nature portfolio

Peer Review File

Prediction of Perovskite Oxygen Vacancies for Oxygen Electrocatalysis at Different Temperatures

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Review of “Prediction of perovskite electrocatalysts for operation at different temperatures” by Li, Z., et al., submitted to Nature Communications

This is an interesting and timely paper. The tradeoff of O vacancy content with catalytic activity is very useful to know about, and the optimization of a new cathode material and its subsequent testing validates the overall approach of the work. Prior to publication, the following major revisions need to be addressed:

The authors state in the abstract “This work can be a starting point for a paradigm shift toward the rational development of catalysts for energy applications.” Many researchers may take issue with such a statement, as rational development of catalysts has been going on for decades, in the form of using physics-based descriptors to understand trends and suggest new materials (e.g., precious metal alloys). More recently, computational-based screening has been used with the aid of such descriptors, as well as machine learning-based work (e.g., Ref 22 of the manuscript)

The plots in Figure 1 show that the training data can be difficult to fit with extremely high accuracy (there are a number of points away from the y=x dashed line). Can the authors describe why this may be the case? Is it due to uncertainty in the experimental data?

Numerous additional details regarding the data are needing to be described and resulting limitations discussed. First, only Co and Fe-containing perovskites were examined. This places very strict domain of applicability limits on the resulting ML model. In short, the model is unlikely to be able to predict perovskites that aren’t mostly Fe- or Co-based. (Another question- what are the composition bounds of the Fe and Co in the data?) Second, how many unique compositions are there? Given that there are thousands of data points, presumably at different temperatures, this indicates that there is a high degree of twinning occurring in the data between train and test sets. In this way, random cross validation is bound to give very good results of the model. The authors should try additional tests, like leave composition out, leave temperature out, leave out Fe/Co, etc. to push the limits of their model.

Related to the above points, does the model give physically sensible results for other perovskites not in the training set. For example, what is the predicted oxygen vacancy content for $LaMnO_3$ or $La_{0.8}Sr_{0.2}MnO_3$?

In addition to reporting R2, it is helpful to give other metrics like mean absolute error and normalized root mean squared error (i.e., RMSE/standard deviation of the data). MAE would be nice to give a sense of absolute error in the delta value.

To better understand the role of features in the model, a feature importance analysis like SHAP (Shapely Additive Explanations) would be helpful and probably worth including.

Authors should probably cite the recent review from Giordano et al. (https://doi.org/10.1021/acs.accounts.1c00509) detailing the role of electronic structure (including oxygen band center and cation redox) on setting many perovskite catalyst properties.

On page 10, authors discuss the wide range of reported values even for the same material for a given catalytic property. This issue has been discussed in detail in the recent paper from Jacobs et al. (https://doi.org/10.1021/acsaem.4c00125), where they found the average experimental deviation of measurements was roughly half a log unit.

In Figure 2, it is a bit confusing why formation energy for O vacancies decreases as more O vacancies are added, especially for delta > 0.5. At delta = 0.5, the perovskite is essentially a Brownmillerite with formula ABO2.5, where, if A is in 3+ oxidation state, then all of the B cations are in the 3+ oxidation state, and further creation of O vacancies would lead to formation of some B in 2+ state, which is typically unfavorable for Fe- and Co-containing perovskites.

The authors have missed a key (albeit recent) reference using machine learning for perovskite catalyst materials discovery. The work from Jacobs et al. (https://doi.org/10.1002/aenm.202303684) used the database from the work cited above, evaluated ML models, and then screened for new catalysts. This work builds on the initial ML screening work done by Zhai et al. (Ref 22 in the manuscript), and should be discussed accordingly.

The use of V and W as dopant elements in perovskite catalysts for fuel cell and electrolyzer applications is rare. Is that a novelty of this study or was it first conceived in a different study? Were these dopants proposed strictly as part of the ML screening to optimize the vacancy content?

All of the data used to train and evaluate the ML model should be freely available, either as a spreadsheet as part of the SI, or hosted in that form on another open-source repository (e.g., Figshare). It would be preferable if the data were in ML-ready format, or at least close to it. This will enable other researchers to easily use your collected O vacancy data to build additional ML models or combine it with other models. Sharing all of the data is essential.

As a final small note, there are numerous typos throughout, especially in the Methods section.

Reviewer #2 (Remarks to the Author):

Applying machine learning techniques to screen high-performance new materials and study the mechanisms of material properties is an emerging and efficient approach. In the manuscript, authors proposed a machine learning-based approach to developing oxygen electrocatalysts for operation at different temperatures. The idea is nice, but the present work seems still need significant improvement.
1. The title is too broad and should focus on the predicted property and the main applications in the manuscript.
2. Screening target oxides through machine learning prediction is a key step in the current manuscript,

but the reliability of predictions made by machine learning model cannot be judged, because the training/selection process of machine learning models seem not be fully described. The important links involved in the machine learning model design process include random division of training/test/validation set, hyperparameter setting, cross-validation of the model on randomized data sets, and sufficient measurement of model fitting effect, but these aspects are not mentioned in the paper.

3. For model selection, a single R-square is insufficient, as it can only characterize the explanatory power of the model prediction variables for the response variables on the training set, but cannot reflect the model's predictive ability on the unknown test set. Other metrics like RMSE is indispensable. In addition, the interpretability of the model is also a very important indicator for model selection.

4. During model building, seven feature were used as predictor variables, but the reasons and rational for this selection were not given. In Fig.1, what does the color represents in the heatmap of descriptors?

5. For (b) and (c) in Fig.1, the model is very likely overfitting. More extensive cross-validation is needed on randomized test/validation datasets.

6. Why do the authors divide the oxides into Co-based and Fe-based oxides for the oxygen vacancy prediction? The reviewer holds the idea that there is no meaning to differ them, since all the elements can be characterized by their properties, such as electronegativity, cation size, etc, that the authors used in this manuscript. The reviewer suggests to train and predict all the oxides with one model.

7. The database should be uploaded.

8. The concentration of oxygen vacancy is closely related to the preparation processes. It is unconvincing not to consider the preparation process, such as sintering temperature, atmosphere, dewelling time. For instance, there is a big difference in oxygen vacancy content for the same material synthesised in Ar or O2 or at high/low temperatures. Such conditions are ignored when the model is trained, which makes the following results are incorrect in the manuscript.

9. Thermogravimetric (TG) analysis was performed to weigh the mass loss of sample at 150-800 oC, and thus calculate the δT at elevated temperature. To eliminate the moisture in the sample, efforts were made to maintain the temperature at 150-200 °C for 1.5 h. However, many oxides are of hydration reaction capability even at 500 oC or higher. That is to say, the weight loss at the temperature below 500 oC may be caused by loss of water or/and oxygen.

10. In Figure S5, there are some data tested by the authors, please provide the original LSV curves for OER of these oxides.

11. The reviewer suggest add a figure for predicted and measured oxygen canvancy shown in Figure 2b and 2c. How is the plot of "ASPR at 450 oC versus oxygen vacancy at 450 oC"?

12. In line 288 and 289, why was SC choosen? As known to us, Sr is easy to react with CO2 and easily segregated, Co based oxides are high TEC.

13. Figure 2d is missing in the main text.

14. In Figure S18, I-V curve of Cell 2 is from 0-OCV, cell-3 is different from it, why?

15. In line 350, please provide the xrd patterns after CO2 tolerance test for SCTV and BSCF.

16. The cells show excellent performances at low temperature, how about the larger cells, such as the diameters of 40 mm or even larger?

Reviewer #3 (Remarks to the Author):

This paper reports a method for predicting oxygen vacancy concentrations without the need for extensive theoretical calculations. The reviewer appreciates that the computational predictions agree with the experiments, resulting in the discovery of electrocatalysts with better performance than the existing materials. However, the present study was only a composition optimization of perovskite-type oxides, which is not surprising in terms of the development of new materials. Although it is interesting to focus on oxygen deficiency, there have been several reports on calculation methods for composition optimization. Regarding the power generation performance of SOFCs, it is worth evaluating that a relatively high power density was achieved even at 500 °C using the developed materials. However, the significance of low-temperature operation and electrode development is lost, as the presented performance is significantly lower than that of existing SOFCs operating at 800°C. In addition, this paper has issues that need to be corrected, which are enumerated below.
In conclusion, the reviewer is compelled to point out that the paper has insufficient impact to be published in Nature Communications, either in terms of computational methods, new material development, or fuel cell properties, and recommends transferring to more specialized journals, such as Communications Materials.

1. The Introduction section was vague. For example, although the authors evaluated new cathode materials at low temperatures, motivation for developing low-temperature solid oxide fuel cell (SOFC) cathode materials has not been described.

2. Although the Methods section should contain all the elements necessary for the interpretation and replication of the results, the Methods section of this literature is incomplete.

3. The authors dried the samples at 150–200 °C before the TG measurements; however, it is concerned that the temperature was insufficient to remove water completely. Proton-conducting perovskite oxides such as BZY retain water within the solid up to approximately 600 °C. Therefore, drying for dehydration should be performed at higher temperatures. Furthermore, the influence of water must be completely eliminated to obtain accurate experimental values that can be compared with the calculated values, and it is recommended that the authors take great care to prevent water during measurement. From the manuscript, the reviewer could not identify any experimental care to accurately estimate the oxygen content.

4. The number of oxygen vacancies calculated from the TG may differ from the actual amount owing to the influence of water. Because measuring the number of oxygen vacancies is one of the important experiments in this paper, the reviewer recommends a double check using a different method (e.g., high-temperature neutron diffraction measurements).

5. The authors claimed to have discovered two new perovskite oxides (SCTW and SCTV). However, the only data available to show that these materials were synthesized are the powder X-ray peak patterns,

which is insufficient evidence to prove the discovery of new materials. High-quality diffraction data should be measured and Rietveld refinement should be performed to determine the structural parameters of these materials. Furthermore, these materials are merely compositional optimizations and are not new materials.

6. Are the lattice parameters of the solid solutions in Fig. S7 and Fig. S16 reasonable?

7. It is intriguing that the ORR activity shows a volcano-trend dependence on the number of oxygen vacancies (Fig. 2b and 2c). This volcanic trend increases linearly with increasing oxygen vacancies and then decreases linearly with increasing oxygen vacancies. The authors should explain why this trend varies "linearly."

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Review of "Prediction of perovskite electrocatalysts for operation at different temperatures" by Li, Z., et al., submitted to Nature Communications

This is an interesting and timely paper. The tradeoff of O vacancy content with catalytic activity is very useful to know about, and the optimization of a new cathode material and its subsequent testing validates the overall approach of the work. Prior to publication, the following major revisions need to be addressed:

The authors state in the abstract "This work can be a starting point for a paradigm shift toward the rational development of catalysts for energy applications." Many researchers may take issue with such a statement, as rational development of catalysts has been going on for decades, in the form of using physics-based descriptors to understand trends and suggest new materials (e.g., precious metal alloys). More recently, computational-based screening has been used with the aid of such descriptors, as well as machine learning-based work (e.g., Ref 22 of the manuscript)

Response:

Thank you for this comment. Accordingly, we have revised this statement in the abstract as follows:

"This work revealed the dominant role of local cationic environment in determining the state of redox-active metal centres and thus the concentration of the oxygen vacancies over perovskites, enabling rational development of new catalysts for energy applications".

The plots in Figure 1 show that the training data can be difficult to fit with extremely high accuracy (there are a number of points away from the y=x dashed line). Can the authors describe why this may be the case? Is it due to uncertainty in the experimental data?

Response:

Yes, you are correct. This represents the uncertainties associated with work across different research groups and testing procedures to obtain the oxygen vacancies. However, as compared to other recent reports (**Refs**: *Science*, **363**, 247 (2019); *Nat. Mater.*, **15**, 1120–1127 (2016); *Phys. Rev. Mater.*, **4**, 063801 (2020); *ACS Appl. Energy Mater.*, **7**, 3366–3377 (2024)), this result represents a far improved accuracy of the prediction over reported experimental data.

