Methods for efficient eutectic screening in CALPHAD databases

K. Firdová[1*], K. E. Borowski[2], R. Tamura[3,4,5], J.-M. Joubert[2], G. Deffrennes[1*]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP, SIMaP, F-38000 Grenoble, France

[2] Univ Paris Est Creteil, CNRS, ICMPE, UMR 7182, 2 rue Henri Dunant, 94320, Thiais, France

[3] Center for Basic Research on Materials, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

[4] Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan

[5] RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

[*] Corresponding authors. E-mail addresses: Katarina.FIRDOVA@cnrs.fr (K. Firdová), Guillaume.DEFFRENNES@cnrs.fr (G. Deffrennes)

Abstract

Eutectic points play a critical role in material science due to their low melting points, absence of a solidification range, and formation of unique microstructures that enhance properties. However, identifying eutectics in calculated ternary and higher-order systems remains computationally challenging due to the lack of a well-established screening approach.

This work addresses this gap by developing a general, efficient methodology for efficient eutectic screening in multicomponent systems. We first propose and benchmark four CALPHAD-based methods for eutectic identification, evaluating their computational speed and reliability. We then investigate the potential of combining CALPHAD with machine learning to further accelerate the process, leveraging two eutectic properties: (1) eutectics are local minima on the liquidus, and (2) they lie at the intersections of primary crystallization regions. Active learning surrogate models are employed to interpolate CALPHAD data for predicting liquidus and primary crystallization phases, and the performance of each approach is evaluated. By benchmarking multiple strategies and releasing code for three approaches that enable rapid and reliable eutectic screening in high-order systems, this work advances the design of materials such as cast alloys, phase change materials, eutectic high entropy alloys, or metallic glasses.

1. Introduction

Eutectic points are of considerable interest across all classes of materials due to their unique properties, including the absence of a solidification range and low melting points. In alloys, eutectics lead to fine microstructures that can enhance mechanical properties [1], notably through composite strengthening [2], but also other properties such as corrosion resistance [3]. Identifying eutectic compositions is important for the design of casting alloys [4], metallic glasses [5,6], soldering and brazing alloys [7], coolants in fast-neutron reactors [8,9], or phase change materials for thermal energy storage [10–14]. Current research efforts increasingly focus on the challenging task of exploring quaternary and higher-order systems [15], including the so-called eutectic high entropy alloys [16], in addition to more conventional binary and ternary alloys.

CALPHAD databases [17–21] offer a wealth of knowledge to identify eutectics. They contain descriptions of the Gibbs energy of phases in binary and ternary systems, from which extrapolations to higher-order systems can be performed. However, there is no well-established computational approach to screen these databases for eutectics, despite their significance. This poses two main challenges. First, a method must be formulated to identify eutectics using the available thermodynamic calculation routines. It must ensure that no false candidates are identified, otherwise an additional verification mechanism must be introduced. Second, screening must be computationally-efficient to navigate large compositional spaces in a reasonable amount of time. For example, screening a septenary system using a grid-search method with 1 at% precision would require 1 705 904 746 calculations, which would take roughly 20 days even if each calculation required only 1ms.

To address the challenge of identifying eutectics, the authors of Ref. [22] have proposed two high-throughput CALPHAD approaches. The first approach selects the points where the calculated solidification range tends to zero, following validation based on the number of solid phases that form to avoid false identification of congruent melting compounds. The second approach compares each composition to its closest neighbors to identify local minima on the liquidus. These screening approaches can prove time-consuming given that liquidus temperatures have to be computed over the entire compositional space at regular steps of 0.5 wt.%. The authors of Ref. [14] used the Mesh Adaptive Direct Search algorithm to optimize search of local minima on the liquidus.

To address the challenge of computational efficiency, machine learning (ML) has been demonstrating its power in various applications in the domain of material science over the last years [23–25] but its potential to accelerate the design of eutectic alloys has not been fully explored yet. Several researches have been done coupling CALPHAD computation with artificial intelligence, however, most show no special interest in eutectics [26–29]. A handful of others have investigated the use of ML for classification of the so-called eutectic high entropy alloys [30–32]. Ref. [33] has summarized recent progress on data-driven design of these alloys, highlighting that the results are promising but still at an infant stage.
The authors of Ref. [34] developed ML models to predict the fraction of primary phases in the Al-Co-Cr-Fe-Ni system using data from CALPHAD calculations and the literature. They used this fraction as a measure of deviation from eutectic composition, where it should tend to zero, to identify three novel eutectic high-entropy alloys. The authors of Ref. [35] applied ML to interpolate liquidus temperatures obtained by thermodynamical computations to identify the eutectic points as the compositions with the lowest temperature within the surrounding compositional space. However, the paper is focused only on the eutectics nearby a high-order slag composition and does not include a time and performance analysis to systematically identify all eutectics in a system.

The main objective of this work is to address the current limitations and to develop a general methodology for fast eutectics screening that is applicable to any multicomponent system. First, we benchmark four methods based on basic CALPHAD calculations for identifying eutectics from a limited number of starting points, evaluating them in terms of computational speed and false-positive rate. Second, we explore whether further acceleration is possible by combining CALPHAD calculations with ML to efficiently identify eutectic candidates that can serve as starting points for validation. To reach this goal, two properties of the eutectics are exploited: (1) eutectics are local minima on the liquidus, (2) eutectics are at the intersection of primary crystallization regions. ML algorithms are used to interpolate data from CALPHAD assessments to predict liquidus or primary crystallization phases, and the performance of each approach is evaluated.

2. Data and implementation

The methods are tested on binary, ternary and quaternary systems composed of the elements Al, Cu, Mg, Si, Zn. These five abundant elements are selected because CALPHAD assessments supported by liquidus data are available in the literature for all ten ternary systems they form [36]. Only the systems containing at least one eutectic point are considered, disregarding equilibria where the amount of one of the solid phases involved is negligible, i.e., when an invariant cannot be distinguished from the melting point of a pure element in a binary system. The systems and corresponding eutectics are listed in Table 1 with permission from Thermo-Calc Software.

All methods are implemented using the TC-Python software development kit of Thermo-calc® [37]. The Thermo-Calc Software TCAL09 Al-based Alloy Database [38] is used to retrieve information required for training, modeling and validation. Scheil simulations used Thermo-Calc's Scheil module with default settings, while all other CALPHAD computations used the single equilibria module. The Global Minimization Technique, which is the most robust routine implemented in the software to find the most stable equilibrium state, was enabled in all cases. Liquidus (resp. solidus) calculations were performed using the *compute_transition* command, starting from a single equilibrium calculation in a fully liquid (resp. fully solid) domain. This approach was the fastest and most reliable. If convergence failed, an equilibrium calculation with the liquid phase fixed at 1 (resp. 0) was used instead. All computations are performed on computer equipped with Intel® Core™ i9-14900K 3.20 GHz, 64.0 GB RAM and NVIDIA GeForce RTX 4080 SUPER (16GB), Intel® UHD Graphics 770 (128 MB).

