ADVANCED MATERIALS
www.advmat.de

Performance Prediction of High-Entropy Perovskites $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_3$ with Automated High-Throughput Characterization of Combinatorial Libraries and Machine Learning

Carlota Bozal-Ginesta, Juande Sirvent, Giulio Cordaro, Sarah Fearn, Sergio Pablo-García, Francesco Chiabrera, Changhyeok Choi, Lisa Laa, Marc Núñez, Andrea Cavallaro, Fjorelo Buzi, Ainara Aguadero, Guilhem Dezanneau, John Kilner, Alex Morata, Federico Baiutti, Alán Aspuru-Guzik,* and Albert Tarancón**

Perovskite oxides form a large family of materials with applications across various fields, owing to their structural and chemical flexibility. Efficient exploration of this extensive compositional space is now achievable through automated high-throughput experimentation combined with machine learning. In this study, we investigate the composition–structure–performance relationships of high-entropy $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_{3\pm\delta}$ perovskite oxides ($0 < x, y, z < 1$; $x+y+z \approx 1$) for application as oxygen electrodes in Solid Oxide Cells. Following the deposition of a continuous compositional map using thin-film combinatorial pulsed laser deposition, compositional, structural, and performance properties are characterized using six different techniques with mapping capabilities. Random forests effectively model electrochemical performance, consistently identifying Fe-rich oxides as optimal compounds with the lowest area-specific resistance values for oxygen electrodes at 700 °C. Additionally, the models identify a statistical correlation between oxygen sublattice distortion—derived from spectral analysis of Raman-active modes—and enhanced performance.

1. Introduction

Mixed ionic electronic conductors (MIECs), such as certain perovskite oxides (ABO_3), are essential for achieving high cathodic oxygen reduction in intermediate-temperature solid oxide fuel cells (IT-SOFCs, T < 700 °C).[1–4] Due to their high ionic conductivity and the consequent increase in the active area, these MIEC materials have progressively replaced electrodes based on noble metals and electron-conducting oxides like $La_{1x}Sr_xMnO_3$ (LSM),[5,6] which were originally employed at higher operating temperatures. Currently, lanthanum strontium perovskite oxide MIECs containing cobalt and iron in the B-site, $La_{1-x}Sr_x(Co_xFe_y)O_{3\pm\delta}$ ($0 < x, y < 1$; $x+y \approx 1$), are the most established cathode material for IT-SOFC applications.[7–10]

C. Bozal-Ginesta, J. Sirvent, F. Chiabrera, L. Laa, M. Núñez, F. Buzi, A. Morata, F. Baiutti, A. Tarancón
Nanoionics and Fuel Cells group
Catalonia Institute for Energy Research
Jardins de Les Dones de Negre 1
Sant Adrià de Besòs, Barcelona 08930, Spain
E-mail: carlota.bozalginesta@empa.ch; atarancon@irec.cat

C. Bozal-Ginesta, S. Pablo-García, C. Choi, A. Aspuru-Guzik
Departments of Chemistry and Computer Science
University of Toronto
Lash Miller Chemical Laboratories
80 St George Street, Toronto, Ontario M5S 3H6, Canada
E-mail: aspuru@utoronto.ca

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adma.202407372

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1002/adma.202407372

G. Cordaro, G. Dezanneau
Laboratoire Structures
Propriétés et Modélisation des Solides
UMR8580
CentraleSupelec
CNRS
Université Paris-Saclay
3 Rue Joliot-Curie, Gif Sur Yvette 91190, France

S. Fearn, A. Cavallaro, A. Aguadero, J. Kilner
Department of Materials
Imperial College London
London SW7 2BP, United Kingdom

A. Aguadero
Instituto de Ciencia de Materiales de Madrid (ICMM-CSIC)
Sor Juana Ines de la Cruz 3, Madrid 28049, Spain

J. Kilner
International Institute for Carbon-Neutral Energy Research (I2CNER)
Kyushu University
Fukuoka 819-0395, Japan

The electrochemical performance of MIEC electrodes toward oxygen reduction is mainly governed by the oxide-ion diffusivity through the lattice (D^*) and the oxygen surface exchange coefficient (k^*). More precisely, Adler et al.[11] proposed the product $D^* \cdot k^*$ as a figure of merit for porous MIEC electrodes operating as cathodes in SOFCs, based on the solution to the continuum model for one-dimensional electrodes. This model directly correlates higher values of $D^* \cdot k^*$ with lower area-specific resistances (ASRs), which indicate better performance. However, the fundamental relationship between these two parameters (D^*, k^*) and the intrinsic properties of the materials remains poorly understood. Some theoretical studies relate D^* and k^* to the formation energy of oxygen vacancies,[12] which in turn is related to the B—O bond strength and the electron exchange driving forces in transition metals.[13] In particular, the ASR and k^* of perovskites were found to be strongly correlated to the oxygen p-band center,[7,14] while the number of d electrons, the charge-transfer energy and the e_g occupancy, among other factors, were identified as relevant descriptors of the oxygen evolution reaction activity in aqueous media.[15,16] Notwithstanding the efforts to understand these phenomena, the obtained correlations are not always quantitative or measurable, and thus not sufficient for guiding the design of new materials and identifying the optimal composition within a family of materials. The main difficulty is most likely the multidimensional and non-linear nature of these relationships. For such complex relationships, machine learning techniques represent a promising and well-suited tool for performance modeling and prediction.

To effectively apply machine learning models in materials science, it is essential to have sufficiently large and consistent experimental datasets, where data for each composition is ideally measured under identical conditions.[17] In materials research, large theoretical datasets can be generated through high-throughput ab initio simulations, but these are often constrained to simplistic models.[18–21] In the extensive field of perovskite oxides and materials for SOFC oxygen electrodes, the application of machine learning is further limited by the lack of coherence in the reported experimental results, due to the variety of synthesis methods, measurement conditions, and figures of merit.[4,22–24] Combinatorial synthesis methods[22,25] can help overcome this problem by enabling the generation of material libraries consisting of continuous compositional maps fabricated in a single process. In particular, combinatorial pulsed laser deposition has been previously used to prepare thin-film perovskite-based cathode materials,[26,27] including binary libraries of LSMC[28] and $La_{0.8}Sr_{0.2}Mn_xO_3$.[6,29] The resulting compositional maps can then be more rapidly and consistently characterized with high-throughput techniques under similar experimental conditions.[30–32] For example, the seminal work by Papac *et al.* reported the electrochemical characterization of a $Ba(Co,Fe,Zr,Y)O_{3-\delta}$ library for protonic ceramic fuel cells utilizing automated impedance spectroscopy measurements and distribution of relaxation times processing.[33–35] However, complete and systematic experimental characterization datasets remain rare, and include only a small variety of characterization techniques.[18,36–43] As a result, it is difficult to gain a comprehensive insight into the intrinsic properties of the materials and their relationship to the electrochemical performance.

Herein, we delve deeper into the experimental composition-structure-performance relationships of $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_{3\pm\delta}$ (LSMCF) ($0 < x, y, z < 1$; $x+y+z \approx 1$) using big data and machine learning. In particular, we prepared a compositional map of LSMCF (**Figure 1a**) through combinatorial thin-film deposition. This map was then characterized through a high-throughput approach, employing multiple techniques: synchrotron X-ray fluorescence (XRF), X-ray diffraction (XRD), Raman spectroscopy and ellipsometry, electrochemical impedance spectroscopy (EIS), and isotopic exchange depth profiling coupled to secondary ion mass spectrometry (IEDP-SIMS) (Figure 1b). The following physicochemical information was obtained from these techniques: cationic stoichiometry from XRF, crystal orientation and unit cell dimensions from XRD, structural symmetry from Raman spectroscopy, electronic transitions and film thickness from ellipsometry, and the electrochemical performance from EIS and IEDP-SIMS. We first processed the raw data to derive characteristic features and developed a solution to match the samples from different measurements. After feature extraction, a variety of machine learning methods with uncertainty quantification were applied to build highly generalizable models correlating experimental features and electrochemical performance properties of the materials.

2. Results and Discussion

2.1. High Throughput Characterization of the $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_{3\pm\delta}$ Materials Library

Thin film compositional maps of LSMCF (Figure 1a) were prepared by combinatorial pulsed laser deposition on 4 inch Si (100) and 3 inch 8% mol YSZ (100) wafer substrates covered with $Ce_{0.8}Gd_{0.2}O_2$ (CGO) buffer layer. The parent compounds of $La_{0.8}Sr_{0.2}MnO_{3\pm\delta}$ (LSM), $La_{0.8}Sr_{0.2}FeO_{3\pm\delta}$ (LSF) and $La_{0.8}Sr_{0.2}CoO_{3\pm\delta}$ (LSC) were sequentially ablated in spatially separated layers with low thickness, allowing for the rearrangement of the B-cation into a single perovskite phase (Figure S1 and Section S1, Supporting Information).[44] The LSMCF library was then comprehensively characterized with high-throughput characterization methodologies (Experimental Section, Table S1, Supporting Information). **Figure 2** provides an overview of the compositional, structural, and functional properties analyzed for this material library.