Numerous additional details regarding the data are needing to be described and resulting limitations discussed. First, only Co and Fe-containing perovskites were examined. This places very strict domain of applicability limits on the resulting ML model. In short, the model is

unlikely to be able to predict perovskites that aren't mostly Fe- or Co-based. (Another question- what are the composition bounds of the Fe and Co in the data?)

Response:

We used the Co- and Fe-containing perovskites as the model catalysts, because Co and Fe represent the most catalytically active metal centers for the targeted oxygen reduction reaction in perovskite materials. For this reason, this work sets the lower bounds for Co or Fe in these perovskites based on whether they are the major components in the B-site of perovskites. Due to the high tunability of the perovskite composition, a massive compositional space can be explored using this reported approach. Importantly, this approach can be used to predict other metal-based catalysts if sufficient size of the data training set is available.

Second, how many unique compositions are there? Given that there are thousands of data points, presumably at different temperatures, this indicates that there is a high degree of twinning occurring in the data between train and test sets. In this way, random cross validation is bound to give very good results of the model. The authors should try additional tests, like leave composition out, leave temperature out, leave out Fe/Co, etc. to push the limits of their model.

Response:

To ensure a fair evaluation, we adopted a 5-fold cross-validation method, as described in the Methods section of machine learning in the manuscript. In this approach, we split the data into five subsets, using 80 % of the data for training and the remaining 20 % for testing, repeating this process five times. This method minimizes the risk of overfitting and ensures that our model is evaluated on different subsets of the data, reducing the likelihood of twinning issues between the train and test sets.

The unique feature of this proposed approach is to leave the Fe/Co out, meaning that there is no active metal centre considered in the machine learning model. Through this way, we show that we could avoid the challenges to predict the electronic structures of the Co/Fe at elevated temperatures. This further highlights the importance of local lattice cationic environment in determining the states of the active metal centres, and therefore the content of the oxygen vacancies.

Additionally, we have conducted partial dependence analysis for each dimension, *e.g.*, polarization, charges, electronegativity, ionic radii, temperature, and tolerance factor, for Co-based and Fe-based perovskites respectively in **Fig. R1**. These analyses provide a comprehensive understanding of the importance and the influence of each feature on the model's predictions. By examining the partial dependence plots, we can see how each feature impacts the predicted outcomes, offering deeper insights into the model's behaviour beyond mere accuracy metrics.

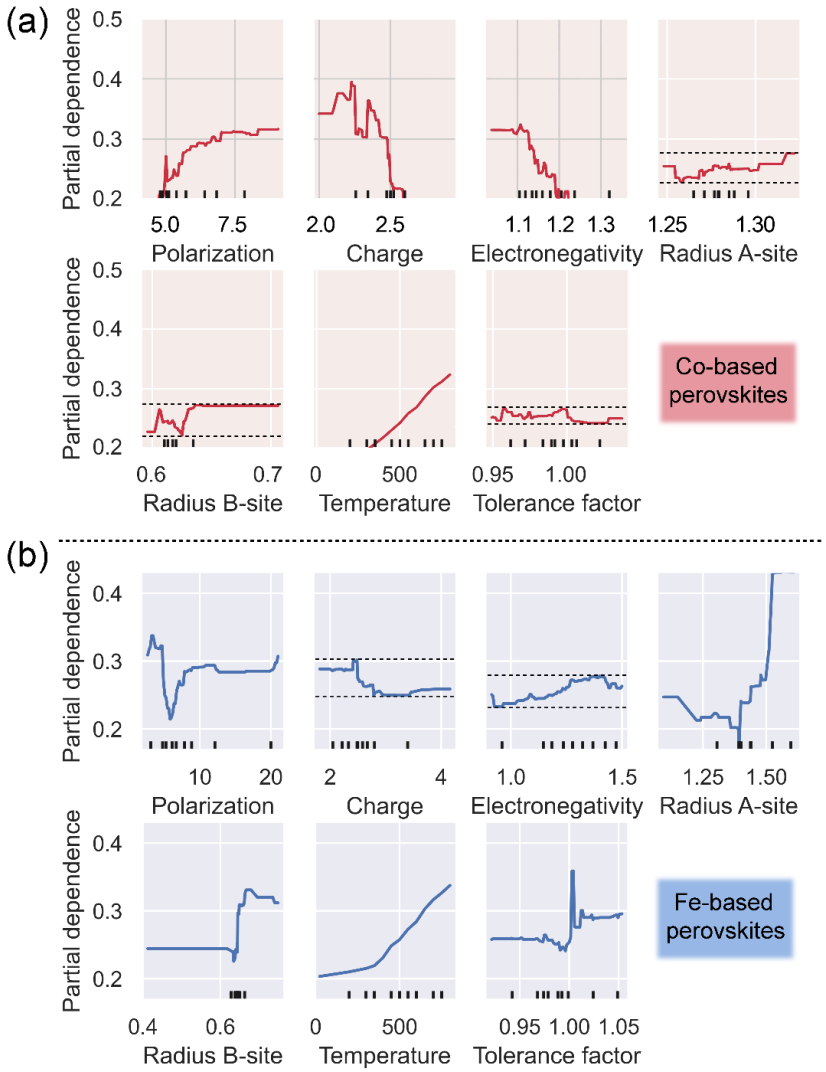

Fig. R1 Partial dependence from LGBM Regressor machine learning model over polarization power, charge, electronegativity, radius A-site, radius B-site, temperature, and tolerance factor for (**a**) Co-based and (**b**) Fe-based perovskites. The dashed lines are guidance for showing the features with low dependence on machine learning model.

Related to the above points, does the model give physically sensible results for other perovskites not in the training set. For example, what is the predicted oxygen vacancy content for LaMnO3 or La0.8Sr0.2MnO3?

Response:

Sorry for the confusion raised from the unclarity of our text. This work needs to first determine the active metal centers and relies on comprehensive existing datasets for model training and predictions. Should these prerequisites be fulfilled, we are confident that the model could also provide physically sensible predictions over Mn-based materials.

Actions: we have added more statements in the Discussion section to illustrate the factors that may affect the accuracy of the prediction as follows:

"Besides, this work needs to determine the active metal centres first and relies highly on comprehensive existing datasets for model training and predictions."

In addition to reporting R2, it is helpful to give other metrics like mean absolute error and normalized root mean squared error (i.e., RMSE/standard deviation of the data). MAE would be nice to give a sense of absolute error in the delta value.

To better understand the role of features in the model, a feature importance analysis like SHAP (Shapely Additive Explanations) would be helpful and probably worth including.

Response:

We have now included the Root Mean Squared Error (RMSE) of regression performance for all machine learning models using 5-fold cross-validation, as suggested (**Fig. R2**). Additionally, we have provided a comprehensive Shapley Additive Explanations (SHAP) feature importance analysis to better understand the role of different features in the model.

The SHAP summary plot shown in the figure provides detailed insights into the impact of each feature on the model output. Each dot represents a SHAP value for a feature, with colors indicating the feature value (blue for low values and red for high values). The x-axis represents the SHAP value, which indicates the impact of the feature on the model's output (oxygen vacancy). Key observations from the SHAP analysis include:

- ***Charge:*** This feature has a significant impact on the model, with higher values generally decreasing the model output.
- ***Electronegativity:*** Electronegativity has a significant effect, with higher values generally leading to lower model outputs.
- ***Temperature:*** There is a notable influence of temperature on the model's predictions, with a mix of positive and negative impacts depending on the specific temperature values.
- ***Polarization:*** Polarization also shows a notable effect, with high polarization values tending to increase the model output.
- ***Radius A-site:*** The A-site radius shows a moderate impact, with both high and low values affecting the model output.
- ***Tolerance Factor:*** This feature exhibits a balanced impact, with higher values contributing positively to the model output.
- ***Radius B-site:*** The B-site radius also demonstrates a balanced impact similar to the tolerance factor.

This SHAP analysis provides a clear understanding of how each feature contributes to the model's predictions, thereby offering a deeper insight into the feature importance within our machine learning model.

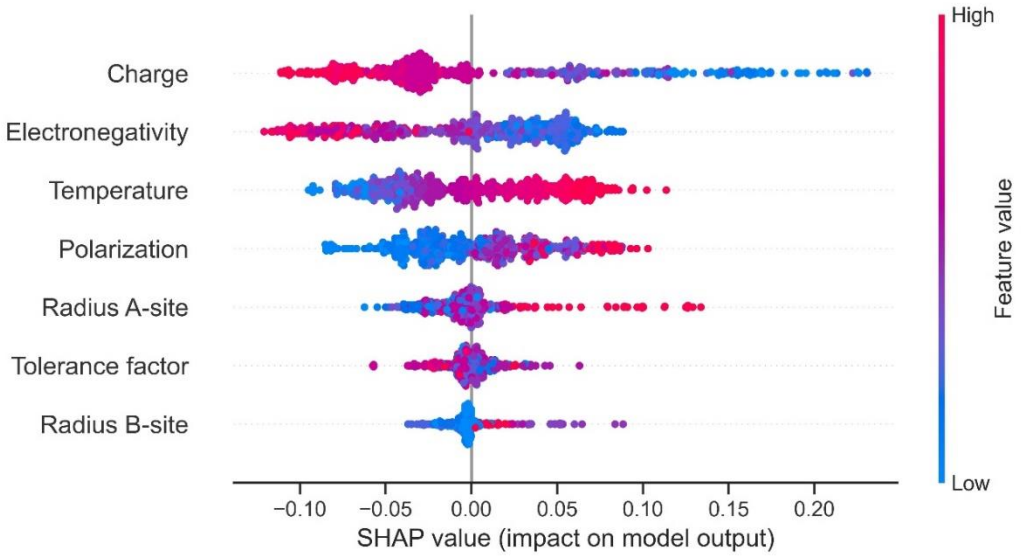

Fig. R2 Distribution of all SHAP values for the seven features.

Action: **Fig. R2** has been added as a new **Fig. S3** in the revised Supplementary Information. We have added relevant statement about SHAP in the revised manuscript as follows:

*"To further elucidate the role of the lattice cationic environment in the oxygen vacancy formation, we evaluated the correlation of each inputted feature with the oxygen vacancies by using partial dependence analysis from the LGBM regressor (**Fig. S2**) and Shapley Additive Explanations (SHAP) feature importance analysis (**Fig. S3**)."*

Authors should probably cite the recent review from Giordano et al. (https://doi.org/10.1021/acs.accounts.1c00509) detailing the role of electronic structure (including oxygen band center and cation redox) on setting many perovskite catalyst properties.

Response:

Thank you for this kind suggestion. This review paper is useful for us to understand existing descriptors in predicting new catalysts and catalytic activity of oxides. So that we have added this reference in the revised manuscript as ***Ref. 23***.

On page 10, authors discuss the wide range of reported values even for the same material for a given catalytic property. This issue has been discussed in detail in the recent paper from Jacobs et al. (https://doi.org/10.1021/acsaem.4c00125), where they found the average experimental deviation of measurements was roughly half a log unit.

Response:

Thanks. We found that the results in this paper is helpful for us to understand the

discreteness of experimental data. We have added this reference in the revised manuscript as ***Ref. 50***.

In Figure 2, it is a bit confusing why formation energy for O vacancies decreases as more O vacancies are added, especially for delta > 0.5. At delta = 0.5, the perovskite is essentially a Brownmillerite with formula ABO2.5, where, if A is in 3+ oxidation state, then all of the B cations are in the 3+ oxidation state, and further creation of O vacancies would lead to formation of some B in 2+ state, which is typically unfavorable for Fe- and Co-containing perovskites.

Response:

We ascribe the increase in O vacancies to the decrease in the formation energy of O vacancies in the materials, because the formation energy is calculated based on the material compositions and structures, which predetermine the concentration of the O concentrations. The result is similar to that of Zhang et al. (**Ref**: *Solid State Ionics*, **316**, 20-28 (2018)).