The chemical composition is the only input information to the models. No physico-chemical properties of the elements are regarded.

Table 1. Systems used for benchmarking with the corresponding rounded eutectic compositions calculated using TCAL9 [38]

Order of system	Studied systems	Eutectics (at%)
Binary	Al-Cu	82Al-18Cu, 18Al-82Cu
	Al-Mg	31Al-69Mg, 58Al-42Mg, 64Al-36Mg
	Al-Si	88Al-12Si
	Al-Zn	12Al-88Zn
	Cu-Mg	16Cu-84Mg, 41Cu-59Mg, 79Cu-21Mg
	Cu-Si	69Cu-31Si, 19Cu-81Si
	Mg-Si	87Mg-13Si, 46Mg-54Si
	Mg-Zn	71Mg-29Zn, 7Mg-93Zn

Total	8	16
Ternary	Al-Cu-Mg	22Al-63Cu-15Mg, 31Al-2Cu-67Mg, 75Al-16Cu-9Mg, 1Al-17Cu-82Mg
	Al-Cu-Si	80Al-13Cu-7Si, 6Al-70Cu-24Si
	Al-Mg-Si	83Al-5Mg-12Si
	Al-Mg-Zn	59Al-34Mg-6Zn, 5Al-70Mg-25Zn, 9Al-6Mg-85Zn, 51Al-43Mg-6Zn
	Cu-Mg-Si	79Cu-6Mg-15Si, 36Cu-22Mg-41Si, 74Cu-6Mg-21Si, 23Cu-34Mg-43Si, 63Cu-7Mg-31Si
	Cu-Mg-Zn	53Cu-13Mg-34Zn, 14Cu-77Mg-9Zn
	Cu-Si-Zn	69Cu-27Si-4Zn
Total	7	19
Quaternary	Al-Cu-Mg-Si	3Al-65Cu-5Mg-27Si, 6Al-42Cu-17Mg-35Si, 77Al-12Cu-3Mg-7Si, 6Al-24Cu-30Mg-39Si, 77Al-14Cu-5Mg-4Si
	Cu-Mg-Si-Zn	21Cu-31Mg-35Si-13Zn, 64Cu-5Mg-28Si-3Zn, 72Cu-5Mg-21Si-2Zn
Total	2	8

3. CALPHAD-based methods

The approaches for eutectics identification presented in this section rely fully on CALPHAD computations. The objective is to address the lack of a straightforward numerical method to identify eutectics. To begin with, the methods are compared to select the fastest and most performant one. For this comparison, starting points are selected as equidistant points across the entire compositional space of selected binary and ternary systems.

The selection of the best method serves two purposes. First, it provides a baseline for comparing the time and performance of the ML-assisted approaches presented in the following section. Second, it is used as a last verification step to eliminate false eutectic candidates identified by ML-assisted approaches. From this perspective, ML-assisted approaches can be seen as models for efficiently selecting starting points to further accelerate eutectic identification. However, the present CALPHAD-based methods can also be used independently.

3.1. Algorithms

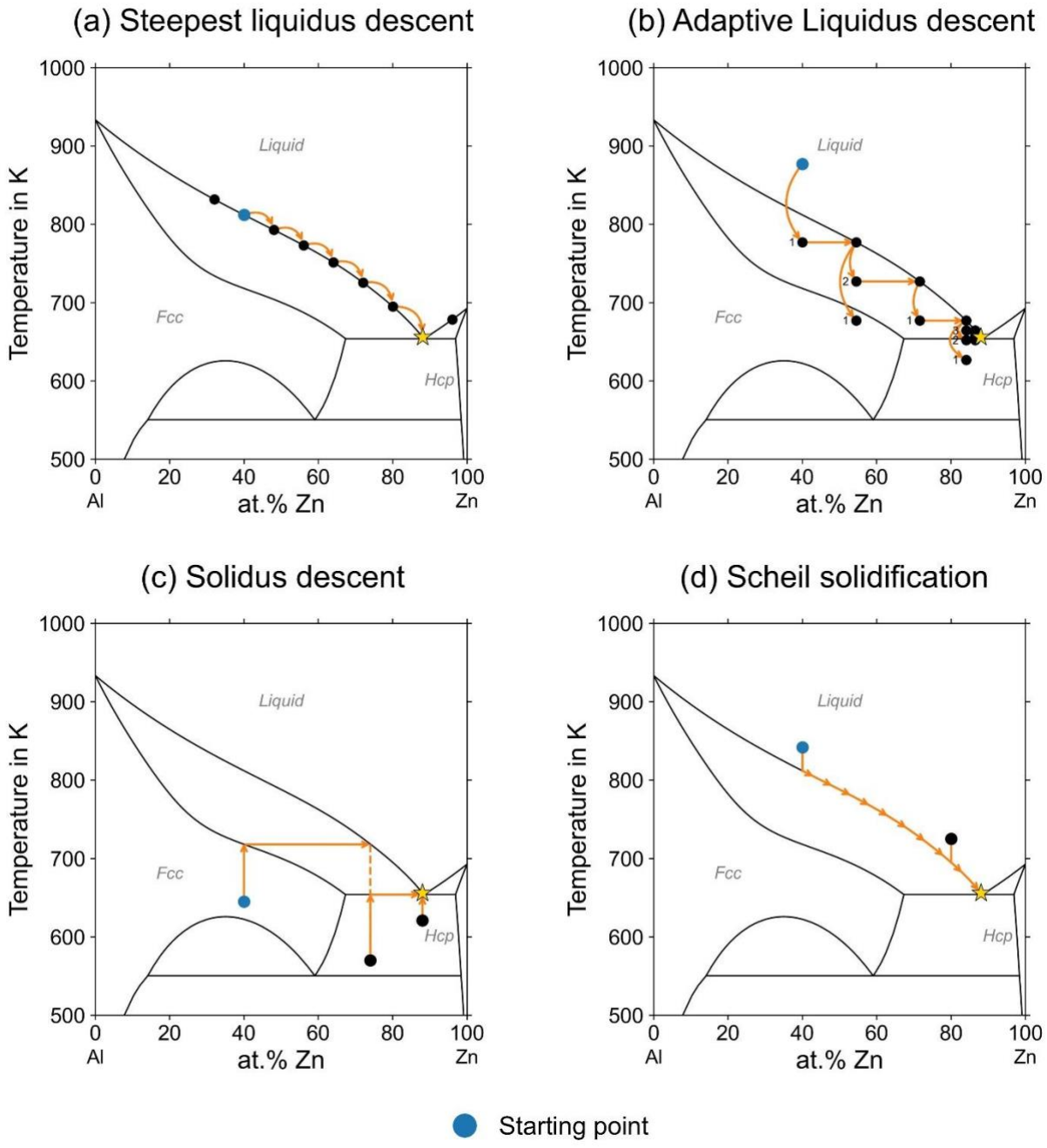


Figure 1. Four CALPHAD-based method to find eutectics illustrated in the Al-Zn system [38]

3.1.1. Steepest liquidus descent

Steepest descent is a specific version of a gradient descent optimization algorithm that allows the identification of local minima in a discretized space. Its idea is based on iteratively moving in the direction of the greatest decrease among the neighboring points in the

(hyper)space. In the context of eutectic identification, the steepest descent is applied to liquidus temperatures computed using Thermo-calc.

