Discovery of high-entropy perovskite oxygen carriers for chemical looping applications via an autonomous active learning protocol

Materials Today Energy 57 (2026) 102239

Contents lists available at ScienceDirect

Materials Today Energy

journal homepage: www.sciencedirect.com/journal/materials-today-energy

Discovery of high-entropy perovskite oxygen carriers for chemical looping applications via an autonomous active learning protocol☆

Joakim Brorsson [a], Henrik Klein Moberg [a], Jonatan Gastaldi [b], Joel Hildingsson [a], Tobias Mattisson [b], Anders Hellman [a,c,*]

[a] Department of Physics, Chalmers University of Technology, Chalmersplatsen 4, Gothenburg, SE-41296, Sweden
[b] Department of Space, Earth and Environment, Chalmers University of Technology, Chalmersplatsen 4, Gothenburg, SE-41296, Sweden
[c] Competence Centre for Catalysis, Chalmers University of Technology, Chalmersplatsen 4, Gothenburg, SE-41296, Sweden

ARTICLE INFO

Keywords:
Materials discovery
Active learning
Oxygen carriers
Chemical looping
High entropy oxides

ABSTRACT

The discovery and design of new materials are paramount for advancing green technologies. High-entropy oxides represent one such group that has been only tentatively explored, mainly due to the inherent problem of navigating vast compositional spaces. Here, oxygen carriers for chemical looping processes have been identified using active learning-based strategies and first-principles-informed calculations. The proposed approaches were validated using an established computational framework for identifying high-entropy perovskites suitable for chemical looping air separation and dry reforming. The central insight gained was the identification of effective strategies, including greedy and Thompson-based sampling, informed by uncertainty estimates from Gaussian processes. Building on this knowledge, the concept was applied to the challenge of discovering high-entropy oxygen carriers for chemical-looping oxygen uncoupling. This resulted in both qualitative and quantitative outcomes, including lists of materials with high oxygen transfer capacities and configurational entropies. The top candidates were based on the known oxygen carrier $CaMnO_3$ and included expected elements such as titanium, cobalt, and copper, as well as unexpected ones such as yttrium and samarium. The results suggest that adopting active learning approaches is critical for materials discovery, as these methods are already reshaping research practice and will soon become the norm.

1. Introduction

Throughout human history, advances in materials discovery have powered technological innovation and shaped societal change. Obvious examples include the successive identification of better metals and alloys, which began in prehistoric times with copper, followed by bronze and then iron [1]. Skipping forward to the modern era, numerous alloys have now been produced that incorporate a wide variety of different elements and have been tailored to suit specific applications. Even so, progress is still being made, as is exemplified by the relatively recent identification of high entropy alloys (HEAs) as a novel class of materials that possess unique properties [2]. Later, this concept was extended to include other types of inorganic compounds, including oxides, carbides, and pnictides, which are together labeled as high entropy materials (HEMs) [3].

A fundamental issue related to the exploration of new materials is the vastness of the compositional, or configurational, space. This is especially true for HEMs, in which five or more elements typically occupy the same sublattice [4]. Due to the enormous number of variations, it is impossible to sample a significant fraction of them. As such, a conventional trial-and-error approach is almost guaranteed to fail to identify the best possible candidate for a given application. While existing knowledge and data can help guide the search, this would likely reduce the chance of discovering a completely novel material.

In recent years, machine learning (ML) has emerged as a potent tool for solving problems involving enormous amounts of data. Active learning (AL), in particular, is a technique that has only begun to be explored, especially in the context of material development. Yet, the few studies conducted so far have revealed the great potency of this approach, underscoring the likelihood that it will become more widespread, or even the new norm, in a not so distant future. Such strategies have, in fact, allowed the millions or even billions of candidate materials to be generated and tested, leading to the discovery of several hundred thousand stable compounds [5–7]. It is worth noting, however, that this type of brute-force screening requires enormous computational resources while still only probing a tiny fraction of the

☆ This article is part of a Special issue entitled: 'AI4EnergyMat' published in Materials Today Energy.

* Corresponding author.
E-mail address: anders.hellman@chalmers.se (A. Hellman).

https://doi.org/10.1016/j.mtener.2026.102239
Received 2 December 2025; Received in revised form 29 January 2026; Accepted 12 February 2026
Available online 20 February 2026

Acronyms

AL	active learning
BO	Bayesian optimization
ARDR	automatic relevance detection regression
BFGS	Broyden–Fletcher–Goldfarb–Shanno
BCC	body-centered cubic
BECCS	bio-energy with carbon capture and storage
CCS	carbon capture and storage
CE	cluster expansion
CL	chemical looping
CLC	chemical looping combustion
CLAS	chemical looping air separation
CLDR	chemical looping dry reforming
CLOU	chemical looping oxygen uncoupling
CV	cross-validation
DFT	density functional theory
DFT+U	density functional theory with Hubbard correction
ECI	effective cluster interactions
FCC	face-centered cubic
FC	force constant
FCP	force constant potential
GP	Gaussian processes
GGA	generalized gradient approximations
GMM	Gaussian mixture model
HEA	high entropy alloy
HEM	high entropy material
HEO	high entropy oxide
HEOC	high entropy oxygen carrier
ICSD	Inorganic Crystal Structure Database
LN	layer normalization
MAE	mean absolute error
MMD	maximum mean discrepancy
MC	Monte Carlo
ML	machine learning
MLP	multi-layer perceptron
MLIP	machine learning interatomic potential
NN	neural network
OC	oxygen carrier
OLS	ordinary least squares
OQMD	Open Quantum Materials Database
OTC	oxygen transfer capacity
PAW	projector augmented wave
PBE	Perdew, Burke, and Ernzerhof
ReLU	rectified linear unit
RFE	recursive feature elimination
RMSE	root mean square error
SI	Supporting Information
SQS	special quasi-random structure
UE	uncertainty estimation
VASP	Vienna ab initio simulation package
WAE	Wasserstein auto-encoder

compositional space spanned by HEMs. In addition, the resulting collection of materials is only known to fulfill a limited set of requirements, chiefly related to their thermodynamic stability at 0 K. Consequently, some type of high-throughput screening process would be required [8,9], together with additional, possibly demanding, calculations of the properties of interest, in order to identify material candidates for a specific application. For this reason, balanced and more streamlined approaches, such as the one originally presented by Rao et al. appear more promising [10]. Specifically, they introduced a procedure that utilizes AL to minimize the amount of data required for training and used it to search for high entropy Invar alloys with high thermal stabilities and low thermal expansion coefficients, thereby discovering several previously untested materials.

Due to the limited use of AL strategies in the study of inorganic systems [10], the aim of this study was to develop such an approach for discovering new, and potentially complex, materials, which are optimized for a particular application, based on sparse data. More precisely, this involves training a neural network (NN)—more precisely, a Wasserstein auto-encoder (WAE)—on compositions together with a target property, which could, in principle, be obtained by any method, such as experiments or first-principles calculations. In this work, we estimate the properties of interest using an machine learning interatomic potential (MLIP), which has been trained on first-principles data and, thus, represents a validated computational method that is both efficient and effective.