It is important to note that this work is investigating a wide range of perovskite compositions that are not only limited to the compositions with A in 3+. Further, most of the active perovskites consist of alkali-earth elements that are in 2+ (e.g., Sr^{2+}, Ba^{2+}). Therefore, delta> 0.5 is possible at elevated temperatures without the formation of brownmillerite materials. For example, the $Sr_{0.95}Ag_{0.05}Nb_{0.1}Co_{0.9}O_{3-\delta}$ (SANC, **Ref**: *Nano Lett.*, **16**, 512–518 (2016)), $SrSc_{0.175}Ta_{0.025}Co_{0.8}O_{3-\delta}$ (SSTC, **Ref**: *Compos. Part B: Eng.*, **178**, 107491 (2019)) and $Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$ (BSCF, **Ref**: *Solid State Ionics*, **177**, 1737-1742 (2006)) exhibit cubic structure at high temperature, but the contents of oxygen vacancies in them can exceed 0.5.

The authors have missed a key (albeit recent) reference using machine learning for perovskite catalyst materials discovery. The work from Jacobs et al. (https://doi.org/10.1002/aenm.202303684) used the database from the work cited above, evaluated ML models, and then screened for new catalysts. This work builds on the initial ML screening work done by Zhai et al. (Ref 22 in the manuscript), and should be discussed accordingly.

Response:

Many thanks for this suggestion. This recently published paper from Jacobs et al. (**Ref**: *Adv. Energy Mater.*, **14**, 2303684 (2024)) provides in-depth insights into predicting perovskite catalytic properties for SOFC/SOEC by using machine learning models. Also, this work reports the feasibility of temperature-dependent predictions using ASR machine learning model. We have cited this reference (***Ref. 24***) in the Introduction section.

The unique part of our proposed approach from this paper and other reports is the highlight of local lattice environment in determining the states of the active metal centers, and the use of ML model to predict oxygen vacancies, which is a collective and important material property for many applications involving oxygen kinetics. This allows us to capitalize on literature data to train the ML models and predict precisely the oxygen vacancies without heavy first principal

calculations at elevated temperatures, as stated in this paper from Jacobs et al.

The use of V and W as dopant elements in perovskite catalysts for fuel cell and electrolyzer applications is rare. Is that a novelty of this study or was it first conceived in a different study? Were these dopants proposed strictly as part of the ML screening to optimize the vacancy content?

Response:

Not necessarily. This work just used these two examples (i.e., SCTW and SCTV) to prove the utility of the proposed approach to accelerate the development of perovskite catalysts for SOFC applications. In our previous reports (**Refs:** *Nat. Commun.*, **8**, 13990 (2017); *J. Am. Chem. Soc.*, **143**, 9507–9514 (2021)), as also discussed in this work, Nb or Ta-doped materials also exhibit notably improved catalytic activity towards oxygen reduction reaction. These examples with these uncommon elements further support the validity of the proposed approach to design of future catalysts that are beyond conventional compositions.

Action: Accordingly, we have included discussions to further clarify this point as follows:

"The two examples with uncommon dopants at B-sites (e.g., W and V) are proposed to support the validity of the ML-assisted prediction process to design catalysts that are beyond conventional compositions."

All of the data used to train and evaluate the ML model should be freely available, either as a spreadsheet as part of the SI, or hosted in that form on another open-source repository (e.g., Figshare). It would be preferable if the data were in ML-ready format, or at least close to it. This will enable other researchers to easily use your collected O vacancy data to build additional ML models or combine it with other models. Sharing all of the data is essential.

Response:

The data and codes for this work can be found in *https://github.com/Luoyadan/MLtoolbox-perovskite-oxygen-vacancy-pred*. We will share these data when the work will be accepted.

As a final small note, there are numerous typos throughout, especially in the Methods section.

Response:

They have been corrected accordingly.

Reviewer #2 (Remarks to the Author):

Applying machine learning techniques to screen high-performance new materials and study the mechanisms of material properties is an emerging and efficient approach. In the manuscript, authors proposed a machine learning-based approach to developing oxygen electrocatalysts for operation at different temperatures. The idea is nice, but the present work seems still need significant improvement.

1. The title is too broad and should focus on the predicted property and the main applications in the manuscript.

Response:

Thanks for this comment. We have now revised the title to be:

"*Prediction of Perovskite Oxygen Vacancies for Oxygen Electrocatalysis at Different Temperatures*".

2. Screening target oxides through machine learning prediction is a key step in the current manuscript, but the reliability of predictions made by machine learning model cannot be judged, because the training/selection process of machine learning models seem not be fully described. The important links involved in the machine learning model design process include random division of training/test/validation set, hyperparameter setting, cross-validation of the model on randomized data sets, and sufficient measurement of model fitting effect, but these aspects are not mentioned in the paper.

Response:

Thank you for your valuable feedback regarding the screening of target oxides through machine learning predictions. The reliability of our model selection can be well justified as we conducted 5-way cross-validation and selected the best-performing models across different train/test splits. This approach ensures that our model is evaluated on various subsets of the data, providing a robust measure of its performance.

We used R^2 and RMSE to quantify model performance. Initially, we focused on R^2 because it indicates the proportion of variance explained by the model, while RMSE (**Fig. R3**), being its complement, was not included in the first draft. However, to address your concerns, we have refined the manuscript to include additional figures that present different regression metrics, providing a more comprehensive evaluation of model performance.

Regarding the hyperparameters of our machine learning models, we adopted the default values as described in the scikit-learn libraries (*https://scikit-learn.org/*). To minimize the impact of hyperparameter fine-tuning, we did not fully examine different values of hyperparameters. This choice was made to focus on the inherent predictive power of the models rather than optimization through extensive tuning.

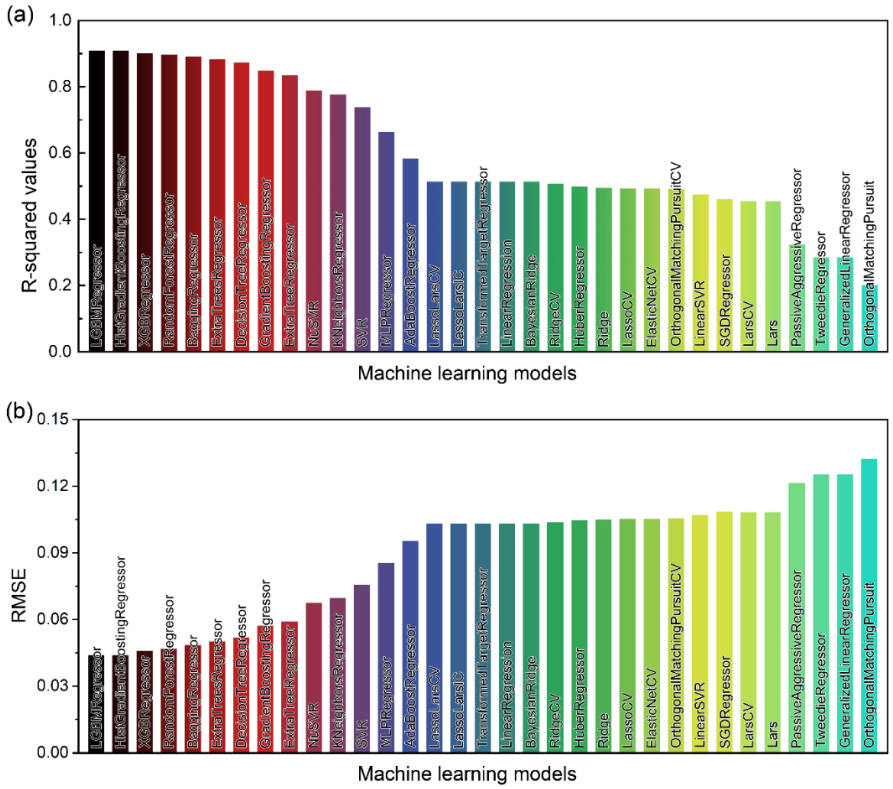

Fig. R3 (a) R-squared and **(a)** RMSE values of top 33 machine learning models. The R-squared and RMSE scores are the averaged mean of the 5 random tests (5-fold cross-validation).

Action: **Fig. R3** has been added as an upgradation of **Fig. S1** in the revised Supplementary Information. To further clarify the details for ML training and validations, we have included the following texts in the revised manuscript:

*"**Fig. 1b** and **Fig. 1c** present an excellent correlation between the predicted and reported oxygen vacancy concentrations for both cobalt- and iron-based perovskite oxides, which is evidenced by the close-to-unity slope of the trendline, high R^2 value (0.92 for cobalt-based and 0.92 for iron-based materials) (**Fig. S1a**) as well as a low value of Root Mean Square Error (RMSE = 0.044), which refers to the average magnitude of the difference between predicted and actual values (**Fig. S1b**). Interestingly, other regression algorithms, such as the Hist-Gradient-Boosting regressor, Random Forest Regressor and XGB-regressor (**Fig. S1a**), can also achieve a good prediction (R^2 > 0.90, RMSE < 0.05)."*

3. For model selection, a single R-square is insufficient, as it can only characterize the explanatory power of the model prediction variables for the response variables on the training set, but cannot reflect the model's predictive ability on the unknown test set. Other metrics like RMSE is indispensable. In addition, the interpretability of the model is also a very important indicator for model selection.

Response:

We acknowledge that a single R^2 value is insufficient to fully characterize a model's performance, as it primarily reflects the explanatory power of the model prediction variables for the response variables on the training set. To address this, we have now incorporated RMSE to provide a more comprehensive evaluation of the model's predictive ability on unknown test sets. The additional metric ensures that our model's performance is thoroughly assessed and its predictive reliability is validated.

Regarding the interpretability of the model, we agree that it is an important aspect of model selection. Tree-based machine learning models, for example, can be interpreted better as each decision can be visualized to understand the conditions leading to specific predictions. However, based on internal discussions, we have observed that model interpretability does not necessarily reflect the physical meaning and can sometimes appear counterintuitive. This discrepancy is why we chose not to disclose the interpretability details in the main text, leaving it as an open question for further investigation.

While interpretability is crucial, it is also important to acknowledge the complexity of the relationships captured by machine learning models, which might not always align with intuitive or physical expectations. Our primary focus has been to ensure the robustness and reliability of the model's predictions, supported by rigorous evaluation metrics like R^2 and RMSE.

4. During model building, seven feature were used as predictor variables, but the reasons and rational for this selection were not given. In Fig.1, what does the color represents in the heatmap of descriptors?

Response:

Based on the inductive effects, the states of the redox-active transition metals (e.g., Co and Fe) can be affected by the cations at A-sites and dopants at B-sites. These features are intrinsic properties of A-site cations and B-site dopants, reflecting the lattice cationic environment in perovskites and thus being applied as predictor variables.

Actions: We have further included the justification for these features in the revised manuscript as follows:

"Six of the features, reflecting the intrinsic properties of cations at A-sites and dopants at B-sites, are used to map the properties of the lattice cationic environment of perovskites, including tolerance factor (τ), electronegativity (χ), polarization power (P), charge (Z), and cation sizes for A-site (r_A) and B-site cations (r_B) [16,31-33]*."*

The heatmap carries no physical meaning. To avoid misconception, we have now upgraded **Fig. R4** as the new **Fig. 1** in the revised manuscript.