A binary eutectic in an A-B system can be found with a precision of 1at% using the following procedure:

1. Compute the liquidus temperature T_{start} at the initial composition (A_x, B_{100-x}).
2. Compute liquidus temperatures T_{n1} and T_{n2} at the neighboring compositions (A_{x+1}, $B_{100-(x+1)}$) and (A_{x-1}, $B_{100-(x-1)}$), respectively.
3. If $T_{n1} > T_{start}$ and $T_{n2} > T_{start}$, then (A_x, B_{100-x}) is a eutectic candidate.
   Else: set the new starting composition to the one where the temperature drop from T_{start} is the largest, and repeat steps 2-3.

The strategy is illustrated on Figure 1(a). The software is queried to compute the liquidus temperature only at the starting points and at the neighboring points after each iteration. This saves time compared to systematically computing the liquidus temperature for all compositions.

3.1.2. Adaptive liquidus descent

This strategy is based on a sequence of single equilibrium calculations to find a minimum on the liquidus.

Each calculation is performed at a lower temperature than the liquidus. If the new equilibrium is in a fully solid region, the temperature decrement is cut in half. Else, if it is between the solidus and the liquidus, the next calculation is performed from this new liquidus point. When the temperature step becomes smaller than the desired precision, an isothermal melting point has been found at a local minimum.

A binary eutectic in an A-B system can be found using the following procedure:

1. At the initial composition, start at an arbitrary temperature (here, the mean melting point of the elements), and increase the temperature in steps of 100K until the liquid phase is present.
2. Perform a calculation at the composition of the last obtained liquid phase and a lower temperature, T – *d* (here, *d* is initially set at 100K).
3. If a liquid phase is still present, repeat steps 2-3. If this occurs four times in a row, double the temperature decrement *d*.

Else (only solids are present), $d = d/2$. If d becomes smaller than a threshold (here, 0.05K), the last calculated point is a eutectic candidate. Else, repeat steps 2-3.

The strategy is illustrated on Figure 1(b). The optimized temperature decrement ensures a rapid descent along the liquidus surface. Besides, this approach only relies on simple single equilibrium calculations with fixed composition and temperature, which are both fast and easy to converge.

3.1.3. Solidus descent

Similar to the Adaptive liquidus descent (Section 3.1.2.), this method relies on iterative single equilibria calculations. For each iteration, the solidus temperature and the composition of the corresponding liquid phase are obtained. The next calculation is performed at the composition of the previously obtained liquid phase, which ensures an optimized descent along the liquidus. Convergence to the eutectic point is achieved when the liquid phase composition is identical to the solidus composition. By default, the algorithm also converges to the special invariant points discussed in Section 3.2. when at least two solid phases are involved. This behavior can be modified by adjusting a constraint on the number of stable phases. If the algorithm encounters a peritectic plateau, it will continue until convergence, since the liquid composition at the peritectic reaction differs from that at the solidus at a lower temperature.

A binary eutectic in an A-B system can be found with the following procedure:

1. Compute the solidus at the starting composition (A_x, B_{100-x}) as described in Section 2.
2. Compute the composition (A_y, B_{100-y}) of corresponding liquid phase.
3. If x=y, the composition is a eutectic candidate, unless there is only one solid phase, in which case it corresponds to a congruently melting compound, a pure element, or a liquidus minimum above a single solid solution, and the algorithm stops without returning a result.

   Else: Set (A_y, B_{100-y}) as the new starting point and repeat from step 1.

The strategy is illustrated on Figure 1(c). While a computation to find the solidus is slower and can diverge more often than those performed in the Adaptive liquidus descent method, fewer computations are necessary to identify a eutectic candidate.

3.1.4. Scheil descent

Scheil calculations are solidification simulations assuming infinite liquid diffusion and no solid diffusion [39,40]. Under these conditions, the liquid composition follows the liquidus until solidification completes at an isothermal melting point, often a eutectic, provided the convergence criterion is small enough.

A binary eutectic in an A-B system can be found with the following procedure:

1. Perform a Scheil simulation from the initial composition (A_x, B_{100-x}) and above the liquidus.
2. Perform a second Scheil simulation from the final liquid composition of the first simulation which may have stopped early, before a eutectic was reached.
3. The final liquid composition after the second simulation is a eutectic candidate.

The strategy is illustrated on Figure 1(d). It is relatively easy to implement, as the Scheil calculation routine is already built into most thermodynamic calculation software.

3.2. Results and discussion

Figure 2 compares the discussed CALPHAD-based methods across 8 binary and 7 ternary systems as a function of the number of initial points for each composition axis, the grid size. The capacity of the methods to identify eutectics improves with an increasing number of initial points, as expected. The distinguishing factors are computational time and number of false candidates.

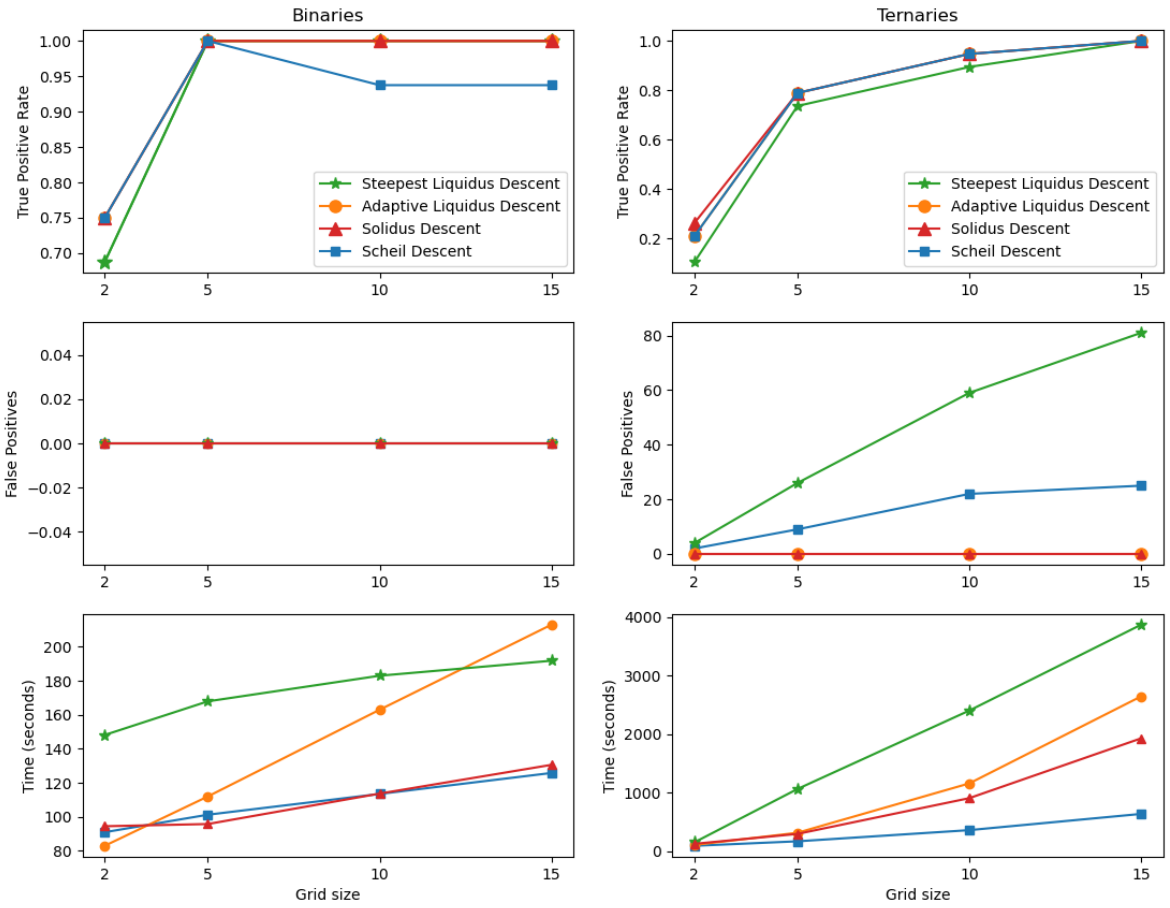
The Scheil descent (Section 3.1.4.) uses two Scheil simulations to avoid premature solidification termination before reaching a eutectic point. However, the method still leads to false eutectics in ternary systems. Running more simulations with a smaller convergence criterion and step size than the default values could improve the issue, at the cost of higher computational cost.

