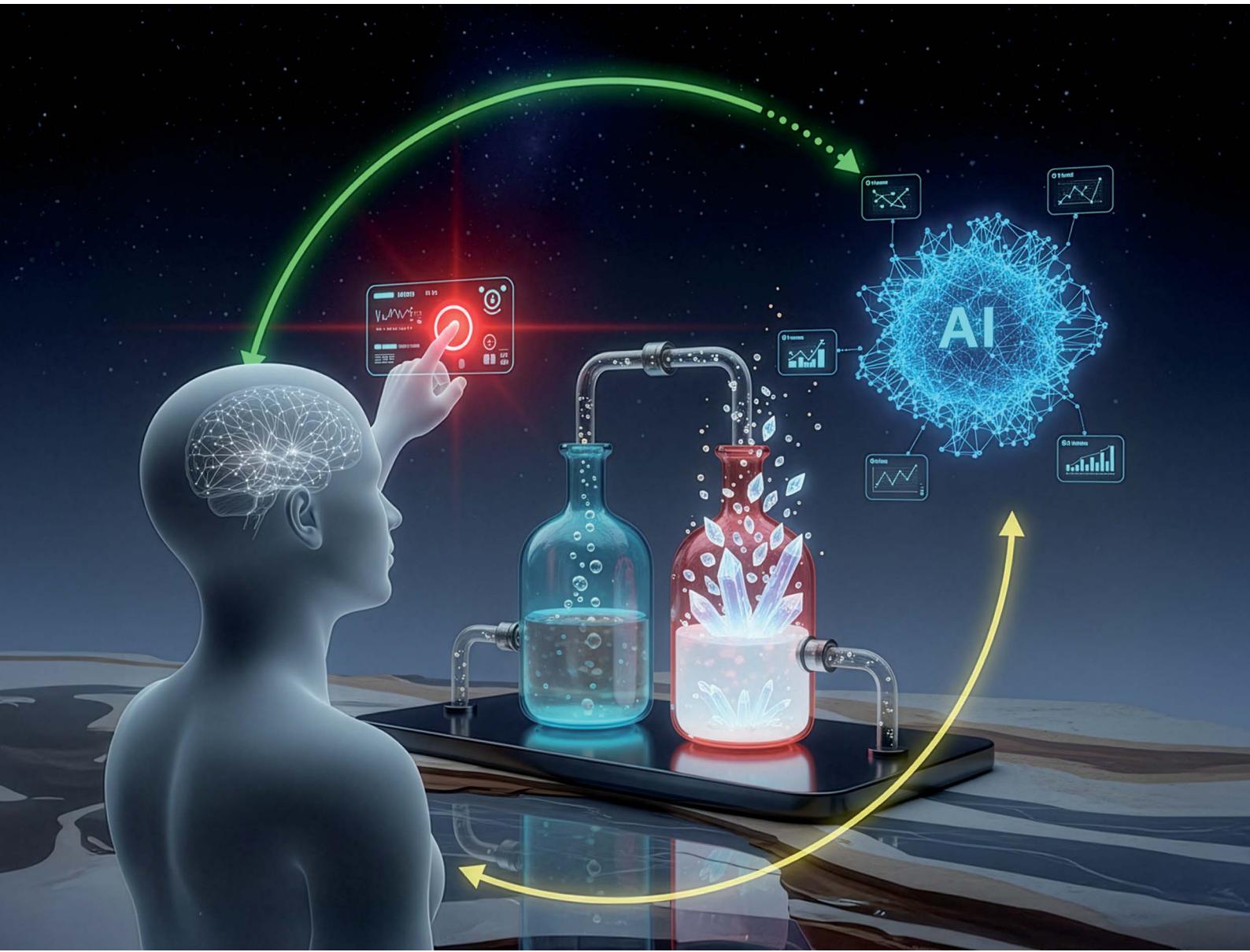


Digital Discovery

rsc.li/digitaldiscovery



ISSN 2635-098X



Cite this: *Digital Discovery*, 2025, 4, 3078

Human-AI synergy in adaptive active learning for continuous lithium carbonate crystallization optimization

Shayan Mousavi Masouleh, ^{ab} Corey A. Sanz, ^c Ryan P. Janssonius, ^c Cara Cronin, ^c Jason E. Hein ^c and Jason Hattrick-Simpers ^{*a}

As the demand for high-purity lithium surges, primarily fueled by the adaptation of the electric vehicle (EV) industry, the need for cost-effective extraction and purification technologies intensifies. The Smackover Formation in southern Arkansas, recently identified as one of the world's largest lithium resources, offers vast potential. This formation is part of a broader array of lithium resources across North America, many of which possess lower-grade lithium compared to renowned sources like South American brines. These alternative formations, while presenting significant opportunities, require innovative purification techniques to make their exploitation economically viable. Continuous crystallization is a promising method to produce battery-grade lithium carbonate from these lower-grade sources. Yet, the optimization of this process is challenging due to its complex parameter space, often constrained by scarce data. This study introduces a Human-in-the-Loop (HITL) assisted active learning framework aimed at adapting and optimizing the continuous crystallization process of lithium carbonate. By integrating human expertise with data-driven insights, this approach significantly accelerates the optimization of lithium extraction from challenging sources. Our results demonstrate the framework's ability to rapidly adapt to new data, improving the process's tolerance to critical impurities, such as magnesium, by industry practices at a few hundred ppm, and extending it to handle contamination levels as high as 6000 ppm. This makes the use of low-grade lithium resources contaminated with such impurities feasible, potentially reducing overhead processes. By leveraging artificial intelligence, we not only refined the operational parameters but also demonstrated a potentially reduced need for extensive pre-refinement, promoting the use of lower-grade materials without sacrificing product quality. This advancement marks a significant step towards economically harnessing North America's lithium reserves, particularly those in the Smackover Formation, thereby contributing to the sustainability of the lithium supply.

Received 28th June 2025
Accepted 21st September 2025

DOI: 10.1039/d5dd00285k
rsc.li/digitaldiscovery

1 Introduction

Demand for high-purity lithium (Li), essential for batteries in electronics and vehicles, is projected to outstrip supply by 2035.¹⁻³ As the demand continues to escalate, lithium-ion battery demand is expected to grow by 27% annually.⁴ This volatility underscores the urgency for more economically viable extraction methods as the traditional techniques remain cost-intensive and environmentally taxing.⁵ While total global lithium resources are estimated at 527 million tonnes—far exceeding the annual consumption rate of 0.5 million tonnes—much of this lithium is sourced from low-grade deposits that

are expensive to exploit due to their complex geochemistry and the presence of impurities.^{1,2}

Among these low-grade sources, the Smackover Formation in southern Arkansas stands out as a significant resource. Identified as potentially one of the world's largest lithium reserves, the Smackover Formation features high concentrations of lithium in brines—over 400 milligrams per liter—which are currently brought to the surface as waste streams from the oil, gas, and bromine industries.⁶⁻¹⁰

However, exploiting these resources is not straightforward. The Smackover's lithium-bearing brines are characterized by a high ratio of impurities to lithium, approximately 1000 atoms of impurity for every atom of lithium. These impurities include elements such as sodium (Na), potassium (K), magnesium (Mg), and calcium (Ca), closely related to lithium on the periodic table, which complicates their separation due to their similar solubility, charge, and mass properties.⁶ This challenge necessitates a shift from traditional methods to innovative extraction

^aCanmet MATERIALS, Natural Resources Canada, 183 Longwood Rd S, Hamilton, ON, Canada. E-mail: jason.hattrick-simpers@nrcan-rncan.gc.ca

^bClean Energy Innovation, National Research Council of Canada, 2620 Speakman Dr, Mississauga, ON, Canada

^cTelescope Innovations, 301-2386 E Mall, Vancouver, BC, Canada



techniques that are both economically viable and environmentally sustainable.

Traditional methods for producing battery-grade Li_2CO_3 are costly and involve multiple steps, including lengthy evaporation processes with high water usage.^{11–14} Exploiting lower-grade brines would be even more expensive due to their dilution, necessitating process improvements to reduce reagent, solvent, and water usage.^{2,15} In contrast, continuous crystallization—a technique commonly used in the pharmaceutical industry but rarely in mining—is well-suited for this purpose.^{16,17} This method achieves metal salt purities of 90–99.9% and eliminates the need for evaporative concentration,^{18–24} thus, reducing water and land use as well as production time.¹⁵ According to a United States Department of Energy assessment, direct lithium extraction using continuous crystallization can cut production costs by 24% compared to traditional evaporative methods, enabling cost-effective purification of Li_2CO_3 from low-grade brines in a single step.^{2,15}

Nonetheless, the continuous crystallization process is not without its challenges, particularly when applied to the low-grade lithium brines such as those found in the Smackover Formation.^{6,7,19} In this regard, the primary challenge is not merely concentrating the lithium, which can be achieved through methods like evaporation, reverse osmosis, or solar concentration, but effectively rejecting closely related impurities (Na, K, Ca, Mg). To achieve high-purity lithium carbonate (Li_2CO_3), the crystallization process must be finely tuned: a dilute solution might lead to sparse crystal formation and low mass recovery, while a too concentrated solution can cause rapid nucleation, trapping impurities within the crystals. Adding to the complexity, Li_2CO_3 exhibits inverse solubility, decreasing from 13 g L^{-1} at 20°C to 8.6 g L^{-1} at 80°C . Thus, continuous crystallization needs to be adapted to be either impurity-tolerant or highly selective for lithium, minimizing the need to remove every challenging impurity. Optimizing such intricate chemical operations is a formidable task due to the high dimensionality of the control space and the sophisticated chemistry involved.

Traditional optimization strategies, such as conventional design of experiment (DOE) methodologies, often involve exploring a vast experimental space. In our study, we initially identified 10 critical variables, which, under a full factorial DOE, would necessitate about more than approximately 1024 experiments. Given our experimental throughput constraints, limited to about four per week, conducting such a large number of experiments was impractical. This limitation highlighted the need for an alternative approach that could achieve optimization with fewer experiments.