A. Aspuru-Guzik
Departments of Chemical Engineering and Applied Chemistry, and Materials Science and Engineering
University of Toronto
Toronto M5S3E5, Canada
A. Aspuru-Guzik
Acceleration Consortium
University of Toronto
Toronto M5G1X6, Canada
A. Aspuru-Guzik
Canadian Institute for Advanced Research (CIFAR) and Vector Institute for Artificial Intelligence
Toronto, Ontario M5G1M1, Canada
A. Tarancón
ICREA
Passeig Lluís Companys 23
Barcelona 08010, Spain

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

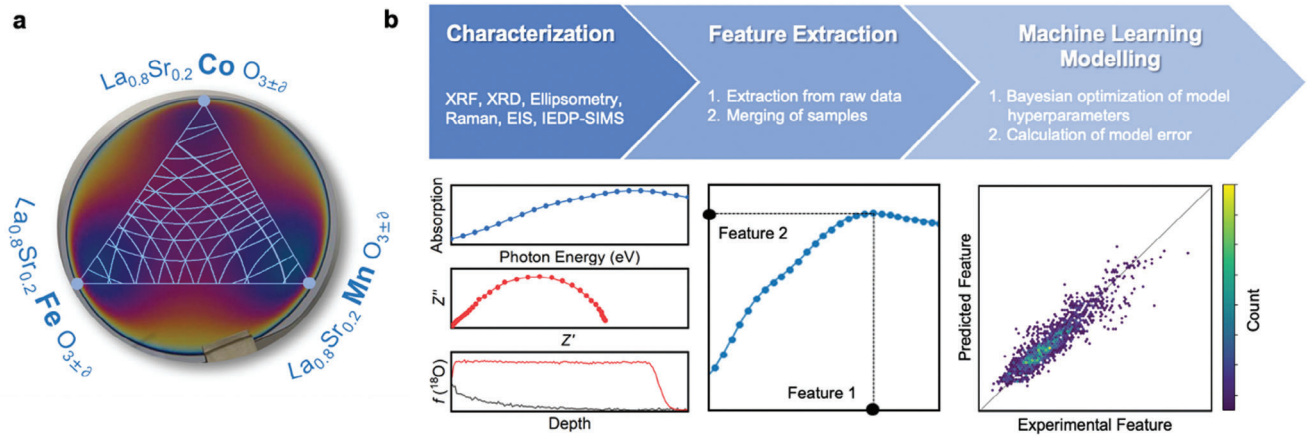

Figure 1. a) $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_{3\pm\delta}$ combinatorial 4" wafer with different Mn:Co:Fe ratios. b) Data analysis workflow.

To evaluate the composition of the LSMCF system, XRF measurements were performed at the SOLEIL synchrotron. The resulting B-cation distribution is presented in Figure 2a, which shows the normalized proportion of Mn, Co, and Fe obtained from fitting the XRF spectra (Figure S2a, Supporting Information). The distribution of these elements on the Si wafer substrate is representative of the contribution of LSM, LSF, and LSC during the deposition and agrees with the thickness distribution observed (Figure S1, Supporting Information). Note that none of the B-cation stoichiometry reaches a value of 1, indicating that a degree of B-cation intermixing is always present. The distribution of Mn, Co, and Fe cations allowed the correlation of each sample position on the Si wafer substrate with a specific LSMCF composition.

The crystallographic structure of the combinatorial LSMCF thin film was characterized by XRD in an automatized diffractometer. The diffraction patterns confirmed the presence of a single perovskite phase all along the compositional space, with the absence of secondary phases (Figure S2b, Supporting Information). The XRD patterns of the perovskite phase could be indexed in the cubic ideal perovskite phase, with well-defined single peaks, but deviations from the cubic phase or widening of diffraction peaks like those observed in the rhombohedral form of the perovskite phase could not be resolved. The pseudo-cubic lattice parameter of the materials was calculated from the diffractograms measured along the sample. Figure 2b shows the variation of the pseudo-cubic lattice parameter of the perovskite phase in the 3.85–3.96 Å range. The smallest lattice parameter, a = 3.85 Å, is found in the Co-rich region, as expected from the lower ionic radius of Co^{3+} when compared to the other two cations, and in line with the literature on Sr-doped and undoped cobaltite thin films.[45,46] A progressive increase of the lattice parameter is observed with the addition of Mn, reaching a relative maximum at a = 3.93 Å, as expected for LSM.[47,48] Finally, the largest lattice parameter in the LSMCF film corresponds to the Fe-rich region, reaching a value of ≈3.96 Å, in agreement with values found in other LSF thin film studies.[49,50]

To provide additional insight into the structural properties influenced by the B-cation distribution, the LSMCF sample was further analyzed by Raman spectroscopy. Figure 2c shows all the Raman spectra obtained from LSMCF. The peaks observed at 300, 520, and 940–970 cm^{-1} are characteristic of the Si(100) substrate (Figure S3, Supporting Information).[46,47] It is evident that the region more susceptible to changes throughout the mapping lies within the 550–700 cm^{-1} range (Figure S2c, Supporting Information). The Mn-rich portion presents a characteristic peak centered at 630 cm^{-1}, which shifts toward 650 cm^{-1} when the Fe content in the LSMCF film increases. Finally, transitioning to the Co-rich region results in a further shift of the Raman peak toward ≈665 cm^{-1}, while exhibiting a small shoulder at ≈620 cm^{-1}. This peak in the 630–650 cm^{-1} range has been typically assigned to the A_g–type mode in the distorted rhombohedral symmetry D^6_{3d} ($R\overline{3}c$) of Sr-doped lanthanum-based perovskites[51–54] and, in iron- and cobalt-doped manganites, it is related to the stiffness of the breathing mode of the oxygen octahedra (B′—O—B motif in $AB_{1-x}B'_xO_3$).[55,56] According to atomistic simulations,[13] the strength of this bond seems to be directly related to the formation of oxygen vacancies in perovskites, which, in turn, drives the oxygen mass transport properties of a material.

Since optical properties relate to redox reactions, as previously demonstrated for some parent compounds of the LSMCF family,[49,50] the optical properties of the combinatorial LSMCF thin film were also studied by means of spectroscopic ellipsometry. Figure 2d shows all the optical absorption spectra obtained during the mapping acquisition (see Figure S2d, Supporting Information for individual spectra in the Mn, Co, and Fe-rich regions). The Co-rich area dominantly absorbs in the low energy range below 3 eV with a band at around 2.7 eV, while the Mn- and Fe-rich regions have greater absorption in the high energy range. In the case of the Mn-rich region, a low-intensity absorption band is present below 1.5 eV, in agreement with previous studies on Sr-doped LSM.[57] The Fe-rich spectra also show a partial increase of the optical absorption coefficient around 1.5 eV, related to Sr-doping but less evident than in the Mn-rich region. In addition, an optical feature is observed at ≈3 eV, which is related to the relevant electronic transition O 2p-Fe t_{2g}.[49,50]

Regarding the functional properties of the LSMCF family, an unprecedented study of the oxygen mass transport properties of a complete ternary compositional map was carried out by using isotope exchange depth profiling – secondary ion mass spectrometry (IEDP-SIMS). In this study, the LSMCF sample was