Because of the urgent need for developing technologies that can help combat climate change, the proposed method is used to identify oxygen carriers (OCs) for chemical looping (CL) in order to demonstrate its versatility. Such a focus is further validated by the fact that complex materials, including high entropy oxides (HEOs), have shown promise for this type of process, which not only allows efficient fuel conversion with low emissions, but also represents a breakthrough concept for carbon capture [11]. In addition, it has previously been demonstrated that ML—but not AL—can be an effective tool in such endeavors [12,13]. To address this knowledge gap, three distinct types of CL processes have been considered: hemical looping dry reforming (CLDR), for CO and H_2 generation; chemical looping air separation (CLAS), for O_2 production; and chemical looping oxygen uncoupling (CLOU), for fuel combustion (see Figure S1, Figure S2, and Figure S3). It is worth noting that strategies relying on high-throughput screening of existing databases, instead of AL, would be more efficient for considering more conventional and well-known OC materials, which release oxygen via phase transitions, as has been demonstrated by Lau et al. [8] and Brorsson et al. [9].

In the section that follows, an elaborate description of the computational method is provided, in terms of training the NN; sampling the latent space; and evaluating the thermodynamic properties. The results obtained when applying this approach to the problem of finding perovskite OCs are presented afterward, first for CLDR and then CLOU; comparable data for CLAS are, for the sake of brevity, presented and discussed in the Supporting Information (SI) (see Fig. 4, Fig. 5, Figure S9, Table S2, and Section 3.2). Finally, the key advantages and future prospects of the proposed methodology are highlighted.

2. Method

The AL-based strategy employed in this study has been inspired by the work of Rao et al. [10]. Specifically, it is based on the same concept for identifying promising materials for specific applications and is especially well suited for handling intricate compositional spaces, such as those encountered when considering HEMs. Even so, the approaches differ in many ways, for instance, in terms of the NN design, sampling procedure, and process for evaluating candidates. In addition, the goal has been to find OCs, in the form of complex oxides, for CL, as this is a potential breakthrough energy conversion technology.

Before the AL cycle can be initiated, one must first generate an initial dataset, here in the form of a structure database (see Note S2). When the latter is in place, one can proceed with the first step, which involves using the NN, in the form of a WAE, to compress composition data, as well as target properties, into a multi-dimensional latent space, represented by a Gaussian mixture model (GMM) (see Note S3 and Figure S5). Thereafter, a suitable method is utilized for generating

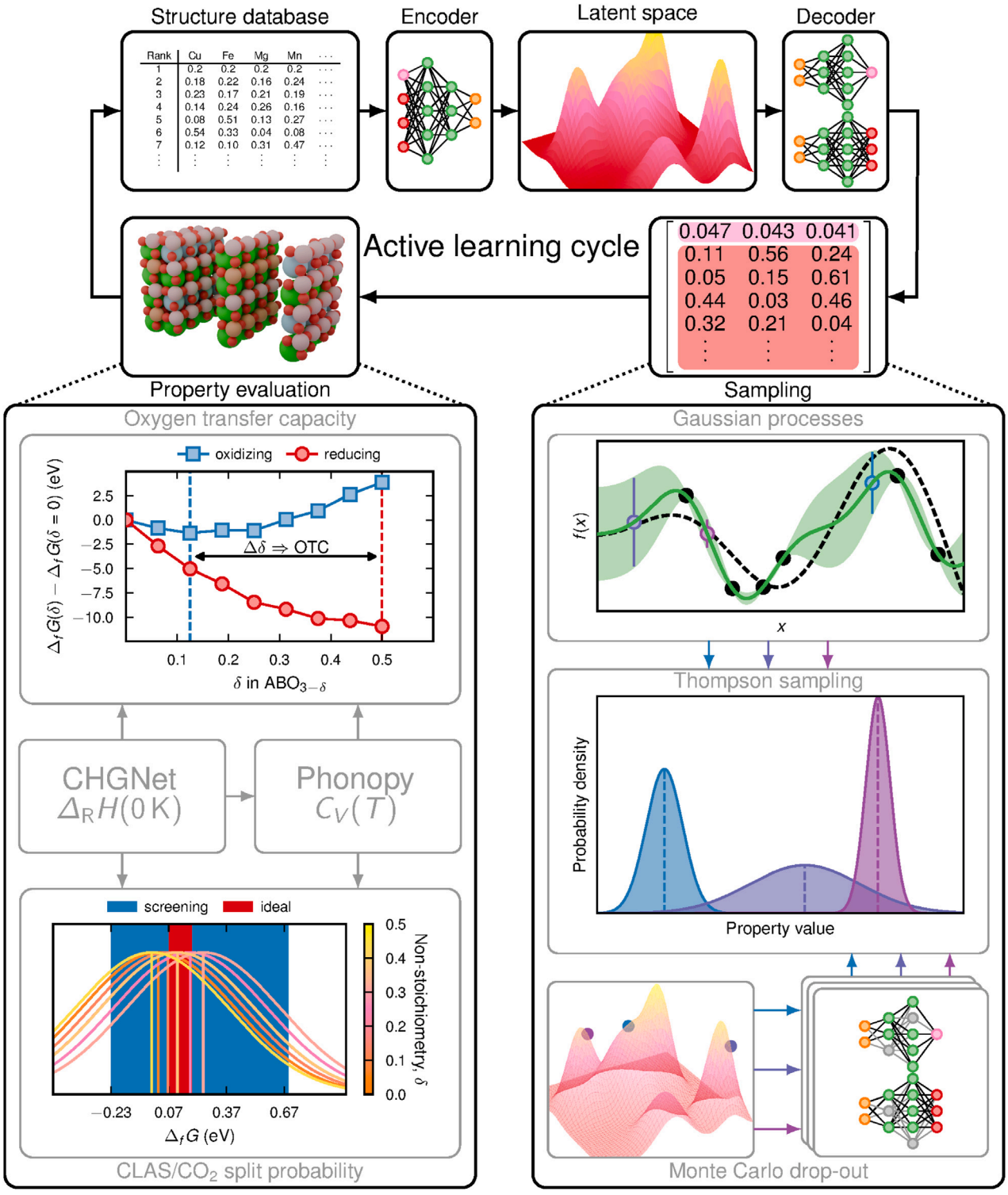

Fig. 1. Conceptual sketch of the main steps in a single AL cycle. First, a WAE is used to encode compositions and target properties, contained in a database, into a latent space, represented by a GMM. Next, promising candidates are selected using a chosen method, such as Greedy or Thompson sampling, based on property values and uncertainty estimates obtained either via GP or MC dropout. As a validation, either the oxygen transfer capacity (OTC) or the probability that the material is a suitable OC for a given type of CL process is, subsequently, calculated with the help of CHGNet. The loop is closed by feeding the resulting compositions, as well as the computed values of the target property, into the database.

candidates; preferably one that balances exploration and exploitation (see Note S4). The structural and thermodynamic properties of the sampled compositions are determined using first-principles-informed calculations (see Note S5 and Figure S4). To be precise, the latter were performed with the help of a MLIP (CHGNet) that has been trained on density functional theory (DFT) data and, thus, achieves almost