A similar problem arose with the Steepest liquidus descent (Section 3.1.1.). In the current implementation, neighbors are selected as the two nearest compositions along each composition axis, with a fixed step of 1at%, leading to 2 and 6 neighbors in binary and ternary systems, respectively. This appears to be insufficient in ternary systems where there is an infinite number of nearest neighbors in a circle around each point, particularly in domains with a flat liquidus surface. Consequently, a high number of false candidates is found. However, adding more neighbors or decreasing the discretization step negatively impacts computational time, which will increase exponentially with the number of elements. This excludes any approach based on nearest-neighbor analysis, such as the second approach in Ref. [22], from the competition.

No false positives were obtained in either binary or ternary systems with both the Adaptive liquidus descent (Section 3.1.2.) and the Solidus descent (Section 3.1.3.). The Solidus descent is the method that offers the best trade-off between performance and required time resources. It is therefore selected as a baseline for the following part of the paper.

All the approaches considered here, except the Scheil descent, can identify local minima in the liquidus with no solidification range. These points can either be eutectics involving as many solid phases as independent components or other special points [41], such as minima in the liquidus above a single solid-solution phase, as observed in the Co-Pd or Ni-Pd systems [42]. Among the 17 systems considered in this study, only one such special point is found in the Cu-Si-Zn system [38]. It corresponds to a minimum on the monovariant line between Diamond-Si and the γ solid solution, as illustrated in Figure 3 for the general case involving two solid solutions. While both eutectics and these special points share key characteristics, only eutectics were considered for our benchmark, so the Cu-Si-Zn special point was excluded. None of the above CALPHAD-based approaches can readily identify a non-invariant point that would result in a eutectic-like microstructure forming over a finite solidification temperature range. However, they may be adapted for this purpose, as their calculations naturally traverse the boundary between two crystallization domains, such as a monovariant line in a ternary system, on the path to the eutectic.

Figure 2. Performance of CALPHAD-based methods on binary (left) and ternary (right) systems as a function of grid size (x-axis) (see more on grid size in Appendix A). True Positive Rate displays the ratio of correctly



identified eutectics. False Positives display the number of compositions that were incorrectly identified as eutectics. All plotted values are cumulative after execution of the algorithms on all systems.

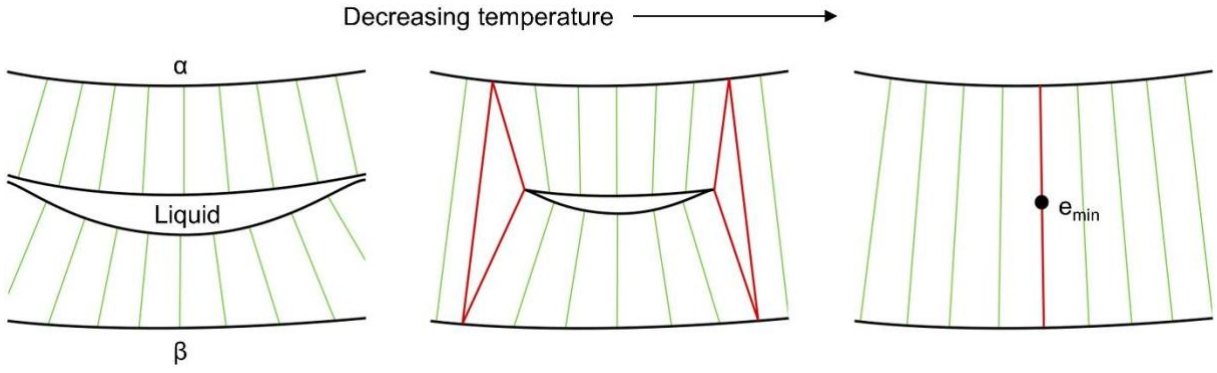


Figure 3. Local minimum in the liquidus with no solidification range in a ternary system. At the point e_{min}, the alloy exhibits eutectic-like solidification

4. Machine learning integrated with CALPHAD

The main motivation for using ML is its lower computational cost compared to software computations. For instance, once trained, the Gaussian model predicts liquidus temperature in approximately 50 µs vs Thermo-Calc's 180 ms (tested on 75Cu25Si). Further acceleration can be achieved by building surrogate models using active learning, an approach where new training points are strategically added to improve performance.



Figure 4. Two ways of determination of eutectics in a binary phase diagram: eutectics as local minima on liquidus curve (left) and eutectics as the intersection of primary crystallization regions (right). These properties are exploited when constructing ML-assisted approaches for eutectics identification.

Figure 4 displays two ways of determining eutectics, illustrated here for a binary phase diagram. The first is to identify eutectics as local minima on the liquidus. The second is to identify eutectics as the intersection of primary crystallization regions.
We build ML surrogate models to predict either liquidus temperatures or primary crystallization regions, by interpolating training data obtained from the software. We use algorithms capable of providing uncertainty estimates to efficiently select training data via active learning, targeting regions of highest prediction uncertainty. All eutectic candidates identified on the predicted space are then verified with the Solidus descent (Section 3.1.3.), which was chosen as the best performing CALPHAD-based algorithm in Section 3.2.