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

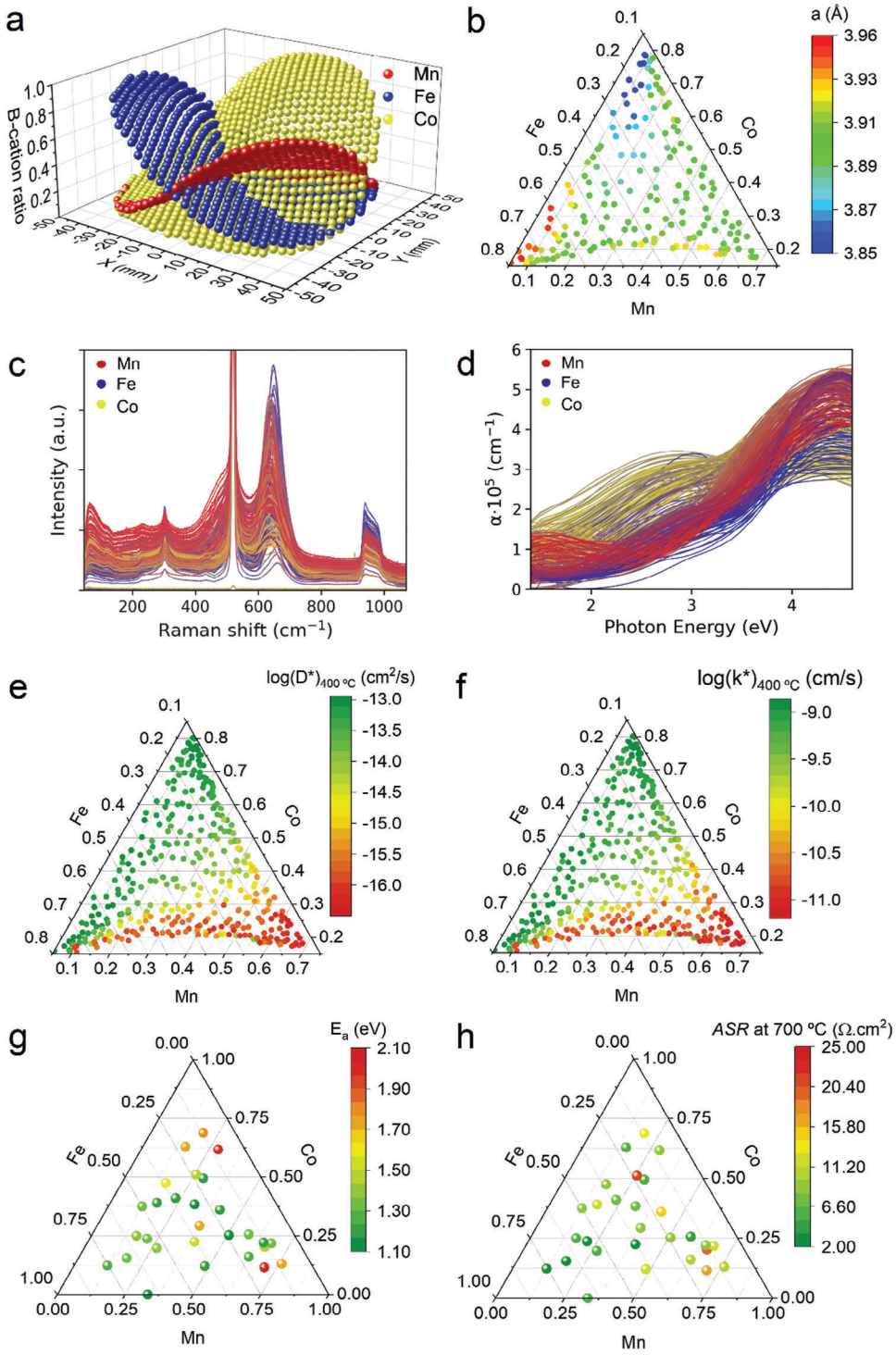

Figure 2. Characterization of the LSMCF combinatorial film. *Composition*: a) normalized B-site stoichiometry (B-cation ratio distribution) obtained by XRF. *Structure*: b) Pseudo-cubic lattice parameter obtained from the XRD analysis; c) Raman spectra measured along the combinatorial sample; d) optical absorption spectra from ellipsometry mapping. *Performance*: e,f) oxygen mass transport parameter maps from IEDP-SIMS at 400 °C; e) oxygen diffusion coefficient D*; f) oxygen surface exchange coefficient k*. g,h) Electrochemical characterization resulting from the EIS mapping after the conditioning period: g) activation energy and h) ASR calculated at 700 °C.

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

exposed to an atmosphere enriched with ≈90% ^{18}O isotopes at 400 °C for 30 min before measuring the corresponding isotope exchange and induced diffusion profiles by SIMS. A collection of ^{18}O fraction self-diffusion profiles is represented in Figure S2e (Supporting Information). In this figure, two oxygen isotope diffusion regimes can be differentiated along the compositional map. In regions with high manganese content, the profile exhibits an exponential-like decay in the ^{18}O fraction along the *z*-direction. This behavior is characteristic of mixed ionic electronic conductors with poor ionic conduction, and it is often observed in chromite-[58,59] and manganite-based[6,60] perovskites. On the other hand, in the Fe- and Co-rich regions, the ^{18}O concentration profile appears flat. This behavior is characteristic of transport mechanisms dominated by surface exchange with no limitation on oxygen diffusion through the film. After fitting these isotopic oxygen diffusion profiles (Section S2, Supporting Information), the most important magnitudes determining the oxygen reactivity and mass transport properties of a material, the oxygen diffusion coefficient (D*) and surface exchange coefficient (k*), were obtained for the complete library of materials (Figure 2e,f; Figure S4, Supporting Information). The values of D* fall within the range of 10^{-13}–10^{-17} cm^2 s^{-1}, while the k* varies from values on the order of 10^{-12} cm s^{-1} for the Mn-rich compounds to 2×10^{-9} cm s^{-1} for both the Co- and Fe-rich regions. The k*·D* product, a figure of merit of the electrochemical performance,[61,11] thus varies in several orders of magnitude from a lower limit of 10^{-28} cm^3 s^{-2} for the Mn-rich region until a maximum value of 4×10^{-22} cm^3 s^{-2} for higher amounts of Co and Fe. Estimated values of k*·D* on parent compounds are higher than the ones reported in the literature for bulk materials in the same temperature range. For example, the Mn-rich region of our combinatorial map presents values of ca. 10^{-28} cm^3 s^{-2} while LSM reported in the literature is typically below 10^{-33} cm^3 s^{-2}.[62] Similarly, k*·D* values of 10^{-22} cm^3 s^{-2} are calculated here for Co-rich regions while the ones reported in the literature for LSC (20% Sr doping)[62] are in the range of 10^{-25} cm^3 s^{-2}. This substantial enhancement of several orders of magnitude in the activity of the thin film LSMCF library is likely attributable to the presence of significant contributions to the oxygen mass transport along highly diffusive defective grain boundaries.[29,44,47,63,64] The origin of this enhancement was elucidated in our previous works[29,44,49,65] for specific compositions and can now be generalized with this work to the whole LSMCF family.

Finally, the electrochemical performance was mapped using a realistic electrochemical cell where LSMCF serves as the working electrode, the YSZ substrate acts as an electrolyte and a highly performing silver layer is the counter electrode. The electrochemical activity as a function of temperature was measured for the whole LSMCF library in an automated probe station capable of controlling the temperature and atmosphere while spatially scanning the library. It is important to note that final measurements were carried out after a conditioning period of 10 h at 750 °C to avoid the evolution of these materials in the initial stages.[65,66] The most relevant parameter of an electrode operating in a solid oxide cell is the associated area-specific polarization resistance (ASR), which can be obtained from EIS spectra using equivalent circuit fitting. In this work, all the impedance spectra were fitted with a single generalized modified Jamnik–Maier equivalent circuit (Figure S5, Supporting Information) irrespective of the composition. The temperature dependence of the resulting values of ASR showed a typical Arrhenius behavior, which allowed us to derive the preexponential factor (ASR_0) and the activation energy (E_a) as temperature-independent performance descriptors (Figure 2g; Figure S6, Supporting Information respectively) from fitting $ASR = ASR_0 \cdot T^{-1} \cdot exp(E_a \cdot k^{-1} \cdot T^{-1})$, where T and k are the temperature and the Boltzmann constant, respectively. Moreover, using the Arrhenius relation, the ASR at the specific temperature of 700 °C was also calculated (Figure 2h) to facilitate a direct comparison of the relative performance of the materials compiled in this study while avoiding the observed variability of temperatures along the library at fixed conditions (Figure S7, Supporting Information). According to these results, the polarization resistance is overall larger in Mn-rich regions than in the Co,Fe-rich regions, consistent with Mn-rich perovskites poor ionic conductivity and sluggish oxygen exchange at the surface, even in thin-film form.[26] Nonetheless, the ASR values observed in the Mn-rich regions are not as high as that of bulk LSM,[26,67] probably due to the presence of Co and Fe, even at relatively low amounts ($(Co+Fe)_{min.} = 0.3$). The Mn and Co-rich regions exhibit E_a values in the 1.6–2.1 eV range, while compositions with higher cation intermixing show lower E_a values, approximately in the 1.1–1.50 eV range. These values are considerably lower than those found for the parent materials (i.e., 1.60, 1.72, and 2.35 eV for LSC, LSF, and LSM, respectively). The lowest ASR values down to $\approx 2\,\Omega\,cm^2$ are localized in the Fe-rich region, with relative B-cation compositions in the ranges of 12–27% Mn, 12–25% Co and 50–75% Fe. Increasing the content of Co and Mn leads to progressively higher ASR, with the highest values observed in the Mn-rich compounds. Despite these valuable general guidelines, it is clear from the reduced number of electrochemical measurements (limited by the complexity and duration of the automated tests) that an adequate material optimization is not completely possible only based on such direct and discrete measurements.

2.2. Machine Learning Modelling

2.2.1. Feature Extraction

After the compositional map of LSMCF was synthesized and characterized using high-throughput methods (Figures S8 and S9, Supporting Information for a summary of the dataset), we extracted several parameters from the raw data to use as features in machine learning models. These features were grouped into three sets called *Composition*, *Structure*, and *Performance* (**Table 1**). The main goal of the work was to model the *Performance* group of features, used as output parameters. The models used either the *Composition* or the *Structure* group of features as input parameters, labeled as *Performance (Composition)* and *Performance (Structure)* models, respectively.

Regarding the *Composition* set, stoichiometries of Mn, Fe, and Co from XRF synchrotron measurements were employed as the only compositional parameters. Regarding the *Structure* set, four types of features were extracted from XRD, Raman, and ellipsometry measurements. We employed the lattice parameter derived from the XRD measurements described above. With respect to Raman, we employed parameters from Gaussians fitted to the

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

RIGHTSLINK
15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1. Summary of the characterization techniques used, the features extracted from the raw data, the number of features per sample, and the number of samples with different compositions measured.

Group	Characterization technique	Extracted features	Number of features	Samples
Composition	X-Ray Fluorescence (XRF)	• Stoichiometry of Co, Fe and Mg	3	850
Structure	X-Ray Diffraction (XRD)	• Lattice Parameter	1	178
	Raman	• Fitted Gaussians (Area, FWHM & Raman shift at $I_{max.}$)	36	139
	Ellipsometry	• Thickness and Roughness • Absorption coefficient at Fixed Photon Energies	6	321
Performance	Electrochemical Impedance Spectroscopy (EIS)	• Activation Energy (E_a) • Preexponential Factor	2	28
	Exchange Depth Profiling – Secondary Ion Mass Spectrometry (IEDP-SIMS)	• Oxygen Diffusion Coefficient (D*) • Surface Exchange Coefficient (k*)	2	482

spectra (Figure 2c). A combination of twelve Gaussians was necessary to satisfactorily fit each Raman spectra, with five Gaussians being fitted in the 50–483 cm^{-1} range and seven Gaussians being fitted in the 569–1071 cm^{-1} range (Section S3, Figure S10, and Table S2, Supporting Information). The 483–569 cm^{-1} shift range containing the peak belonging to the Si substrate was omitted because its larger intensity and contribution to the fitting error disrupted the correct fitting of the smaller peaks. The total area, the full width at maximum intensity (FWHM), and the Raman shift corresponding to the maximum intensity were calculated within the fitting range for each fitted Gaussian, resulting in 36 features. Finally, we employed two types of features from ellipsometry. From the absorption spectra (Figure 2d), the absorption coefficient at photon energies of 1.6, 2.7, 3.5, and 4.4 eV, where broad bands were observed, was obtained for further modeling (Section S3, Figure S11, and Table S2, Supporting Information). The thickness and the roughness of the film were also obtained from the ellipsometer raw data (see Experimental Section). As a result, a total of six features were extracted from Ellipsometry. Regarding the *Performance*, four different parameters were obtained from IEDP-SIMS and EIS measurements. In particular, IEDP-SIMS delivered the oxygen diffusion and oxygen surface exchange coefficient at 400 °C, while EIS measurements at different temperatures were represented by the preexponential factor and the activation energy characteristic of the observed Arrhenius behavior (Figure 2e,f). The extracted feature groups, summarized in Table 1, were then used either as input or target output parameters in machine learning models.