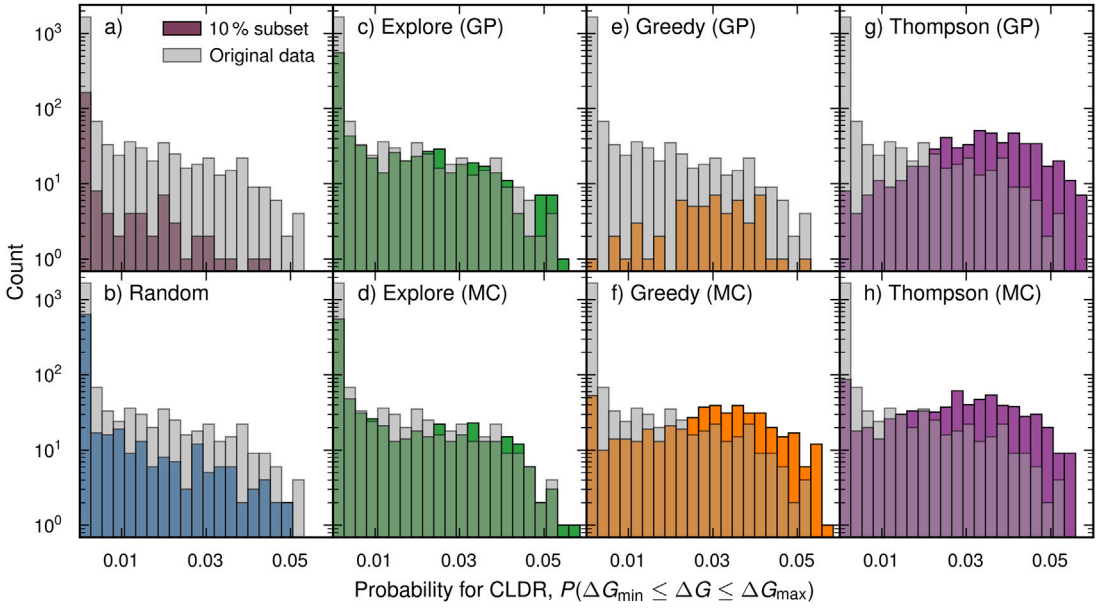

Fig. 2. Histograms based on the probability that a given OC candidate is valid for CLDR, showing the original dataset (light gray) together with the 10% subset used for training (a) and the results obtained with different AL approaches. Specifically, this includes random selection (b) as well as sampling based on exploration (c,d), a greedy strategy (e,f), and the Thompson method (g,h) combined with either GP (c,e,g) or MC dropout (d,f,h). Note that all panels, except (a), show results obtained from five individual runs after duplicate entries have been removed.

the same level of accuracy at a fraction of the cost. As a final step, the training database is updated with the resulting dataset, thereby completing the AL cycle (see Fig. 1).

3. Results and discussion

This section begins with a brief comparison of the training data used in this study and that of Wang et al. [13], respectively. Thereafter, the main results of the calculations outlined in Section 2 are presented and discussed, beginning with those pertaining to OCs for CLDR, followed by CLAS and, finally, HEOCs.

3.1. Chemical looping dry reforming

As explained by Wang et al. [13], the CLDR process places rather strict requirements on the OCs, making the task of finding suitable materials difficult. It would, therefore, be expected that this is a problem for which the previously outlined AL strategy would be well suited. This hypothesis is supported by a comparison of the counts of the number of candidates identified using a variety of different approaches (Fig. 2).

The initial 10% subset selected from the full dataset, which was used to train the surrogate model, has a similar pattern as the overall distribution of probabilities (notice the log scale in Fig. 2a). To be specific, most candidates are concentrated at values close to zero, while only a few are in the higher-probability range. Similar results were obtained when applying random sampling—most candidates cluster at very low probabilities, and only a small number appear in the higher range—confirming that unguided selection is an unwise strategy for materials with a vast composition space (see Fig. 2b). Explorative sampling, by contrast, shifts the distribution somewhat, though not in a way that favors the discovery of promising OCs. By construction, this strategy focuses on regions of the latent space where the model is most uncertain. This does expand coverage but also leads to fewer high-probability candidates, although slightly less so when using MC dropout method for uncertainty estimation (UE) compared to GP.

A greedy strategy could, in principle, be more effective, in particular if the entire phase space has been mapped or all the best materials are located within the surveyed area. It is therefore quite interesting that much fewer data points are added when such an approach is combined with GP compared to MC dropout (see Fig. 2e,f). Since this might initially seem to favor the latter UE method, the opposite is in fact true. The reason is that a greedy strategy, if efficiently implemented, is bound to get trapped, as it is designed to focus only on the region that, based on the available information, is the most likely to contain the best candidates. Hence, combining it with MC dropout apparently leads to a more uncertain model that consequently samples a larger area and, thus, adds more points. The clearest improvement arises from Thompson sampling (see Fig. 2g,h). By drawing from the full posterior distribution, this strategy achieves a balance between exploration and exploitation. As a result, the distributions are strongly skewed towards a high probability and generate far fewer candidates on the lower end of the scale than random sampling, pure exploration, or a greedy strategy. The resulting data are instead centered around a higher probability, close to $P \approx 0.04$. In practice, the combination of Thompson sampling and Gaussian processes (GP) emerges as the most efficient approach for discovering promising OCs for CLDR. Even so, it is worth noting that Monte Carlo (MC) dropout is nearly as effective for generating UEs, as the upper end of the distribution reaches comparably high values.

While the aggregated data revealed distinct differences in efficacy depending on the strategy, these trends become even clearer when considering the evolution of the CLDR probability (see Fig. 3). For instance, a greedy strategy combined with GP initially identifies many high-performing materials relatively quickly, but ceases to find any new compositions—only duplicates—after just 12 cycles (see Fig. 3b,e). When the uncertainty is estimated via MC dropout, meanwhile, multiple data points with a relatively high average ($P \approx 0.030$) are added in each cycle. A very similar trend is observed when using Thompson sampling, but in this case the performance actually improves markedly when switching to GP (see Fig. 3c,f). Specifically, doing so means that the average probability not only rises more sharply in the beginning but also flattens out at a higher value ($P \approx 0.035$). It is also worth mentioning that while exploration generates candidates with a much lower mean, the maximum value in each cycle is close to those obtained with the other approaches (see Fig. 3a,d). Consequently, there exist promising candidates within unexplored areas of the latent space,