4.1. ML-assisted methods

4.1.1. ML for liquidus prediction

For liquidus temperature prediction, a Gaussian process is used, as implemented in Scikit learn [43]**.** Initial training points are compositions selected across the compositional space using a grid size of 5 (Appendix A), with their corresponding liquidus temperature as target. New training points are added one by one based on the model's highest prediction uncertainty to improve predictions.

Once training is complete, the liquidus temperature is predicted across the entire compositional space with a regular step of 1 at%. Steepest liquidus descent (Section 3.1.1.) is then applied to identify local minima on the liquidus, starting from equidistant points on a grid size of 10.

Identified local minima are then used as starting points to Solidus descent (Section 3.1.3.) to confirm the eutectic point, refine the identified composition if necessary, or refute it. The method is illustrated in Figure 5.

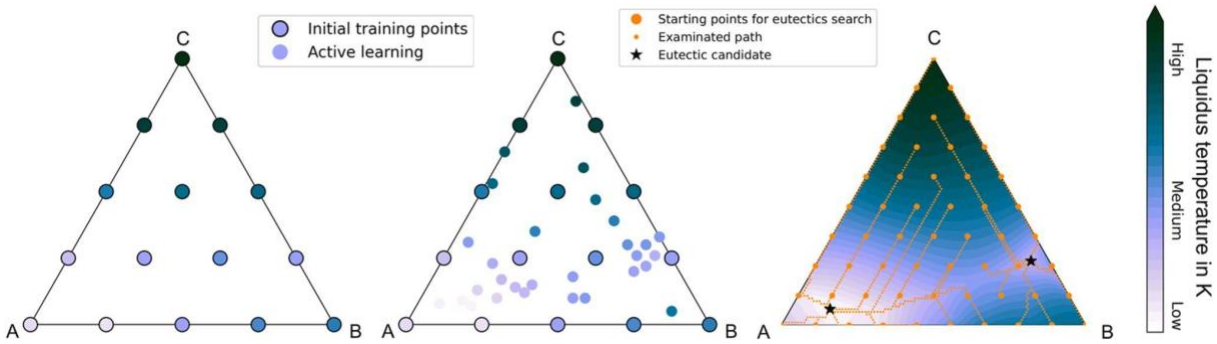


Figure 5. Example of surrogate liquidus prediction in a ternary system. Initial training points are selected on a grid size of 5 (left). 30 additional training points are proposed via active learning (middle). Once the liquidus is predicted, a steepest descent algorithm is applied on equidistant starting points (right). Eutectic candidates are the points with the lowest liquidus temperature in their surrounding compositional space.

4.1.2. ML for primary crystallization regions prediction

For primary crystallization region prediction, AIPHAD (Artificial Intelligence technique for PHAse Diagram [44]) is used. As in Section 4.1.1., the initial training data are compositions selected from the compositional space using a grid size of 5, with their corresponding primary crystallization phase as the target. New training points are added one by one where prediction uncertainty is the highest.

The primary crystallization phase is predicted using the Label Spreading method implemented in AIPHAD across the entire compositional space with a regular step of 4 at%. Eutectic candidates are then identified as points where *n* different phases meet among their first neighbors, with *n* the number of elements in the system. The identified compositions are then used as starting points for Solidus descent (Section 3.1.3.) to confirm, refine or refute the eutectic point. The method is illustrated in Figure 6.

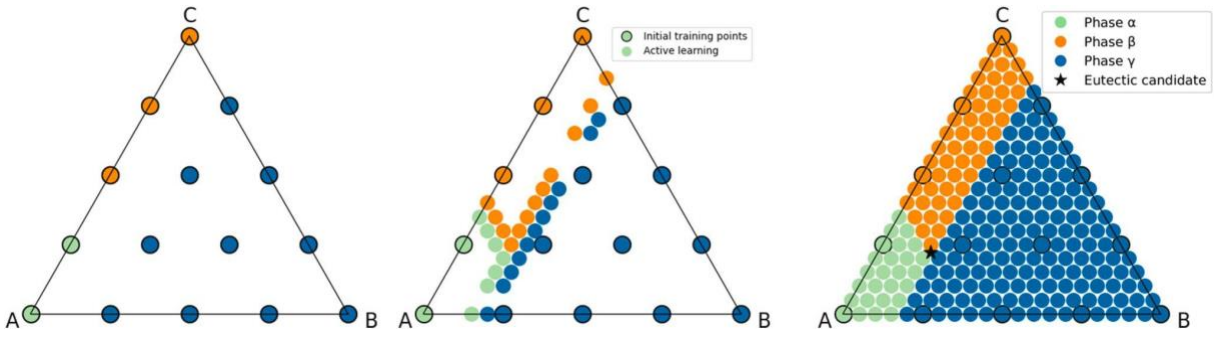


Figure 6. Example of surrogate primary crystallization region prediction in a ternary system. Initial training points are selected on a grid size of 5 (left). 30 additional training points are proposed by active learning to improve the prediction (middle). AIPHAD employs the Label Spreading method to propagate the labels of the training points to adjacent regions (right). Eutectic candidates are determined as the points at the intersection of three primary crystallization regions.

4.2. Results and discussion

Figure 7 compares ML-assisted approaches with the Solidus descent (Section 3.1.3.) based solely on CALPHAD computation.

The number of training points serves as the basis for the comparison. In the case of ML-assisted methods, training points include an initial grid of size 5 and new training points added after active learning cycles. Since eutectic verification is already integrated into the strategy of ML-assisted methods, the approaches do not produce false candidates.

A time gain can be observed for ML-assisted methods, confirming the interest to couple ML algorithms with CALPHAD computations. Eutectic identification via primary crystallization region prediction outperformed identification via liquidus prediction. In binary systems, liquidus prediction could perform better by further optimizing the Gaussian process hyperparameters or using a different algorithm. However, in systems with three or more elements, this can be partially explained by issues with the steepest descent algorithm discussed in Section 3.2. ML for liquidus prediction (Section 4.1.1.) is therefore not evaluated on quaternary systems.

Time gain of ML-assisted method on quaternary systems is less significant compared to the gain in ternary systems, presumably due to large number of candidates that needs to be verified. The current strategy for the candidate search is based on number of different neighboring phases computed for every composition but this may lead to an increased number of false candidates due to mistakes in the prediction such as isolated phases.

Moreover, the model seems to struggle when multiple eutectic points with similar compositions are present. For instance, only one quaternary eutectic from the pair 77Al-12Cu-3Mg-7Si and 77Al-14Cu-5Mg-4Si is detected and the ML-assisted model fails to achieve full performance on the tested system. While more training points could be added to address this limitation, this would come at the cost of increased computation time.

In the context of design of eutectics, a trade-off between the computational costs and ratio of the identified eutectics must be considered. Although the detection rate of ML-assisted methods is inferior than Solidus descent for certain configurations, the proposed ML-assisted methods enable faster identification of the majority of the eutectics.