2.2.2. Machine Learning Method Screening

After the feature extraction, we subsequently evaluated the accuracy of different machine learning methods in modeling different combinations of *Composition*, *Structure*, and *Performance* relationships. Due to variations in the samples measured with the different techniques across the wafer (Figure S9, Supporting Information), the analysis focused on the data from the 321 coordinates measured by ellipsometry, and the missing data was calculated as a weighted average of the closest samples (see Methods). Seven supervised machine learning methods known for their effectiveness with small datasets were tested, including k-nearest neighbors (KNN), least absolute shrinkage and selection operator regression (LassoR), kernel ridge regression (KRR), a support vector machine (SVM), a random forest (RF), a multiple layer perceptron (MLP) neural network and Gaussian Processes (GP).[68–70] The average prediction score and uncertainty, calculated via five-fold cross-validation on the training set and representing the prediction capacity on unseen data,[71] are similar to the score against a test set comprising 10% of the dataset (Figure S12, Supporting Information, Experimental Section), indicating limited overfitting.

Among all the ML methods, random forests exhibit some of the best scores and robustness across various models, similar to KNNs, and closely followed by SVMs (Figure S12, Supporting Information). The good performance of KNNs, which rely on proximity to make predictions, may stem from the continuous nature of the compositional map and the similarity between the ML method and the approach applied to merge data from different techniques. The *Performance (Composition)* model display strong prediction and generalization capacities, facilitating the estimation of the performance for any composition, and the identification of compositions with optimal performance. Notably, the *Structure* feature group is also shown to reliably model *Performance*, which will enable the correlation of the functional features with the intrinsic properties of the materials. Random forests were then selected for further investigation based on their good prediction errors and their interpretability. The prediction accuracy of Random forests was further improved by exploring different *Structure* parameters, including three approaches to process Raman and Ellipsometry spectra, different combinations of features, and various scaling treatments (Sections S3 and S5, Figures S13 and S14, Supporting Information). We proceeded with random forests and the *Structure* features yielding the best scores: fitted Gaussians from the Raman spectra and the absorption coefficient at fixed photon energies from the Ellipsometry spectra.

2.2.3. Performance Prediction ML Models

The performance of the whole LSCMF family was predicted by a Random Forest model using the B-cation stoichiometry as the sole input, i.e., labeled as the *Performance (Composition)* model.

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

RIGHTSLINK

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

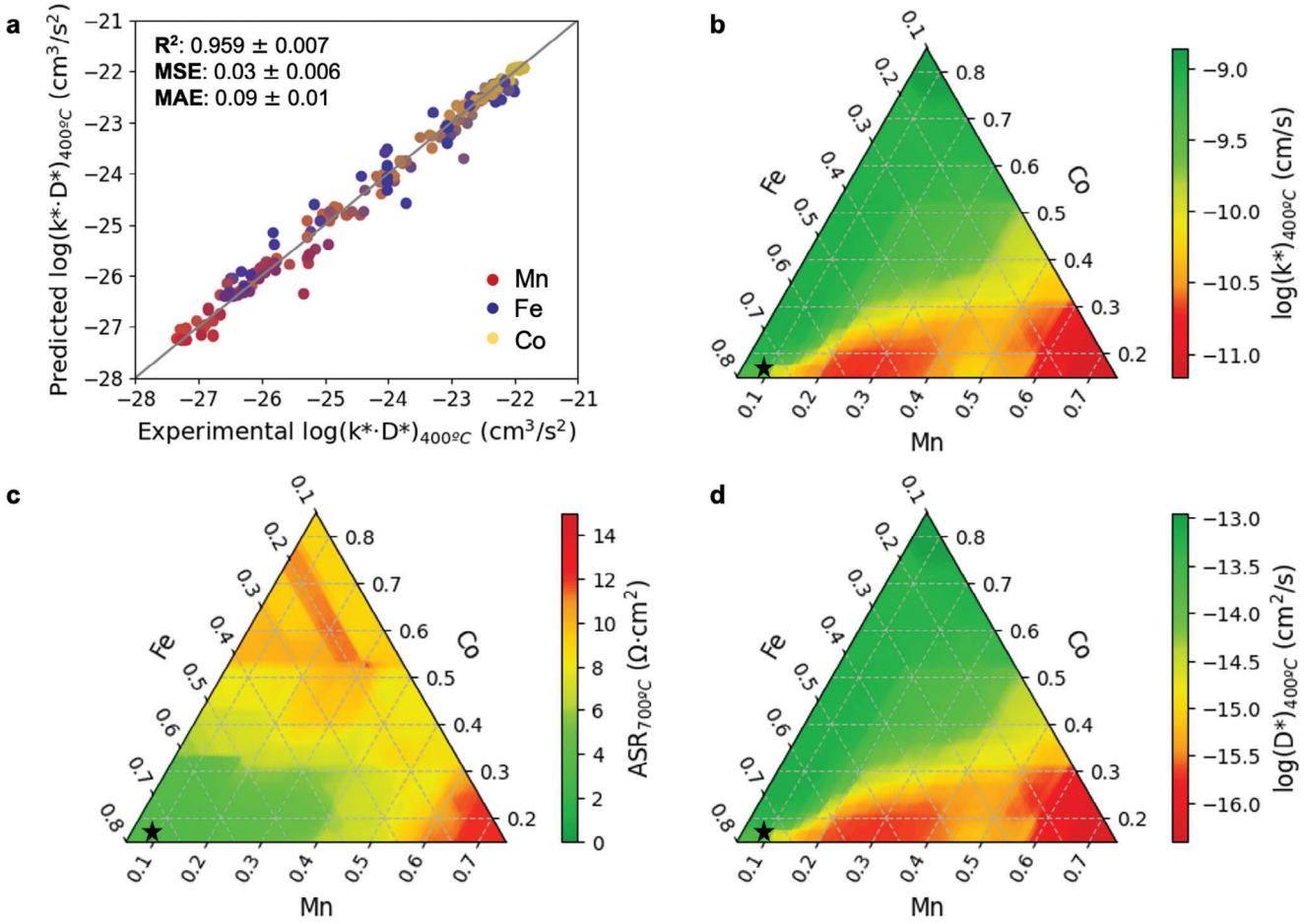

Figure 3. Prediction by the *Performance (Composition)* random forest model. The output parameters of the model are the *Performance* features k*, D*, E_a, and preexponential factor, separately, while the input parameters are the Co, Fe, and Mn ratios. a) Experimental values versus the values predicted by the model. ML-predicted b) log(k*) at 400 °C, c) area-specific polarization resistance (ASR) at 700 °C, and d) log(D*) at 400 °C, starting from the B-cation elemental stoichiometry as input. The star in (b–d) corresponds to the composition $La_{0.8}Sr_{0.2}Mn_{0.09}Fe_{0.75}Co_{0.16}O_{3\pm\delta}$ with the optimal $ASR_{700°C}$ of 3. Ω cm², as predicted by the model.

The output parameters of the model were the *Performance* features k*, D*, E_a, and preexponential factor, separately, while the input parameters were the Co, Fe, and Mn ratios. The R^2 of the model is 0.959 ± 0.006, showing an excellent match between the predicted and the experimental values (see **Figure** 3a for the figure-of-merit k*·D*). Despite the good fitting, the divergence of the predicted values with respect to the experimental values is larger than with respect to the training data due to the coordinate merging (Figures S15 and S16, Supporting Information). The mean absolute error of the model for the experimental values is 0.2 eV, $10^{1.2}$ Ω cm² K^{-1}, $10^{0.6}$ cm s^{-1} , and $10^{0.9}$ cm² s^{-1}, for the activation energy, the preexponential factor, k* and D*, respectively. Nevertheless, this does not prevent the *Performance (Composition)* from successfully identifying the composition range with the best electrochemical performance and providing an insight into the trends across the LSCMF compositional space.

Using this model, ternary continuous maps covering the full range of Co:Fe:Mn ratios with infinitesimal resolution were predicted for $k^*_{400\,°C}$ (Figure 3b), $D^*_{400\,°C}$ (Figure 3d), the activation energy and the preexponential factor (Figure S15a,b, Supporting Information). Additionally, the predicted ASR at the relevant operation temperatures of 700 and 600 °C calculated from E_a and the preexponential factor were also plotted (Figure 3c; Figure S15c, Supporting Information respectively). All the predicted and experimental datasets have been made available to the scientific community in an open repository.