Over the past decade, various AI-driven active learning solutions have emerged, treating the optimization task as a search through a “black box.” Among these, Bayesian optimization methods have been widely used as a principled framework for guiding experimental selection in data-efficient ways. These methods have led to more efficient exploration and optimization of complex systems.^{25–31} Recent contributions, such as M. Lazin *et al.*³² and K. Yang *et al.*,³³ have extended Bayesian optimization approaches to better handle cases where

high dimensionality is the dominant challenge under minimal constraints. However, our process involves moderate-to-high dimensionality coupled with strong practical constraints, interdependent variables, and nonlinear process responses, which present different challenges that can still limit optimization speed and efficiency in practice. These limitations reflect a broader challenge: AI-driven active learning approaches rely heavily on machine learning models, which build correlations without necessarily providing causal understanding or integrating well-established, yet hard-to-quantify, heuristics of the physical and chemical nature of the system under study. This reliance can result in a process that remains time-consuming and resource-intensive. Furthermore, design biases may arise from a limited understanding of the system’s complexities, leading to skewed outcomes and compounding errors. These biases can hinder the model’s ability to generalize, impacting its overall performance and reliability.

To mitigate biases and address complexity challenges, the field of human-in-the-loop AI has emerged as a promising solution.^{34–36} This approach leverages the collaboration between human intelligence and artificial intelligence.³⁴ Human cognitive abilities and domain expertise play a crucial role in enhancing AI’s predictive capabilities by interpreting data-driven correlations and offering intuition-driven insights. These insights help in refining the evaluation process of AI models, ensuring that the models align more closely with real-world complexities and expectations. Human-in-the-loop AI involves human input in tasks such as data collection, algorithm selection, and model tuning, creating a feedback cycle that helps reduce biases in both human decision-making and AI predictions.³⁵ This collaborative framework allows for quick adjustments and a deeper understanding of the machine learning workflow, ensuring that AI-driven systems become more adaptive, efficient, and effective in managing complex optimization tasks. This integration not only addresses the limitations of each approach but also combines their strengths to improve overall outcomes in high-dimensional optimization environments.^{34–36}

In this paper, we present a human-in-the-loop-assisted active learning (HITL-AL) AI framework specifically designed to optimize a continuous crystallization technique for producing high-quality, battery-grade lithium from diverse low-grade brines. This approach directly addresses the challenges posed by feedstocks like the Smackover Formation’s complex geochemistry, where high impurity levels closely related to lithium complicate traditional extraction methods. Our objective extends beyond merely adjusting control parameters such as reactor temperatures and flow rates. It also involves investigating the interactions between brine compositions and system controls, particularly the interplay of major contaminants like Na, Ca, Mg, and K, to determine how much contamination and dilution can be tolerated while still producing battery-grade outcomes. By doing so, we aim to unlock the potential of these abundant but challenging resources, reducing reliance on high-quality sources and advancing cost-effective, sustainable lithium extraction methods.



In our HITL-AL process, human experts play a central role in refining machine learning-suggested experiments, using their judgment to focus on those most likely to yield meaningful results. This strategic selection is crucial for conducting experiments within practical throughput constraints while exploring promising pathways that models might overlook. Moreover, human experts are integral to evaluating outcomes and adjusting both hypotheses and workflows. This involves developing new hypotheses from emerging data and rigorously testing these ideas, helping to identify and correct biases in design and chemical assumptions, such as the difficulty of impurity removal and the ranges of control parameters. The methods section provides detailed insights into these processes and how different hypotheses are tested.

This iterative approach not only reduces the number of experiments required but also uncovers critical insights that drive innovation. For instance, through this collaborative and adaptive process, we discovered that adjusting cold reactor temperatures significantly reduces magnesium impurities. This counterintuitive breakthrough was achieved with minimal observations, demonstrating the effective synergy of human intuition and AI analysis. This finding significantly enhances the production of battery-grade lithium by expanding the acceptable range of magnesium contamination levels. We explore these discoveries further, illustrating the framework's effectiveness in optimizing lithium production. Additionally, we highlight the broader implications of our findings, showing how the integration of human expertise and AI not only improves experimental efficiency but also provides a sustainable and cost-effective solution to challenges posed by low concentration and polluted brines, such as those found in the Smackover Formation.

2 Methods

This study introduces a Human-in-the-Loop Active Learning (HITL-AL) framework designed to optimize a continuous lithium crystallization process from low-grade brines, specifically brines with high levels of sodium (Na), magnesium (Mg), calcium (Ca), and potassium (K). These impurities, prevalent in the Smackover Formation's brine sources, significantly complicate the extraction and purification of lithium. The primary objective is to ascertain operational conditions within our experimental setup that can efficiently convert these challenging brines into battery-grade lithium carbonate, even when heavily contaminated with these impurities (see Table S1 for specifications of battery-grade lithium carbonate used in this study). Given our limited capacity to conduct experiments—restricted to approximately four per week due to resource and time constraints—it is crucial to maximize the efficiency and impact of each experimental run. Our integrated workflow synergizes experimental and computational efforts, facilitated by AI and human expertise, in a cyclic process that progressively refines our understanding and control of the crystallization system; see Fig. 1. Detailed descriptions of the experimental setup are provided in Section 2.1.

Upon initiating the HITL-AL cycle, as illustrated in Fig. 1, the “data and results inspection module” acts as the initial juncture. This module integrates cumulative experimental observations into our workflow. It features a statistical dashboard that compiles and analyzes incoming data. Detailed information on these statistical analyses can be found in Section 2.2.1. This dashboard provides human experts with insights into parameter interdependencies and correlations, crucial for real-time monitoring and timely response to experimental findings.

As outlined in Fig. 1, following the inspection step, the “human interpretation” phase occurred. In this phase, experts apply their scientific and intuitive insights to the results of the statistical analysis. This step emphasizes the integration of human expertise into the cycle, thus avoiding the exhaustive need to fully automate complex decision-making in R&D setups.

Leading directly from the interpretation step, is the “informed design and adjustment” phase. In this phase, the insights obtained from human interpretation are employed to meticulously adjust and define the configurations, ranges, and constraints within the surrogate space—a multidimensional representation of experimental conditions. Each point within this space specifies operational controls such as temperature and flow rate, as well as reactant concentrations of lithium, sodium, calcium, and magnesium. Experts selectively refine input features and output targets based on these insights, ensuring that both machine learning model settings and experimental parameters are meticulously aligned with our objectives to maximize lithium yield and purity.

Building on the “informed design and adjustment” phase, the ML model prediction phase leverages these refined inputs to train machine learning models. These models predict crystallization outcomes, ensuring the configurations within the surrogate space are optimized for maximum lithium yield and purity. Details about the model configurations, training, and operational parameters are provided in Section 2.2.4.

In the “data acquisition” phase, predictions from the machine learning models guide the generation of a list of potential experiments, as detailed in Section 2.2.5. These suggestions, coupled with model predictions, are rigorously reevaluated within the “data and results inspection and interpretation modules.” Here, statistical analyses help human experts assess the performance of the models and the efficacy of the data acquisition process. During the interpretation phase, the viability of the experiments is evaluated. If the assessments determine that the experiments are feasible the process advances to the experimentation phase. If not, the HITL-AL cycle proceeds with further iterations, adjusting the surrogate space, data acquisition strategies, and model parameters until the experimental plan is finalized.

Once experiments are authorized for “experimentation,” experts review the suggestions, select the most feasible experiments, and conduct them along with reproducibility tests to verify data quality.

This continuous, iterative cycle of review and refinement, depicted in Fig. 1, enhances the capacity to efficiently produce battery-grade lithium from low-grade brines. Subsequent sections will delve deeper into the experimentation phase and



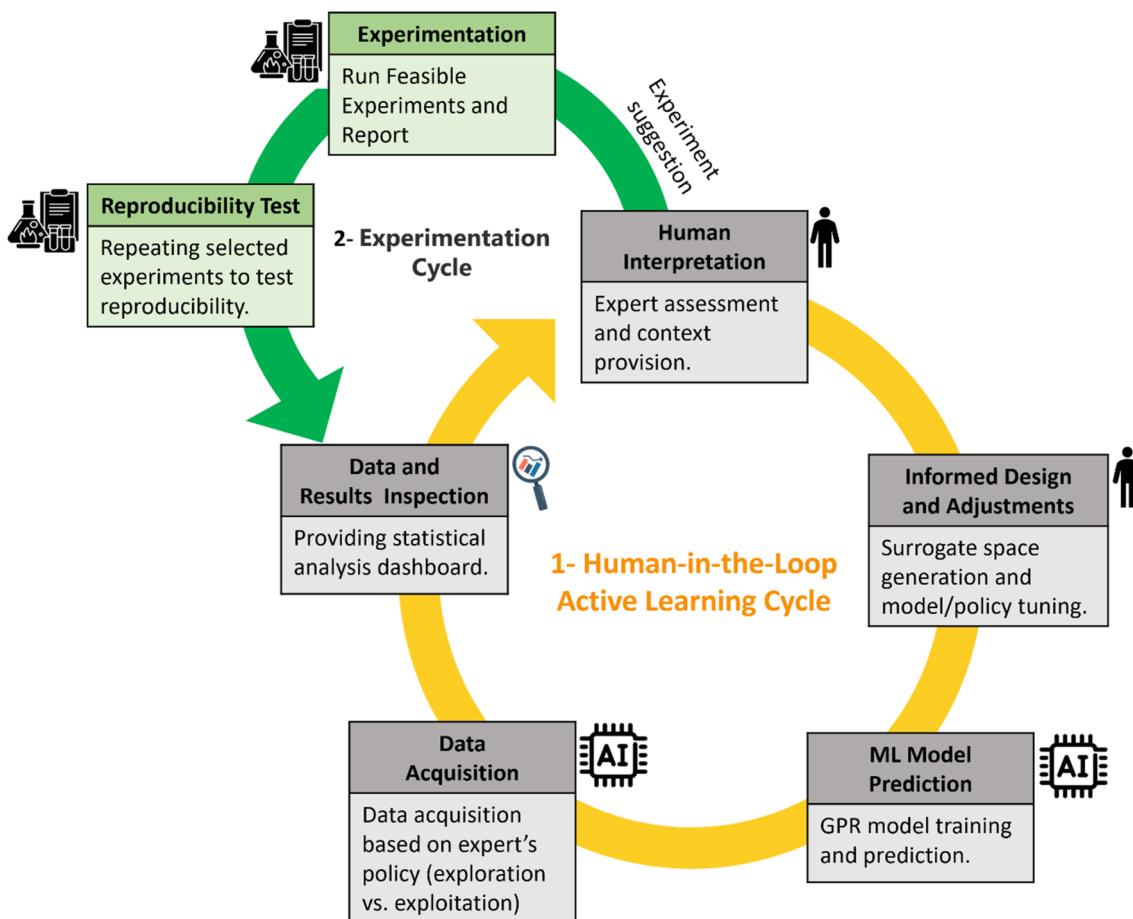


Fig. 1 Overview of the integration between the experimentation cycle and the human-in-the-loop active learning cycle, highlighting the different steps or modules within each phase.

detail the integration of the HITL-AL cycle, illustrating how each step contributes to achieving optimal outcomes.

