 1401/9-1). F.J. gratefully acknowledges support by the DFG in the frame of its Emmy Noether program (grant no. 528649696).

## References

- (1) McCabe, W. L.; Thiele, E. W. Graphical Design of Fractionating Columns. *Ind. Eng. Chem.* **1925**, *17*, 605–611.
- (2) Buchheim, G.; Sonnemann, R. *Geschichte der Technikwissenschaften*; Birkhäuser Basel, 1990.
- (3) Reynolds, T. S. *75 Years of Progress: A History of the American Institute of Chemical Engineers, 1908-1983*; American Institute of Chemical Engineers.: New York, 1983.
- (4) Divall, C.; Johnston, S. F. *Scaling Up. The Institution of Chemical Engineers and the Rise of a New Profession*; Springer: Dordrecht, 2000.

- (5) Walker, W. H.; Lewis, W. K.; McAdams, W. H. *Principles of Chemical Engineering*, 1st ed.; McGraw-Hill Book Company, Incorporated, 1923.
- (6) Liddell, D. M. *Handbook of Chemical Engineering*; McGraw-Hill Book Company, Incorporated, 1922; Vol. 1.
- (7) Kohn, J. P. *Ernest W. Thiele (1895-1993)*. In: *Memorial Tributes. National Academy of Engineering*; The National Academies Press: Washington, DC, 1996; Vol. 8; pp 193–196.
- (8) Thiele, E. W. Relation between Catalytic Activity and Size of Particle. *Ind. Eng. Chem.* **1939**, *31*, 916–920.
- (9) McCabe, W. L.; Smith, J. C. *Unit Operations of Chemical Engineering*; McGraw-Hill Education: New York, 1956.
- (10) Fadum, R. E. *Warren Lee McCabe (1899-1982)*. In: *Memorial Tributes. National Academy of Engineering*; The National Academies Press: Washington, DC, 1984; Vol. 2; pp 193–196.
- (11) Van Antwerpen, F. J.; Fourdrinier, S. *Highlights: The First Fifty Years of the American Institute of Chemical Engineers*; American Institute of Chemical Engineers: New York, 1958.
- (12) Vincenti, W. G. *What Engineers Know and How They Know It. Analytical Studies from Aeronautical History*; The Johns Hopkins University Press, 1990.
- (13) Johnson, A.; Lenhard, J. *Cultures of Prediction. How Engineering and Science Evolve with Mathematical Tools*; The MIT Press: Cambridge, MA, 2024.
- (14) Kister, H. Z. *Distillation Design*; McGraw-Hill , Incorporated: New York, 1992.
- (15) Kister, H. *Distillation Diagnostics: An Engineer's Guidebook*; Wiley, 2025.

- (16) Gorak, A.; Olujic, Z. *Distillation: Equipment and Processes*, 1st ed.; Academic Press, 2014.
- (17) Gorak, A.; Olujic, Z. *Distillation: Fundamentals and Principles*, 1st ed.; Academic Press, 2014.
- (18) Stichlmair, J. G.; Fair, J. R. *Distillation: Principles and Practice*, 1st ed.; Wiley VCH: New York, 1998.
- (19) Coulson, J. M.; Richardson, J. F. *Chemical Engineering. Volume 2 Particle Technology & Separation Processes*, 5th ed.; Butterworth-Heinemann: Oxford, 2013.
- (20) Kiss, A. A. *Advanced Distillation Technologies: Design, Control and Applications*; Wiley: Chichester, 2013.
- (21) Breyse, J. Ernest Sorel (1850-1904): un pionnier méconnu de la distillation et du Génie des procédés. *Bull. SABIX* **2011**, 63–70.
- (22) Sorel, E. Sur la rectification de l'alcool. *Comptes rendu* **1889**, *108*, 1128–1131, 1204–1207, 1317–1320.
- (23) Sorel, E. Sur la rectification de l'alcool. *Comptes rendu* **1894**, *118*, 1213–1215.
- (24) Hausbrand, E. *Die Wirkungsweise der Destillier- und Rektifizierungsapparate. Mit Hilfe einfacher mathematischer Betrachtungen.*; Julius Springer: Berlin, 1893.
- (25) Lewis, W. K. The Efficiency and Design of Rectifying Columns for Binary Mixtures. *Ind. Eng. Chem.* **1922**, *14*, 492–496.
- (26) Ponchon, M. Études graphiques de la distillation fractionnée industrielle. *La Technique Moderne* **1921**, *XIII(1)*, p. 20 et *XIII(2)*.