Regarding the oxygen mass transport properties of the LSMCF family, predicted values for k* and D* indicate that the best performances (green areas in Figure 3b,d) are reached across the entire range of Fe concentration (z ≈ 0.1–0.9), for moderate contents of Mn ($x < 0.4$) combined with values of Co higher than Mn ($y >$ x). This highlights that manganese strongly decreases the oxygen mass transport properties of the LSMCF compounds, especially those based on Co. Interestingly, this trend is not preserved in electrochemical measurements (i.e., ASR after 10 h of stabilization in Figure 3d), where better performances correspond to materials with relevant amounts of Mn (up to $x = 0.4$). Indeed, contrary to k*·D*, an excess of Co ($x > 0.3$) represents a penalty for the ASR. This behavior is likely related to the deleterious surface evolution of Co-rich compounds due to Sr segregation often

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

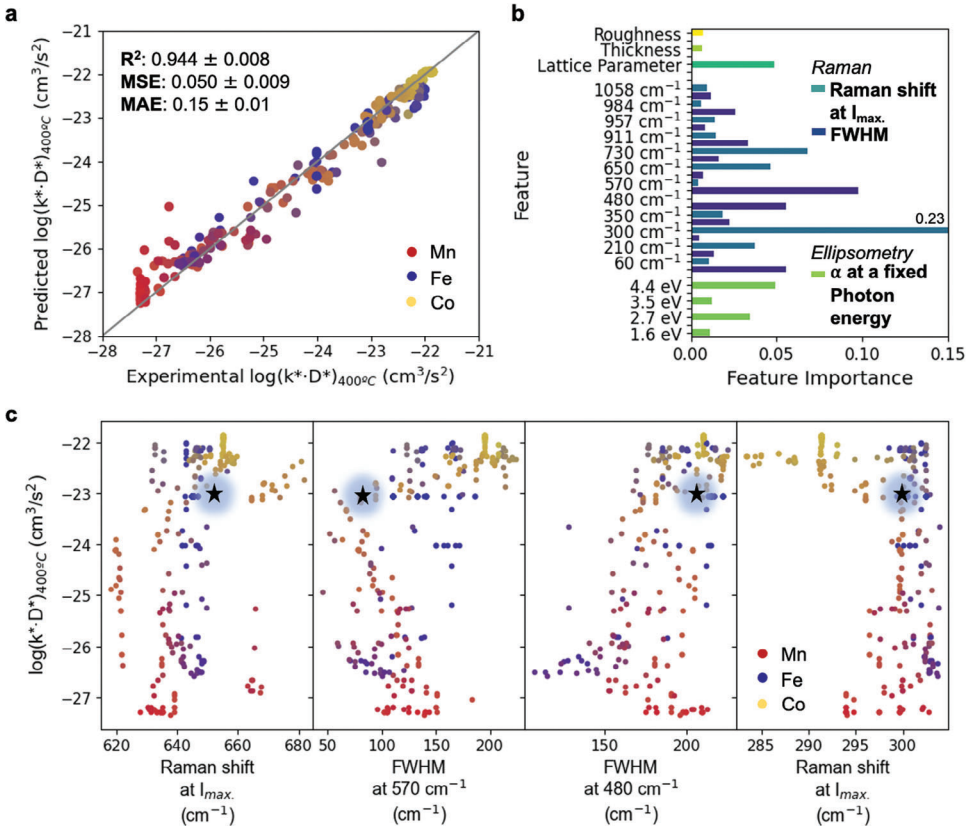

Figure 4. *Performance (Structure)* random forest models. The output parameters of the model are the *Performance* features k* and D*, separately, while the input parameters correspond to the Gaussian fitting of the Raman spectra, the lattice parameter, the thickness, the roughness, and the absorption coefficient at fixed photon energies in the ellipsometry spectra. a) Experimental values versus the values predicted by the model and b) the corresponding feature importance. c) *Performance* features log (k*·D*) as a function of key *Structure* features: the Gaussian Raman shift at $I_{max.}$ around $\nu \approx 650$ cm^{-1}, the FWHM at $\nu \approx 570$ cm^{-1} and at $\nu \approx 480$ cm^{-1}, and the Gaussian Raman shift at $I_{max.}$ around $\nu \approx 300$ cm^{-1}. The labels of the features in (b), corresponding to the Gaussian fit of the Raman spectra, are calculated as the average wavenumber of the n^{th} Gaussian across all the samples, where n represents the position of the Gaussian when ordered according to the average wavenumber at $I_{max.}$. The star in **c** corresponds to the composition $La_{0.8}Sr_{0.2}Mn_{0.08}Fe_{0.84}Co_{0.09}O_{3\pm\delta}$ within the optimal $ASR_{700°C}$ range of 3.4 ± 0.5 Ω cm^2.

reported in the literature,[72–74] which could be partially compensated by adding manganese according to our model. The optimal ASR values at 700 °C correspond to compositions with $x = 0$–0.4, $y = 0$–0.3, and $z = 0.4$–0.9, following a similar trend to that at 600° (Figure S15, Supporting Information). The best $ASR_{700\,°C}$ performance values of 3.4 ± 0.5 Ω cm^2 correspond to the composition range $La_{0.8}Sr_{0.2}Mn_{0.08\pm0.04}Fe_{0.84\pm0.07}Co_{0.09\pm0.05}O_{3\pm\delta}$ with log(k*·D*) of −23.5 ± 0.7 cm^3 s^{-2}, while the composition with the best predicted performance is $La_{0.8}Sr_{0.2}Mn_{0.09}Fe_{0.75}Co_{0.16}O_{3\pm\delta}$.

Similarly to the *Performance (Composition)* model, the accuracy of *Performance (Structure)* random forest models was very high (see the scores and the excellent matching between experimental and predicted values in **Figure** 4a, and Figures S17–S21, Supporting Information). In this model, the output parameters of the model are the *Performance* features k* and D*, separately, while the input parameters correspond to the Gaussian fitting of the Raman spectra, the lattice parameter, the thickness, the roughness, and the absorption coefficient at fixed photon energies in the ellipsometry spectra. The relative importance of the different input features in the final prediction is shown in Figure 4b. As observed in the plot, the input features from the Raman spectra have the largest relative importance. Eventually, it is possible to predict the performance of the LSMCF family with R^2 of 0.962 ± 0.004 using models trained with Raman features as the only input features, as opposed to an R^2 of 0.63 ± 0.05 when the features from the ellipsometry spectra are used instead. In this *Performance (Structure)* model, the Raman features were extracted from the Gaussian functions fitted to the Raman spectra and yielded a similar feature importance distribution as Raman features extracted from another approach based on the detection of Raman peak maxima (Figure S22, Supporting Information). In parallel, the lattice parameter is observed to display a secondary role in the prediction, as it represents the long-range structure but fails to capture the short-range distortions. Other features, such as the thickness of the layer at each composition or the peaks observed in ellipsometry, do not govern the prediction accuracy of the model either.

According to Figure 4b, Raman bands in the regions of 200–300, 450–500, and 600–700 cm^{-1} are especially relevant for the electrochemical performance, which agrees with the results of the correlation analysis (Figures S23–S24, Supporting Information). For instance, the peak at ≈280–300 cm^{-1} shifts to lower Raman shifts with increasing k*·D* and Co concentration

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Figure 4c), which could be due to the appearance of a vibration mode <290 cm^{-1} overlapping with the Si substrate mode at 300 cm^{-1} (Figure S3, Supporting Information). In parallel, the peak at ≈650 cm^{-1} shifts to larger Raman shifts, and the peaks at ≈570 and ≈200 cm^{-1} become wider (Figure 4c). These Raman regions have been previously assigned in the literature to lattice distortions present in nearly cubic Sr-doped lanthanum transition-metal perovskites such as $La_{1-x}Sr_xMnO_3$ and $La_{1-x}Sr_xFeO_3$.[51–56,75] These Raman-active modes in $La_{1-x}Sr_xMO_3$ ($x < 0.3$, M = Fe, Mn, and Co) have been associated to disorder and distortions in the oxygen sublattice (tilting, bending, and stretching of the oxygen octahedra) due to a different ionic radius of the mixed-valence B-site substituents and their specific electronic properties.[76] For instance, the peak at 630–650 cm^{-1} has been previously assigned to the oxygen stretching in the extended B′—O—B motif of $AB_{1-x}B'_xO_3$ in iron- and cobalt-doped lanthanum strontium manganites, as described above.[49–54] Therefore, the *Performance (Structure)* model presents an unambiguous statistical correlation between the distortions of the oxygen sublattice and the electrochemical performances. These distortions can be linked to mass transport properties using theoretical models that correlate the formation energy of oxygen vacancies with the B—O bond strength and electron exchange driving forces.[13] However, previous works were not able to quantify the impact of the different transition metals on the oxygen sublattice distortions by establishing monotonic relations between the Raman shifts/intensities and relevant physical magnitudes of the B-cation, e.g. B—O bond distance and stiffness or the covalency of the bond. Despite remarkable efforts dedicated to establishing such relationships,[75–79] the link between vibrational modes and oxygen mass transport properties has remained unclear, hindering the use of Raman spectroscopy for straightforwardly anticipating performance in perovskite oxides. In this work, we were able to overcome this barrier and capture the relevant information contained in the Raman spectra of the LSMCF family to predict the final performance of highly complex materials using machine learning models.

3. Conclusion

Following a high-throughput approach, this study integrates various experimental techniques with distinct limitations to comprehensively characterize a thin film compositional map of $La_{0.8}Sr_{0.2}Mn_xCo_yFe_zO_{3\pm\delta}$ perovskite oxides deposited via combinatorial pulsed laser deposition. Compositional, structural, and functional properties were retrieved through XRF, XRD, Raman spectroscopy, ellipsometry, IEDP-SIMS, and EIS mapping. Spectral deconvolution techniques in tandem with machine learning methods, particularly Random Forest, enable excellent prediction of mass transport properties (oxygen diffusion and surface exchange coefficients) and electrochemical performance of the whole family of materials while offering novel insights into the significance of individual spectral features. In this regard, we conclude that Raman spectra alone contain all the relevant information for the prediction of electrochemical properties of LSMCF perovskites. Additionally, the application of machine learning has facilitated the creation of unprecedented continuous functional property maps that effectively capture trends in experimental samples despite the inherent variability in experimental conditions, delivering optimal compositions for relevant SOFC applications. To boost further development, we have made the complete dataset available to the research community. Our integrative approach not only advances the understanding of LSMCF compounds themselves but also offers a promising framework for optimizing the electrochemical properties of alternative sets of materials, thereby accelerating the identification and development of novel high-performance compositions.