2.1. Experimental procedures

The experimental portion of this study consisted of a dual-reactor continuous crystallization setup, inspired by mixed suspension mixed product removal crystallizer (MSMPR) designs used in the pharmaceutical industry.^{16,37} The crystallization setup consisted of two, 400 mL reactors set at different temperatures (e.g. 25 °C, and 70 °C respectively; Fig. 2a and b). Lithium carbonate crystallizes in the higher temperature reactor due to its lower solubility in hot solutions, while crude lithium carbonate solids in the colder reactor slowly dissolve and maintain a steady concentration of lithium carbonate in solution. Liquid was continuously circulated between the two reactors with peristaltic pumps while the flow of solids between reactors was restricted by polyethylene filters slotted into the tubing between the reactors.

At the start of the process, the cold reactor was loaded with the initial solution (e.g., low-grade Li₂CO₃ brine). The concentration of various elements in the input solution was monitored before loading the cold reactor. At the end of the crystallization process, the solution was unloaded and filtered from the hot

reactor, where the crystallized purified lithium was collected. After unloading the product from the hot reactor, the concentration of different elements in the resulting crystals was measured using Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES).

The continuous crystallization process was designed to isolate battery-grade Li₂CO₃ from a Li₂CO₃ impure feedstock; see Fig. 1a for a schematic overview. The initial crude material for this study was primarily synthesized to simulate various real-world lithium carbonate crudes with typical contaminant levels, including Mg, Ca, K, and Na. To ensure reproducibility, we also used industry-grade feedstock contaminated with a similar set of elements. Control parameters managed within the system included the temperatures of the cold and hot reactors, defining the temperature differential, as well as the pH, stirring rate, slurry concentration, flow rate between reactors, and retention time.

All experiments conducted were systematically recorded in tables and passed to the HITL cycle. Each experiment was assigned a quality score (1, 2, or 3) based on the process expert's observations. A score of 1 indicated no anomalies, such as sedimentation in containers. A score of 2 was given for minor occurrences of such issues, while a score of 3 was assigned to



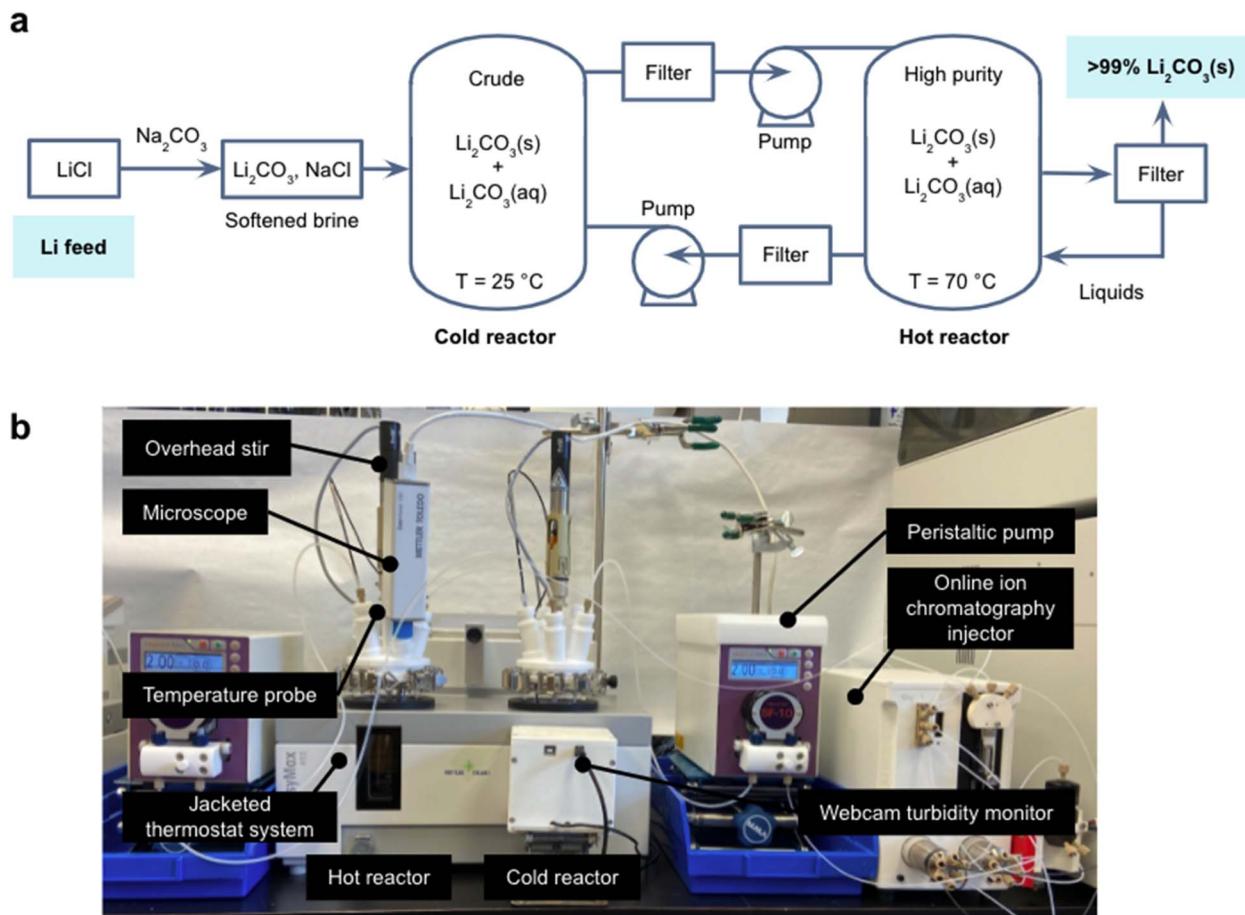


Fig. 2 (a) Simplified process flow diagram of the continuous hot-cold crystallization setup. (b) Image of benchtop setup used for processing LiCl brines into battery grade Li_2CO_3 .

conditions where noticeable anomalies, like sedimentation on container walls, were observed. Additionally, specialists provided comments describing any anomalies noticed during the experimentation.

Throughout various cycles of active learning and experimentation, specialists conducted reproducibility tests. These tests primarily focused on experiments that showed anomalies or unusual results, including a random selection of previous experiments. If contradictory results were observed, the scores could be adjusted. Experiments with significantly divergent outcomes were deemed failed and excluded from further analysis.

2.2. Human-in-the-loop active learning framework

The HITL-AL framework designed for this study consisted of five iterative steps: data and results inspection, human interpretation, informed design and adjustments, machine learning model prediction, and data acquisition. The interactions among these steps were discussed earlier in the opening of Section 2. This section delves deeper, providing more detailed information about each step.

2.2.1. Data and results inspection. In the data and results inspection step, experimental (observed) and simulated (ML-

predicted surrogate space) data were structured into comprehensive tables, including ICP-OES elemental concentrations results (Ca, K, Li, Mg, Na in ppm) before and after crystallization, alongside operational parameters including reactor temperatures ($^\circ\text{C}$), flow rate (mL min^{-1}), and slurry concentration (g per 100 mL).

Statistical analyses were systematically performed to explore correlations and identify key parameters influencing process outcomes. Initially, observed experimental data were analyzed to establish baseline relationships. After the first active learning iteration, analyses were extended to incorporate simulated surrogate space conditions and machine learning predictions, thereby elucidating model predictions and highlighting important signals identified by the models.

Pearson correlation analyses quantified linear relationships between input parameters and final impurity levels. SHAP (SHapley Additive exPlanations) values clarified the relative contributions of each feature to model predictions, while sensitivity analyses evaluated the robustness of these predictions by perturbing parameters individually by one standard deviation. The SHAP and sensitivity analyses utilized Random Forest Regressors (RFR) due to their interpretability and straightforward hyperparameter tuning. Hyperparameters—



including the number of estimators, maximum depth, minimum samples split, and minimum samples per leaf—were optimized using the Tree-structured Parzen Estimator (TPE) algorithm *via* Optuna. Collectively, these analyses provided insights into the underlying data relationships and clarified model behaviors.

2.2.2. Human interpretation. Human experts assessed analytical outcomes guided by structured questions across three key domains: process understanding, impurity management, and model evaluation. For process understanding, experts investigated how variations in operational parameters—such as reactor temperatures, flow rates, and slurry concentration— influenced impurity removal. They sought to identify parameter ranges at which these influences became significant. Under impurity management, assessments clarified which impurities (*e.g.*, Mg and Ca) were effectively managed under existing experimental conditions and which remained resistant or challenging. In model evaluation, experts critically examined whether ML models produced unrealistic or non-physical predictions, such as outputs significantly deviating from observed experimental distributions or predictions lacking expected variability, given the anticipated suboptimal model performance typical of low-data active learning scenarios.