- (27) Savarit, P. Études graphiques des colonnes à distiller les mélanges binaires et ternaires. *Chim. Ind.* **1923**, N° spécial mai 1923 (compte rendu du congrès international des combustibles liquides, octobre 1922).
- (28) Hasse, H.; Lenhard, J. *Boon and bane: On the role of adjustable parameters in simulation models. In: Mathematics as a tool. Tracing new roles of mathematics in the sciences*; Springer International Publishing, 2017; pp 93–115.
- (29) Ferguson, E. S. *Engineering and the Mind's Eye*; The MIT Press: Cambridge, MA, 1992.
- (30) Lenhard, J. *Calculated Surprises. A Philosophy of Computer Simulation*; Oxford University Press: New York, 2019.
- (31) McMullin, E. Galilean idealization. *Stud. Hist. Philos. Sci. A* **1985**, *16*, 247–273.
- (32) Hertz, H. *The Principles of Mechanics Presented in a New Form*; MacMillan and Co.: London and New York, 1899.
- (33) Batterman, R. W.; Rice, C. C. Minimal Model Explanations. *Philos. Sci.* **2014**, *81*, 349–376.
- (34) Rice, C. Models Don't Decompose That Way: A Holistic View of Idealized Models. *Br. J. Philos. Sci.* **2019**, *70*, 179–208.
- (35) Frigg, R.; Hartmann, S. Models in Science. The Stanford Encyclopedia of Philosophy, 2024; <https://plato.stanford.edu/archives/fall2024/entries/models-science/>.
- (36) Fenske, M. R. Fractionation of Straight-Run Pennsylvania Gasoline. *Ind. Eng. Chem.* **1932**, *24*, 482–485.
- (37) Underwood, A. J. V. Fractional Distillation of Multicomponent Mixtures. *Ind. Eng. Chem.* **1949**, *41*, 2844–2847.

- (38) Kenig, E. Y.; Blagov, S. *Modeling of Distillation Processes In: Distillation: fundamentals and principles.*; Academic Press, 2014; pp 383–436.
- (39) Taylor, R.; Krishna, R. Modelling reactive distillation. *Chem. Eng. Sci.* **2000**, *55*, 5183–5229.
- (40) Sundmacher, K.; Kienle, A. *Reactive Distillation: Status and Future Directions*; Wiley-VCH: Weinheim, Germany, 2002.
- (41) Malone, M. F.; Doherty, M. F. Reactive Distillation. *Ind. Eng. Chem. Res.* **2000**, *39*, 3953–3957.
- (42) Waibel, Y.; Trescher, L.; Ranger, L.-M.; Grutzner, T. First multiple dividing wall column: Design and operation. *Chem. Eng. Res. Des.* **2023**, *193*, 132–144.
- (43) Mikyřka, J. Robust and efficient methods for solving the Rachford–Rice equation in flash equilibrium calculation. *Fluid Ph. Equilib.* **2023**, *571*, 113803.
- (44) Monroy-Loperena, R. An Iterated Eigenvalue Algorithm for Vapor–Liquid Split in Flash Calculations. *Ind. Eng. Chem. Res.* **2012**, *51*, 8488–8494.
- (45) Seidel, T.; Hoffmann, A.; Bortz, M.; Scherrer, A.; Burger, J.; Asprion, N.; Kufer, K.-H.; Hasse, H. A novel approach for infeasible path optimization of distillation-based flowsheets. *Chem. Eng. Sci.:X* **2020**, *7*, 100063.
- (46) Bortz, M.; Burger, J.; von Harbou, E.; Klein, M.; Schwientek, J.; Asprion, N.; Bottcher, R.; Kufer, K.-H.; Hasse, H. Efficient Approach for Calculating Pareto Boundaries under Uncertainties in Chemical Process Design. *Ind. Eng. Chem. Res.* **2017**, *56*, 12672–12681.
- (47) Kiss, A. A.; Muthia, R.; Pazmino-Mayorga, I.; Harmsen, J.; Jobson, M.; Gao, X. Conceptual methods for synthesis of reactive distillation processes: recent developments and perspectives. *J. Chem. Technol. Biotechnol.* **2024**, *99*, 1263–1290.

