A Computational Workflow for Cost-Effective Synthesis of Inorganic Materials: Integrating Thermodynamics, Cellular Automata, Machine Learning, and Commercial Databases

César Alonso Garrido
HI iberia
Madrid, Spain
calonso@hi-iberia.es

David García
HI iberia
Madrid, Spain

Eduardo Abenza
HI iberia
Madrid, Spain

Roberto Gómez-Espinosa
HI iberia
Madrid, Spain

María Carmen Asensio
Instituto de Ciencia de Materiales de Madrid (ICMM/CSIC)
MATINÉE: CSIC Associated Unit Between ICMUV and ICMM
Madrid, Spain

Abstract

This work presents a workflow designed to enhance economic efficiency in industrial processes by optimizing material synthesis through thermodynamic modeling, cellular automata, machine learning (ML), and commercial databases. The methodology enables the systematic identification and evaluation of synthesis routes that balance thermodynamic performance with economic profitability. As a proof of concept, the workflow was applied to the production of Ca_2SiO_4, a key refractory material for boilers and ship engines. Three efficient synthesis recipes and four optimal commercial reagents (CaO, SiO_2, Ca_3SiO_5, and Si_3N_4) were identified, demonstrating both improved process efficiency and cost-effectiveness. Beyond this case study, the approach is broadly applicable to a wide range of inorganic systems, offering a scalable path toward maximizing economic efficiency in solid-state synthesis. These results highlight the potential of data-driven workflows to accelerate the development of sustainable and competitive industrial manufacturing.

1 Introduction

Silicon-based compounds have gained increasing relevance in the naval and defense sectors due to their unique combination of mechanical strength, thermal stability, and corrosion resistance — properties that are essential for critical structural components. Materials such as $ZrSiO_4$, Al_2SiO_5, and Ca_2SiO_4 are widely employed in thermal coatings and heat barriers for naval engines and high-performance propulsion systems, owing to their high refractory capacity and excellent oxidation resistance. However, the synthesis of these compounds often entails high energy consumption and significant raw material demand. In this context, the development of novel and more cost-effective synthesis routes represents a strategic advancement for the sector.

Optimizing the synthesis of these compounds in terms of both profitability and performance is inherently a multifactorial challenge. It requires balancing thermodynamic efficiency, commercial adaptability, and technical effectiveness. Furthermore, the criteria and constraints for defining an optimal synthesis pathway are diverse, as summarized in Table 1. Even when focusing exclusively on

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

thermodynamic efficiency, the optimization problem is highly complex: it involves exploring millions of possible combinations of reactants and environmental conditions within the vast physicochemical space. This complexity increases further when economic considerations are introduced, since the profitability of a synthesis route depends not only on its energy and material efficiency but also on the price and commercial availability of the precursors. These economic variables are dynamic, being influenced by technological advances, regulatory frameworks, fiscal incentives, and geopolitical factors.

Table 1: Criteria and Constraints Governing the Optimal Synthesis of Materials.

	Technical	**Commercial**
Criteria	Thermodynamic performance, energy efficiency, isolation feasibility	Profitability, economic efficiency
Constraints	Purity and stability of reactants	Commercial availability, logistical capacity, legislation

To address this challenge, this work presents a developed workflow capable of systematically proposing the most cost-effective synthesis recipes by integrating thermodynamic modeling of chemical reactions with the commercial evaluation of reactants. The workflow is designed to be both versatile and customizable, allowing adaptation to different technical objectives and logistical conditions. This makes it a valuable tool for both manufacturing industries and academic research groups. Figure 1 illustrates the modular structure of the workflow, whose individual components are detailed in the following section.