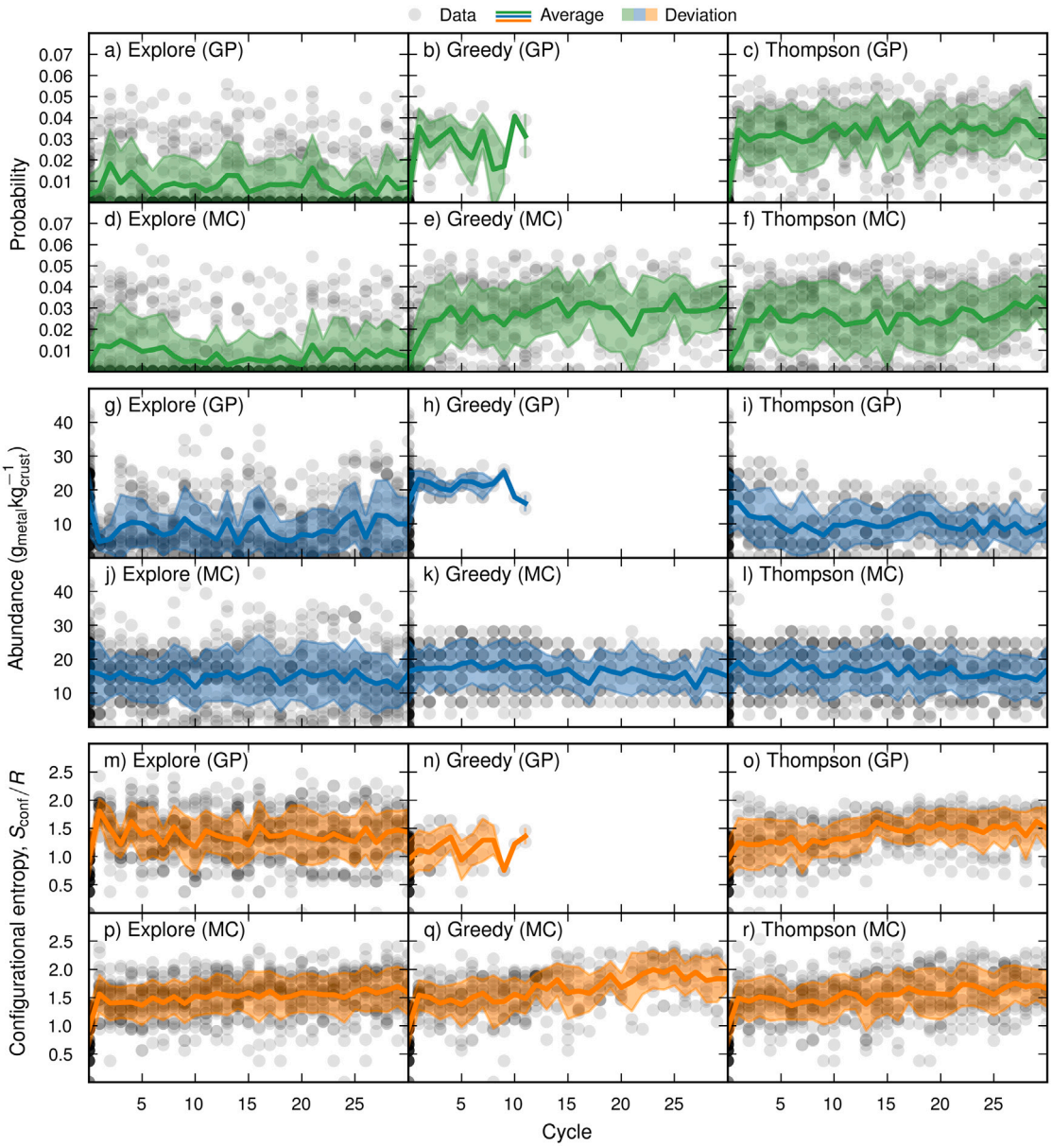

Fig. 3. Scatter plots with the probability that a given OC candidate is useful for CLDR (a–f); average abundance of the metallic elements in the earth's crust (g–l); and configurational entropy, S_{conf}/R, (m–r) in the original dataset (light gray) as well as those added in each cycle (purple) during five individual runs, after removing duplicates. The corresponding averages (solid line) and standard deviations (filled curve) are also shown.

which explains why Thompson sampling, which balances exploration and exploitation, achieves better results than a greedy strategy.

An interesting observation is that the configurational entropy increases continuously for most method combinations, even though this property is not included in the training of the WAE (see Fig. 3g–l). A likely reason, however, is that the original training data is not optimally constructed for AL. More precisely, it was constructed in compliance with the study by Wang et al. [13] and, thus, only covers a small portion of the latent space, namely binary, tertiary, and quaternary oxides obtained via the substitution of Fe and Sr in $Sr_{1-x}A_xFe_{1-y}B_yO_{3-\delta}$. As such, the increasing trend speaks to the inherent robustness of AL, especially when combined with an effective sampling method, and indicates an inherent tendency to discover HEMs, in spite of the fact that no specific restrictions or conditions have been applied to ensure such an outcome. The average abundance of the metallic element, meanwhile, remains relatively constant, even though a decrease would be expected since the added candidates should successively be made up of a higher number of distinctive, and therefore rarer, atomic species. In fact, this is only observed when using Thompson sampling in combination with GP, which therefore seems to explore other regions of the latent space.

As the models are trained to maximize the CLDR probability, the average compositions should give an indication of which elements are the most likely to provide favorable qualities (see Figure S8). In this context, it is worth remembering that Sr and Fe are the most common species in the original data set, since the latter was constructed through elemental substitutions in $Sr_{1-x}A_xFe_{1-y}B_yO_{3-\delta}$ (see Figure S8i–j). For a random sampling, however, all elements appear with about the same frequency except Mg and, to a lesser extent, K (see Figure S8g–h). As mentioned by Wang et al. [13], the reason is that the former (latter) is the only alkali (alkaline earth) metal on the A (B) site, which means that it is less likely to form stable perovskite structures, based on the Bartel tolerance factor as well as the requirement of charge neutrality (see Figure S6). The AL models, on the other hand, seem to favor La followed by Sr, Sm, and Y together with Mn as well as Fe and, to some extent, Co (see Figure S8a–f). Even so, there are quite prominent differences when comparing the individual approaches. As

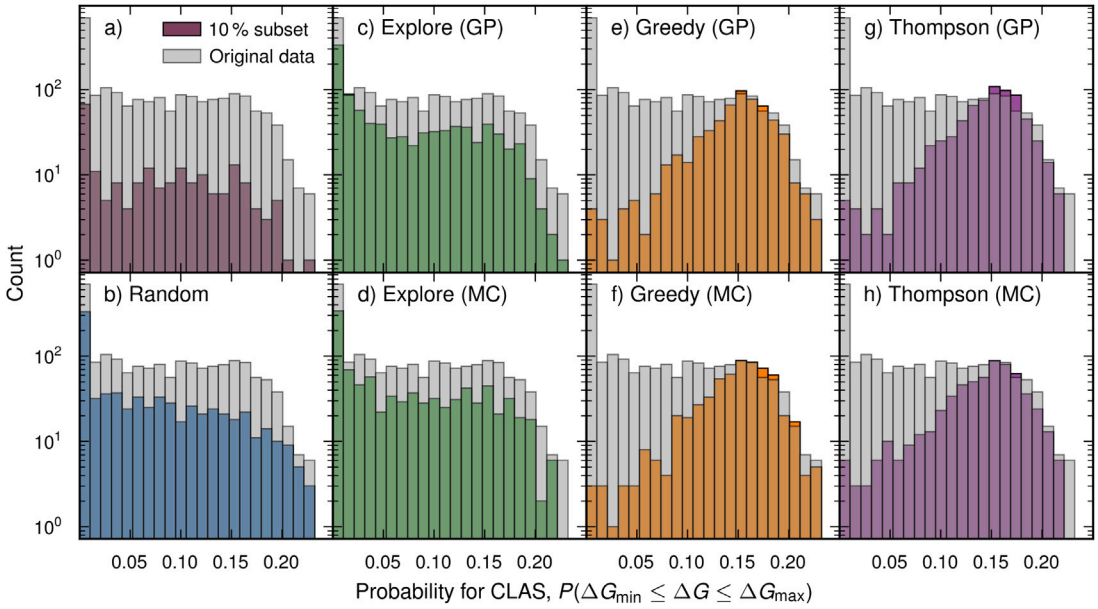

Fig. 4. Histograms based on the probability that a given OC candidate is useful for CLAS, showing the original dataset (light gray) together with the 10% subset used for training (a) and the results obtained with different AL approaches. Specifically, this includes random selection (e) as well as sampling based on exploration (b,f), a greedy strategy (c,g), and the Thompson method (d,h) combined with either GP (b–d) or MC dropout (f–h). Note that all panels, except (a), show results obtained from five individual runs after duplicate entries have been removed.

anticipated, exploration leads to a more uniform distribution, even if the trends mentioned earlier are still visible. Another observation is that greedy sampling based on GP gives a narrow spread of candidates that primarily contain La and Fe, which is indicative of the inherent "nearsightedness". Yet when combined with MC dropout, the average concentrations of the former two elements drop slightly to the benefit of Y, Sr, and Sm, in the former case, as well as Mn and Co, in the latter. Surprisingly, the values remain more or less the same for the Thompson method. Slight shifts can be observed, however, if a GP-based UE is used, which yields more La, Mn, and Co but less Y, Sm, Sr, and Fe.