If it was the initial active learning iteration, experts proceeded directly to the informed design and adjustments step. However, in subsequent iterations, experts only moved forward to the informed design and adjustments step when unsatisfactory outcomes or unexpected ML behaviors were identified. These actions included retraining ML models with adjusted hyperparameters, refining surrogate space parameter boundaries, or conducting targeted verification experiments—such as deploying random walkers (described further in Section 2.2.3)—to clarify anomalies and potential biases. These iterative refinements, guided by human insights, ensured continuous alignment among model predictions, experimental design, and chemical process understanding.

2.2.3. Informed design and adjustments. The surrogate space was constructed using Latin hypercube sampling to uniformly distribute multidimensional combinations of initial elemental concentrations (Ca, K, Li, Mg, Na) and operational parameters (reactor temperatures, flow rates, slurry concentrations). Feature selection and feasible parameter ranges were primarily informed by statistical analyses, including SHAP feature importance scores and sensitivity analysis, alongside operational constraints. Explicit constraints—such as maintaining a minimum temperature differential between reactors ($T_{\text{hot}} \geq T_{\text{cold}} + 20 \text{ }^{\circ}\text{C}$, later reduced to $2 \text{ }^{\circ}\text{C}$, as detailed in the Results section) and normalizing total elemental concentrations to one million ppm—were strictly enforced. When human interpretation indicated insufficient exploration or unexpected predictive anomalies, surrogate space boundaries were adjusted accordingly.

Experts evaluated feasible parameter ranges by identifying biases, model inconsistencies, and persistent trends suggesting areas for further investigation. If surrogate space coverage was deemed inadequate, a random walk algorithm was activated for targeted verification, deploying 100 000 random walkers.

Walkers were initialized near flagged regions and took randomized steps within $\pm 25\%$ of identified Pareto frontier boundaries, systematically probing overlooked parameter ranges.

Decisions regarding machine learning model strategies—whether to continue exploratory analyses or transition toward exploitation—were informed by empirical results and experimental observations rather than purely computational predictions. The exploratory analysis phase primarily identified key parameters and their critical ranges necessary for achieving battery-grade outcomes. Once sufficient understanding was reached, the strategy shifted to exploitation, focusing on defining decision boundaries between battery-grade and non-battery-grade conditions, particularly in relation to Mg impurity tolerance. The decision to transition between strategies or initiate targeted verification was systematically informed by human interpretation, ensuring continuous alignment between the experimental workflow and evolving process insights.

2.2.4. ML model predictions. Gaussian Process (GP) were selected for their model-agnostic nature, flexibility, and Bayesian properties, providing uncertainty quantification valuable for process optimization in low-data active learning settings.^{25,38-41} They are also widely used as surrogate models in Bayesian optimization, including in recent work addressing high-dimensional optimization challenges in the physical sciences.^{32,33} GP Regressor (GPR) models predicted post-crystallization impurity concentrations, while GP Classifier (GPC) models classified experimental conditions as battery-grade or non-battery-grade. As detailed in Section 2.2.3, input features and prediction targets were selected during the informed design and adjustments step. Thus, GP models were trained using elemental concentrations (Ca, K, Li, Mg, Na) and process control parameters (reactor temperatures, temperature differences, flow rates, slurry concentrations) as input features. GPR models targeted post-refinement impurity concentrations, and GPC models used a binary battery-grade label (true/false) as prediction targets.

Separate GPR models were developed for each target outcome, employing a differentiable Matern kernel, tuned using the Tree-structured Parzen Estimator (TPE) algorithm. Expert evaluations of predictive accuracy informed further hyperparameter refinements and model configurations.

During the initial exploration, GPR models predicted post-crystallization elemental concentrations, aiding experts in identifying key parameters that influence impurity reduction. Battery-grade label predictions using GPC were subsequently incorporated to delineate boundaries between battery-grade and non-battery-grade outcomes, providing insights into how impurity levels impact lithium carbonate purity. Standard scaling was applied to both the training dataset (features only) and surrogate simulation datasets to ensure consistent model performance by maintaining uniform feature magnitudes and preventing numerical inconsistencies in predictions.

2.2.5. Data acquisition. The role of the data acquisition step is to suggest experiments based on predictions of the ML models. However, it should be noted that these suggestions are not guaranteed to be directly implemented by the



experimentalists. In fact, in high dimensional experimental space and low data regimes, sub-optimal predictions are inevitable. Therefore, these suggestions may be adjusted according to the knowledge of experimentation experts.

The data acquisition in this study has two strategies. The first strategy emphasized exploratory analysis in the initial active learning cycles. After gaining some insight an information decision boundary exploitation was followed.

2.2.5.1 Exploratory analysis. The initial phase involved methodical exploration of the experimental space using GPR predictions of the surrogate space using algorithms like non-dominant sorting of genetic algorithm (NSGA-II), Pareto frontiers were extracted to pinpoint conditions that GPR models believed to minimize key impurities such as calcium and magnesium. These predictions were reviewed by human experts who adjusted experimental parameters to ensure they were practically and theoretically viable.

2.2.5.2 Decision boundary exploitation. GPC model was used to classify experimental outcomes as either non-battery-grade or battery-grade. The model calculated the probability of each experimental condition yielding a battery-grade product, with the decision boundary defined at a class probability of 0.5. To interrogate this boundary, a ray-tracing-inspired algorithm was implemented. This algorithm systematically paired each non-battery-grade data point with its nearest battery-grade counterpart to calculate their mathematical midpoint. These midpoints, particularly those located closest to the decision boundary, were prioritized for subsequent experiments to efficiently refine the conditions for producing battery-grade lithium.

3 Results and discussions

Optimization of the continuous lithium crystallization process began with analysis of 16 preliminary experiments conducted before active learning integration. These initial experiments, selected based on expert judgment, consistently produced lithium carbonate (Li_2CO_3) with purity exceeding 99%, as shown in Table S2.

Post-refinement analyses showed that Na impurities consistently fell below the project's set battery-grade threshold of 500 ppm and for most tests even under the more restrictive limit of 250 ppm, as introduced in Table S1, demonstrating this without additional purification treatments. K concentrations, although occasionally exceeding the strict battery-grade threshold of 10 ppm, were deemed manageable due to the high water solubility of potassium salts. Literature indicates typical removal efficiencies of 97–99% for K through simple multi-stage water washes, allowing reduction from initial concentrations around 500 ppm to within battery-grade specifications (Table S1).^{1,2} Similarly, Ca impurities, despite exceeding the battery-grade threshold of 70 ppm in some preliminary outcomes, were also considered manageable. Literature supports mild supplementary treatments such as dilute acid washes or carbonation-decarbonation recrystallization, which reliably reduce calcium from initial concentrations of a few hundreds to a thousand ppm to below battery-grade

limits, typically achieving removal efficiencies between 70–85%.^{1,3,4}

Given the satisfactory refinement consistently achieved for Na and Li impurities, further optimization efforts for these elements were deprioritized. Although post-refinement K concentrations occasionally exceeded battery-grade thresholds, optimization for potassium was similarly deprioritized due to the ease and effectiveness of post-process water washing. This decision was corroborated by statistical analyses conducted within the HITL framework. Specifically, SHAP analysis demonstrated that post-refinement K concentrations exhibited lower feature importance than a randomly generated control variable (Fig. S1). Additionally, Pearson correlation matrices indicated no significant correlations between K concentrations and process control parameters after refinement (Fig. S2). Although sensitivity analysis suggested a slight influence from flow rate on K reduction, this was statistically insignificant, exhibiting comparable magnitude to random variations (Fig. S1). Together, these analyses justified assigning K a lower priority for subsequent experimental optimization.

Conversely, Mg impurity removal posed significant challenges. In scenarios where initial Mg concentrations exceeded approximately 80 ppm—the conventional battery-grade threshold—none of the preliminary experiments successfully reduced Mg below this desired limit (Table S1). Furthermore, elevated initial Mg concentrations corresponded with increased post-refinement Ca concentrations, occasionally surpassing the battery-grade threshold, though still manageable through previously described post-processing methods. Consequently, informed by initial observations and HITL-driven analyses, subsequent experimental designs specifically prioritized optimizing Mg and Ca removal.

Exploratory analysis began by employing Pareto frontier extraction, aiming to reduce post-refinement Mg and Ca concentrations. Gaussian Process Regression (GPR) predictions were generated across a surrogate experimental space consisting of 10 000 data points, as detailed fully in Tables S3. At each exploratory iteration, approximately 30 candidate conditions identified by Pareto frontier analysis were presented to experts, from which a subset was selected for experimental execution. After four active learning iterations, a cumulative total of 36 successful, reproducible experiments—including the initial 16 experiments—were collected for further analysis (Table S4). The iterative refinement and optimization achieved through these cycles are illustrated in Fig. 3.

As depicted in Fig. 3, despite inherent complexity causing occasional sub-optimal model predictions—including non-physical outcomes such as negative concentrations—GPR predictions remained valuable for identifying meaningful trends and signals.