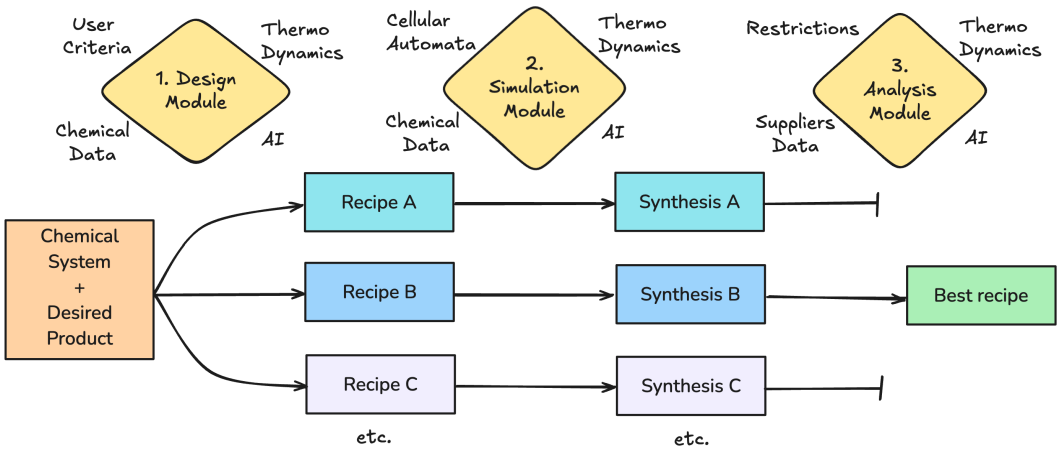

Figure 1: Modular flowchart for the identification of optimal synthesis recipes.

2 Methods

2.1 Design Module

Based on predefined conditions, the design module constructs a chemical space of materials from both computational and experimental origins, using databases such as Materials Project Jain et al. [2013] and JANAF Chase [1985]. Once this is completed, the module applies the SISSO-learned descriptor Bartel et al. [2018] to estimate the energetic properties of the materials at different synthesis temperatures, as extensively done in the pymatgen library Ong et al. [2013]. The training details of this standard method can be found in the cited reference. Next, the module uses the reaction-network library Kitchin [2018] to select the best synthesis recipes based on thermodynamic principles. Finally, the best proposed recipes are cross-checked against the PubChem database to discard those whose reactants are commercially unavailable. As a result, the module outputs a set of recipes that proceed to the next module.

2.2 Simulation Module

The simulation module is built on the React-CA cellular automaton framework Gallant et al. [2024]. The first task of this framework is to identify all possible reactions in the chemical system corresponding to a given synthesis recipe 'x' and to score them according to their probabilities of success. Once this is done, the framework executes the simulation of synthesis recipe 'x' as a series of preprogrammed thermodynamic stages. Essentially, at each stage of the simulation, the different reactants may undergo a change of state and position, or a chemical reaction from among those previously identified and weighted. After a significant number of generations, the system tends to evolve toward the products with the highest likelihood of occurrence. Both the scoring function for the possible reactions and the stopping criteria of the simulation are described in detail in the original paper. To enhance the interpretability of the results, the simulation module can generate comparative graphs that depict the entire evolutionary process of the simulated chemical systems using React-CA. Regarding the characteristic uncertainty of these processes, the authors account for it both at the reaction level and at the phase transition level. However, this uncertainty is not reflected in any way in the final yield results, which we acknowledge as a limitation of the workflow.

2.3 Analysis Module

The analysis module gathers the thermodynamic results generated by the simulation module and complements them with various customizable criteria and constraints to produce a range of descriptors tailored to the predefined needs and objectives. For example, while a manufacturing company can typically afford to purchase reactants in tons, an academic research group may not be able to acquire quantities above ten kilograms. Such a seemingly trivial constraint alters both the potential price of the reactants and the profitability analysis metrics, which can significantly change the final list of optimal recipes depending on the context. To achieve this versatility, the analysis module leverages commercial supplier databases such as ChemSpace. In addition, predictive AI models such as CrabNet Wang et al. [2021] will be applied to calculate descriptors related to the technical feasibility of both the synthesis itself and the subsequent isolation of the target products. In this way, the analysis module will be able to estimate the solubility or volatility of the reaction products, enabling recommendations at both the recipe level and the purification method level.

3 Results

To evaluate the applicability of the proposed workflow, a synthesis recipe recommendation experiment was conducted for Ca_2SiO_4, a compound of strategic importance in the naval industry. This material is widely used as a coating in boilers and marine engines due to its high refractory capacity. The search for improved synthesis routes was restricted to the chemical space defined by the elements Ca, Si, O, C, N, H, Al, and B, and to a temperature range of 800–1600 °C.

Upon execution, the analysis module suggested 261 synthesis recipes involving 36 commercially accessible reactants. The average execution time of the 261 simulations was 311.4 seconds, with a standard deviation of 121.26 seconds. The time distribution for these recipes is shown in detail in Figure 2. Additionally, Table 2 summarizes the most relevant statistics of the 36 reactants analyzed in terms of commercial availability. Figure 3 illustrates the dynamic evolution of the three best-performing synthesis recipes, selected on the basis of economic efficiency. This metric was calculated by jointly considering the thermodynamic performance of each synthesis process and the average market price of the required reactants (minimum purity 95%). The final results are presented in Table 3, which compares the three recipes across their main evaluation metrics.