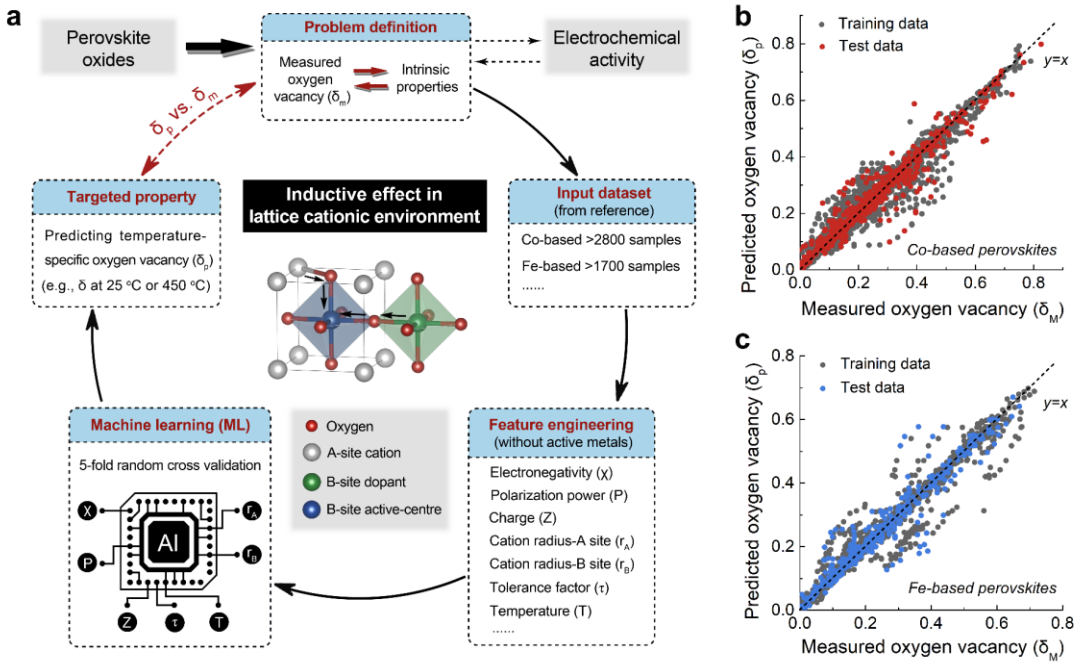

Fig. R4 Machine learning (ML) assisted prediction of the oxygen vacancy concentration of perovskite oxides. (**a**) Workflow of the prediction of the oxygen vacancies of the $ABO_{3-\delta}$ perovskites at different temperatures. Comparison of the measured (δ_M) and predicted oxygen vacancy (δ_P) for (**b**) cobalt-based and (**c**) iron-based perovskite oxides.

5. For (b) and (c) in Fig.1, the model is very likely overfitting. More extensive cross-validation is needed on randomized test/validation datasets.

Response:

Sorry for the confusion here. We implemented 5-fold random cross validation to avoid the overfitting issue in the previously submitted manuscript.

6. Why do the authors divide the oxides into Co-based and Fe-based oxides for the oxygen vacancy prediction? The reviewer holds the idea that there is no meaning to differ them, since all the elements can be characterized by their properties, such as electronegativity, cation size, etc, that the authors used in this manuscript. The reviewer suggests to train and predict all the oxides with one model.

Response:

The unique feature of this proposed approach is to leave the Fe/Co out, meaning that there are no active metal centers considered in the machine learning model. Through this way, we show that we could avoid the challenges to predict the electronic structures of the Co/Fe at elevated temperatures. This further highlights the importance of local lattice cationic environment in determining the states of the active metal canters, and therefore the content of the oxygen vacancies.

Besides, different B-site hosts (e.g., Co-/Fe-based perovskites) contribute to different oxygen vacancy contents and different ASPR values at the same temperature. For example, $SrSc_{0.175}Nb_{0.025}Fe_{0.8}O_{3-\delta}$ (SSNF) and $SrSc_{0.175}Nb_{0.025}Co_{0.8}O_{3-\delta}$ (SSNC) are very similar except for the B-site host, but the oxygen vacancy content of the former one (δ = ~0.231 at 450 °C and δ = ~0.256 at 500 °C, **Ref:** *Energy*, **95**, 137-143 (2016)) is much smaller than those of the latter one (δ = ~0.450 at 450 °C and δ = ~0.462 at 500 °C, **Ref:** *Ceram. Int.*, **42**, 12894-12900 (2016)). Considering the unique feature of this ML approach, Co-based and Fe-based oxides are trained and predicted regarding to the oxygen vacancy, separately.

7. The database should be uploaded.

Response:

The data and codes for this work can be found in *https://github.com/Luoyadan/MLtoolbox-perovskite-oxygen-vacancy-pred*. We will share these data when the work will be accepted.

8. The concentration of oxygen vacancy is closely related to the preparation processes. It is unconvincing not to consider the preparation process, such as sintering temperature, atmosphere, dewelling time. For instance, there is a big difference in oxygen vacancy content for the same material synthesised in Ar or O2 or at high/low temperatures. Such conditions are ignored when the model is trained, which makes the following results are incorrect in the manuscript.

Response:

We tried to constrain these factors, such as the response of materials to their operating conditions, by selecting most of reported samples collected here were prepared and conditioned at elevated temperatures in air. However, it is extremely challenging to rule out these factors given different approaches used by different research groups. Nevertheless, the model could still perform well in predicting the oxygen vacancies across these samples, meaning that these variations might not play a significant role in determining oxygen vacancies if proper materials are selected.

9. Thermogravimetric (TG) analysis was performed to weigh the mass loss of sample at 150-800 oC, and thus calculate the δT at elevated temperature. To eliminate the moisture in the sample, efforts were made to maintain the temperature at 150-200 °C for 1.5 h. However, many oxides are of hydration reaction capability even at 500 oC or higher. That is to say, the weight loss at the temperature below 500 oC may be caused by loss of water or/and oxygen.

Response:

It is important to note that the materials we selected here are mostly designed for oxygen reduction in dry air for SOFC application. Most of them are not ideal for proton-conducting applications. Meanwhile, we included new results over TG for weight loss vs. time, and there is no significant decrease of mass before proceeding towards higher temperatures. Meanwhile, the temperature-programmed desorption mass spectrometry (TPD-MS) did not show any peaks

relevant to CO_2 or water at temperatures beyond 300 °C (**Fig. R5**). Meanwhile, the samples prepared in this report have been sintered in dry air at much higher temperatures (>1200 °C) before sending out for TG measurement. Therefore, there is no significant hydration step taking place in our test conditions

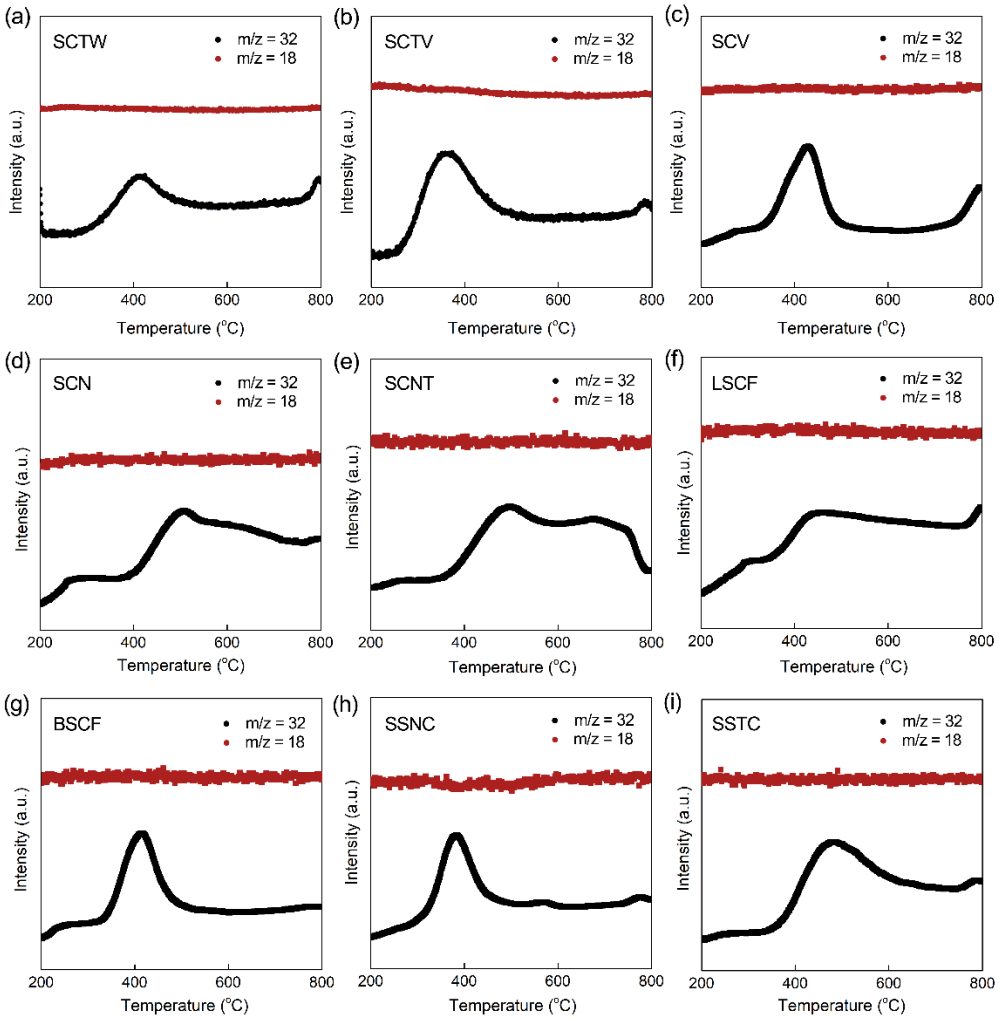

Fig. R5 TPD-MS analysis tested from 200 – 800 °C for (**a**) SCTW, (**b**) SCTV, (**c**) SCV, (**d**) SCN, (**e**) SCNT, (**f**) LSCF, (**g**) BSCF, (**h**) SSNC, and (**i**) SSTC oxides.

Action: **Fig. R5** has been added as a new **Fig. S13** in the revised Supplementary Information. Relevant statement has been added in the revised manuscript as follows:

*"The temperature-programmed desorption mass spectrometry (TPD-MS) shows no peak relevant to CO_2 or water at temperature above 300 °C (**Fig. S13**), illustrating that the oxygen vacancy formation is the main result of the mass loss."*

10. In Figure S5, there are some data tested by the authors, please provide the original LSV curves for OER of these oxides.

Response:

Based on your suggestion, we have added the LSV curves for OER of the perovskite oxides (**Fig. R6**) in the revised SI file as follows:

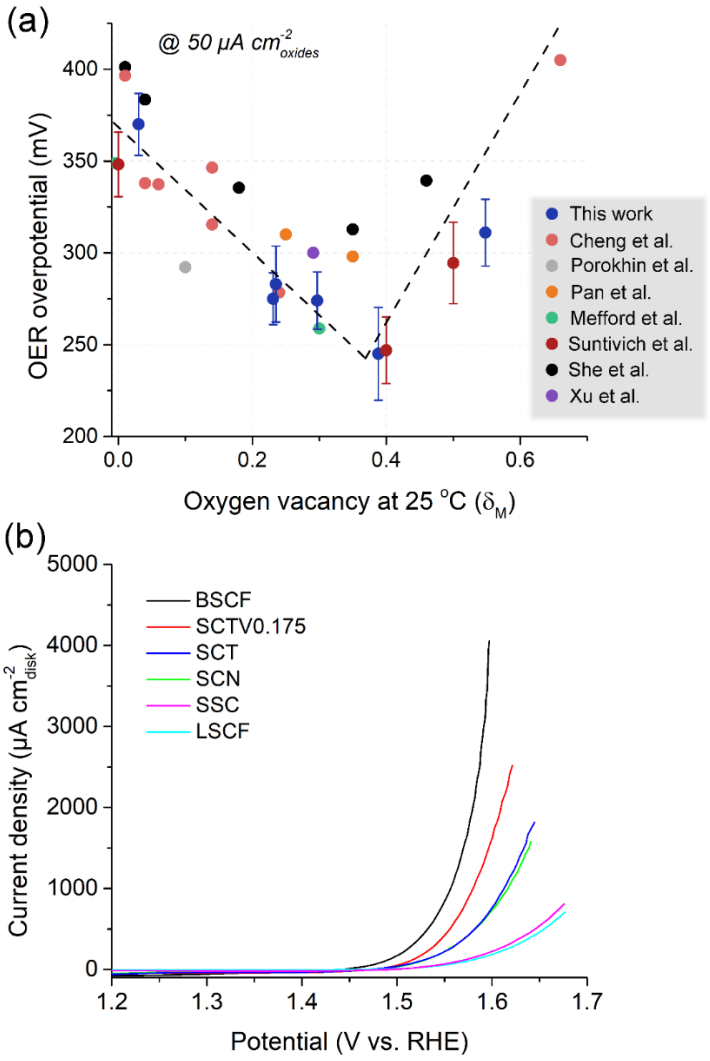

Fig. R6 (a) Volcano relationship between OER overpotentials and oxygen vacancy for various perovskites. The data were collected from the references (Cheng et al. (**Ref:** *ACS Catalysis,* **8**, 9567-9578 (2018)), Porokhin et al. (**Ref:** *ACS Catalysis,* **11**, 8338-8348 (2021)), Pan et al. (**Ref:** *Nat. Commun.,* **11**, 2002 (2020)), Mefford et al. (**Ref:** *Nat. Commun.,* **7**, 11053 (2016)), Suntivich et al. (**Ref:** *Science,* **334**, 1383-1385 (2011)), She et al. (**Ref:** *Interfaces,* **10**, 11715-11721 (2018)), and Xu et al. (**Ref:** *Adv. Sci.,* **3**, 1500187 (2016)) and tested by ourselves (blue ones) in 0.1 M KOH. The dashed lines are demonstrated for guidance. (b) OER polarization profiles of perovskite catalysts synthesized in this work.