Throughout the initial exploratory active learning phase, informed by insights from the first 16 expert-designed experiments, it was observed that the continuous crystallization process consistently met battery-grade lithium carbonate specifications for Na (Tables S1 and S4). Potassium and calcium impurities were effectively managed with straightforward, cost-efficient secondary treatments, ensuring they did not obstruct



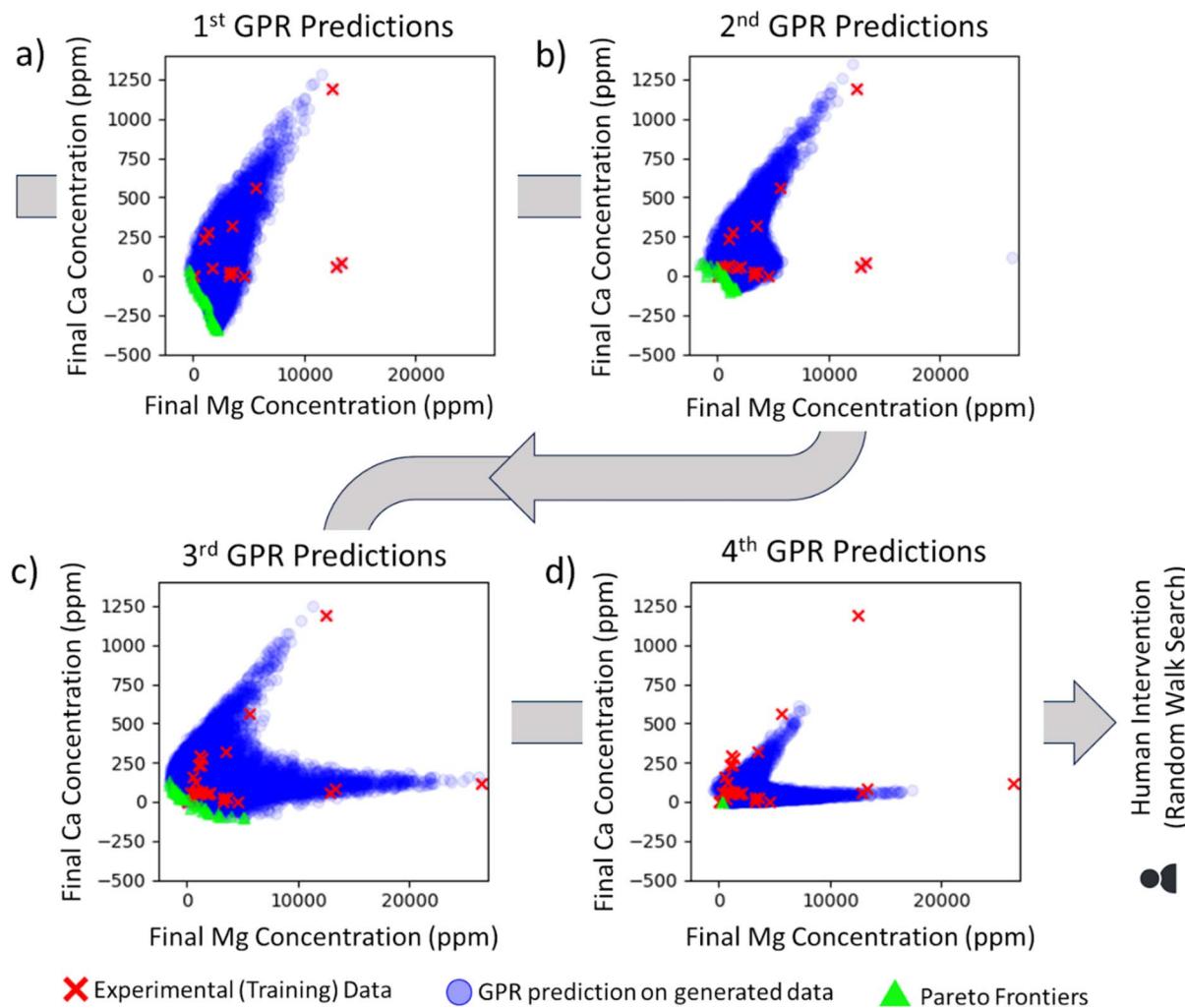


Fig. 3 GPR predictions and Pareto frontier extraction during the first four cycles of Bayesian active learning interactions on the 10 000 points in the surrogate space: (a) cycle 1, (b) cycle 2, (c) cycle 3, and (d) cycle 4.

overall lithium purification. However, across the first 36 experiments conducted, when initial Mg concentrations exceeded 200 ppm, reducing Mg below the battery-grade threshold of 80 ppm was consistently unattainable (Tables S1 and S4). This persistent challenge led to a strategic shift in optimization efforts, concentrating specifically on improving magnesium impurity removal to reliably achieve battery-grade lithium.

To address the persistent challenge of effectively reducing Mg impurities, human experts examined potential biases in the experimental design—specifically concerning the selected parameter ranges. A targeted random walk algorithm was deployed along the boundaries defined by the last identified Pareto frontier, systematically exploring adjacent regions. By allowing deviations up to $\pm 25\%$ from these Pareto frontier boundaries, this approach generated a new surrogate space containing 5000 additional experimental conditions. GPR models, trained on all experimentally observed data available up to that point (the first 36 data points, Table S4), were subsequently employed to predict outcomes across this newly

created surrogate space, enabling the detection of overlooked parameter spaces or latent biases.

Analysis of GPR-predicted data from this random walk surrogate space revealed an unexpected inverse correlation between the cold reactor's temperature and predicted post-refinement Mg concentration. This surprising observation, identified through Pearson correlation, SHAP feature importance, and sensitivity analyses performed on the model-predicted dataset (Fig. 4), starkly contrasted prevailing assumptions derived from prior literature and scientific heuristics. The statistical analyses indicated that higher cold reactor temperatures could significantly enhance Mg impurity removal efficiency—directly challenging established recommendations of maintaining cold reactor temperatures below 60 °C with a 20 °C differential between reactors. Prompted by these model-derived insights, experimental trials were subsequently proposed and conducted to validate this observation.

To experimentally validate this correlation, human experts conducted targeted trials, systematically increasing the cold reactor temperature while holding other operational

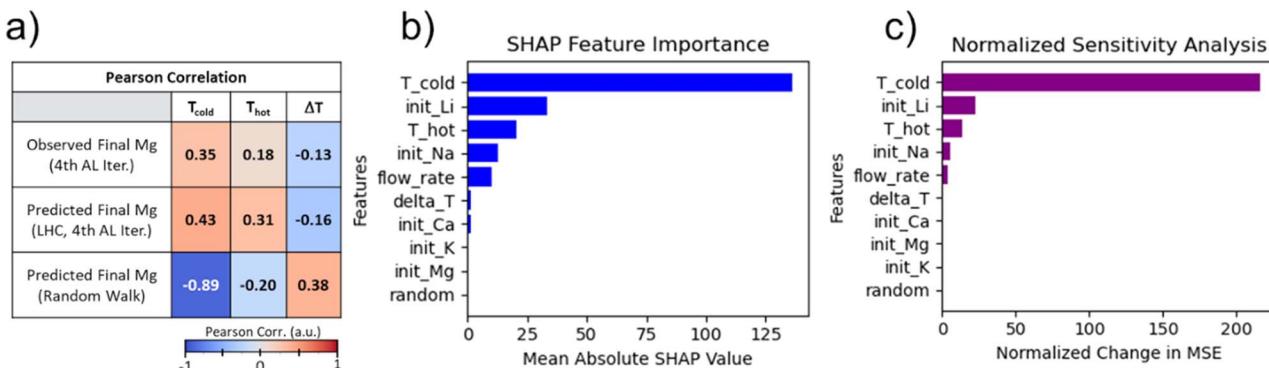


Fig. 4 (a) Pearson correlation analysis illustrating the inverse relationship between cold reactor temperature and magnesium concentration across observed, LHC, and random walk-generated data. (b) SHAP feature importance analysis revealing the impact of cold reactor temperature on the final Mg concentration. (c) Normalized sensitivity analyses for the final Mg target from the random walk-generated data, demonstrating the significant influence of cold reactor temperature on reducing final Mg concentrations.

parameters constant. These validation experiments confirmed the initial GPR-derived predictions, substantiating the inverse relationship between cold reactor temperature and post-refinement Mg concentration. Fig. 5 presents one such validation experiment, highlighting the direct impact of cold reactor temperature on Mg impurity reduction (detailed experimental conditions provided in Tables S4 and S5).

Following the experimental validation of the inverse relationship between cold reactor temperature and Mg impurity concentration, the study proceeded to the exploitation phase. The objective in this phase was to clearly delineate the decision

boundary between battery-grade and non-battery-grade lithium outcomes, specifically quantifying the maximum permissible initial Mg concentration that could yield battery-grade lithium under optimized temperature conditions. Fig. 6 visualizes this boundary by plotting experimentally observed outcomes—classified as battery-grade or non-battery-grade—against initial Mg concentration and cold reactor temperature. The resulting visualization reveals a distinct and actionable decision boundary.

Subsequently, a GPC model was trained to formalize the experimentally identified decision boundary (Fig. 6). Utilizing

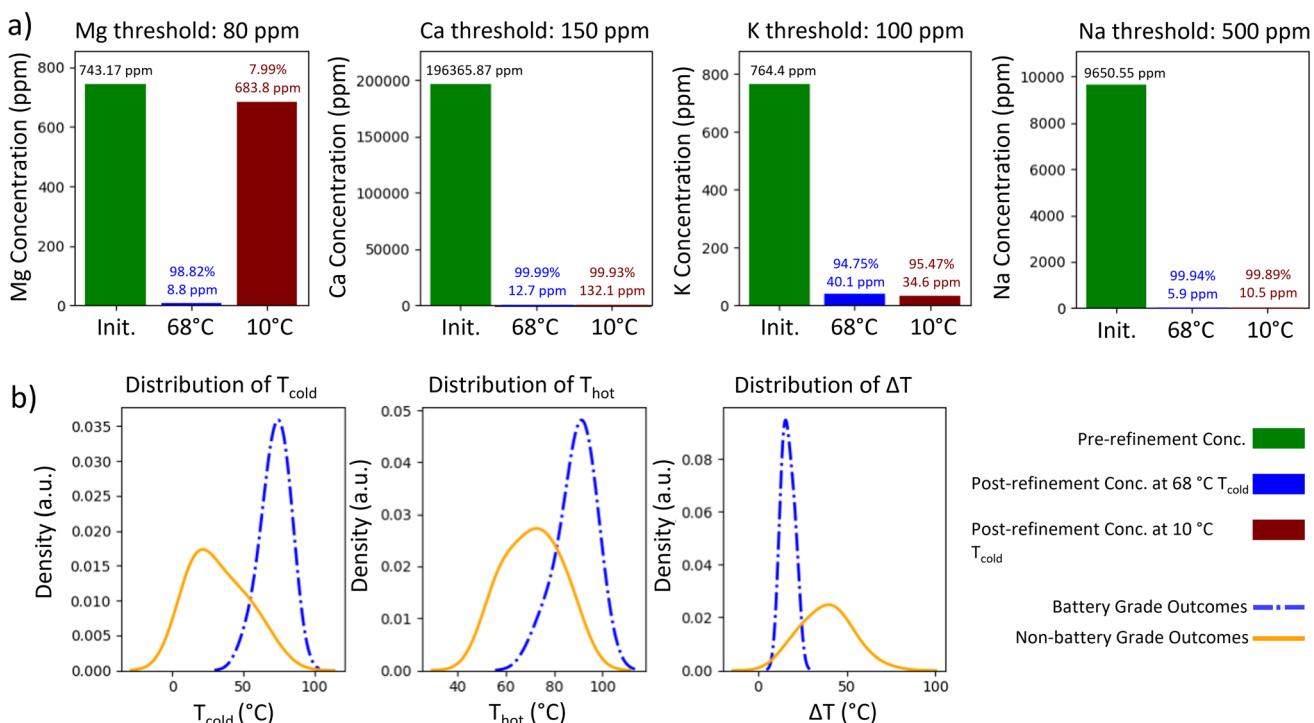


Fig. 5 An experiment showcasing the impact of reactor temperatures on reducing impurity concentration, especially magnesium concentration (Table S5). (a) Comparison of final impurity concentrations at cold reactor temperatures of 68 °C and 10 °C. (b) Kernel Density Estimate (KDE) plots of reactor temperatures (T_{cold} , T_{hot} , and ΔT) for battery-grade versus non-battery-grade outcomes, based on the data in Table S4.