In conclusion, it would seem that an ideal candidate is likely to be based on $LaMnO_3$, where La (Mn) has, in part, been substituted by Sr and Sm (Fe and Co). It could be argued that the relatively high percentage of Fe and Sr stems from the skewness of the original training data. This appears less likely, however, given the formulations that appear at the top of the leaderboard; obtained via a combination of Thompson sampling and GP (see Table S3). In fact, Fe; Sr; Ti; and Co are, respectively, found in as many as 17; 16; 17; and 12 of the 20 best candidates, including the first four: $CoFe_4La_7Mn_3SrO_{24}$; $Co_3Fe_4La_7SrTiO_{24}$; $Fe_3La_6Mn_3NiSr_2TiO_{24}$; and $CoCuFe_2La_7Mn_3SrTiO_{24}$.

3.2. Chemical looping air separation

While identical computational workflows and training datasets were applied in the search for suitable candidates for both CLAS and CLDR, the underlying optimal conditions vary significantly between the two processes. For the sake of brevity, however, the discussion that follows will focus on the areas where distinctive differences are observed. Beginning with the candidate counts after the final AL cycle, the trends are very similar to those described at the beginning of the previous section (see Fig. 4). Nevertheless, there are two key differences. Firstly, the choice of UE seems to be inconsequential and, secondly, greedy sampling apparently outperforms the Thompson method, since the upper end of the distribution is slightly shifted towards higher probabilities. Still, neither is able to identify any new compositions with a significantly better performance than those in the initial training set.

When considering the evolution over the cycles, the strong similarities in the efficacy of the alternative approaches are equally apparent (see Fig. 5). In particular, the average CLAS probability remains almost constant throughout, albeit at a higher level when using a greedy strategy or Thompson sampling compared to exploration. If anything, the trend is somewhat decreasing. The metal abundance, on the other hand, displays a steady decline while the configurational entropy increases, although more tentatively.

Given the aforementioned results, it is not astounding that the almost negligible variations are observed when comparing the compositions of the candidates (see Figure S9). With the exception of a somewhat broader distribution when using pure exploration, it is evident that Sr (Fe) followed by Ba (Co and Cu) are the most common elements on the A (B) site. Naturally, this holds true for the highest ranking candidates, among those sampled using the Thompson method combined with GP: $Co_6Fe_2La_2Sr_6O_{24}$, $Fe_6La_3Mg_2Sr_5O_{24}$, $Fe_5Mg_3Sm_4Sr_4O_{24}$, and $Ba_4Fe_8Sr_4O_{24}$ (see Table S2). Interestingly, these four, as well as nine others among the 16 that follow, all stem from the initial set of $Sr_{1-x}A_xFe_{1-y}B_yO_{3-\delta}$ training structures. While this may, at first glance, appear as a failure, it should be noted that these candidates are correctly identified via AL when only a subset of the original data is used for training. The lack of improvement is therefore an indication that the problem of finding OCs based on the CLAS probability, at least when defined by such a broad energy interval, is not specific enough to warrant the utilization of advanced search methods.

While the AL seems to be more effective at discovering OCs CLDR than CLAS the opposite is true for the ML model trained by Wang et al. [13], which was able to produce about four times as many viable candidates in the latter case. Taken together with the superiority of greedy sampling, this suggests that the problem of finding CLAS OCs, at least based on the broad criteria utilized here and by Wang et al. [13], is handled equally well by conventional ML approaches. One should, nonetheless, remember that there exist other potential causes, such as a poor choice of initial training data and lax ranges for the vacancy formation energies. As advertised earlier, some blame might also fall on the choice of target property. To be specific, any probability-based measure will, by definition, have a maximum value, which, in the present case, is attained by members in the original training set. As such, reaching even higher values could be an impossible task, and the apparent failure to do so should thus not be attributed to the AL strategy.

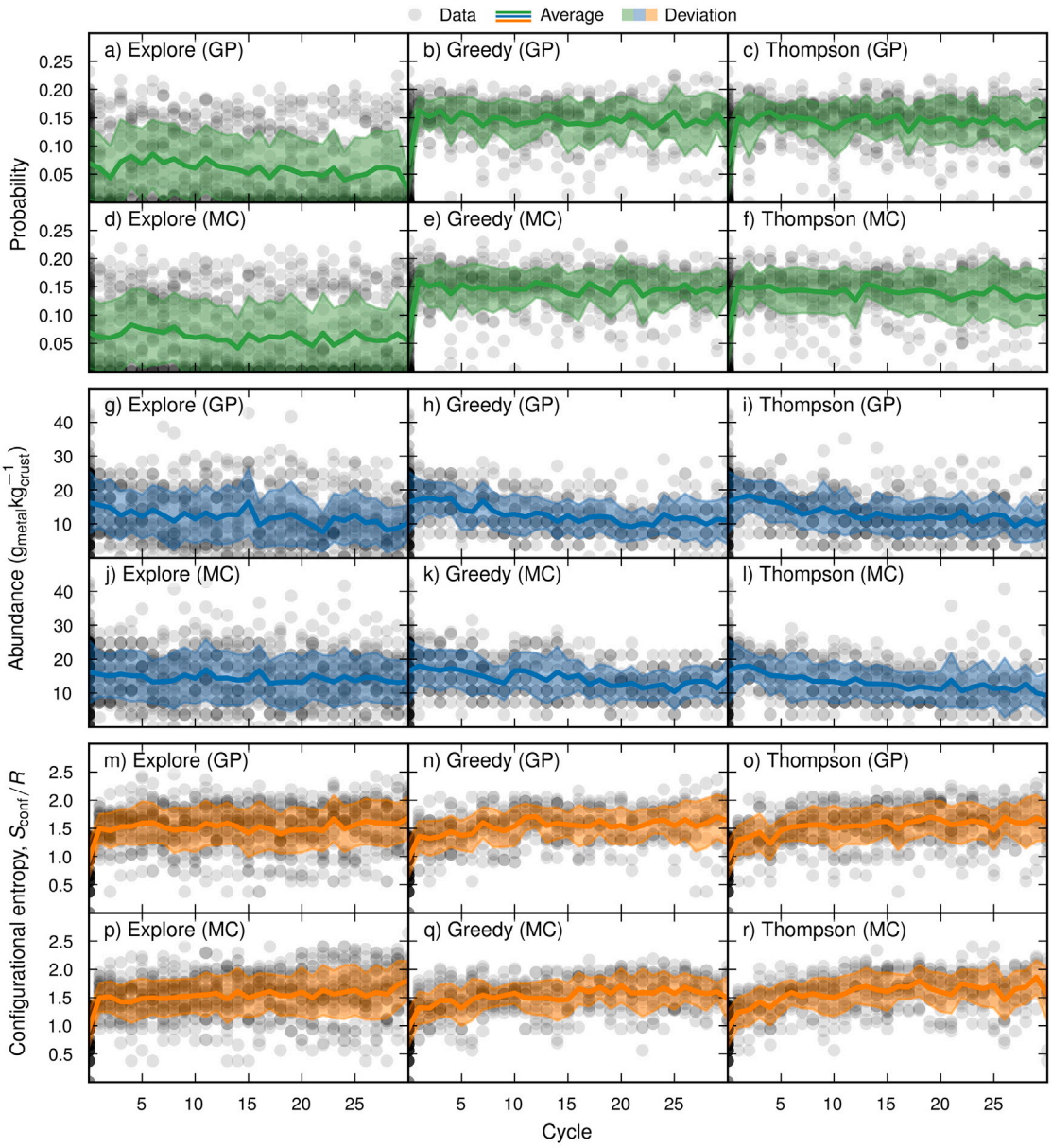

Fig. 5. Scatter plots with the probability that a given OC candidate is useful for CLAS (a–f); average abundance of the metallic elements in the earth's crust (g–l); and configurational entropy, S_{conf}/R, (m–r) in the original dataset (light gray) as well as those added in each cycle (purple) during five individual runs, after removing duplicates. The corresponding averages (solid line) and standard deviations (filled curve) are also shown.