Action: We have upgraded **Fig. R6** as the new **Fig. S6** in the revised Supplementary Information.

11. The reviewer suggest add a figure for predicted and measured oxygen vacancy shown in

Figure 2b and 2c. How is the plot of "ASPR at 450 oC versus oxygen vacancy at 450 oC"?

Response:

Thank you for your suggestion. Accordingly, we have added a new plot (**Fig. R7**) to illustrate the linear correlation between predicted and measured oxygen vacancy.

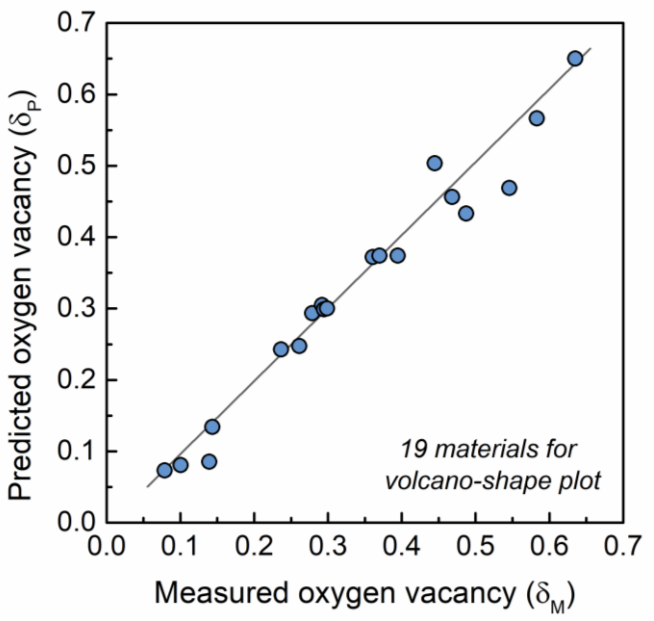

Fig. R7 Linear correlation of predicted oxygen vacancy with measured oxygen vacancy.

Action: **Fig. R7** has been added as a new **Fig. S11** in the revised Supplementary Information. We have relevant statement in the revised manuscript as follows:

*"**Fig. 2b** and **Fig. 2c** show the trend of the ASPR values of the model materials as a function of the predicted and tested concentration of oxygen vacancies, respectively. The measured oxygen vacancies are linearly correlated with the predicted ones (**Fig. S11**)."*

12. In line 288 and 289, why was SC choosen? As known to us, Sr is easy to react with CO2 and easily segregated, Co based oxides are high TEC.

Response:

We chose this material because $SrCoO_{3-\delta}$ represents one of the most active candidates for ORR such as BSCF, LSCF, SSC, and SCNT. We could use this candidate as an example to prove our concept. Yes, there are issues associated with the Sr and Co, but our previous work also indicates that there are opportunities to address these issues while sustaining their high ORR activities. To further illustrate this point, we include new discussions in the main text.

Action: We have added statement to explain why we chose SC-based oxides as examples in this work as follows:

"We chose SC based oxides as the model perovskites to prove our concept because they represent one of the most active candidates for ORR, such as BSCF, LSCF, SSC, and $SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$ (SCNT)."

13. Figure 2d is missing in the main text.

Response:

Thank you for the careful checking. We have added the "**Fig. 2d**" in the revised manuscript.

14. In Figure S18, I-V curve of Cell 2 is from 0-OCV, cell-3 is different from it, why?

Response:

The voltage of **Cell-3** falls rapidly in the high current density region, known as concentration polarization (**Refs**: *Renew. Sust. Energy Rev.*, **36**, 149-179 (2014); *J. Power Sources*, **302**, 53-60 (2016)). In this region, the fuel and oxidant consumption exceed their diffusion limit, leading to increased concentration polarization to **Cell-3** especially at above 500 °C (**Fig. R8**). Accordingly, we have updated this figure in the revised Supplementary Information.

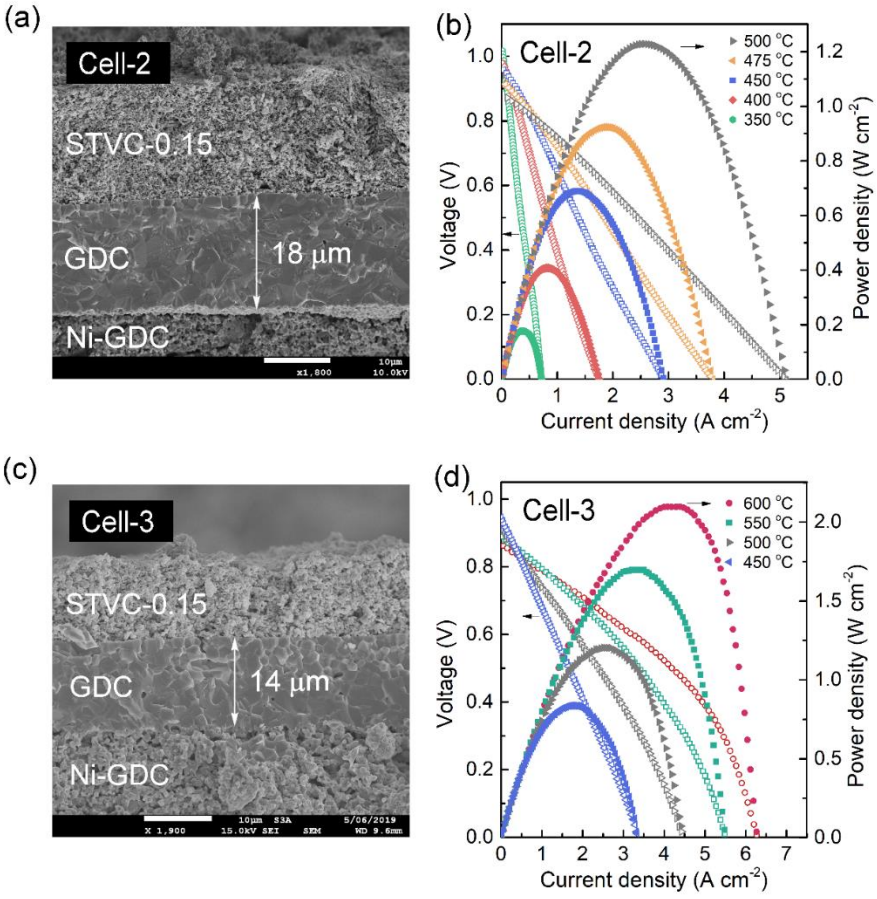

Fig. R8 (**a**) Cross-sectional SEM image and (**b**) Voltage-current density curves and power densities of STVC catalyzed single-cell-2 (i.e., **Cell-2**). (**c**) Cross-sectional SEM image and (**d**) Voltage-current density curves and power densities of STVC catalyzed single-cell-3 (i.e., **Cell-3**).

Action: We have upgraded **Fig. R8** as the new **Fig. S22** in the revised Supplementary Information.

15. In line 350, please provide the xrd patterns after CO2 tolerance test for SCTV and BSCF.

Response:

After CO_2 resistance tests at 550 °C, we found the appearance of carbonates (2θ = ~ 24.7° and ~ 43.7 °, **Ref**: *ACS appl. Mater. Interfaces*, **8**, 3003-3011 (2016)) for BSCF, while SCTC shows no carbonate phase (**Fig. R9**). The results, together with the ASPR tests in revised Supplementary **Figs. S24** and **S25,** indicate the better CO_2 resistance of SCTV than BSCF.

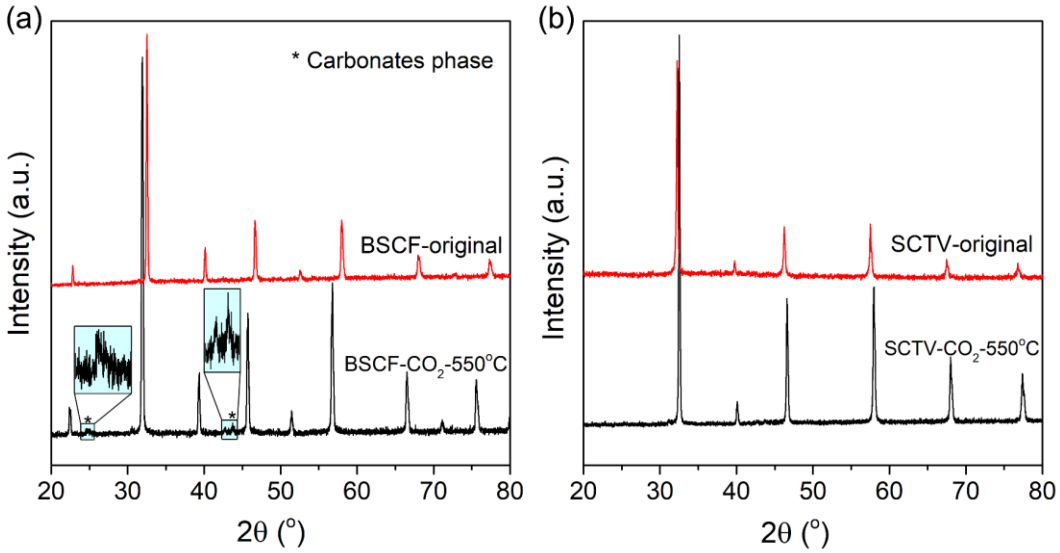

Fig. R9 XRD patterns for (**a**) BSCF and (**b**) SCTC before and after CO_2 treatments at 550 °C

Action: We have added **Fig. R9** as the new **Fig. S25** in the revised Supplementary Information. Statement about the CO2 resistance has been revised as follows:

*"Although the SCTV is not completely resistant to 1 % CO_2 due to the presence of Sr, it still shows a much better CO_2 tolerance than the benchmark BSCF cathode (**Fig. S26** and **Fig. S27**)."*

16. The cells show excellent performances at low temperature, how about the larger cells, such as the diameters of 40 mm or even larger?

Response:

Thanks for this comment. It is important to understand and address challenges associated with the scaling up, which is outside the scope of this work. Instead, the focus of this article is to prove the concept of the proposed ML-assisted approach to design catalyst materials.

Reviewer #3 (Remarks to the Author):

This paper reports a method for predicting oxygen vacancy concentrations without the need for extensive theoretical calculations. The reviewer appreciates that the computational predictions agree with the experiments, resulting in the discovery of electrocatalysts with better performance than the existing materials. However, the present study was only a composition optimization of perovskite-type oxides, which is not surprising in terms of the development of new materials. Although it is interesting to focus on oxygen deficiency, there have been several reports on calculation methods for composition optimization.