4. Experimental Section

Thin Film Deposition: Thin film libraries of $La_{0.8}Sr_{0.2}Mn_xCo_yFe_{1-x-y}O_{3\pm\delta}$ were prepared by combinatorial pulsed laser deposition on a Si (100) (*Siegert Wafer*) and yttria-stabilized zirconia (YSZ) 8YSZ (100) (MSE Supplies) wafer substrates using a large area pulsed laser deposition equipment (*model 5000 PVD Products*) with a 248 nm KrF excimer laser (*Lambda Physik, COMPex PRO 205*). Commercial targets (*Kceracell*) of $La_{0.8}Sr_{0.2}CoO_3$, $La_{0.8}Sr_{0.2}FeO_3$, and $La_{0.8}Sr_{0.2}MnO_3$ were employed for the deposition of the thin films. The oxygen partial pressure was set to 0.007 mbar and the temperature was 700 °C. Depositions were done with an ablation frequency of 10 Hz, a laser fluency of ≈0.8 $J{\cdot}cm^{-2}$ and a target-substrate distance of 90 mm. The total number of pulses applied to each target in a single cycle was adjusted to deposit a ≈1 nm layer of each parent compound. After the ablation of each parent compound, the substrate holder was rotated by 120°. The final thickness of the film was controlled by adjusting the number of deposition cycles. The sample deposited on YSZ included a $Ce_{0.8}Gd_{0.2}O_2$ (CGO) interlayer deposited by large area PLD. During the CGO deposition, the substrate was rotated at 10 rpm and the position of the laser spot on the target was rastered to obtain a homogeneous layer. The rest of the deposition parameters remained unchanged.

Compositional, Structural, and Optical Characterization: The composition of the films was analyzed by X-ray Fluorescence performed using synchrotron radiation at the *DiffAbs* beamline of the *SOLEIL* facility. XY-resolved XRF elemental characterization was carried out using a primary beam energy of 12 keV and an incident angle of 10° between the source beam and the sample surface. The beam spot was a 0.3×0.3 mm^2 square, but the projected area onto the sample surface increased to 1.25 mm in the vertical direction due to the incident angle. The fluorescence signal was collected with a 4-element silicon drift detector placed 50 cm away from the sample. The acquisition was performed within a coordinate system consisting of a 110×110 mm^2 squared mesh with 1 mm lateral steps in both the horizontal and vertical directions. Each XRF spectrum was collected with an integration time of 0.2 s, and the entire map acquisition took ≈40 min. A total of 6305 XRF patterns were obtained on the sample. To enhance the signal/noise ratio of the XRF data, binning with 3 mm intervals was applied, resulting in ≈850 patterns to study. A Matlab code was developed to analyze each XRF spectrum and calculate the elemental Mn:Fe:Co ratio. The composition of the samples characterized by Raman and Electrochemical Impedance Spectroscopies was analyzed by Energy-dispersive X-ray spectroscopy (EDX) with a *Carl ZEISS Auriga* scanning electron microscope operating at 18 kV.

In-plane (XY) resolved structural characterization was performed using a Panalytical X'pert MRD X-ray diffractometer in the ω-2θ Bragg-Brentano configuration. The instrument was equipped with a copper X-ray tube in line-focus mode and a PIXcel3D detector in scanning-line mode, with an active length of 2.511°2θ. The incident beam was modulated through 1/8° divergence, 4 mm vertical, and 0.02 rad Soller slits, leading to a spot size of ≈2 mm^2. Data acquisition was carried out within a coordinate system consisting of a 90×90 mm^2 squared mesh with 5 mm lateral steps both in the horizontal and vertical directions, resulting in a total of 289 diffractograms. Each XRD pattern was collected with a 1.1° tilt angle (ω_0) to reduce the intensity of the YSZ (200) and (400) reflections over the 20–80° 2θ range with 0.02° step and 140 s/step. The acquisition of each diffractogram took ≈30 min, resulting in a total measurement time of about 142 h. A custom Matlab code was developed to efficiently analyze the large number

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH
RIGHTSLINK

of diffractograms and obtain a sample map with the lattice parameters of LSMCF. The code identifies the presence of LSMCF and CGO peaks in each XRD pattern. All the identified peaks are fitted using a function composed of two pseudo-Voigt profiles, simulating the $K_{\alpha 1}$ (1.5405929 Å) and $K_{\alpha 2}$ (1.5444274 Å) Cu lines. The angular position in °2θ is calculated for each peak and used in a minimization routine to determine the sample displacement and the lattice parameter of the LSMCF deposition.

Raman spectroscopy measurements were carried out in films deposited on Si (100) with an *XploRA Nano* instrument (*Horiba*). A x100 objective, a grating of 1800 (450–850 nm range), a 638 nm laser with a power filter of 25%, a slit of 200 µm, and a hole of 500 µm were used. Spectra were measured with an acquisition time of 2–5 s and 15–30 accumulations. Samples in the combinatorial films on Si (100) were measured with a spacing of 7 mm. The Raman spectra were smoothed with a Savitzky-Golay filter (*savgol_filter* from *scipy*) with a 3rd order polynomial. Filter window lengths of 21 and 11 were used in the identification of maxima and in the Gaussian fitting, respectively. Spectroscopic ellipsometry was carried out with a multiwavelength light source (*UVISEL, Horiba*) and an automatized XY stage. Within the combinatorial films on Si (100), 321 points distributed in a cartesian grid were measured in the 1.5–5 eV energy range. The absorption spectra and the thickness were obtained by fitting four oscillators to the raw data with a batch processing functionality of the *DeltaPsi2* analysis software (*Horiba*).

Functional Characterization: IEDP-SIMS was employed for studying the oxygen kinetics of the LSMCF combinatorial library. A combinatorial LSMCF film grown directly on a 3″ diameter Si (100) substrate was selected for the analysis. The sample was enclosed in a quartz tube and pumped down to 10^{-7} mbar. A first annealing was performed at 400 °C in 200 mbar of pure oxygen (99.999%) with $^{18}O_2$ in the normal isotopic abundance. After the annealing, the tube was pumped down and re-filled with a ≈90% $^{18}O_2$ enriched gas (200 mbar). The sample was quenched to room temperature once the exchange was completed. The conditions selected for the isotope exchange were 400 °C and 30 min. The ^{18}O diffusion profile signal was recorded using a TOF-SIMS 5 instrument (*IONTOF GmbH*). The mapping was carried out by scanning the sample with a 3 mm spacing between the measured points. A total of 480 points were measured by the end of the experiment.

A large-size high-temperature probe station (*MicroXact CPS-HT*) was employed to carry out electrochemical impedance spectroscopy (EIS) mappings. The chamber remained closed during the measurements and synthetic air flowed through at 500 mL min^{-1}. A low impedance silver paste electrode (*Sigma-Aldrich*) was brushed on the bottom of the YSZ wafer, while a layer of gold paste (*Fuel Cell Materials*) was applied on top of the combinatorial film to improve the electrical contact. A platinum wire and a gold-plated tip were then used to connect to the bottom and top parts, respectively. To electrically isolate individual spots of the films, the thin films were scribed into 3 × 3 mm^2 squared samples with a custom-made diamond cutter coupled to an automatized robocasting equipment. The films were immobilized on a hot plate with a stainless-steel metallic ring and clamps together with alumina felt insulators to improve the homogeneity of the thermal distribution. EIS measurements were performed with a *Novocontrol* impedance spectrometer in the 1 MHz–0.2 Hz frequency range, at open circuit potential and with 50 mV of amplitude. A total of 50 frequency points were measured in each impedance spectra. The wafers were kept under constant temperature for at least 15 min. prior to the measurement, and several impedance spectra were measured in each sample to confirm the absence of thermal drift. The impedance spectra were fitted with a modified Jamnik-Maier equivalent circuit model, allowing the extraction of the series resistance (R_s) and the area-specific polarization resistance (ASRp). The activation energy E_a and preexponential factor ASR_0 of a reference YSZ substrate were calculated by fitting the Arrhenius equation ($ASR = ASR_0 \cdot T^{-1} \cdot exp(E_a \cdot k^{-1} \cdot T^{-1})$, where T and k are the temperature and the Boltzmann constant) to temperature-dependance measurements in a *ProboStat* furnace setup with precise temperature control. The real temperature of each spot in the wafer was obtained by measuring the ASR of the reference YSZ in the same probe station as LSMCF at different temperatures. Once the calculated temperature distribution was mapped along the wafer at each temperature set, a linear fit was obtained (T_{real} vs T_{set}) to correct for artifacts coming from parasitic resistances in the setup. With this analysis, the preexponential factors and activation energies across the LSCMCF film were extracted from fitting the Arrhenius equation.

Machine Learning Modelling: Before the modeling, the data from the different techniques was merged based on either the wafer coordinates if the data corresponding to the same batch (Table S1, Supporting Information) or the stoichiometry. The weighted average of each property in the closest neighbors was calculated, with the weights being inversely proportional to the sum of the absolute differences in stoichiometry or wafer coordinates. The datasets used in all the models are built based on wafer coordinates measured by ellipsometry unless otherwise stated.