3.3. Chemical looping oxygen uncoupling

Given the inherent issues related to the use of a probability to assess OC performance as well as the severe restrictions placed on the original training data, it was deemed prudent to go beyond the framework used by Wang et al. [13] and instead consider the problem of finding high entropy oxygen carriers (HEOCs) for CLOU. As detailed in Section 2, the OTC was used as the target property since it governs the ability to transfer oxygen between the air and fuel reactors and, therefore, is a well-defined and common measure of the usability of a given OC. The AL was, moreover, initiated based on a 1.5 times larger database of structures, with random compositions that included as many as seven distinctive metallic elements. These factors were found to have a profound impact on the aggregated results of the AL cycles (see Fig. 6). Specifically, the distribution with respect to the OTC shifts to significantly higher values compared to the original data, in terms of both the average and the maximum, when using Thompson or greedy sampling together with GP. As should be expected, the distribution of the configurational entropy is centered around a higher value compared to the original data—which is not by design even though the explicit goal is to find HEOCs (see Fig. 6b,e). As earlier, the opposite is true for the average abundance (see Fig. 6c,f).

A clearer distinction between the two strategies emerges upon analyzing the averages of the properties over the cycles (see Fig. 7). In particular, the Thompson method gives a relatively monotonous increase in the OTC, modulated by small fluctuations (see Fig. 7b). For a greedy sampling, meanwhile, it is difficult to detect a clear trend in the average OTC, after an initial, sharp rise; in fact, there are substantial dips at irregular intervals, especially when few new candidates are added, and even a decreasing trend during the last 10 cycles. Thompson sampling would, in other words, have been superior if fewer cycles had been performed. Notwithstanding the cases discussed in the two previous subsections, there is no persistent increase in the configurational entropy, beyond the first few cycles (see Fig. 7b–c). Similarly, the average metal abundance remains relatively constant, and even has

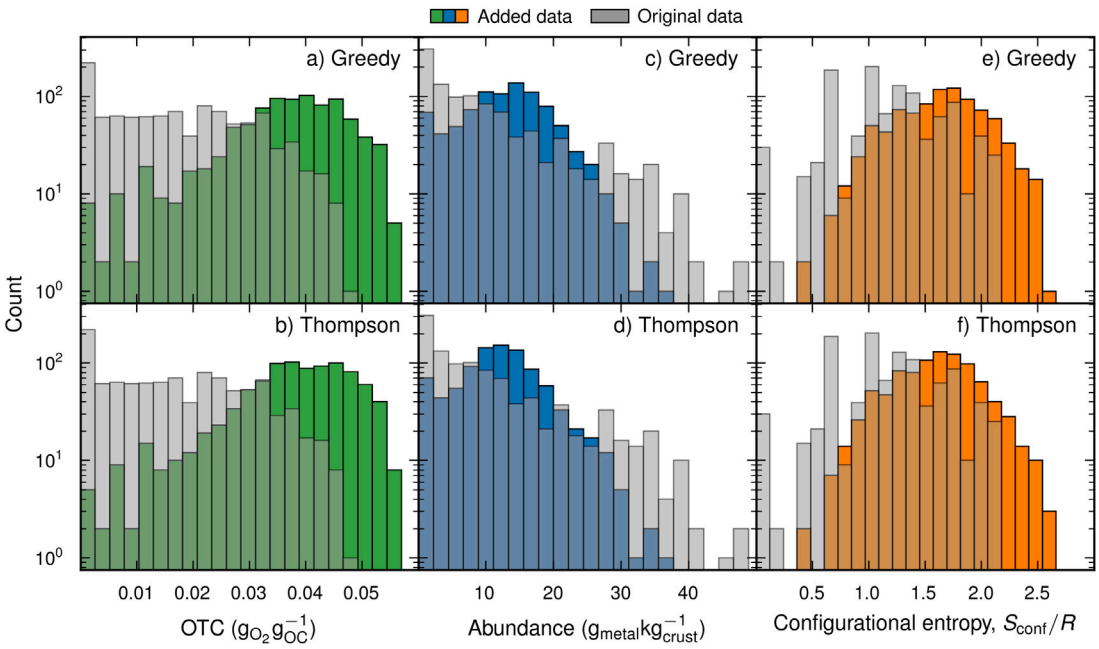

Fig. 6. Histograms based on the OTC (a,d); average abundance of the metallic elements in the earth's crust (b,e); and configurational entropy, S_{conf}/R, (c,f) of HEOCs, in which the original dataset is also shown (light gray). Note that the data from the first 10 cycles were compiled from 10 individual runs, after removing duplicate entries.

a tendency to rise slightly rather than decline (see Fig. 7e–f). Another interesting observation is that the variations in mean values for all three properties are less pronounced during the first 10 cycles, in spite of the large spread among the individual points, due to the fact that the corresponding data were generated in 10 separate runs. Interestingly, this is followed by a substantial rise in the OTC, especially when using the Thompson method. As such, it could be argued that parallel model training based on random seeds can be efficient initially or when using greedy sampling, as this leads to a more stable progression. Even so, it should be stressed that the associated computational demand is substantially higher since many more candidates have to be validated in each cycle. Evidently, this extra effort is associated with a modest learning rate, indicating the need for a sober attitude towards training multiple models in parallel.

When considering the elemental distribution among the candidate HEOCs, it is crucial to remember that the original training data set consists of completely random compositions, in contrast to the cases discussed in the two previous sections (see Figure S10e–f). The average concentrations should therefore provide a direct measure of which atomic species are the most beneficial for CLOU. Specifically, it is found that Ca, Y, and Mn are by far the most common, both when using a greedy and Thompson-based sampling (see Figure S10a-d). This should be regarded as a validation of both the AL strategy as well as the MLIP-based computations since $CaMnO_3$, and doped variants thereof, have been extensively tested in studies of CLOU [14–18]. The fact that the standard deviations are considerably larger during the first 10 cycles is not unexpected, given that parallelized learning, as mentioned above, leads to a wider spread in terms of the key properties. The candidates from the later cycles should, nonetheless, be regarded as the most interesting since their OTC tends to be higher. It is, hence, noteworthy that significant shifts can be observed when comparing the average compositions from cycles 1–10 and 11–50, respectively. In particular, the concentrations of elements other than Ca, Y, and Mn drop significantly, albeit from low levels. The only exception is Ti, whose average actually increases, indicating that this species is a possible dopant, in agreement with reported experiments [15,18].

Before continuing, it is helpful to restate that the OTC stems directly from the difference in oxygen vacancy concentration between the air and fuel reactors, weighted by the molar mass of the perovskite. As the non-stoichiometry (δ) was only allowed to vary between 0.0 and 0.5, the same holds true for the difference ($\Delta\delta$), which means that the HEOCs can be roughly categorized as having a low, medium or high oxygen release capability if $0 \leq \Delta\delta \leq 0.1875$, $0.1875 \leq \Delta\delta \leq 0.375$, and $0.375 \leq \Delta\delta \leq 0.5$, respectively. By analyzing the candidate counts calculated based on these groupings, valuable insights can be gained regarding which atomic species—when added to $Ca_{1-x}Y_xMnO_3$ for the purpose of forming HEOs—allow a high OTC to be maintained (see Figure S11). Although one would anticipate that many candidates containing Ca, Y, and Mn should belong to the "high" category, given the data presented earlier, significant variations can still be observed. For instance, the corresponding count continuously increases with the amount of Mn and reaches its highest value when this element is found on all 16 B sites, while Ca peaks at nine atoms per $A_{16}B_{16}O_{40}$ supercell and Y at six. Of those that remain, only Sm, Cu, and Ti display a maximum above the lowest limit, more precisely at three, two, and four atoms per supercell, respectively. While a successive decline is observed for the other species, more than a few of the best candidates do contain a limited concentration of La, Sr, Co, Mg, and Fe, in fact, there are quite many "high" examples in which either of these, except Sr and Mg, occupy all available sites.