Response:

Sorry that our originally submitted work did not state the novelty of the work clearly. Instead of mere compositional optimization, the key novelty here is the identification of the dominant role of the local cationic environment in determining the state of the redox-active metal centers, and thus the concentrations of the oxygen vacancies over perovskite oxides. This is proved by the successful prediction of oxygen vacancies using the ML models that leave out the active metal centers such as Co and Fe. Importantly, this lattice cationic environment can be quantified through the existing database, requiring no further theoretical computations. The new perovskite development is a demonstration of the utility of the proposed approach in guiding the design of the material. To further clarify this point, we further revise the main text as follows.

Action:

We have revised statements in the Abstract section to make the novelty of this work much clear as follows:

"This work revealed the dominant role of local cationic environment in determining the state of redox-active metal centres and thus the concentration of the oxygen vacancies over perovskites, enabling rational development of new catalysts for energy applications.".

Regarding the power generation performance of SOFCs, it is worth evaluating that a relatively high power density was achieved even at 500 °C using the developed materials. However, the significance of low-temperature operation and electrode development is lost, as the presented performance is significantly lower than that of existing SOFCs operating at 800°C. In addition, this paper has issues that need to be corrected, which are enumerated below.

Response:

It seems unfair to compare the power densities of SOFCs at different temperatures, particularly these reactions are thermally driven. Lowering the operating temperature of ceramic fuel cells is a meaningful research direction to pursue considering the challenges faced by the existing high temperature fuel cell technologies, such as the thermal cycling stability, sealing, and sluggish start-up/shut-down procedures. All these challenges limit the viability of the fuel cells at a scale. Again, this paper is unintended to pursue performance, as the power

density can be significantly impacted by the geometry and microstructure of single cells such as the thickness of the electrolyte. For example, reducing the thickness of electrolyte (e.g., 80 nm, ~5 μm as shown in **Figs. R10a** and **R10b**) significantly improves the SOFC performance. Different from these prior reports, we show the improvement of the SOFC performance based on cells with thick electrolyte (13-18 μm in this work as shown in **Fig. R10c**). We would anticipate a much-improved power density of the SOFC (e.g., ~ 2.8 W cm^{-2} at 500 °C) if the electrolyte thickness can be reduced to ~ 5 μm (**Fig. R10d**). However, the key novelty of this study lies in the successful prediction of oxygen vacancies at a wide range of temperatures without demands for heavy computation loads and the use of this insight for rational design of electrocatalyst for oxygen electrolysis.

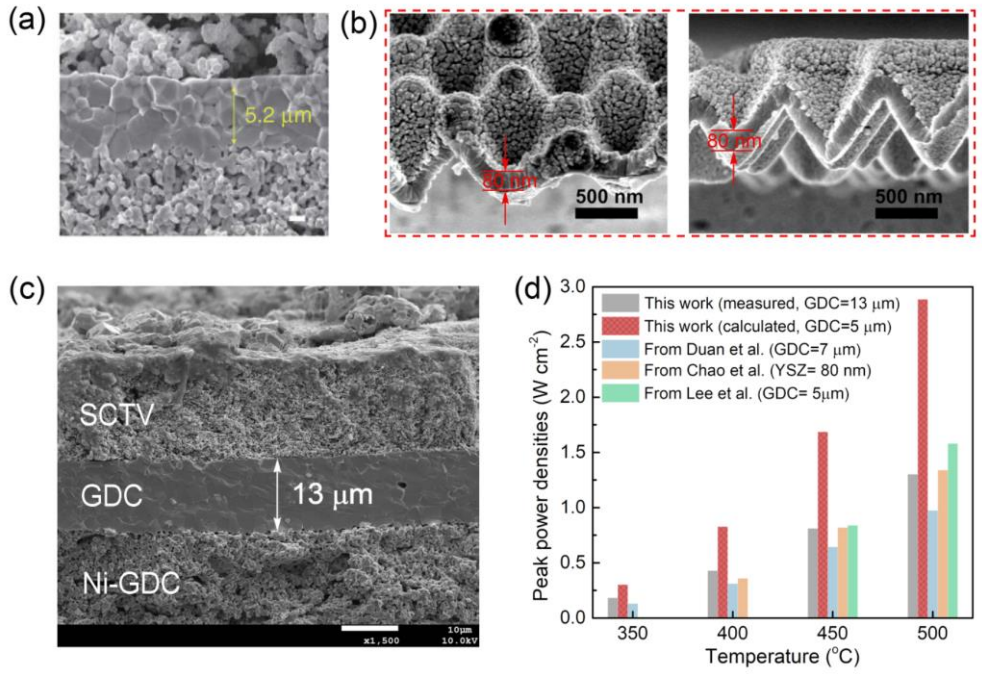

Fig. R10 (**a**) GDC thickness of ~ 5 μm from Lee et al. (**Ref:** *Nat. Commun.*, **5**, 4045 (2014)), (**b**) YSZ thickness of ~ 80 nm from Chao et al. (**Ref:** *ACS Nano*, **5**, 5692-5696 (2011)), (**c**) GDC thickness of ~ 13 μm from this work, and (**d**) performance comparison of this work with other reported studies.

Action: Again, we believe this concern might arise from the clarity of original text. Accordingly, we included new discussions in the main text to illustrate the reason why we would like to develop perovskites that were utilized at reduced temperatures as follows:

"Lowering the operating temperatures (e.g., ⩽ 500 ℃) is meaningful to address the challenge faced by the existing high temperature SOFC technologies, such as the thermal cycling stability, sealing, and sluggish start-up/shut-down procedures [30]."

In conclusion, the reviewer is compelled to point out that the paper has insufficient impact to

be published in Nature Communications, either in terms of computational methods, new material development, or fuel cell properties, and recommends transferring to more specialized journals, such as Communications Materials.

1. The Introduction section was vague. For example, although the authors evaluated new cathode materials at low temperatures, motivation for developing low-temperature solid oxide fuel cell (SOFC) cathode materials has not been described.

Response:

SOFCs operated at high temperature (e.g., ⩾ 800 °C) are usually suffering from many technological challenges. Relevant statements about the motivation for developing new perovskite materials for low-temperature SOFCs are added in the revised manuscript.

Action: The following statements have been added in the main text:

"We then used this machine learning based strategy to develop new perovskite materials for low-temperature solid oxide fuel cells (SOFCs). Lowering the operating temperatures (e.g., ⩽ 500 ℃) is meaningful to address the challenge faced by existing high temperature SOFC technologies, such as thermal cycling stability, sealing, and sluggish start-up/shut-down procedures [30]."

2. Although the Methods section should contain all the elements necessary for the interpretation and replication of the results, the Methods section of this literature is incomplete.

Response:

Thank you for the careful check. We have added more statements in the Methods section, including Electrical conductivity of the perovskite oxides, Electrochemical measurements of the perovskite oxides, Calculation of oxygen vacancy at room temperature, and calculation of oxygen vacancy at elevated temperature.

Actions: The following statements have been added in the Methods section of revised manuscript:

*"**Materials characterization:** Scanning electron microscopic images (SEM, JEOL JSM-7001F) was applied at 15 kV to investigate the thickness of perovskite materials on electrolyte."*

*"**Electrical conductivity of the perovskite oxides:** The perovskite powders were pressed and then sintered at 1200 ℃ for 5 h to form dense bars. The relative densities of the samples were over 95 %. After polish, the silver wires were attached to the sample bars, followed by the electrical conductivity by using a DC 4-probe method on an Autolab PGSTAT30 workstation. To ensure the accuracy, the measurements were conducted in the presence of flow air (120 mL min^{-1})."*

*"**Electrochemical measurements of the perovskite oxides:** 10 mg perovskite powder was ultrasonically mixed with 10 mg carbon black (Super C65, TIMCAL C'NERGY) in 1.0 mL ethanol with 100 μL Nafion 117 solution (5 wt.% in ethanol) for 40 min. Then, 5 μL suspension*

was drop-cast onto a glassy carbon disk electrode (diameter of 5.0 mm, area of 0.196 cm^2) and dried in a vacuum oven at room temperature. The electrochemical cyclic voltammograms (CV) measurements of the perovskite materials were conducted in 0.1 M KOH (saturated with 99.999 % O_2) at a scan rate of 5 $mV\ s^{-1}$ by using a rotating disk electrode (RDE) device. The Pt wire was used as the counter electrode and an Hg/HgO electrode was used as the reference electrode. The potentials tested in this work are iR-corrected according to the following equation:

$$E(RHE) = E(Hg/HgO) + 0.110V + 0.059 \times pH - 90\ \% \times iR \qquad \textbf{Eq. 1}$$

where i and R represent current and ohmic resistance of 0.1 M KOH, respectively, and E(Hg/HgO) is the measured potentials."

*"**Calculation of oxygen vacancy at room temperature:** Iodometric titration was applied to determine the oxygen vacancy in the perovskite-type oxides [79]. Generally, 0.1 g pre-dried oxide was mixed with 2.0 g KI in an Erlenmeyer flask, followed by sealing with Ar gas for 10 min and adding 10 mL HCl (3.0 mol/L). The ultrasound treatment was applied to better dissolve the oxide in HCl acid. The solution was then titrated with 0.1 mol/L $Na_2S_2O_3$.*

Taking $SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$ (SCNT) oxide as an example, only Co ion participates in the redox reaction. The oxidation state of Co is assumed as +x, and the redox reactions during the titration can be expressed as:

$$Co^{x+} + (x-2)I^- = Co^{2+} + \frac{x-2}{2} I_2 \qquad \textbf{Eq. S2}$$

$$I_2 + 2Na_2S_2O_3 = 2NaI + 2Na_2S_4O_6 \qquad \textbf{Eq. S3}$$

So that the stoichiometric relationship between SCNT and $Na_2S_2O_3$ should be:

$$SrNb_{0.1}Ta_{0.1}Co_{0.8}O_{3-\delta} \sim \frac{x-2}{2} \times 0.8 I_2 \sim (x-2) \times 0.8 Na_2S_2O_3 \qquad \textbf{Eq. S4}$$

The molar relationship between SCNT and $Na_2S_2O_3$ is calculated as:

$$\frac{m}{M} \times 0.8 \times (x-2) = cV \qquad \textbf{Eq. S5}$$

where m and M are the exact mass and molar mass of SCNT cathode, while c and V represent the concentration and volume of the $Na_2S_2O_3$ solution. The presence of oxygen vacancy will inevitably lead to a decrease in oxygen non-stoichiometry $\boldsymbol{\delta_0}$ at room temperature. So that M can be expressed as:

$$M = M_{Sr} + M_{Nb} \times 0.1 + M_{Ta} \times 0.1 + M_{Co} \times 0.8 + M_O \times (3 - \delta_0) \qquad \textbf{Eq. S6}$$

where M_{Sr}, M_{Nb}, M_{Ta}, M_{Co}, and M_O represent the molar atomic mass of Sr, Nb, Ta, Co, and O, being 87.62, 92.91, 180.95, 58.93, and 16.00 g mol^{-1}, respectively.

Besides, the conservation of charge in SCNT can provide another equation:

$$X_{Sr} + X_{Nb} \times 0.1 + X_{Ta} \times 0.1 + X_{Co} \times 0.5 = X_O \times (3 - \delta_0) \qquad \textbf{Eq. S7}$$

where X_{Sr}, X_{Nb}, X_{Ta}, X_{Co}, and X_O are the valences of Sr, Nb, Ta, Co, and O, being +2, +5, +5, +x, and -2, respectively.