To train and test the machine learning models, the dataset was split using train_test_split from scikit-learn, saving 10% as test set and with the random state set to 42. The prediction performance of the models was evaluated in two main ways: through 5-fold cross-validation on the training set (289 datapoints), and against the test set (32 datapoints). The optimal hyperparameters for each machine learning model were identified using a Bayesian optimization approach (*BayesSearchCV* from *scikit-optimize*) on the training set, with a five-fold cross-validation and the random state equal to 42. The supervised machine learning models tested were:*KNeighborsRegressor*, *Lasso*, *KernelRidge*, *SVR*, *RandomForestRegressor*, *MLPRegressor*, and *GaussianProcessRegressor* (all from *scikit-learn*). The hyperparameters optimized in each method are listed in Table S3 (Supporting Information). The input data used in the method screening (Figure S12, Supporting Information) were scaled using the *MinMaxScaler* from *scikit-learn*; no scaling was applied during the rest of the modeling. The performance scores in the training set and their uncertainty were calculated as the average score across the five cross-validation folds and the standard deviation, respectively. Finally, principal component analysis was carried out with *PCA* from *scikit-learn*, while the correlation matrix was calculated using the functionality.*Dataframe.corr* from *pandas*.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

C.B.-G. acknowledges funding from a Marie Skłodowska Curie Actions Postdoctoral Fellowship grant (101064374). This project received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 824072 (HARVESTORE) and No. 101017709 (EPISTORE) and under the Marie Skłodowska-Curie grant agreement No 840787 (Thin-CATALYzER). C.F.C. acknowledges funding from a Marie Skłodowska Curie Actions Postdoctoral Fellowship grant (101107093). The authors acknowledge support from the Generalitat de Catalunya (2021-SGR-00750, NANOEN) and the French National Research Agency (ANR) as part of the AAPG2021 – CES50 call (ANR-20-CE05-0001). C.C. and S.P.-G. acknowledge that this material is based upon work supported by the U.S. Department of Energy, Office of Science, Subaward by ′University of Minnesota, Project title: Development of Machine Learning and Molecular Simulation Approaches to Accelerate the Discovery of Porous Materials for Energy-Relevant Applications under Award Number DE-SC0023454. AAG. thanks Anders G. Frøseth for his generous support. AAG also acknowledges the generous support from the Acceleration Consortium, the Natural Resources Canada, and the Canada 150 Research Chairs program.

Conflict of Interest

The authors declare no conflict of interest.

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

RIGHTSLINK

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Author Contributions

C.B.-G., and J.S. contributed equally to this work. Conceptualization was done by C.B., J.S., G.D., J.K., F.B., A.A., A.T. Investigation was carried out by C.B., J.S., A.T. Experimental work was carried out by J.S. (sample preparation, XRF, Raman, Ellipsometry, EDX, EIS), G.C. (XRF, XRD), S.F. (IEDP-SIMS), L.L. (sample preparation, Raman, EDX, EIS), M.N. (EIS), A.C. (IEDP-SIMS), F.B. (XRF), and G.D. (XRF). Experimental data analysis was performed by J.S., G.C., and L.L. Machine learning analysis was performed by C.B., S.P., C.C. Writing of original draft was carried out by C.B., J.S., A.T. Writing, editing and discussion was performed by S.P., F.C., C.C., A.A., G.D., J.K., A.M., F.B., and A.T. Funds were acquired by C.B., G.D., J.K., F.B., A.A., and A.T.

Data Availability Statement

The data that support the findings of this study are openly available in Github at https://nanoionicshub.github.io/LSMCF_database/, reference number 0.

Keywords

high entropy oxides, high-throughput experimentation, machine learning, perovskite oxides, solid oxide fuel cells

Received: May 23, 2024
Revised: September 9, 2024
Published online:

[1] D. J. Brett, A. Atkinson, N. P. Brandon, S. J. Skinner, *Chem. Soc. Rev.* **2008**, *37*, 1568.
[2] E. D. Wachsman, K. T. Lee, *Science* **2011**, *334*, 935.
[3] S. A. Kumar, P. Kuppusami, In *Intermediate Temperature Solid Oxide Fuel Cells*, (Ed: G. Kaur), Elsevier, Amsterdam **2020**, pp 113–163.
[4] S. J. Skinner, *Int. J. Inorg. Mater.* **2001**, *3*, 113.
[5] S. P. Jiang, *J. Power Sources* **2003**, *124*, 390.
[6] F. M. Chiabrera, F. Baiutti, J. M. Börgers, G. F. Harrington, L. Yedra, M. O. Liedke, J. Kler, P. Nandi, J. Sirvent, D. de, J. Santiso, M. López-Haro, J. J. Calvino, S. Estradé, M. Butterling, A. Wagner, F. Peiró, R. A. De Souza, A. Tarancón, *J. Phys. Energy* **2022**, *4*, 044011.
[7] Y.-L. Lee, J. Kleis, J. Rossmeisl, Y. Shao-Horn, D. Morgan, *Energy Environ. Sci.* **2011**, *4*, 3966.
[8] Z. Gao, L. V. Mogni, E. C. Miller, J. G. Railsback, S. A. Barnett, *Energy Environ. Sci.* **2016**, *9*, 1602.
[9] Y. Takeda, R. Kanno, M. Noda, O. Yamamoto, *Bull. Inst. Chem. Res., Kyoto Univ.* **1986**, *64*, 157.
[10] K. Wincewicz, J. Cooper, *J. Power Sources* **2005**, *140*, 280.
[11] S. B. Adler, J. A. Lane, B. C. Steele, *J. Electrochem. Soc.* **1996**, *143*, 3554.
[12] R. Merkle, J. Maier, H. J. Bouwmeester, *Angew. Chem., Int. Ed.* **2004**, *43*, 5069.
[13] M. Pavone, A. M. Ritzmann, E. A. Carter, *Energy Environ. Sci.* **2011**, *4*, 4933.
[14] L. Giordano, K. Akkiraju, R. Jacobs, D. Vivona, D. Morgan, Y. Shao-Horn, *Acc. Chem. Res.* **2022**, *55*, 298.
[15] W. T. Hong, R. E. Welsch, Y. Shao-Horn, *J. Phys. Chem. C* **2016**, *120*, 78.
[16] Z. Li, L. E. K. Achenie, H. Xin, *ACS Catal.* **2020**, *10*, 4377.
[17] A. Ludwig, *npj Comput. Mater.* **2019**, *5*, 1.
[18] T. A. A. Batchelor, T. Löffler, B. Xiao, O. A. Krysiak, V. Strotkötter, J. K. Pedersen, C. M. Clausen, A. Savan, Y. Li, W. Schuhmann, J. Rossmeisl, A. Ludwig, *Angew. Chem., Int. Ed.* **2021**, *60*, 6932.
[19] Z. W. Ulissi, M. T. Tang, J. Xiao, X. Liu, D. A. Torelli, M. Karamad, K. Cummins, C. Hahn, N. S. Lewis, T. F. Jaramillo, K. Chan, J. K. Nørskov, *ACS Catal.* **2017**, *7*, 6600.
[20] K. Lin, R. Gómez-Bombarelli, E. S. Beh, L. Tong, Q. Chen, A. Valle, A. Aspuru-Guzik, M. J. Aziz, R. G. Gordon, *Nat. Energy* **2016**, *1*, 16102.
[21] J. Hachmann, R. Olivares-Amaya, A. Jinich, A. L. Appleton, M. A. Blood-Forsythe, L. R. Seress, C. Román-Salgado, K. Trepte, S. Atahan-Evrenk, S. Er, S. Shrestha, R. Mondal, A. Sokolov, Z. Bao, *Energy Environ. Sci.* **2014**, *7*, 698.
[22] J. C. H. Rossiny, J. Julis, S. Fearn, J. A. Kilner, Y. Zhang, L. Chen, S. Yang, J. R. G. Evans, *Solid State Ion.* **2008**, *179*, 1085.
[23] H. Anderson, *Solid State Ionics* **1992**, *52*, 33.
[24] A. Mai, V. A. Haanappel, S. Uhlenbruck, F. Tietz, D. Stöver, *Solid State Ionics* **2005**, *176*, 1341.
[25] J. Rossiny, S. Fearn, J. A. Kilner, Y. Zhang, L. Chen, S. Yang, J. R. Evans, T. Zhang, K. Yates, L. F. Cohen, *ECS Trans.* **2007**, *7*, 1005.
[26] R. E. Usiskin, S. Maruyama, C. J. Kucharczyk, I. Takeuchi, S. M. Haile, *J. Mater. Chem. A* **2015**, *3*, 19330.
[27] D. Sari, F. Piskin, Z. C. Torunoglu, B. Yasar, Y. E. Kalay, T. Ozturk, *Solid State Ion.* **2018**, *326*, 124.
[28] A. M. Saranya, A. Morata, D. Pla, M. Burriel, F. Chiabrera, I. Garbayo, A. Hornes, J. A. Kilner, A. Tarancón, *Chem. Mater.* **2018**, *30*, 5621.
[29] F. Chiabrera, I. Garbayo, L. López-Conesa, G. Martín, A. Ruiz-Caridad, M. Walls, L. Ruiz-González, A. Kordatos, M. Núñez, A. Morata, S. Estradé, A. Chroneos, F. Peiró, A. Tarancón, *Adv. Mater.* **2019**, *31*, 1805360.
[30] M. L. Green, C. L. Choi, J. R. Hattrick-Simpers, A. M. Joshi, I. Takeuchi, S. C. Barron, E. Campo, T. Chiang, S. Empedocles, J. M. Gregoire, A. G. Kusne, J. Martin, A. Mehta, K. Persson, Z. Trautt, J. Van Duren, A. Zakutayev, *Appl. Phys. Rev.* **2017**, *4*, 011105.
[31] A. Corma, J. Serra, P. Serna, M. Moliner, *J. Catal.* **2005**, *232*, 335.
[32] L. A. Baumes, P. Serna, A. Corma, *Appl. Catal., A* **2010**, *381*, 197.
[33] M. C. Papac, J. Huang, A. Zakutayev, R. O'Hayre, *J. Mater. Chem. A* **2023**, *11*, 5267.
[34] J. Huang, M. Papac, R. O'Hayre, *Electrochim. Acta* **2021**, *367*, 137493.
[35] M. C. Papac, K. R. Talley, R. O'Hayre, A. Zakutayev, *Rev. Sci. Instrum.* **2021**, *92*, 065105.
[36] S. Langner, F. Häse, J. D. Perea, T. Stubhan, J. Hauch, L. M. Roch, T. Heumueller, A. Aspuru-Guzik, C. J. Brabec, *Adv. Materi.* **2020**, *32*, 1907801.
[37] J. M. Serra, V. B. Vert, *Catal. Today* **2011**, *159*, 47.
[38] S. Chen, Y. Hou, H. Chen, X. Tang, S. Langner, N. Li, T. Stubhan, I. Levchuk, E. Gu, A. Osvet, C. J. Brabec, *Adv. Energy Mater.* **2018**, *8*, 1701543.
[39] B. Rohr, H. S. Stein, D. Guevarra, Y. Wang, J. A. Haber, M. Aykol, S. K. Suram, J. M. Gregoire, *Chem. Sci.* **2020**, *11*, 2696.
[40] A. Kafizas, I. P. Parkin, *J. Am. Chem. Soc.* **2011**, *133*, 20458.
[41] D. Seley, K. Ayers, B. A. Parkinson, *ACS Comb. Sci.* **2013**, *15*, 82.
[42] J. E. Katz, T. R. Gingrich, E. A. Santori, N. S. Lewis, *Energy Environ. Sci.* **2008**, *2*, 103.
[43] A. Corma, J. M. Serra, E. Argente, V. Botti, S. Valero, *Chem. Phys. Chem.* **2002**, *3*, 939.
[44] A. M. Saranya, A. Morata, D. Pla, M. Burriel, F. Chiabrera, I. Garbayo, A. Hornés, J. A. Kilner, A. Tarancón, *Chem. Mater.* **2018**, *30*, 5621.
[45] M. Kubicek, G. M. Rupp, S. Huber, A. Penn, A. K. Opitz, J. Bernardi, M. Stöger-Pollach, H. Hutter, J. Fleig, *Phys. Chem. Chem. Phys.* **2014**, *16*, 2715.
[46] Z. Cai, Y. Kuru, J. W. Han, Y. Chen, B. Yildiz, *J. Am. Chem. Soc.* **2011**, *133*, 17696.
[47] E. Navickas, T. M. Huber, Y. Chen, W. Hetaba, G. Holzlechner, G. Rupp, M. Stöger-Pollach, G. Friedbacher, H. Hutter, B. Yildiz, J. Fleig, *Phys. Chem. Chem. Phys.* **2015**, *17*, 7659.
[48] F. Chiabrera, I. Garbayo, D. Pla, M. Burriel, F. Wilhelm, A. Rogalev, M. Núñez, A. Morata, A. Tarancón, *APL Mater.* **2019**, *7*, 013205.