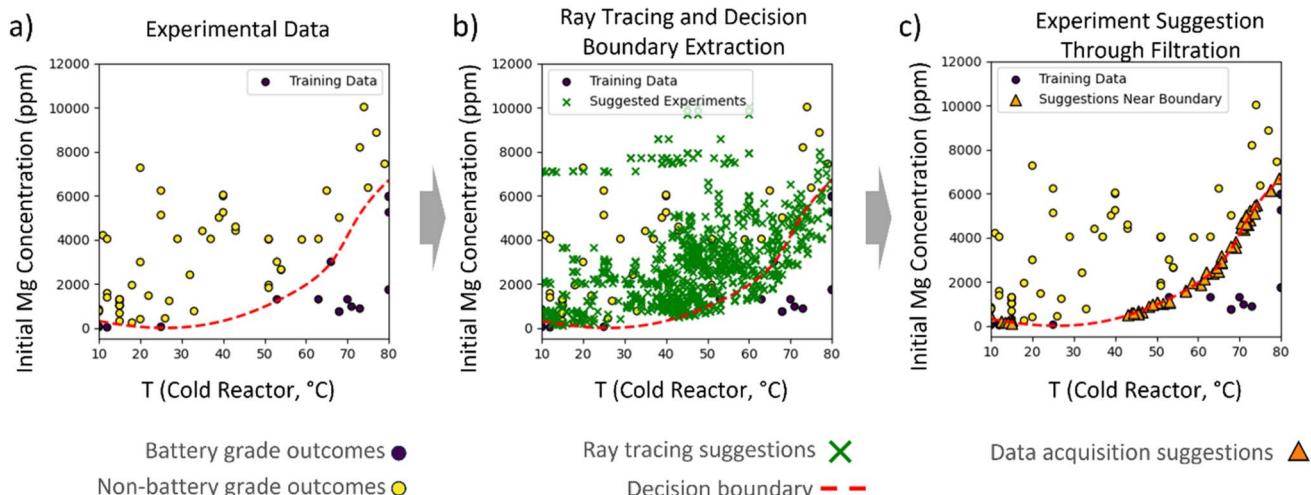


Fig. 6 (a) Experiment refinement outcomes focused on the impact of initial Mg and temperature of cold reactor focused on initial Mg ranges below 12 000 ppm. (b) Demonstration of ray tracing and decision boundary extraction process. (c) Filtered experiments by selecting those closest to the decision boundary.

a ray tracing algorithm, experimental candidates with predicted probabilities closest to the critical 0.5 threshold—indicating equal likelihood of achieving battery-grade or non-battery-grade lithium—were selected for further verification experiments.

Fig. 7 presents these GPR-derived predictions, distinguishing battery-grade from non-battery-grade outcomes, overlaid with all observed experimental data (totaling 80 experiments). This visualization highlights a substantial improvement in permissible initial Mg contamination levels. Historically, industry standards limited acceptable Mg contamination to approximately 80 ppm. However, our HITL-driven optimization demonstrates that increasing the cold reactor temperature beyond the initially recommended limit of

60 °C can significantly elevate the tolerance of initial Mg contamination to several thousand ppm. This finding implies a potential reduction in the extent of pre-refinement processes and lessens reliance on higher-purity lithium sources to achieve battery-grade products.

Fig. 8 provides additional analytical evidence supporting the significant influence of cold reactor temperature and initial Mg concentration on process outcomes. Specifically, Fig. 8a shows the distribution of cold reactor temperatures for experiments with initial Mg concentrations above 200 ppm, categorized by battery-grade or non-battery-grade outcomes. This distribution clearly demonstrates that higher reactor temperatures strongly correlate with successful Mg impurity reduction. Further supporting these observations, Fig. 8b presents SHAP analysis results, highlighting the prominent impact of cold reactor temperature on achieving battery-grade lithium. Lastly, Fig. 8c illustrates sensitivity analysis outcomes, reinforcing that both initial Mg concentration and cold reactor temperature are key parameters determining the purity of the final lithium carbonate product.

Including the initial 16 historic experiments, a total of 38 experiments (22 additional experiments) were required to obtain clear experimental evidence of the influence of cold reactor temperature on achieving battery-grade lithium carbonate. In total, 80 experiments were conducted throughout the study to validate these insights and precisely identify the decision boundary between battery-grade and non-battery-grade outcomes (as represented in Fig. 6–8). Given the complexity of the crystallization process and the extensive dimensionality of the parameter space, identifying critical parameters within just 38 experiments demonstrates the effectiveness of integrating human expertise, statistical analyses, and machine learning-driven active learning methods. Conducted at an accelerated pace of approximately one experiment per day, this human-in-the-loop approach substantially reduced

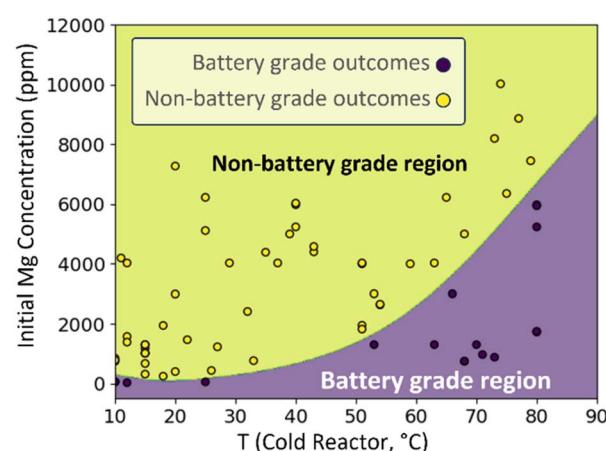


Fig. 7 Projection of GPC model predictions on the initial Mg concentration and cold reactor temperature plane, illustrating the decision boundary between battery-grade and non-battery-grade outcomes. This visualization emphasizes improved Mg contamination tolerance, demonstrating that higher initial impurity levels can still yield battery-grade lithium by increasing the temperature of the cold reactor.



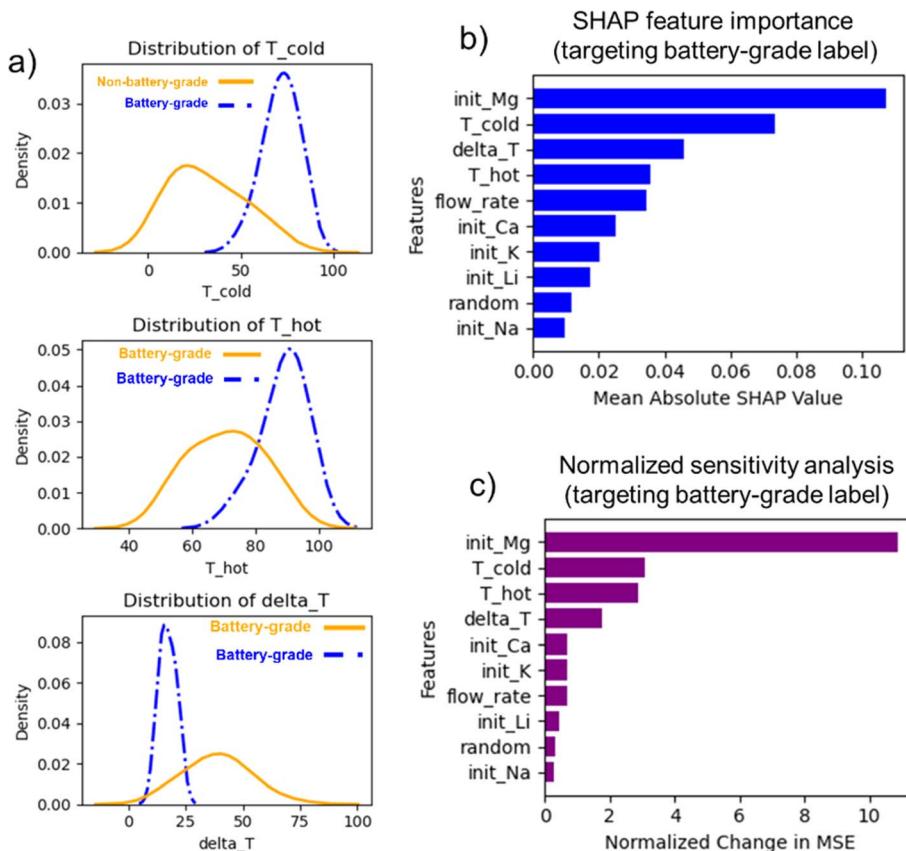


Fig. 8 (a) Shows the kernel density estimate distribution of reactor temperatures against initial magnesium concentrations. (b) Presents SHAP analysis indicating the influence of temperature on achieving battery-grade outcomes. (c) Illustrates sensitivity analysis results, highlighting critical temperature-related parameters.

the timeline traditionally required for such process optimization, facilitating rapid decision-making and iterative refinement despite the challenges inherent in low-data machine learning scenarios.

