

Cite this: *J. Mater. Chem. C*, 2020, **8**, 10352

Interactive human–machine learning framework for modelling of ferroelectric–dielectric composites†

Ning Liu,[‡]^a Achintha Ihalage,[‡]^a Hangfeng Zhang,^b Henry Giddens,[‡]^a Haixue Yan[‡]^b and Yang Hao[‡]^{*a}

Data driven materials discovery and optimization require databases that are error free and experimentally verified. Performing material measurements is time-consuming and often restricted by the fact that material sample preparations are non-trivial, labour-intensive and expensive. Numerical modelling of materials has been studied over the years in order to address these issues and nowadays it has been developed at multi-scale and multi-physics levels. However, numerical models for nano-composites, especially for ferroelectrics, are limited due to multiple unknowns including oxygen vacancy densities, grain sizes and domain boundaries existing in the system. In this work, we introduce a human–machine interactive learning framework by developing a scalable semi-empirical model to accurately predict material properties enabled by deep learning (DL). MgO-Doped BST ($\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$) is selected as an example ferroelectric–dielectric composite for validation. The DL model transfer-learns the experimental features of materials from a measurement database which includes data for over 100 different ferroelectric composites collected by screening the published data and combining our own measurement data. The trained DL model is utilized in providing feedback to human researchers, who then refine computer model parameters accordingly, hence completing the interactive learning cycle. Finally, the developed DL model is applied to predict and optimise new ferroelectric–dielectric composites with the highest figure of merit (FOM) value.

Received 6th November 2019,
Accepted 18th February 2020

DOI: 10.1039/c9tc06073a

rsc.li/materials-c

1 Introduction

Material modelling is a key pre-design strategy used to eliminate trial-and-error loops in the new materials development process. A range of modelling approaches have been proposed for atomistic and theoretical modelling of materials, such as density functional theory (DFT),^{1,2} molecular dynamics (MD),^{3,4} Monte Carlo method,^{5,6} semi-empirical physical models^{7,8} and finite element models.^{9,10} Modelling of spontaneous polarization, dielectric properties, ferroelectric to paraelectric phase transition taking effect at the Curie point (T_c) and structural properties of ferroelectric materials has been extensively scrutinized with both atomic level and numerical simulations.^{11–13} With the advancement of material characterization techniques in recent years, some of the specific behaviours such as domain-wall motion, defects causing the pinning effect, negative capacitance and the

presence of local dipole components in the paraelectric region have been experimentally identified and have been considered in the theoretical models.^{14–18} However, the first principle calculations are computationally expensive and theoretical modelling requires a profound knowledge of the physical phenomena of the material. Hence, machine learning (ML) is increasingly being used to effectively bypass these calculations.^{19–21} Thus, these ML frameworks are used to predict material properties^{22,23} and to design and discover novel materials.^{24,25} The integration of ML with theoretical models, despite being useful for materials discovery and optimization, still remains a less-explored research field.

The real challenge is therefore to develop models that comply well with measurement data.²⁶ In this work, we propose a human–machine interactive learning framework where a semi-empirical model of ferroelectrics is improved to model ferroelectric–dielectric composites by feeding ‘machine-learned’ experimental features in order to significantly boost the modelling accuracy. The proposed framework is embraced by the power of inherent natures of learning abilities where ML algorithms are good at rapid learning from mass data while capturing even the slightest variation, and humans are empowered by their analytical knowledge to anticipate new scenarios by abstracting different domains.

To demonstrate the concept as shown in Fig. 1, we select a semi-empirical model for ferroelectrics, known as the Vendik

^a School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS, UK. E-mail: y.hao@qmul.ac.uk

^b School of Engineering and Materials Science, Queen Mary University of London, E1 4NS, UK

† Electronic supplementary information (ESI) available: Full model derivation and analysis. Simulated and measurement materials databases are also provided.

See DOI: 10.1039/c9tc06073a

‡ These authors contributed equally to this work.