Surrogate liquidus prediction, coupled with an optimization algorithm to quickly find local minima, can identify both eutectics and other special invariant points discussed in Section 3.2. Surrogate primary crystallization region prediction using AIPHAD was used here only to

identify eutectics, but it can also efficiently map boundaries between two crystallization domains. Along these boundaries, alloys forming a binary eutectic microstructure during a non-invariant transformation can be found, which is of interest for designing eutectic high-entropy alloys.

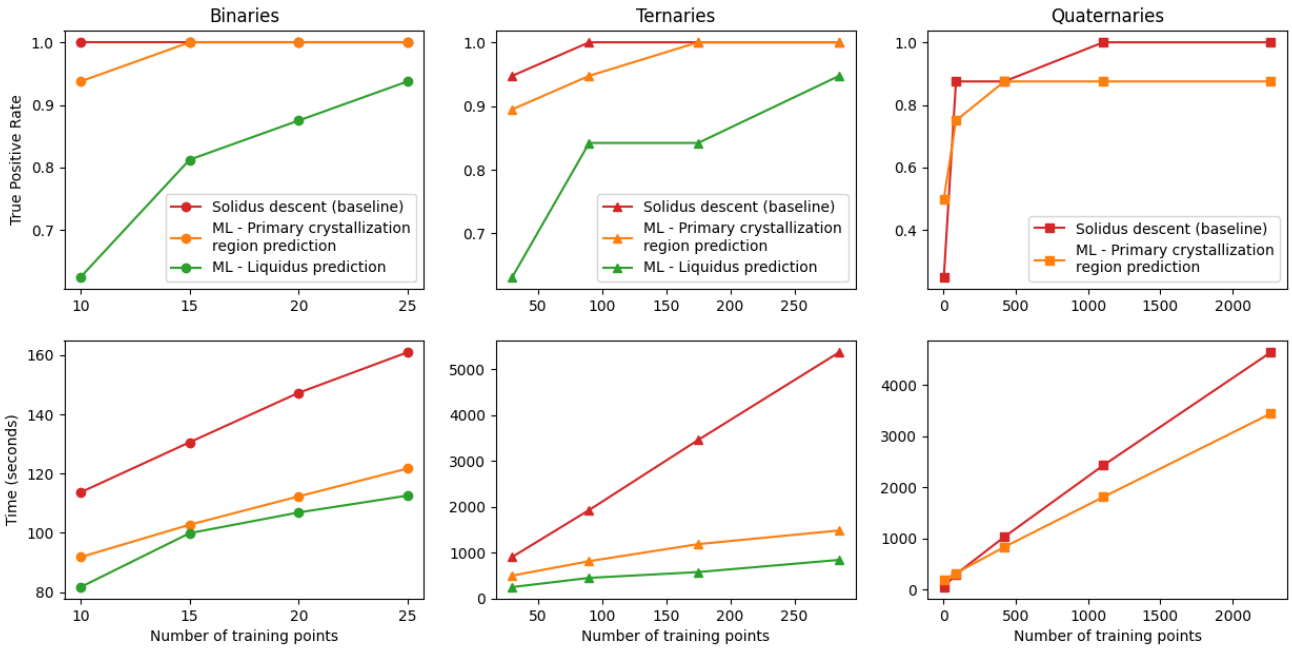


Figure 7. Comparison of ML-assisted approaches with CALPHAD-based Solidus descent. The x-axis shows the number of training points. For ML-assisted approaches, the initial grid points and new points added by active learning are accumulated. For the Solidus descent, training points are equidistant starting points on the grid.

5. Conclusions

In this work, several approaches for eutectic identification were introduced and examined. First, approaches based solely on CALPHAD computations were compared. Second, two machine learning assisted approaches were proposed and compared with the most performant CALPHAD approach on binary to quaternary systems.

The CALPHAD-based Solidus descent method can be employed independently as a reliable and computationally efficient eutectics identification method when applied on equidistant points in the compositional space. It also enables the identification of other special points that

share key characteristics with eutectics (Section 3.2.) and are thus relevant to the broad range of applications outlined in the Introduction. This study demonstrates improved trade-offs between computational cost and the ratio of identified eutectics by coupling this approach with an active learning surrogate model to optimize starting point selection. This can be particularly relevant for exploring high-order systems, as the size of the composition space becomes prohibitively large.

Further improvements of the ML surrogate models are possible. As discussed in Section 3.2., the steepest descent algorithm used in this work to identify local minima on the predicted liquidus (hyper)space hits a limitation in ternary and high-order systems. The method was selected as a gradient-free and interpretable approach for the discretized non-convex space, but other optimization algorithms could also be explored. Another perspective is to optimize candidate identification to reduce the computational time in the approach involving primary crystallization region prediction in higher-order systems. To efficiently identify where they intersect, one possible alternative is to consider points with high uncertainty as new candidates, assuming the algorithm is most uncertain about predictions at the boundaries of the regions. Alternatively, a candidate-filtering strategy could be implemented to reduce the number of candidates requiring verification. As discussed in Section 4.2., the current process generates several false candidates. Filtering in dense candidate regions could eliminate redundancies and improve the algorithm's efficiency.

Code and Data availability

The data used in this work were generated using Thermo-Calc® software and the TCAL9 database. Python code for the Solidus descent and the Adaptive liquidus descent are provided as supplementary materials. Python code for Primary crystallization regions predictions is available at https://github.com/katkaf/eutectics-pcr-prediction.

Acknowledgement

This work benefited from a grant managed by the Agence Nationale de la Recherche under the France 2030 program with the reference ANR-23-PEXD-0002. We acknowledge the support of Thermo-Calc Software for making an exception to their EULA and kindly allowing us to publish the data generated using TCAL9. We extend our thanks to Alisson Kwiatkowski da Silva and Nicholas Grundy from Thermo-Calc for their help and advices.

References

[1] N.A. Belov, Effect of eutectic phases on the fracture behavior of high-strength castable aluminum alloys, Met. Sci. Heat Treat. 37 (1995) 237–242. https://doi.org/10.1007/BF01152226.

[2] C.S. Tiwary, P. Pandey, S. Sarkar, R. Das, S. Samal, K. Biswas, K. Chattopadhyay, Five decades of research on the development of eutectic as engineering materials, Prog. Mater. Sci. 123 (2022) 100793. https://doi.org/10.1016/j.pmatsci.2021.100793.

[3] J. Ma, Y. Zhang, M. Ma, C. Qin, F. Ren, G. Wang, Corrosion and discharge performance of a magnesium aluminum eutectic alloy as anode for magnesium–air batteries, Corros. Sci. 170 (2020) 108695. https://doi.org/10.1016/j.corsci.2020.108695.

[4] C. Suwanpreecha, P. Pandee, U. Patakham, C. Limmaneevichitr, New generation of eutectic Al-Ni casting alloys for elevated temperature services, Mater. Sci. Eng. A 709 (2018) 46–54. https://doi.org/10.1016/j.msea.2017.10.034.

[5] H.-J. Fecht, W.L. Johnson, Thermodynamic properties and metastability of bulk metallic glasses, Mater. Sci. Eng. A 375–377 (2004) 2–8. https://doi.org/10.1016/j.msea.2003.10.254.