The previous analysis suggests that the most promising materials are based on $Ca_{1-x}Y_xMnO_3$, with $x < 0.5$, in which a few Mn atoms have been replaced by Ti; Co; Fe; Mg; or Cu and some Ca/Y by Sm; La; or Sr. This is perfectly reflected in the 20 highest-ranking materials, identified via Thompson sampling based on GP. In fact, this list is almost exclusively made up of $(Ca_{16-y}Y_y)(Mn_{16-x}Ti_x)O_{48}$, where $3 \leq y \leq 7$ and $1 \leq x \leq 6$, with small additions of K as well as Co, Cu, Mg, and Fe; i.e., no more than three and four substitutions on the A and B site, respectively (see Table S4). Moreover, the vast majority have a configurational entropy above the 1.5R threshold and should, thus, be regarded as HEOCs, including the top four candidates: $Ca_{13}CoMgMn_9Ti_5Y_3O_{48}$, $Ca_{11}Mg_3Mn_8Ti_5Y_5O_{48}$, $Ca_{11}FeMg_3Mn_8Ti_4Y_5O_{48}$, and $Ca_{13}Co_2FeMn_9Ti_4Y_3O_{48}$.

4. Conclusions

The study at hand has not only demonstrated the capability of AL to consistently identify promising materials based on specific conditions.

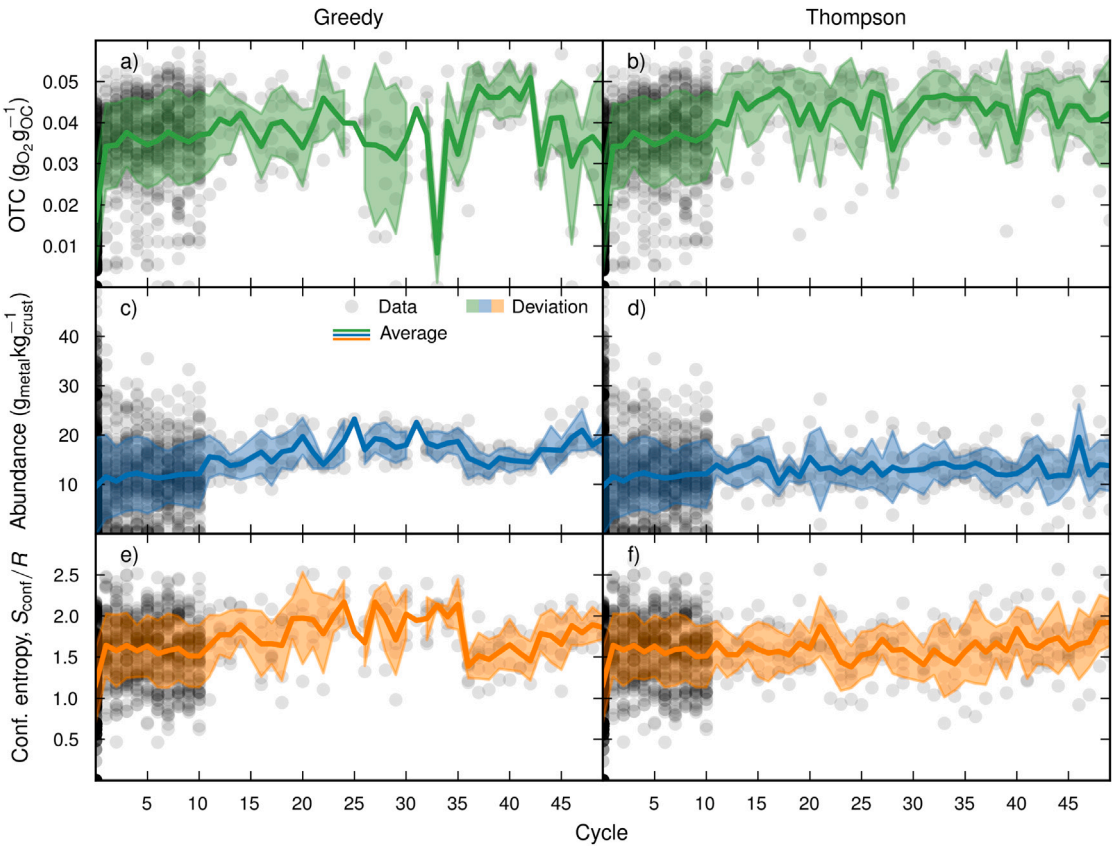

Fig. 7. Scatter plots with the OTC (a–b); average abundance of the metallic elements in the earth's crust (c-d); and configurational entropy, S_{conf}/R, (e–f) of the HEOCs in the original dataset (light gray) as well as those added in cycles 1–10, which includes data from 10 individual runs, (purple) and 11–50 (pink). The corresponding averages (solid line) and standard deviations (filled curve) are also shown.

Convincingly, new candidates with equal or enhanced performance were found when a 10 % subset was used as a starting point, in spite of the artificial measure, in the form of the probability that a given OC is useful for CLAS or CLDR, and a non-optimal set of initial training structures. Especially impressive results were observed in the latter, and presumably more difficult, case. Even so, it is assumed that a similar level of improvement could have been obtained when considering CLAS if more stringent requirements could had been implemented. The fact that a greedy strategy fares better than Thompson sampling, which balances exploration and exploitation, is deemed the direct consequence of either a lucky coincidence or, more probably, prior knowledge: the original training data already covers the limited portion of configurational space with the best materials. It would seem that conventional methods can rival or even excel their more advanced counterparts under such circumstances, namely when the area of interest can be limited based on *a priori* information. Indeed, the initial grid search performed by Wang et al. [13] was evidently sufficient, even though said authors trained a ML model using the resulting data to facilitate the sampling of compositions with additional components.

While the results presented here indicate that the relative efficacy of different AL strategies is strongly correlated with the exact definition and scope of the problem, the ability to balance exploration and exploitation is beneficial, especially in terms of robustness. As the search for CLDR OCs showed, UEs based on GP, as opposed to MC dropout, provides a higher learning rate and average value on the target property. In addition, greedy sampling has a tendency to become trapped and consequently find fewer promising candidates compared to the Thompson method. Even though this issue can be circumvented by training multiple models in parallel, this is accomplished at the cost of slower progress and greater computational demand.

The insights gained from the initial trial studies were applied to the problem of finding HEOCs for CLOU, which proved to be a very successful enterprise. More precisely, a plentiful number of candidates with better performance and larger configurational entropy compared to the randomly generated initial training data were identified, thereby showcasing the true potential of AL-based strategies. In agreement with available experimental data, $CaMnO_3$-based materials displayed the highest OTC, especially when doped by species such as Y, Sm, Ti, Fe, and Mg.