- (48) Ryll, O.; Blagov, S.; Hasse, H.  $\infty/\infty$ -Analysis of heterogeneous distillation processes. *Chem. Eng. Sci.* **2013**, *104*, 374–388.
- (49) Petlyuk, F.; Danilov, R.; Burger, J. A novel method for the search and identification of feasible splits of extractive distillations in ternary mixtures. *Chem. Eng. Res. Des.* **2015**, *99*, 132–148.
- (50) Skiborowski, M.; Harwardt, A.; Marquardt, W. Conceptual Design of Distillation-Based Hybrid Separation Processes. *Annu. Rev. Chem. Biomol. Eng.* **2013**, *4*, 45–68.
- (51) Lucia, A.; Amale, A.; Taylor, R. Distillation pinch points and more. *Comput. Chem. Eng.* **2008**, *32*, 1342–1364.
- (52) Blagov, S.; Hasse, H. Topological analysis of vapor–liquid equilibrium diagrams for distillation process design. *Phys. Chem. Chem. Phys.* **2002**, *4*, 896–908.
- (53) Ryll, O.; Blagov, S.; Hasse, H. Thermodynamic analysis of reaction-distillation processes based on piecewise linear models. *Chem. Eng. Sci.* **2014**, *109*, 284–295.
- (54) Dyga, M.; Keller, A.; Hasse, H. Distillation behavior of the system (formaldehyde + water + isoprenol). *Chem. Eng. Sci.* **2022**, *263*, 118082.
- (55) Bortz, M.; Burger, J.; Asprion, N.; Blagov, S.; Küfer, K.-H.; Hasse, H. On the effort of approaching pure components and azeotropes in distillation. *Chem. Eng. Sci.* **2015**, *127*, 253–259.
- (56) Burger, J. L.; Schneider, N.; Bruno, T. J. Application of the Advanced Distillation Curve Method to Fuels for Advanced Combustion Engine Gasolines. *Energy Fuels* **2015**, *29*, 4227–4235.
- (57) Bortz, M.; Burger, J.; Asprion, N.; Blagov, S.; Böttcher, R.; Nowak, U.; Scheithauer, A.; Welke, R.; Küfer, K.-H.; Hasse, H. Multi-criteria optimization in chemical process de-

- sign and decision support by navigation on Pareto sets. *Comput. Chem. Eng.* **2014**, *60*, 354–363.
- (58) Schwientek, J.; Teichert, K.; Schröder, J.; Höller, J.; Asprion, N.; Schäfer, P.; Wlotzka, M.; Bortz, M. Pareto-optimal treatment of uncertainties in model-based process design and operation. *arXiv* **2025**, arXiv:2503.19680 [math.OC].
- (59) Asprion, N.; Benfer, R.; Blagov, S.; Böttcher, R.; Bortz, M.; Berezhnyi, M.; Burger, J.; Harbou, E. v.; Küfer, K.; Hasse, H. INES – An Interface Between Experiments and Simulation to Support the Development of Robust Process Designs. *Chem. Ing. Tech.* **2015**, *87*, 1810–1825.
- (60) Kenig, E. Y.; Górak, A.; Pyhälähti, A.; Jakobsson, K.; Aittamaa, J.; Sundmacher, K. Advanced rate-based simulation tool for reactive distillation. *AIChE J.* **2004**, *50*, 322–342.
- (61) Mueller, I.; Kenig, E. Y. Reactive Distillation in a Dividing Wall Column: Rate-Based Modeling and Simulation. *Ind. Eng. Chem. Res.* **2007**, *46*, 3709–3719.
- (62) Yang, P.; Wu, J.; Luo, Y.; Kang, Y.; Jia, S.; Yuan, X. Optimization-based design of structure and operation parameters simultaneously for real-world distillation column based on rate-based model. *Sep. Purif. Technol.* **2025**, *355*, 129480.
- (63) Whitman, W. G. The two film theory of gas absorption. *Chem. Metall. Eng.* **1923**, *29*, 146–148.
- (64) Roshdi, S.; Kasiri, N.; Hashemabadi, S. H.; Ivakpour, J. Computational fluid dynamics simulation of multiphase flow in packed sieve tray of distillation column. *Korean J. Chem. Eng.* **2013**, *30*, 563–573.
- (65) Gozálvarez-Zafrilla, J. M.; García-Díaz, J. C.; Santafé-Moros, A. Uncertainty quantifica-