[6] L.-L. Shi, J. Xu, E. Ma, Alloy compositions of metallic glasses and eutectics from an idealized structural model, Acta Mater. 56 (2008) 3613–3621. https://doi.org/10.1016/j.actamat.2008.04.002.

[7] M. Way, J. Willingham, R. Goodall, Brazing filler metals, Int. Mater. Rev. 65 (2020) 257–285. https://doi.org/10.1080/09506608.2019.1613311.

[8] C.F. Smith, L. Cinotti, Lead-cooled fast reactor, in: Handb. Gener. IV Nucl. React., Elsevier, 2016: pp. 119–155. https://doi.org/10.1016/B978-0-08-100149-3.00006-9.

[9] B. Garcinuño, R. Fernández-Saavedra, T. Hernández, M.B. Gómez, A. Quejido, D. Rapisarda, Establishing technical specifications for PbLi eutectic alloy analysis and its relevance in fusion applications, Nucl. Mater. Energy 30 (2022) 101146. https://doi.org/10.1016/j.nme.2022.101146.

[10] J. Jiao, B. Grorud, C. Sindland, J. Safarian, K. Tang, K. Sellevoll, M. Tangstad, The Use of Eutectic Fe-Si-B Alloy as a Phase Change Material in Thermal Energy Storage Systems, Materials 12 (2019) 2312. https://doi.org/10.3390/ma12142312.

[11] N. Gokon, C.S. Jie, Y. Nakano, S. Okazaki, T. Kodama, T. Hatamachi, S. Bellan, Phase Change Material of Copper–Germanium Alloy as Solar Latent Heat Storage at High Temperatures, Front. Energy Res. 9 (2021) 696213. https://doi.org/10.3389/fenrg.2021.696213.

[12] J. Jiao, J. Safarian, M. Tangstad, The Use of Fe-26Si-9B Alloy as Phase Change Material in Si3N4 Container, Crystals 12 (2022) 376. https://doi.org/10.3390/cryst12030376.
[13] Y. El Karim, Y. Grosu, A. Faik, R. Lbibb, Investigation of magnesium-copper eutectic alloys with high thermal conductivity as a new PCM for latent heat thermal energy storage at intermediate-high temperature, J. Energy Storage 26 (2019) 100974. https://doi.org/10.1016/j.est.2019.100974.
[14] A.E. Gheribi, A.D. Pelton, J.-P. Harvey, Determination of optimal compositions and properties for phase change materials in a solar electric generating station, Sol. Energy Mater. Sol. Cells 210 (2020). https://doi.org/10.1016/j.solmat.2020.110506.
[15] I. Chang, Q. Cai, From simple binary to complex multicomponent eutectic alloys, Prog. Mater. Sci. 123 (2022) 100779. https://doi.org/10.1016/j.pmatsci.2021.100779.
[16] Y. Lu, Y. Dong, H. Jiang, Z. Wang, Z. Cao, S. Guo, T. Wang, T. Li, P.K. Liaw, Promising properties and future trend of eutectic high entropy alloys, Scr. Mater. 187 (2020) 202–209. https://doi.org/10.1016/j.scriptamat.2020.06.022.
[17] C.W. Bale, E. Bélisle, P. Chartrand, S.A. Decterov, G. Eriksson, A.E. Gheribi, K. Hack, I.-H. Jung, Y.-B. Kang, J. Melançon, A.D. Pelton, S. Petersen, C. Robelin, J. Sangster, P. Spencer, M.-A. Van Ende, FactSage thermochemical software and databases, 2010–2016, Calphad 54 (2016) 35–53. https://doi.org/10.1016/j.calphad.2016.05.002.
[18] H.-L. Chen, H. Mao, Q. Chen, Database development and Calphad calculations for high entropy alloys: Challenges, strategies, and tips, Mater. Chem. Phys. 210 (2018) 279–290. https://doi.org/10.1016/j.matchemphys.2017.07.082.
[19] R. Schmid-Fetzer, F. Zhang, The light alloy Calphad databases PanAl and PanMg, Calphad 61 (2018) 246–263. https://doi.org/10.1016/j.calphad.2018.04.006.
[20] A. Van De Walle, C. Nataraj, Z.-K. Liu, The Thermodynamic Database Database, Calphad 61 (2018) 173–178. https://doi.org/10.1016/j.calphad.2018.04.003.
[21] B. Hallstedt, The SGTE collection of binary datasets, Calphad 89 (2025) 102833. https://doi.org/10.1016/j.calphad.2025.102833.
[22] J. Zhang, R. Wang, Y. Zhong, Identification of the Eutectic Points in the Multicomponent Systems with the High-Throughput CALPHAD Approach, J. Phase Equilibria Diffus. 43 (2022) 844–857. https://doi.org/https://doi.org/10.1007/s11669-022-01003-1.
[23] J. Schmidt, M.R.G. Marques, S. Botti, M.A.L. Marques, Recent advances and applications of machine learning in solid-state materials science, Npj Comput. Mater. 5 (2019) 83. https://doi.org/10.1038/s41524-019-0221-0.
[24] J.E. Saal, A.O. Oliynyk, B. Meredig, Machine Learning in Materials Discovery: Confirmed Predictions and Their Underlying Approaches, Annu. Rev. Mater. Res. 50 (2020) 49–69. https://doi.org/10.1146/annurev-matsci-090319-010954.
[25] J. Fang, M. Xie, X. He, J. Zhang, J. Hu, Y. Chen, Y. Yang, Q. Jin, Machine learning accelerates the materials discovery, Mater. Today Commun. 33 (2022) 104900. https://doi.org/10.1016/j.mtcomm.2022.104900.
[26] Y. Zeng, M. Man, K. Bai, Y.-W. Zhang, Revealing high-fidelity phase selection rules for high entropy alloys: A combined CALPHAD and machine learning study, Mater. Des. 202 (2021) 109532. https://doi.org/https://doi.org/10.1016/j.matdes.2021.109532.
[27] Y. Zeng, M. Man, K. Bai, Y.-W. Zhang, Explore the full temperature-composition space of 20 quinary CCAs for FCC and BCC single-phases by an iterative machine learning + CALPHAD method, Acta Mater. 231 (2022) 117865. https://doi.org/https://doi.org/10.1016/j.actamat.2022.117865.
[28] Y. Xian, P. Dang, Y. Tian, X. Jiang, Y. Zhou, X. Ding, J. Sun, T. Lookman, D. Xue, Compositional design of multicomponent alloys using reinforcement learning, Acta Mater. 274 (2024) 120017. https://doi.org/https://doi.org/10.1016/j.actamat.2024.120017.
[29] G. Vazquez, S. Chakravarty, R. Gurrola, R. Arróyave, A deep neural network regressor for phase constitution estimation in the high entropy alloy system Al-Co-Cr-Fe-Mn-Nb-Ni, Npj Comput. Mater. 9 (2023) 68. https://doi.org/10.1038/s41524-023-01021-8.