Table 2: Statistics of the 36 commercial reactants analyzed for the synthesis of Ca2SiO4.

Min. price	Avg. price	Max.price	Avg. purity	Avg. delivery	Avg. suppliers
0.015 (€/g)	100.74 (€/g)	8526.31 (€/g)	91.79 (%)	17.19 (days)	6.18

Notably, the recipe combining CaO and SiO_2 emerges as nearly twenty times more cost-effective than the alternatives, despite being approximately 30% less productive in thermodynamic terms. This trade-off arises from the currently low price of CaO relative to compounds such as Ca_3SiO_5 or Si_3N_4.

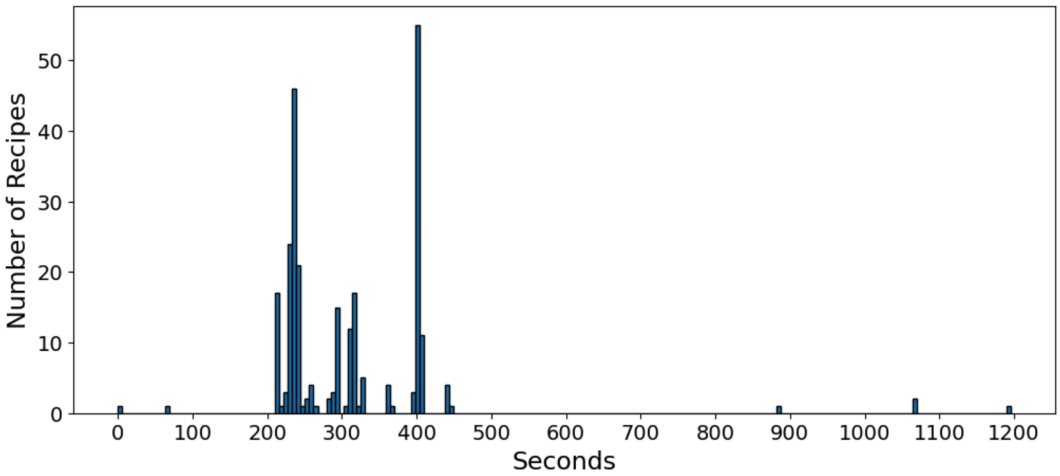

Figure 2: Distribution of simulation times for the Ca2SiO4 recipes. The simulations were run on an Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz.

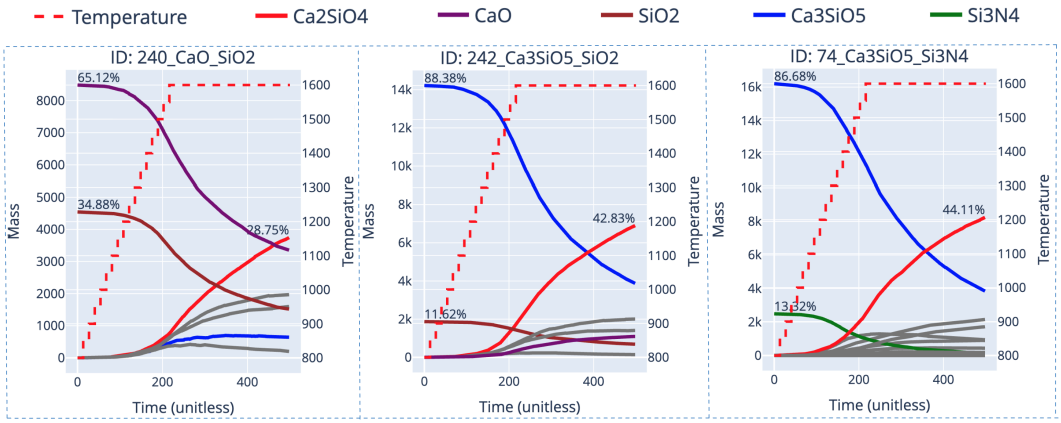

Figure 3: Simulation of the three best Ca2SiO4 recipes in terms of grams produced per euro invested. Gray lines correspond to undesired synthesis byproducts, and the percentages are calculated by mass. The time variable is representative and does not reflect actual values, as the simulations consider thermodynamic principles rather than kinetics.