To further illustrate how the HITL framework accelerated the identification of experimental conditions conducive to producing battery-grade lithium carbonate with initial Mg concentrations exceeding 200 ppm, we compared its efficacy against two alternative active learning frameworks without human intervention. Specifically, we compared the performance of our HITL-assisted active learning (HITL-AL) approach, which required 38 experiments to identify key conditions leading to battery-grade outcomes, against the performance of purely computational active learning methods without expert guidance. These methods included a random-sampling approach and a simplified Bayesian optimization approach, both conducting one experiment per active learning cycle.

For this comparative evaluation, a surrogate GPC model—trained on data and insights obtained from the HITL approach—was used to predict battery-grade outcomes. Two distinct experimental datasets, each comprising 10 000 000 simulated experimental scenarios, were generated for this comparison. The first dataset, termed “uninformed,” strictly adhered to initial parameter ranges without incorporating insights from the HITL process. The second dataset, termed

“informed,” was explicitly constrained by parameter ranges informed by the HITL-identified optimal temperature settings and impurity conditions.

Each active learning method underwent 100 computational simulations (instances) employing different random seeds to statistically evaluate the frequency at which battery-grade conditions were identified. The first active learning approach applied a random sampling data acquisition policy. In contrast, the second method used a simplified Bayesian approach explicitly adapted for classification tasks, predicting experimental outcomes as either battery-grade or non-battery-grade. The classification approach was adopted due to challenges in reliably simulating detailed regression predictions of impurity concentrations with the surrogate GPC model. For condition selection, the Bayesian method relied on the upper confidence bound (UCB) metric to target experimental conditions with high uncertainty.

Fig. 9 summarizes the comparative results of the simplified Bayesian and random active learning methods using both informed and uninformed datasets. The simplified Bayesian approach leveraging the informed dataset (integrating optimized parameter ranges derived from prior HITL insights) identified battery-grade conditions within 40 experiments at a success rate of approximately 67%, significantly outperforming the uninformed dataset scenario, which achieved



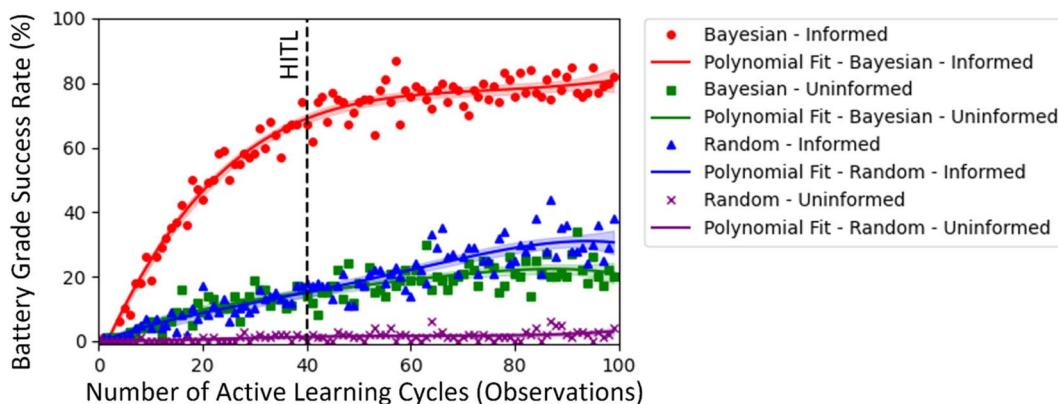


Fig. 9 Comparison of success rates in identifying battery-grade conditions using HITL active learning, human-independent Bayesian active learning, and random sampling methods, with and without prior information.

only 14% success. Similarly, the random sampling approach achieved success rates of 14% with the informed dataset and only 1% with the uninformed dataset. These results clearly illustrate the substantial advantage provided by incorporating human-derived knowledge and insights into active learning optimization strategies.

Collectively, these findings underscore the practical effectiveness and importance of integrating human expertise with machine learning analyses. Targeted human interventions, guided by nuanced insights from ML-driven data exploration, notably accelerated the discovery of optimal process parameters, demonstrating a highly effective synergy for refining complex chemical systems.

4 Conclusions

In this study, we presented a Human-in-the-Loop (HITL)-assisted active learning framework tailored specifically for the optimization of continuous lithium carbonate crystallization processes from low-grade brine sources, such as those found in the Smackover Formation. By seamlessly integrating human domain expertise with artificial intelligence-driven methods, we effectively addressed critical challenges associated with complex impurity management, especially regarding magnesium, which has historically constrained battery-grade lithium production.

Through iterative experimentation and informed interpretation of machine learning model predictions, our approach notably accelerated the identification of optimal process conditions. A critical insight emerged from this collaborative optimization: contrary to traditional assumptions recommending lower cold reactor temperatures (below 60 °C with a 20 °C temperature differential), our framework identified and experimentally validated that elevated cold reactor temperatures significantly enhanced magnesium impurity removal efficiency. This finding permitted an unprecedented increase in magnesium impurity tolerance—from conventional limits of approximately 80 ppm up to several thousand ppm—markedly reducing the need for intensive pre-refinement stages.

Overall, this research demonstrates the pivotal role human expertise can play in refining AI-driven optimization processes, particularly in high-dimensional chemical systems constrained by limited experimental throughput. By effectively balancing human intuition, domain knowledge, and machine learning analysis, our approach not only expedited critical discoveries but also facilitated agile responses to emergent insights, thereby optimizing experimental efficiency and accuracy. Consequently, the methodological and practical advancements presented herein represent significant progress toward the sustainable, economically viable extraction of lithium carbonate from complex, impurity-rich brines. This advancement contributes directly to the broader objective of responsibly harnessing North America's abundant but challenging lithium resources, supporting the growth and sustainability of lithium-based technologies vital for the global energy transition.

Conflicts of interest

The authors declare that there are no competing financial interests or conflicts of interest.

Data availability

All experimental data, AI-generated experiment suggestions, executable Python scripts, and resources necessary for reproducing the paper's figures and analyses are hosted on our dedicated GitHub repository and archived on Zenodo. For full access to these materials and further details, please visit our GitHub repository at https://github.com/shmouses/HITL_Adaptive_Active_Learning_Lithium and the Zenodo archive at DOI: <https://doi.org/10.5281/zenodo.17122531>.

The experimental data supporting this article are provided in Table S4 of the SI, which also includes additional information on experimental procedures and complementary analyses. Supplementary information is available. See DOI: <https://doi.org/10.1039/d5dd00285k>.



Acknowledgements

The authors gratefully acknowledge funding from the Critical Minerals Research, Development, and Demonstration (CMRDD) Program administered by Natural Resources Canada. We also extend our gratitude to Standard Lithium for their financial support and valuable discussions that have guided the project's design. Additionally, Telescope Innovations thanks the Mining Innovation Commercialization Accelerator (MICA) for financial support related to this project. We sincerely appreciate the reviewers for their constructive feedback, which has contributed to the refinement of the manuscript.