[30] Y. Zeng, M. Man, C.K. Ng, Z. Aitken, K. Bai, D. Wuu, J.J. Lee, S.R. Ng, F. Wei, P. Wang, D.C.C. Tan, Y.-W. Zhang, Search for eutectic high entropy alloys by integrating high-throughput CALPHAD, machine learning and experiments, Mater. Des. 241 (2024) 112929. https://doi.org/https://doi.org/10.1016/j.matdes.2024.112929.
[31] M. Wu, B. Guan, J. Wang, S. Wang, C. Yang, C. Zhang, D. Shu, C. Xiao, B. Sun, Accelerated design of eutectic high-entropy alloys using high-throughput phase diagram calculations and machine learning, Mater. Des. 254 (2025) 114125. https://doi.org/https://doi.org/10.1016/j.matdes.2025.114125.
[32] Z.Q. Chen, Y.H. Shang, X.D. Liu, Y. Yang, Accelerated discovery of eutectic compositionally complex alloys by generative machine learning, Npj Comput. Mater. 10 (2024) 204. https://doi.org/10.1038/s41524-024-01385-5.
[33] Z. Chen, Y. Yang, Data-driven design of eutectic high entropy alloys, J. Mater. Inform. 3 (2023). https://doi.org/10.20517/jmi.2023.06.
[34] X. Li, X. Chen, Data-Driven Discovery of New Eutectic High-Entropy Alloys via Calculation of Phase Diagrams and Machine Learning Integration, Adv. Eng. Mater. 27 (2025) 2500879. https://doi.org/10.1002/adem.202500879.
[35] Y. Li, J. Tian, Z. Jiang, An advanced approach to directly calculate eutectic points in quaternary and higher-order slags: A case study on CaF2–Al2O3–CaO–MgO–SiO2 slag systems, J. Mater. Res. Technol. 35 (2025) 3432–3439. https://doi.org/10.1016/j.jmrt.2025.01.119.
[36] G. Deffrennes, K. Terayama, T. Abe, R. Tamura, A machine learning–based classification approach for phase diagram prediction, Mater. Des. 215 (2022) 110497. https://doi.org/10.1016/j.matdes.2022.110497.
[37] J.-O. Andersson, T. Helander, L. Höglund, P. Shi, B. Sundman, Thermo-Calc & DICTRA, computational tools for materials science, Calphad 26 (2002) 273–312. https://doi.org/10.1016/S0364-5916(02)00037-8.
[38] The Thermo-Calc Software TCAL09 Al-based Alloy Database, (n.d.). https://thermocalc.com/products/databases/aluminum-based-alloys/ (accessed April 3, 2026).
[39] G.H. Gulliver, The Quantitative Effect of Rapid Cooling Upon the Constitution of Binary Alloys, J Inst Met 9 (1913) 120–157.
[40] E. Scheil, Bemerkungen zur Schichtkristallbildung, Z Met. 34 (1942) 70–72.
[41] G. Kaptay, The Generalized Phase Rule, the Extended Definition of the Degree of Freedom, the Component Rule and the Seven Independent Non-Compositional State Variables: To the 150th Anniversary of the Phase Rule of Gibbs, Materials 17 (2024) 6048. https://doi.org/10.3390/ma17246048.
[42] G. Ghosh, C. Kantner, G.B. Olson, Thermodynamic modeling of the Pd-X (X=Ag, Co, Fe, Ni) systems, J. Phase Equilibria 20 (1999) 295–308. https://doi.org/10.1361/105497199770335811.
[43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine Learning in Python, (2012). https://doi.org/10.48550/ARXIV.1201.0490.
[44] R. Tamura, H. Morito, G. Deffrennes, M. Naito, Y. Nose, T. Abe, K. Terayama, AIPHAD, an active learning web application for visual understanding of phase diagrams, Commun. Mater. 5 (2024). https://doi.org/https://doi.org/10.1038/s43246-024-00580-7.

Appendix

Appendix A: Compositional Grid Size Details

In the context of selecting equidistant starting points, a grid size of k means that k points are considered along each compositional axis to construct the grid. The compositional space is then discretized accordingly, as in Figure A. When analyzing the methods, grid size is an independent variable that allows performance comparison based on an increasing number of starting points for CALPHAD-based methods, or training data for ML-based methods.

Due to their design, Adaptive liquidus descent, Solidus descent and Scheil descent (Sections 3.1.2., 3.1.3. and 3.1.4., respectively) exclude pure elements and lower-order systems from their starting points as it leads to early convergence or degenerate eutectics. These approaches consider only interior points, which are given by the following formula: $\frac{k!}{(k-(n-1))!(n-1)!}$, for $\boldsymbol{n \leq k}$ where k is the grid size and n the number of elements in the systems. An example for $k = n = 3$ is shown in Figure A (left). Note that for consistency, results in Figure 2. were computed with interior starting points for all methods, including the Steepest liquidus descent.

The design of ML-assisted approaches (Section 4.) does not impose any constraint for initial training points. The total number of grid points is given as: $\frac{(k+(n-2))!}{(n-1)!(k-1)!}$. An example for $k = n = 3$ is shown in Figure A (right).

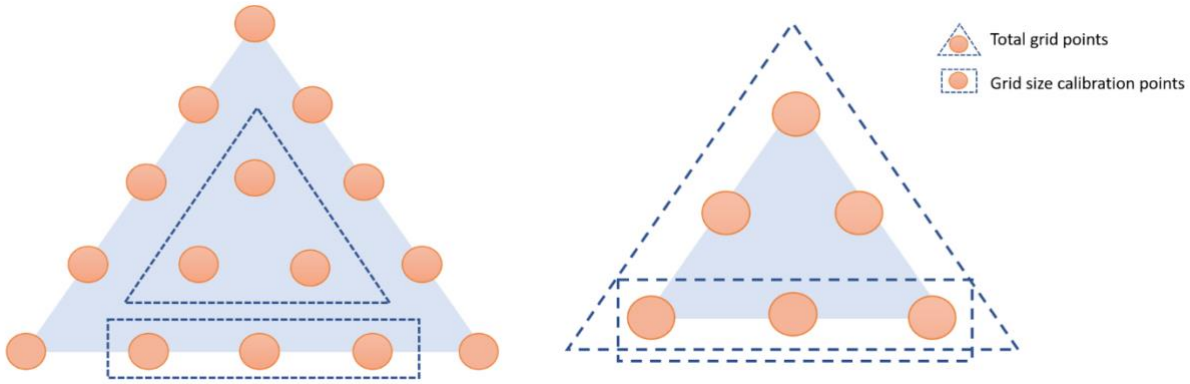


Figure A. Example of a discretized compositional space for a ternary system with a grid size of 3: three compositions are taken along each compositional axis, while the interior points are generated from all combinations of these compositions that sum to 100 at%.

When pure elements and lower-order systems are not considered as the starting points, grid size is defined by 3 points along the compositional axis, vertices excluded (left). This results in 3 grid points.

When all points on the axis are considered (right), there are 6 grid points in total.