However, shifts in raw material prices could reverse this situation, making the latter reactants more profitable in the future. In an era of geopolitical turbulence and volatile supply chains, access to such predictive insights can provide a decisive competitive advantage, potentially determining the economic resilience of manufacturing companies. Importantly, the methodology described here can be extended to generate similar analyses for thousands of technologically relevant materials.

Despite its promise, the current implementation of the tool is subject to certain limitations. First, the enumeration of all possible reactions within a chemical system is computationally demanding, with costs scaling steeply with the number of chemical elements considered. This constraint reduces the breadth of the search space and highlights the need for further research into computational acceleration strategies, such as process parallelization or GPU-based implementations. Moreover, the present economic efficiency metrics do not yet account for factors such as the energy consumption associated with heating (heat capacities of reactants and products) or the purification costs required to isolate the target compound from byproducts. Future developments of the workflow will incorporate predictive models capable of estimating key physicochemical properties (e.g., solubility, volatility, and phase stability), thereby enabling the construction of more accurate and robust descriptors. Such improvements will enhance the precision of economic assessments for Ca-based silicates, but still remain limited to solid-state synthesis and do not apply to methods like sol-gel or hydrothermal.

Table 3: Performance descriptors of the 3 best synthesis recipes for Ca2SiO4.

Recipe	Economic efficiency (avg. market price) (g/€)	Economic efficiency (min. market price) (g/€)	Thermodynamic yield (mass %)
CaO+SiO2	0.977	1.313	28.75
Ca3SiO5+SiO2	0.051	0.170	42.83
Ca3SiO5+Si3N4	0.045	0.177	44.11

4 Conclusion

This study demonstrates that the most thermodynamically favorable synthesis pathway is not always the most economically viable, as profitability can be strongly influenced by market fluctuations in reactant prices. The results emphasize the strategic value of the proposed workflow, which enables the anticipation of optimal synthesis strategies under dynamic economic and geopolitical conditions. By systematically combining thermodynamic modeling, cellular automata simulations, and commercial data, the workflow provides a versatile tool for supporting decision-making in both industrial and academic contexts.

Nevertheless, the current implementation presents limitations. The computational cost of exhaustively exploring chemical reaction networks increases sharply with the number of elements considered, constraining the size of the accessible search space. In addition, the present economic metrics do not yet incorporate critical variables such as energy consumption for heating and cooling cycles, or the costs associated with the purification of target compounds relative to synthesis byproducts. Moreover, it would be advisable to add an explicit treatment of uncertainty in both the simulation and analysis modules.

Future work will focus on addressing these challenges by integrating predictive models capable of estimating key physicochemical properties such as solubility, volatility, and phase stability. These enhancements will enable the construction of more accurate and robust descriptors of synthesis feasibility and economic efficiency. Ultimately, the continued development of this workflow has the potential to accelerate the design of cost-effective synthesis routes across a wide range of inorganic materials, contributing to more sustainable and competitive industrial solid-state syntheses.

References

Christopher J. Bartel, Samantha L. Millican, Ann M. Deml, John R. Rumptz, William Tumas, Alan W. Weimer, Stephan Lany, Vladan StevanoviÄ‡, Charles B. Musgrave, and Aaron M. Holder. Physical descriptor for the gibbs energy of inorganic crystalline solids and temperature-dependent materials chemistry. *Nature Communications*, 9(1), 2018.

Malcolm W. Chase. *JANAF Thermochemical Tables, 3rd Edition*. American Chemical Society and American Institute of Physics for the National Bureau of Standards, 1985.

Max C. Gallant, Matthew J. McDermott, Bryant Li, and Kristin A. Persson. Reactca: A cellular automaton for predicting phase evolution in solid-state reactions. *arXiv preprint*, 2407.19124, 2024. `https://doi.org/10.48550/arXiv.2407.19124`.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013.

John R. Kitchin. reaction-network: A python library for constructing and analyzing chemical reaction networks, 2018. Available at `https://github.com/jkitchin/reaction-network`.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis, 2013. Available at `https://github.com/materialsproject/pymatgen`.

Alex Wang, Steven K. Kauwe, Ryan Murdock, and Taylor D. Sparks. Compositionally restricted attention-based network for materials property prediction. *npj Computational Materials*, 7(1):77, 2021.