Unfortunately, HEOs have only been tentatively explored as OCs for CL applications; apparently, the only mention of high entropy perovskites in such a context appears in the recent article by Wang et al. [13] and even then concerns materials in which five distinct metallic species are distributed between two sublattices and thus fail to satisfy the criteria outlined by Aamlid et al. [4]. The results presented here will hopefully spur the CL research community to apply a wider perspective when searching for new OCs, specifically, by exploring and testing the plethora of high-entropy and multi-component oxides such as the ones identified in this study. Indeed, it would be very helpful to build public databases with standardized experimental evaluations of OCs, including informations regarding the OTC, preferably under different conditions, as well as other key mechanical and chemical properties. This kind of data could not only be used to train ML models, for high-throughput studies or AL-based discovery schemes, but also serve as a benchmark or offer the chance of applying corrections when performing investigations that are purely based on first-principles calculations, as in the case at hand.

Approaches such as those used in this study serve a plethora of purposes simultaneously: offering a fully automatic framework for identifying specific, promising candidates and, concomitantly, generating valuable qualitative information regarding favorable chemistries. In

other words, they can be useful for guiding experimental work either as inline tools and possibly incorporated into the workflow of completely autonomous laboratories, or to perform, e.g., a pre-study. Regardless of the specific implementation, it is envisioned that the introduction of AL represents a paradigm shift within materials science, with the prospect of boosting the development of green technologies and thereby aiding the transition to a more sustainable society.

CRediT authorship contribution statement

Joakim Brorsson: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Henrik Klein Moberg:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jonatan Gastaldi:** Writing – original draft, Validation, Investigation, Formal analysis, Data curation. **Joel Hildingsson:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Tobias Mattisson:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Anders Hellman:** Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the Swedish Research Council (2020–03 487) and the computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC and C3SE, partially funded by the Swedish Research Council through grant agreement no. SNIC 2021/3-41, SNIC 2021/5-623, SNIC 2021/5-561, and SNIC 2022/5-156.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.mtener.2026.102239.

References

[1] D.A. Scott, R. Schwab, Metallographyin archaeology and art, in: Metallography in Archaeology and Art, in: Cultural Heritage Science, Springer International Publishing, Cham, 2019, pp. 7–17, http://dx.doi.org/10.1007/978-3-030-11265-3_2.

[2] E.P. George, D. Raabe, R.O. Ritchie, High-entropy alloys, Nat. Rev. Mater. 4 (8) (2019) 515–534, http://dx.doi.org/10.1038/s41578-019-0121-4.

[3] M. Fu, X. Ma, K. Zhao, X. Li, D. Su, High-entropy materials for energy-related applications, IScience 24 (3) (2021) http://dx.doi.org/10.1016/j.isci.2021.102177.

[4] S.S. Aamlid, M. Oudah, J. Rottler, A.M. Hallas, Understanding the role of entropy in high entropy oxides, J. Am. Chem. Soc. 145 (11) (2023) 5991–6006, http://dx.doi.org/10.1021/jacs.2c11608.

[5] A. Merchant, S. Batzner, S.S. Schoenholz, M. Aykol, G. Cheon, E.D. Cubuk, Scaling deep learning for materials discovery, Nature 624 (7990) (2023) 80–85, http://dx.doi.org/10.1038/s41586-023-06735-9.

[6] J. Schmidt, N. Hoffmann, H.-C. Wang, P. Borlido, P.J.M.A. Carriço, T.F.T. Cerqueira, S. Botti, M.A.L. Marques, Machine-learning-assisted determination of the global zero-temperature phase diagram of materials, Adv. Mater. 35 (22) (2023) 2210788.

[7] C. Zeni, R. Pinsler, D. Zügner, et al., A generative model for inorganic materials design, Nature 639 (2025) 624–632, http://dx.doi.org/10.1038/s41586-025-08628-5.

[8] C.Y. Lau, M.T. Dunstan, W. Hu, C.P. Grey, S.A. Scott, Large scale in silico screening of materials for carbon capture through chemical looping, Energy Env. Sci. 10 (2017) 818–831, http://dx.doi.org/10.1039/C6EE02763F.

[9] J. Brorsson, V. Rehnberg, A.A. Arvidsson, H. Leion, T. Mattisson, A. Hellman, Discovery of oxygen carriers by mining a first-principle database, J. Phys. Chem. C 127 (20) (2023) 9437–9451, http://dx.doi.org/10.1021/acs.jpcc.2c08545.

[10] Z. Rao, P.-Y. Tung, R. Xie, Y. Wei, H. Zhang, A. Ferrari, T. Klaver, F. Körmann, P.T. Sukumar, A.K. da Silva, Y. Chen, Z. Li, D. Ponge, J. Neugebauer, O. Gutfleisch, S. Bauer, D. Raabe, Machine learning-enabled high-entropy alloy discovery, Science 378 (6615) (2022) 78–85, http://dx.doi.org/10.1126/science.abo4940, URL https://www.science.org/doi/abs/10.1126/science.abo4940.

[11] Y. De Vos, M. Jacobs, P. Van Der Voort, I. Van Driessche, F. Snijkers, A. Verberckmoes, Development of stable oxygen carrier materials for chemical looping processes—A review, Catalysts 10 (8) (2020) http://dx.doi.org/10.3390/catal10080926, URL https://www.mdpi.com/2073-4344/10/8/926.

[12] Y. Yan, T. Mattisson, P. Moldenhauer, E.J. Anthony, P.T. Clough, Applying machine learning algorithms in estimating the performance of heterogeneous, multi-component materials as oxygen carriers for chemical-looping processes, Chem. Eng. J. 387 (2020) 124072, http://dx.doi.org/10.1016/j.cej.2020.124072, URL https://www.sciencedirect.com/science/article/pii/S1385894720300632.

[13] X. Wang, Y. Gao, E. Krzystowczyk, S. Iftikhar, J. Dou, R. Cai, H. Wang, C. Ruan, S. Ye, F. Li, High-throughput oxygen chemical potential engineering of perovskite oxides for chemical looping applications, Energy Env. Sci. 15 (2022) 1512–1528, http://dx.doi.org/10.1039/D1EE02889H.

[14] J. Adánez, A. Abad, T. Mendiara, P. Gayán, L. de Diego, F. García-Labiano, Chemical looping combustion of solid fuels, Prog. Energy Combust. Sci. 65 (2018) 6–66, http://dx.doi.org/10.1016/j.pecs.2017.07.005, URL https://www.sciencedirect.com/science/article/pii/S0360128517300199.

[15] X. Zhu, K. Li, L. Neal, F. Li, Perovskites as geo-inspired oxygen storage materials for chemical looping and three-way catalysis: A perspective, ACS Catal. 8 (9) (2018) 8213–8236, http://dx.doi.org/10.1021/acscatal.8b01973.

[16] D. Jing, F. Snijkers, P. Hallberg, H. Leion, T. Mattisson, A. Lyngfelt, Effect of production parameters on the spray-dried calcium manganite oxygen carriers for chemical-looping combustion, Energy Fuels 30 (4) (2016) 3257–3268, http://dx.doi.org/10.1021/acs.energyfuels.5b02872.

[17] D. Jing, T. Mattisson, H. Leion, M. Rydén, A. Lyngfelt, Examination of perovskite structure caMnO$_{3-\delta}$ with mgo addition as oxygen carrier for chemical looping with oxygen uncoupling using methane and syngas, Int. J. Chem. Eng. 2013 (1) (2013) 679560, http://dx.doi.org/10.1155/2013/679560, URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2013/679560.

[18] T. Mattisson, Materials for chemical-looping with oxygen uncoupling, ISRN Chem. Eng. 2013 (2013) 526375, http://dx.doi.org/10.1155/2013/526375.