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

RIGHTSLINK

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[49] Y. Tang, F. Chiabrera, A. Morata, I. Garbayo, N. Alayo, A. Tarancón, *Adv. Mater. Interfaces* **2021**, *8*, 2001881.
[50] M. D. Scafetta, Y. J. Xie, M. Torres, J. E. Spanier, S. J. May, *Appl. Phys. Lett.* **2013**, *102*, 081904.
[51] M. N. Iliev, M. V. Abrashev, *J. Raman Spectrosc.* **2001**, *32*, 805.
[52] V. B. Podobedov, A. Weber, D. B. Romero, J. P. Rice, H. D. Drew, *Solid State Commun.* **1998**, *105*, 589.
[53] A. Ishikawa, J. Nohara, S. Sugai, *Phys. Rev. Lett.* **2004**, *93*, 136401.
[54] N. O. Golosova, D. P. Kozlenko, A. I. Kolesnikov, V. Y. Kazimirov, M. B. Smirnov, Z. Jirák, B. N. Savenko, *Phys. Rev. B* **2011**, *83*, 214305.
[55] N. V. R Minh, *J. Phys.: Conf. Ser.* **2009**, *187*, 012011.
[56] A. G. Souza Filho, J. L. Faria, I. Guedes, J. M. Sasaki, P. T. Freire, V. N. Freire, M. Filho, M. M. Xavier, F. A. Cabral, J. H. de Araújo, J. A. da Costa, *Phys. Rev. B* **2003**, *67*, 052405.
[57] L. V. Nomerovannaya, A. A. Makhnëv, A. Y. Rumyantsev, *Phys. Solid State* **1999**, *41*, 1322.
[58] Z. Sha, E. Cali, Z. Shen, E. Ware, G. Kerherve, S. J. Skinner, *Chem. Mater.* **2021**, *33*, 8469.
[59] F. Baiutti, F. Chiabrera, D. Diercks, A. Cavallaro, L. Yedra, L. López-Conesa, S. Estradé, F. Peiró, A. Morata, A. Aguadero, A. Tarancón, *Adv. Mater.* **2021**, *33*, 2105622.
[60] E. Navickas, Y. Chen, Q. Lu, W. Wallisch, T. M. Huber, J. Bernardi, M. Stöger-Pollach, G. Friedbacher, H. Hutter, B. Yildiz, J. Fleig, *ACS Nano* **2017**, *11*, 11475.
[61] A. Tarancón, S. J. Skinner, R. J. Chater, F. Hernández-Ramírez, J. A. Kilner, *J. Mater. Chem.* **2007**, *17*, 3175.
[62] H. J. Bouwmeester, A. J. Burggraaf, In *Fundamentals of Inorganic Membrane Science and Technology*, Elsevier, Amsterdam **1996**, Vol. *4*, pp 435–528.
[63] A. M. Saranya, D. Pla, A. Morata, A. Cavallaro, J. Canales-Vázquez, J. A. Kilner, M. Burriel, A. Tarancón, *Adv. Energy Mater.* **2015**, *5*, 1500377.
[64] V. Thoreton, M. Niania, J. Druce, H. Tellez, J. Kilner, *J. Electrochem. Soc.* **2022**, *169*, 044513.
[65] F. Baiutti, F. Chiabrera, M. Acosta, D. Diercks, D. Parfitt, J. Santiso, X. Wang, A. Cavallaro, A. Morata, H. Wang, A. Chroneos, J. MacManus-Driscoll, A. Tarancón, *Nat. Commun.* **2021**, *12*, 2660.
[66] J. Sirvent, D. de, A. Carmona, L. Rapenne, F. Chiabrera, A. Morata, M. Burriel, F. Baiutti, A. Tarancón, *ACS Appl. Mater. Interfaces* **2022**, *14*, 42178.
[67] F. Chiabrera, A. Morata, M. Pacios, A. Tarancón, *Solid State Ion.* **2017**, *299*, 70.
[68] H. M. Cartwright, *Machine Learning in Chemistry: The Impact of Artificial Intelligence*, Royal Society of Chemistry, London, Burlington House **2020**.
[69] S. Guido, A. C. Müller, *Introduction to Machine Learning with Python: A Guide for Data Scientists*, O'Reilly Media, Sebastopol, CA **2016**.
[70] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media, Sebastopol, CA **2020**.
[71] T. Hastie, R. Tibshirani, J. Friedman, In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, (Eds: T. Hastie, R. Tibshirani, J. Friedman), Springer Series in Statistics, Springer, New York, NY **2009**, pp 219–259.
[72] Z. Feng, Y. Yacoby, M. J. Gadre, Y.-L. Lee, W. T. Hong, H. Zhou, M. D. Biegalski, H. M. Christen, S. B. Adler, D. Morgan, Y. Shao-Horn, *J. Phys. Chem. Lett.* **2014**, *5*, 1027.
[73] B. Koo, K. Kim, J. K. Kim, H. Kwon, J. W. Han, W. Jung, *Joule* **2018**, *2*, 1476.
[74] D. Kim, R. Bliem, F. Hess, J.-J. Gallet, B. Yildiz, *J. Am. Chem. Soc.* **2020**, *142*, 3548.
[75] N. M. Astik, H. Soni, P. K. Jha, V. Sathe, *Phys. B* **2018**, *541*, 103.
[76] A. Dubroka, J. Humlíček, M. V. Abrashev, Z. V. Popović, F. Sapiña, A. Cantarero, *Phys. Rev. B* **2006**, *73*, 224401.
[77] M. Bouzayen, R. Dhahri, A. Benali, S. Chaabouni, K. Khirouni, B. F. Costa, *J. Mater. Sci.: Mater. Electron.* **2021**, *32*, 13000.
[78] F. B. Abdallah, A. Benali, M. Triki, E. Dhahri, K. Nomenyo, G. Lerondel, *J. Mater. Sci: Mater. Electron.* **2019**, *30*, 3349.
[79] W. Rativa-Parada, J. A. Gómez-Cuaspud, M. Schmal, A. F. Cruz-Pacheco, E. Vera-López, *J. Aust. Ceram Soc.* **2021**, *57*, 767.

© 2024 The Author(s). Advanced Materials published by Wiley-VCH GmbH

15214095, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/adma.202407372 by CochraneItalia, Wiley Online Library on [05/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License