In summary, the oxygen non-stoichiometry of $\boldsymbol{\delta_0}$ can be calculated from the association of

Eq. S5-S7. *"*

*"**Calculation of oxygen vacancy at elevated temperature:** Thermogravimetric (TG) analysis was performed to weigh the mass loss of sample at 150-800 °C, and thus calculate the δT at elevated temperature. To eliminate the moisture in the sample, efforts were made to maintain the temperature at 150-200 °C for 1.5 h. Assuming that no oxygen vacancy exists in the SCNT material, the original mass of the sample should be:*

$$m_{original} = \frac{m \times (M + M_O \times \delta_0)}{M} \qquad \textbf{Eq. S8}$$

*where M is the molar mass of SCNT, and M_O represents the molar atomic mass of the oxygen atom. So that the SCNT mole ($M_{original}$) can be calculated from **Eq. S9**, no matter how much oxygen goes away as the raising of temperature.*

$$M_{original} = \frac{m_{original}}{M + M_O \times \delta_0} \qquad \textbf{Eq. S9}$$

Then the δ_T will be obtained from:

$$\delta_T = \frac{(m_{original} - m_T)/M_O}{M_{original}} \qquad \textbf{Eq. S10}$$

where m_T is the sample mass at temperature T."

*"**Cell characterization:** The stability tests were conducted at 375 °C and 500 °C at 0.272 A cm^{-2}, with H_2 flow of 80 ml min^{-1} on the anode side and air flow of 120 ml min^{-1} on the cathode side."*

*"**Density functional theory (DFT) calculations:** To mimic the real experimental condition, the cubic $SrCoO_3$ models were established by substituting Co atoms with co-dopants to form supercells with scales of 2×2×2 (e.g., for $SrCo_{0.8}Nb_{0.2}O_{3-\delta}$, $La_{0.6}Sr_{0.4}CoO_{3-\delta}$, and $Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$) and 2×2×4 (e.g., for $SrCo_{0.8}Ta_{0.1}V_{0.1}O_{3-\delta}$, $SrSc_{0.175}Nb_{0.025}Co_{0.8}O_{3-\delta}$, and $SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$), and a 3×3×3 Monkhorse-pace mesh was chosen for k-sampling of the Brillouin zone."*

3. The authors dried the samples at 150–200 °C before the TG measurements; however, it is concerned that the temperature was insufficient to remove water completely. Proton-conducting perovskite oxides such as BZY retain water within the solid up to approximately 600 °C. Therefore, drying for dehydration should be performed at higher temperatures. Furthermore, the influence of water must be completely eliminated to obtain accurate experimental values that can be compared with the calculated values, and it is recommended that the authors take great care to prevent water during measurement. From the manuscript, the reviewer could not identify any experimental care to accurately estimate the oxygen content.

Response:

It is important to note that the materials we selected here are mostly designed for oxygen reduction in dry air for SOFC application. Most of them are not ideal for proton-conducting application, and Sr-based perovskite oxides show significantly lower hydration capacities than

proton-conducting perovskite (e.g., BZY) (**Ref**: *Adv. Funct. Mater.*, **28**, 1801241 (2018)). Meanwhile, we included new results over TG for weight loss vs. time, and there is no significant decrease of mass before proceeding towards higher temperatures.

Meanwhile, we conducted temperature-programmed desorption mass spectrometry (TPD-MS) tests over the perovskite oxides (e.g., SCTV and SCTW) in the temperature range of 200-800 °C. TPD-MS results show that these perovskites do not undergo any dehydration reaction within this temperature range (**Fig. R11**). As such, the release of lattice oxygen should be mainly responsible for the weight loss observed in the TG results.

Besides, the samples prepared in this report have been sintered in dry air at much higher temperatures (>1200 °C) before sending out for TG measurement. Therefore, there is no significant hydration step taking place in our test conditions.

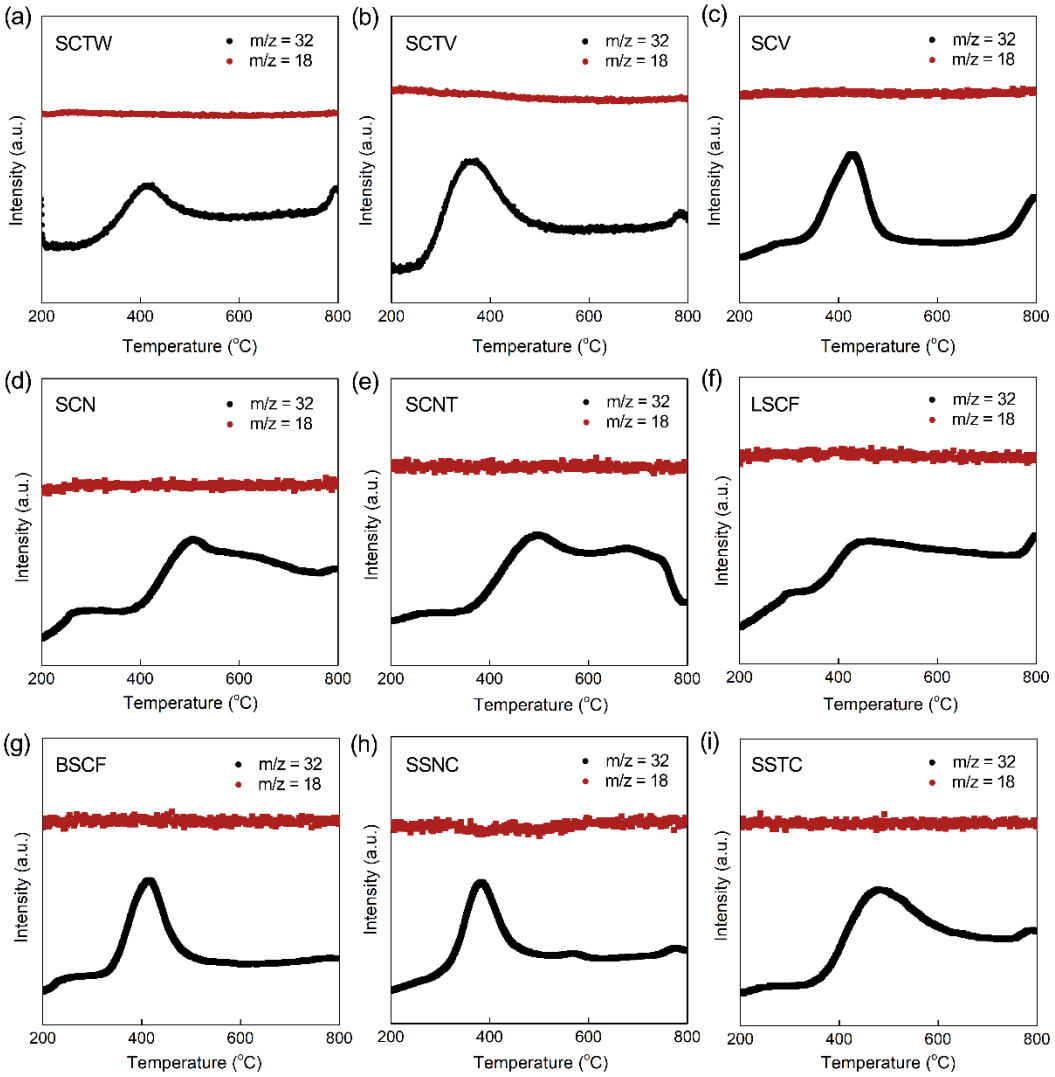

Fig. R11 TPD-MS analysis tested from 200 – 800 °C for (**a**) SCTW, (**b**) SCTV, (**c**) SCV, (**d**) SCN, (**e**) SCNT, (**f**) LSCF, (**g**) BSCF, (**h**) SSNC, and (**i**) SSTC oxides. No peak for m/z = 18 (red) can be observed, indicating that there is no significant hydration step taking place during tests for all materials.

Actions: We have added **Fig. R11** as a new **Fig. S13** in the revised Supplementary Information. Relevant statement has been added in the revised manuscript as follows:

*"The temperature-programmed desorption mass spectrometry (TPD-MS) shows no peak relevant to CO_2 or water at temperature above 300 °C (**Fig. S13**), illustrating that the oxygen vacancy formation is the main result of the mass loss."*

4. The number of oxygen vacancies calculated from the TG may differ from the actual amount owing to the influence of water. Because measuring the number of oxygen vacancies is one of the important experiments in this paper, the reviewer recommends a double check using a different method (e.g., high-temperature neutron diffraction measurements).

Response:

Thanks for this suggestion. Based on the TPD-MS results, we confirm that there is nearly no water in the materials because of high-temperature quench (>1200 °C) before test. Therefore, there is also no significant hydration step taking place in our TG test conditions. Also, our previous results using joint Rietveld refinements using XRD and neutron diffraction measurements (**Ref**: *Adv. Funct. Mater.*, **33**, 2210496 (2023)) showed that increasing temperatures enabled the increase of oxygen vacancy contents in perovskites, which is consistent with the results in this work. Also, the content of oxygen vacancy in SCNT are higher than those of SCN and SCT with either at 25 °C or 600 °C, in accordance with the results in this work. Different methodologies may lead to different results. To verify the trend of oxygen vacancy content, we also checked the oxygen vacancy content of more oxides by using XAS measurements. As shown in **Fig. R12**, the contents of oxygen vacancies in SCTV, SCV, SCN, and SCT from titration and TG method are close those from XAS method.

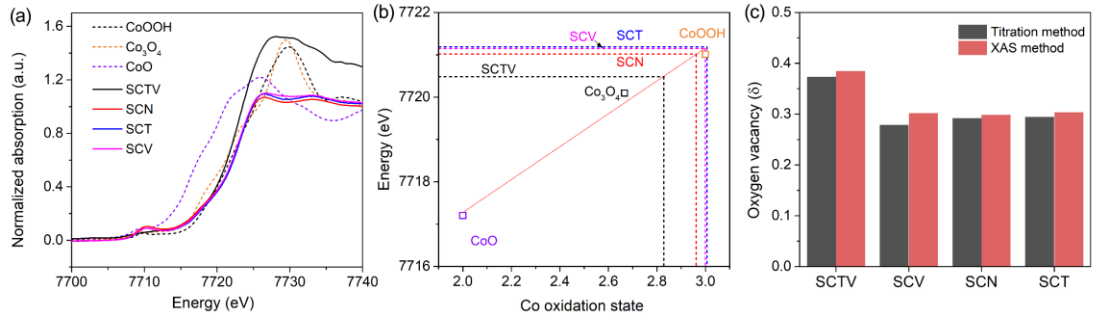

Fig. R12 (**a**) Co K-edge XANES data of SCTV (black line), SCN (red line), SCT (blue line), and SCV (pink line) and standards for CoOOH (black, dashed line), Co_3O_4 (orange, dashed line), CoO (purple, dashed line). Before testes, SCTV, SCN, SCT, and SCV were quenched at 450 °C for 2 h and then sealed in Ar gas. We used the energy position at ordinate 0.6 to compare the oxidation states of Co in oxides. (**b**) Co oxidation states calculated from the absorption edge shift in XANES data of SCTV, SCN, SCT, and SCV. (**c**) Comparison in oxygen vacancy between titration method and XAS method for SCTV, SCN, SCT, and SCV.

Action: **Fig. R12** has been added as the new **Fig. S14** in the revised Supplementary Information. Relevant statements about the confirmation of oxygen vacancy contents have been added in the revised manuscript as follows:

*"The results are consistent with those obtained by using Co-K edge X-ray absorption near-edge spectroscopy (XANES) method (**Fig. S14**)."*

5. The authors claimed to have discovered two new perovskite oxides (SCTW and SCTV). However, the only data available to show that these materials were synthesized are the powder X-ray peak patterns, which is insufficient evidence to prove the discovery of new materials. High-quality diffraction data should be measured and Rietveld refinement should be performed to determine the structural parameters of these materials. Furthermore, these materials are merely compositional optimizations and are not new materials.

Response:

Sorry for the confusion. To improve the clarity, we have now revised the "new materials" to "new compositions".

Accordingly, we have also conducted the Rietveld refinement (**Fig. R13**, **Tables R1** and **R2**) over SCTV and SCTW perovskites to demonstrate their lattice structure. The results show that both SCTV and SCTW are in $Pm\bar{3}m$ space group, with lattice parameters of 3.8916(5) and 3.8969(2), respectively.