- tion and global sensitivity analysis of continuous distillation considering the interaction of parameter uncertainty with feed variability. *Chem. Eng. Sci.* **2021**, *235*, 116509.
- (66) Burger, J.; Asprion, N.; Blagov, S.; Bortz, M. Simple Perturbation Scheme to Consider Uncertainty in Equations of State for the Use in Process Simulation. *J. Chem. Eng. Data* **2016**, *62*, 268–274.
- (67) Wittig, R.; Lohmann, J.; Gmehling, J. Vapor- liquid equilibria by UNIFAC group contribution. 6. Revision and extension. *Ind. Eng. Chem. Res.* **2003**, *42*, 183–188.
- (68) Constantinescu, D.; Gmehling, J. Further development of modified UNIFAC (Dortmund): revision and extension 6. *J. Chem. Eng. Data* **2016**, *61*, 2738–2748.
- (69) Jirasek, F.; Alves, R. A.; Damay, J.; Vandermeulen, R. A.; Bamler, R.; Bortz, M.; Mandt, S.; Kloft, M.; Hasse, H. Machine learning in thermodynamics: Prediction of activity coefficients by matrix completion. *J. Phys. Chem. Lett.* **2020**, *11*, 981–985.
- (70) Hayer, N.; Jirasek, F.; Hasse, H. Prediction of Henry’s law constants by matrix completion. *AIChE J.* **2022**, *68*, e17753.
- (71) Großmann, O.; Bellaire, D.; Hayer, N.; Jirasek, F.; Hasse, H. Database for liquid phase diffusion coefficients at infinite dilution at 298 K and matrix completion methods for their prediction. *Digit. Discov.* **2022**, *1*, 886–897.
- (72) Jirasek, F.; Hayer, N.; Abbas, R.; Schmid, B.; Hasse, H. Prediction of parameters of group contribution models of mixtures by matrix completion. *Phys. Chem. Chem. Phys.* **2023**, *25*, 1054–1062.
- (73) Gond, D.; Sohns, J.-T.; Leitte, H.; Hasse, H.; Jirasek, F. Hierarchical matrix completion for the prediction of properties of binary mixtures. *Comp. Chem. Eng.* **2025**, 109122.

- (74) Hoffmann, M.; Hayer, N.; Kohns, M.; Jirasek, F.; Hasse, H. Prediction of pair interactions in mixtures by matrix completion. *Phys. Chem. Chem. Phys.* **2024**, *26*, 19390–19397.
- (75) Zenn, J.; Gond, D.; Jirasek, F.; Bamler, R. Balancing molecular information and empirical data in the prediction of physico-chemical properties. *Digit. Discov.* **2025**, *4*, 683–693.
- (76) Kohns, M.; Zittlau, P.; Jirasek, F. Prediction of the density of aqueous electrolyte solutions with matrix completion methods. *Fluid Ph. Equilib.* **2025**, 114454.
- (77) Hayer, N.; Wendel, T.; Mandt, S.; Hasse, H.; Jirasek, F. Advancing thermodynamic group-contribution methods by machine learning: UNIFAC 2.0. *Chem. Eng. J.* **2025**, *504*, 158667.
- (78) Hayer, N.; Hasse, H.; Jirasek, F. Modified UNIFAC 2.0-A Group-Contribution Method Completed with Machine Learning. *Ind. Eng. Chem. Res.* **2025**, *64*, 10304–10313.
- (79) Hoffmann, M.; Hasse, H.; Jirasek, F. GRAPPA – A Hybrid Graph Neural Network for Predicting Pure Component Vapor Pressures. *arXiv* **2025**, arXiv:2501.08729 [cs.LG].
- (80) Specht, T.; Nagda, M.; Fellenz, S.; Mandt, S.; Hasse, H.; Jirasek, F. HANNA: hard-constraint neural network for consistent activity coefficient prediction. *Chem. Sci.* **2024**, *15*, 19777–19786.
- (81) Hoffmann, M.; Specht, T.; Hayer, N.; Hasse, H.; Jirasek, F. MLPROP – an open interactive web interface for thermophysical property prediction with machine learning. *arXiv* **2025**, arXiv:2504.05970 [cs.CE].
- (82) Jirasek, F.; Bamler, R.; Mandt, S. Hybridizing physical and data-driven prediction methods for physicochemical properties. *Chem. Comm.* **2020**, *56*, 12407–12410.

- (83) Jirasek, F.; Bamler, R.; Fellenz, S.; Bortz, M.; Kloft, M.; Mandt, S.; Hasse, H. Making thermodynamic models of mixtures predictive by machine learning: matrix completion of pair interactions. *Chem. Sci.* **2022**, *13*, 4854–4862.
- (84) Jirasek, F.; Hasse, H. Combining Machine Learning with Physical Knowledge in Thermodynamic Modeling of Fluid Mixtures. *Annu. Rev. Chem. Biomol. Eng.* **2023**, *14*, 31–51.
- (85) Loubet, R.; Zittlau, P.; Hoffmann, M.; Vollmer, L.; Fellenz, S.; Leitte, H.; Jirasek, F.; Lenhard, J.; Hasse, H. Superstudent intelligence in thermodynamics. *arXiv* **2025**, arXiv:2506.09822 [cs.CE].
- (86) Roose, K. Why an Octopus-like Creature Has Come to Symbolize the State of A.I. The New York Times, 2023; <https://www.nytimes.com/2023/05/30/technology/shoggoth-meme-ai.html>.
- (87) Bender, E. M.; Koller, A. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.
- (88) Smith, B. C. *The Promise of Artificial Intelligence: Reckoning and Judgment*; The MIT Press, 2019.
- (89) Kuhn, T. S. *The Structure of Scientific Revolutions*, 4th ed.; University of Chicago Press, 2012.