References

- 1 F. Meng, J. McNeice, S. S. Zadeh and A. Ghahreman, Review of Lithium Production and Recovery from Minerals, Brines, and Lithium-Ion Batteries, *Miner. Process. Extr. Metall. Rev.*, 2021, **42**(2), 123–141, DOI: [10.1080/08827508.2019.1668387](https://doi.org/10.1080/08827508.2019.1668387).
- 2 H. Ambrose and A. Kendall, Understanding the Future of Lithium: Part 1, Resource Model, *J. Ind. Ecol.*, 2020, **24**(1), 80–89, DOI: [10.1111/jiec.12949](https://doi.org/10.1111/jiec.12949).
- 3 G. Liu, Z. Zhao and A. Ghahreman, Novel Approaches for Lithium Extraction from Salt-Lake Brines: A Review, *Hydrometallurgy*, 2019, **187**, 81–100, DOI: [10.1016/j.hydromet.2019.05.005](https://doi.org/10.1016/j.hydromet.2019.05.005).
- 4 J. Fleischmann, P. Schaufuss, M. Linder, M. Hanicke, E. Horetsky, D. Ibrahim, S. Jautelat, L. Torscht and A. van de Rijt, *Lithium-Ion Battery Demand Forecast for 2030*, McKinsey & Company, 2023, <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/battery-2030-resilient-sustainable-and-circular>.
- 5 Lithium carbonate price 2023, Statista, <https://www.statista.com/statistics/606350/battery-grade-lithium-carbonate-price/>, accessed 2024-08-08.
- 6 K. J. Knierim, M. S. Blondes, A. Masterson, P. Freeman, B. McDevitt, A. Herzberg, P. Li, C. Mills, C. Doolan, A. M. Jubb, S. M. Ausbrooks and J. Chenault, Evaluation of the Lithium Resource in the Smackover Formation Brines of Southern Arkansas Using Machine Learning, *Sci. Adv.*, 2024, **10**(39), eadp8149, DOI: [10.1126/sciadv.adp8149](https://doi.org/10.1126/sciadv.adp8149).
- 7 C. J. Manche, J. Markello, J. C. Laya, M. Pope, K. Fritsche and K. Hill, Diagenetic and Sequence Stratigraphic Controls on Reservoir Quality in the Upper Jurassic Smackover Formation, Southern Arkansas, USA, *Mar. Pet. Geol.*, 2025, **177**, 107363.
- 8 A. Ettehadi, M. Chuprin, M. Mokhtari, D. Gang, P. Wortman and E. Heydari, Geological Insights into Exploration and Extraction of Lithium from Oilfield Produced-Water in the USA: A Review, *Energy Fuels*, 2024, **38**(12), 10517–10541, DOI: [10.1021/acs.energyfuels.4c00732](https://doi.org/10.1021/acs.energyfuels.4c00732).
- 9 P. J. Daitch, Lithium Extraction from Oilfield Brine, PhD thesis, University of Texas at Austin, 2018, <https://repositories.lib.utexas.edu/items/13f127c1-4b41-4669-8a6e-34830b025daf>, accessed 2025-06-13.
- 10 X. Huang, B. Ahmed, M. Mehana, S. Bhattacharya and C. Neil, *Data-Driven Lithium and Geothermal Resource Assessment in the Smackover Formation*.
- 11 N. Linneen, R. Bhave and D. Woerner, Purification of Industrial Grade Lithium Chloride for the Recovery of High Purity Battery Grade Lithium Carbonate, *Sep. Purif. Technol.*, 2019, **214**, 168–173, DOI: [10.1016/j.seppur.2018.05.020](https://doi.org/10.1016/j.seppur.2018.05.020).
- 12 V. Flexer, C. F. Baspineiro and C. I. Galli, Lithium Recovery from Brines: A Vital Raw Material for Green Energies with a Potential Environmental Impact in Its Mining and Processing, *Sci. Total Environ.*, 2018, **639**, 1188–1204, DOI: [10.1016/j.scitotenv.2018.05.223](https://doi.org/10.1016/j.scitotenv.2018.05.223).
- 13 L. Li, V. G. Deshmane, M. P. Paranthaman, R. Bhave, B. A. Moyer and S. Harrison, Lithium Recovery from Aqueous Resources and Batteries: A Brief Review, *Johnson Matthey Technol. Rev.*, 2018, **62**(2), 161–176.
- 14 *Lithium Sector: Production Costs Outlook*, S&P Global Market Intelligence, <https://pages.marketintelligence.spglobal.com/lithium-sector-production-costs-outlook-MS.html>, accessed 2024-08-08.
- 15 I. Warren, *Techno-Economic Analysis of Lithium Extraction from Geothermal Brines*, National Renewable Energy Lab (NREL), Golden, CO (United States), 2021, <https://www.osti.gov/biblio/1782801>, accessed 2024-08-08.
- 16 D. Zhang, S. Xu, S. Du, J. Wang and J. Gong, Progress of Pharmaceutical Continuous Crystallization, *Engineering*, 2017, **3**(3), 354–364, DOI: [10.1016/J.ENG.2017.03.023](https://doi.org/10.1016/J.ENG.2017.03.023).
- 17 Y. Ma, S. Wu, E. G. J. Macaringue, T. Zhang, J. Gong and J. Wang, Recent Progress in Continuous Crystallization of Pharmaceutical Products: Precise Preparation and Control, *Org. Process Res. Dev.*, 2020, **24**(10), 1785–1801, DOI: [10.1021/acs.oprd.9b00362](https://doi.org/10.1021/acs.oprd.9b00362).
- 18 Y. Ma, M. Svärd, X. Xiao, J. M. Gardner, R. T. Olsson and K. Forsberg, Precipitation and Crystallization Used in the Production of Metal Salts for Li-Ion Battery Materials: A Review, *Metals*, 2020, **10**(12), 1609.
- 19 J. Warner and N. Shah, Crystallization for Ultra-Pure Chemicals, *Chem. Eng.*, 2022, **129**(3), 28–32.
- 20 B. Han, O. Böckman, B. P. Wilson, M. Lundström and M. Louhi-Kultanen, Purification of Nickel Sulfate by Batch Cooling Crystallization, *Chem. Eng. Technol.*, 2019, **42**(7), 1475–1480, DOI: [10.1002/ceat.201800695](https://doi.org/10.1002/ceat.201800695).
- 21 X. Chen, Y. Chen, T. Zhou, D. Liu, H. Hu and S. Fan, Hydrometallurgical Recovery of Metal Values from Sulfuric Acid Leaching Liquor of Spent Lithium-Ion Batteries, *Waste Manage.*, 2015, **38**, 349–356, DOI: [10.1016/j.wasman.2014.12.023](https://doi.org/10.1016/j.wasman.2014.12.023).
- 22 G. Leonzio, Recovery of Metal Sulphates and Hydrochloric Acid from Spent Pickling Liquors, *J. Cleaner Prod.*, 2016, **129**, 417–426, DOI: [10.1016/j.jclepro.2016.04.037](https://doi.org/10.1016/j.jclepro.2016.04.037).
- 23 W. M. Haynes, *CRC Handbook of Chemistry and Physics*, CRC press, 2016.
- 24 F. L. Muller, M. Fielding and S. Black, A Practical Approach for Using Solubility to Design Cooling Crystallisations, *Org. Process Res. Dev.*, 2009, **13**(6), 1315–1321, DOI: [10.1021/op9001438](https://doi.org/10.1021/op9001438).

25 S. S. M. Masouleh, C. A. Sanz, R. P. Jansonius, S. Shi, M. J. G. Romero, J. E. Hein and J. Hattrick-Simpers, Artificial Intelligence-Enabled Optimization of Battery-Grade Lithium Carbonate Production, *Digital Discovery*, 2024, **3**(11), 2320–2326.

26 B. L. DeCost, J. R. Hattrick-Simpers, Z. Trautt, A. G. Kusne, E. Campo and M. L. Green, Scientific AI in Materials Science: A Path to a Sustainable and Scalable Paradigm, *Mach. Learn.: Sci. Technol.*, 2020, **1**(3), 033001.

27 M. M. Noack, G. S. Doerk, R. Li, J. K. Streit, R. A. Vaia, K. G. Yager and M. Fukuto, Autonomous Materials Discovery Driven by Gaussian Process Regression with Inhomogeneous Measurement Noise and Anisotropic Kernels, *Sci. Rep.*, 2020, **10**(1), 17663, DOI: [10.1038/s41598-020-74394-1](https://doi.org/10.1038/s41598-020-74394-1).

28 S. Back, A. Aspuru-Guzik, M. Ceriotti, G. Gryn'ova, B. Grzybowski, G. Ho Gu, J. Hein, K. Hippalgaonkar, R. Hormázabal, Y. Jung, S. Kim, W. Youn Kim, S. Mohamad Moosavi, J. Noh, C. Park, J. Schrier, P. Schwaller, K. Tsuda, T. Vegge, O. A. v. Lilienfeld and A. Walsh, Accelerated Chemical Science with AI, *Digital Discovery*, 2024, **3**(1), 23–33, DOI: [10.1039/D3DD00213F](https://doi.org/10.1039/D3DD00213F).

29 X. Bai and X. Zhang, Artificial Intelligence-Powered Materials Science, *Nano-Micro Lett.*, 2025, **17**(1), 135, DOI: [10.1007/s40820-024-01634-8](https://doi.org/10.1007/s40820-024-01634-8).

30 M. Ziatdinov, Y. Liu, K. Kelley, R. Vasudevan and S. V. Kalinin, Bayesian Active Learning for Scanning Probe Microscopy: From Gaussian Processes to Hypothesis Learning, *ACS Nano*, 2022, **16**(9), 13492–13512, DOI: [10.1021/acsnano.2c05303](https://doi.org/10.1021/acsnano.2c05303).

31 A. Biswas, Y. Liu, N. Creange, Y.-C. Liu, S. Jesse, J.-C. Yang, S. V. Kalinin, M. A. Ziatdinov and R. K. Vasudevan, A Dynamic Bayesian Optimized Active Recommender System for Curiosity-Driven Human-in-the-Loop Automated Experiments, *arXiv*, 2023, preprint, arXiv:2304.02484, DOI: [10.48550/arXiv.2304.02484](https://doi.org/10.48550/arXiv.2304.02484), <http://arxiv.org/abs/2304.02484>, accessed 2024-01-28.

32 M. F. Lazin, C. R. Shelton, S. N. Sandhofer and B. M. Wong, High-Dimensional Multi-Fidelity Bayesian Optimization for Quantum Control, *Mach. Learn.: Sci. Technol.*, 2023, **4**(4), 045014.

33 K. Yang, L. Liu and Y. Wen, The Impact of Bayesian Optimization on Feature Selection, *Sci. Rep.*, 2024, **14**(1), 3948.

34 J. Wang, B. Guo and L. Chen, Human-in-the-Loop Machine Learning: A Macro-Micro Perspective, *arXiv*, 2022, preprint, arXiv:2202.10564, DOI: [10.48550/arXiv.2202.10564](https://doi.org/10.48550/arXiv.2202.10564).

35 H. Hysmith, E. Foadian, S. P. Padhy, S. V. Kalinin, R. G. Moore, O. S. Ovchinnikova and M. Ahmadi, The Future of Self-Driving Laboratories: From Human in the Loop Interactive AI to Gamification, *Digital Discovery*, 2024, **3**(4), 621–636, DOI: [10.1039/D4DD00040D](https://doi.org/10.1039/D4DD00040D).

36 E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán and Á. Fernández-Leal, Human-in-the-Loop Machine Learning: A State of the Art, *Artif. Intell. Rev.*, 2023, **56**(4), 3005–3054, DOI: [10.1007/s10462-022-10246-w](https://doi.org/10.1007/s10462-022-10246-w).

37 A. D. Randolph, The Mixed Suspension, Mixed Product Removal Crystallizer as a Concept in Crystallizer Design, *AIChE J.*, 1965, **11**(3), 424–430, DOI: [10.1002/aic.690110311](https://doi.org/10.1002/aic.690110311).

38 C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, 2006.

39 J. Chen, L. Kang and G. Lin, Gaussian Process Assisted Active Learning of Physical Laws, *Technometrics*, 2021, **63**(3), 329–342, DOI: [10.1080/00401706.2020.1817790](https://doi.org/10.1080/00401706.2020.1817790).

40 M. M. Noack, G. S. Doerk, R. Li, J. K. Streit, R. A. Vaia, K. G. Yager and M. Fukuto, Autonomous Materials Discovery Driven by Gaussian Process Regression with Inhomogeneous Measurement Noise and Anisotropic Kernels, *Sci. Rep.*, 2020, **10**(1), 17663.

41 E. Uteva, R. S. Graham, R. D. Wilkinson and R. J. Wheatley, Active Learning in Gaussian Process Interpolation of Potential Energy Surfaces, *J. Chem. Phys.*, 2018, **149**(17), 174114, DOI: [10.1063/1.5051772](https://doi.org/10.1063/1.5051772).