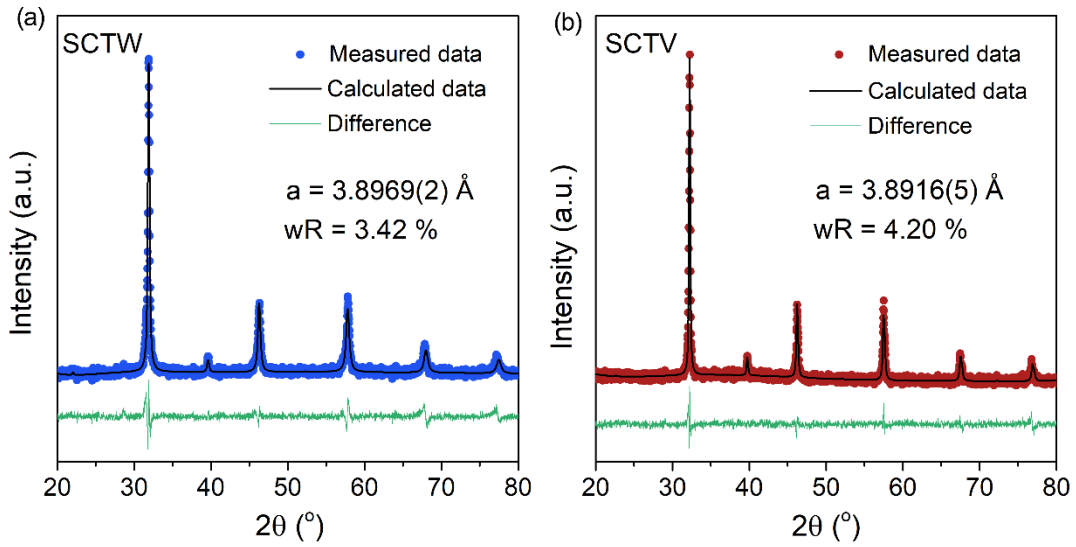

Fig. R13 XRD patterns of (**a**) SCTW and (**b**) SCTV perovskites.

Table R1 Laboratory XRD refinement details for SCTV

Atom	Site Multiplicity	*x*	*y*	*z*	Occupancy	U_{iso} (Å^2)

SCTV, a = 3.8880(5) Å, $Pm\bar{3}m$, phase fraction 100%						
Sr	1	0.5	0.5	0.5	1.03(8)	0. 017(9)
V	1	0	0	0	0.05(4)	0.02(1)[a]
Co	1	0	0	0	0.800(2)	0.02(1)[a]
Ta	1	0	0	0	0.15(4)	0.02(1)[a]
O	3	0.5	0	0	0.9[b]	0.01[c]
Goodness of fit (GOF) = 1.16						
Weighted profile R-factor (wR) = 4.20%						

[a] Atomic displacement parameters constrained to be equal for atoms at the same site.

[b] Occupancy is fixed to the titration results

[c] Atomic displacement fixed to the literature value.

Table R2 Laboratory XRD refinement details for SCTW

Atom	Site Multiplicity	x	y	z	Occupancy	U_{iso} (Å^2)
SCTW, a = 3.8894(2) Å, $Pm\bar{3}m$, phase fraction 100%						
Sr	1	0.5	0.5	0.5	1.01(2)	0. 009(2)
Co	1	0	0	0	0.77(2)	0.008(3)[a]
Ta/W[b]	1	0	0	0	0.23(2)	0.008(3)[a]
O	3	0.5	0	0	0.9[c]	0.01[d]
Goodness of fit (GOF) = 1.37						
Weighted profile R-factor (wR) = 3.42%						

[c] Atomic displacement parameters constrained to be equal for atoms at the same site.

[b] Ta and W are indistinguishable in XRD and combined.

[c] Occupancy is fixed to the titration results

[d] Atomic displacement fixed to the literature value.

Action: **Fig. R13** has been upgraded as **Fig. S20**; **Tables R1** and **R2** were added as new **Tables S2** and **S3** in the revised Supplementary Information.

We have added the statements of Rietveld refinement for both SCTV and SCTW perovskites in the revised manuscript as follows:

*"We experimentally synthesized SCTW and SCTV materials and verified their crystal structure using Rietveld analysis of XRD (**Fig. S20**), with details shown in **Table S2** and **Table S3**. The results reveal that both SCTV and SCTW are in $Pm\bar{3}m$ space group. The lattice parameters are 3.8916(5) for SCTV and 3.8969(2) for SCTW, consistent with the larger size of W^{5+} and Ta^{5+} compared to V^{5+} [34]."*

6. Are the lattice parameters of the solid solutions in Fig. S7 and Fig. S16 reasonable?

Response:

Yes. We have carefully checked the XRD data, and we accordingly calculated the lattice parameters of the revised XRD results via Rietveld refinement over the representative samples (**Fig. S8** and **Fig. S20)**, and compared them with reported values in the **Table R3**. As shown in the table, they are in the same ranges as reported in the literature.

Table R3 Lattice parameter of the perovskite oxides compared to those in references

Materials	Lattice parameter (Å)	References
$Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$	3.9862(3)	*J. Solid State Chem.*, **18**, 576-586 (2008).
$SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$	3.9066(1)	*Nat. Commun.*, **8**, 13990 (2017).
$SrSc_{0.175}Nb_{0.025}Co_{0.8}O_{3-\delta}$	3.904	*Angew. Chem. Int. Ed.*, **52**, 14036-14040 (2013)
$SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$	3.90402(3)	*Adv. Funct. Mater.*, **33**, 2210496 (2023).
$SrCo_{0.8}Nb_{0.2}O_{3-\delta}$	3.90357(3)	*Adv. Funct. Mater.*, **33**, 2210496 (2023).
$SrCo_{0.8}Ta_{0.2}O_{3-\delta}$	3.90820(4)	*Adv. Funct. Mater.*, **33**, 2210496 (2023).
$Sr_{0.95}Ag_{0.05}Nb_{0.1}Co_{0.9}O_{3-\delta}$	3.8733	*Nano letters*, **16**, 512-518 (2016).
$SrCo_{0.8}V_{0.2}O_{3-\delta}$	3.8647(7)	This work
$SrCo_{0.8}Nb_{0.2}O_{3-\delta}$	3.8983(6)	This work
$SrCo_{0.8}Ta_{0.2}O_{3-\delta}$	3.9053(6)	This work
$SrSc_{0.175}Ta_{0.025}Co_{0.8}O_{3-\delta}$	3.9119(4)	This work
$Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$	3.9144(3)	This work
$SrCo_{0.8}Ta_{0.16}W_{0.04}O_{3-\delta}$	3.8969(2)	This work
$SrCo_{0.8}Ta_{0.1}V_{0.1}O_{3-\delta}$	3.8764(3)	This work
$SrCo_{0.8}Ta_{0.15}V_{0.05}O_{3-\delta}$	3.8916(5)	This work
$SrCo_{0.8}Ta_{0.175}V_{0.025}O_{3-\delta}$	4.0759(5)	This work
$(La_{0.2}Sr_{0.8})_{0.95}Co_{0.8}Ta_{0.15}V_{0.05}O_{3-\delta}$	3.8501(3)	This work
$La_{0.1}Sr_{0.9}Co_{0.8}Ta_{0.15}V_{0.05}O_{3-\delta}$	3.8904(2)	This work
$La_{0.2}Sr_{0.8}Co_{0.8}Ta_{0.15}V_{0.05}O_{3-\delta}$	3.8872(1)	This work
$SrSc_{0.175}Ta_{0.025}Co_{0.8}O_{3-\delta}$	3.9349(5)	This work
$SrCo_{0.8}Nb_{0.1}Ta_{0.1}O_{3-\delta}$	3.90147(9)	This work

Action: **Table R3** has been added as a new **Table S1** in the revised Supplementary Information. We have added relevant statement in the revised manuscript as follows:

"The X-ray diffraction (XRD) analysis show that our perovskite oxides are mainly in Pm $\bar{3}$*m cubic structure, with lattice parameters similar to those of reported perovskites (**Fig. S8** and **Table S1**)."*

7. It is intriguing that the ORR activity shows a volcano-trend dependence on the number of oxygen vacancies (Fig. 2b and 2c). This volcanic trend increases linearly with increasing oxygen vacancies and then decreases linearly with increasing oxygen vacancies. The authors should explain why this trend varies "linearly."

Response:

Thanks for this comment. Understanding the trend in detail could be of interest for future research but may not be the focus of this work. The ORR activity is non-linear trend with the content of oxygen vacancies provided that the ASR values are shown in log scale. The lines are

only used to guide the eye. Furthermore, this work is intended to show the important role of oxygen vacancies in determining the overall ORR activity instead of why the trend is in linear or non-linear shape. The former is believed to be sufficient to prove the utility of the proposed approach in guiding the catalyst design, which is the key focus of this study.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have mostly addressed all of my previous comments adequately. Given the heavy focus of the ML model on only Fe- and Co-based perovskites, perhaps it is obvious there may be limitations of the model to other transition metals outside of this set. However, I think an explicit note pointing this out would be beneficial. Along these lines, it is further worth pointing out that the reliance on 5-fold cross validation will give overly optimistic results, particularly if users want to explore O vacancy predictions for different temperatures or transition metals not contained in the training set.

Assuming the above caveats are added and the authors adequately address the concerns raised by the other reviewers, the work is likely suitable for publication

Reviewer #2 (Remarks to the Author):

All my comments have been properly answered.

Reviewer #3 (Remarks to the Author):

First, the reviewer would like to pay tribute to the authors for their efforts.
Additional experiments have dispelled concerns about the accuracy of the oxygen deficit. The reviewers recommend this paper for publication in Nature Communications.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have mostly addressed all of my previous comments adequately. Given the heavy focus of the ML model on only Fe- and Co-based perovskites, perhaps it is obvious there may be limitations of the model to other transition metals outside of this set. However, I think an explicit note pointing this out would be beneficial. Along these lines, it is further worth pointing out that the reliance on 5-fold cross validation will give overly optimistic results, particularly if users want to explore O vacancy predictions for different temperatures or transition metals not contained in the training set.

Assuming the above caveats are added and the authors adequately address the concerns raised by the other reviewers, the work is likely suitable for publication

Response:

We sincerely thank you for the kind suggestions. Sufficient sample size is crucial for machine learning. In this work, Co- and Fe-based oxides are studied because sufficient sample size of Co- and Fe-perovskites are available. Also, we tried to predict the oxygen vacancy of Cr-and Mn-based perovskites, but their datasets are insufficient. 5-fold cross validation is a common machine learning model evaluation technique while sufficient sample size is probably more important for the good prediction of machine learning. Accordingly, we have added statements to clarify this in the revised manuscript.

Action: The following statement is added in the Result section of manuscript:

"Note that Co- and Fe-based oxides are mainly explored in this work because of their large dataset size, and other active metal-containing perovskites should also show similar prediction trend if sufficient data are available. The ML model in this work is thus limited in predicting the oxygen vacancies for other transition metals (e.g., Cr, Mn) containing perovskites."

We added the following texts in the Discussion section of manuscript:

"Besides, this work relies on comprehensive existing datasets for model training and predictions. Due to the limited data available in the literature for other metal-based perovskite oxides, the models reported here are only valid for predicting cobalt- or iron-based perovskite oxides at different temperatures. "

Reviewer #2 (Remarks to the Author):

All my comments have been properly answered.

Response:

Thank you very much for your support on our work.

Reviewer #3 (Remarks to the Author):

First, the reviewer would like to pay tribute to the authors for their efforts.

Additional experiments have dispelled concerns about the accuracy of the oxygen deficit. The reviewers recommend this paper for publication in Nature Communications.

Response:

We sincerely thank you for your support on our work.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response:

We appreciate your contributions on improving the quality of our manuscript.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response:

We sincerely thank you for your support on our work.