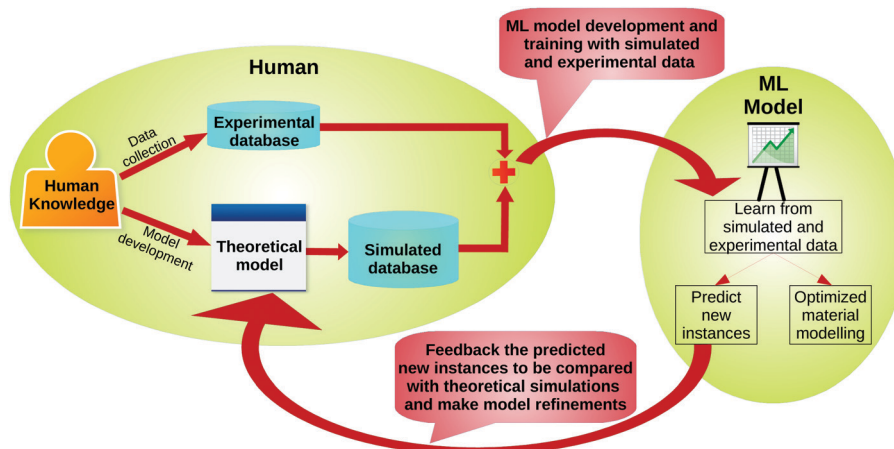


Fig. 1 Human–machine interactive learning framework. Human develops a theoretical model based on the knowledge. The developed model is used to produce a big database which enables “materials by design”. An experimental dataset is also created by screening those from published literature and combining those with our own measurements. A machine learning model successively learns from these two databases and predicts new instances. Predicted results are compared with theoretical calculations and further refinements are then introduced to the theoretical model such that final numerical results reflect the actual behaviour of materials.

model developed from Landau–Ginzburg theory^{7,8,27–29} and $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ (BST) as a modelling example. The Vendik model allows us to calculate the dielectric properties of both incipient and displacive types of ferroelectrics as a function of temperature, biasing field, frequency and material defects. However, by comparing our initial calculations with our measurement data and those from ref. 30 on BST ceramics, we have concluded that the model overestimates dielectric constants at high frequencies (over 1 GHz). Furthermore, the original Vendik model is not designed for ferroelectric–dielectric composites, which are commonly synthesised by research scientists to modify material parameters such as dielectric constant, loss tangent and tunability.

In the present study, we first theoretically develop an improved Vendik model which is valid for high frequencies as well as ferroelectric–dielectric composites such as MgO-doped BST ceramics by considering the new mechanisms brought by the dielectric dopant and improving the analytical equations to reflect the doping effect. Deep learning (DL), which is a branch of ML, will then be used to successively learn from a simulated database and an experimental database. We will receive feedback from the trained DL model and will subsequently make appropriate refinements to the semi-empirical model parameters such that the resulting simulation data adhere to those from measurements. A fully connected deep neural network (DNN) architecture is proposed to avoid the phenomenon known as catastrophic forgetting,³¹ frequently occurring in the context of transfer learning. The tendency of the neural network to forget what it had learned previously upon learning new information is known as catastrophic forgetting. Learning from two databases can be mapped into a transfer learning task, where the same ML model trained on the simulated database is re-purposed by the subsequent training with the experimental database. The refined theoretical model following the above interactive learning process is experimentally validated with different $\text{Ba}_{0.6}\text{Sr}_{0.4}\text{TiO}_3$ (BST64) samples and the trained DL

model is utilized for optimized material modelling to explore the required conditions for designing highly tunable, low loss ferroelectrics operating in the paraelectric state. Combined ML and theoretical modelling supported by experiments demonstrates the applicability and scalability of the proposed interactive learning framework, seemingly having a vast scope of applications in materials modelling.

2 Theory and design

2.1 Original Vendik model

An analytic equation to calculate the complex dielectric constant of ferroelectric materials under different temperatures and electric fields for both ferroelectric and paraelectric states was proposed by Vendik *et al.*^{7,8} The equations are derived based on the conventional Landau theory and four energy dissipation mechanisms are considered in the derivation (see Appendix, Section A.1 for full model derivation in the ESI†). Thus, the proposed equation to calculate the complex permittivity of $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ ceramics can be formulated as

$$\varepsilon(E, T, f, x, \xi_s) = \frac{\varepsilon_{00}(x)}{G(E, T, x, \xi_s)^{-1} + \sum_{q=1}^4 \Gamma_q(E, T, f, x, \xi_s)} \quad (1)$$

where $G(E, T, x, \xi_s)$ is the real part of the Green function for a dielectric response of the ferroelectric, x is the barium proportion, T is the temperature and f is the operating frequency of biasing field E . $\varepsilon_{00}(x)$ is an analogue of Curie–Weiss constant C and can be represented as $\varepsilon_{00} = C/T_c$. ξ_s is the statistical dispersion of the biasing field (also known as the defect factor) which reflects the ‘quality’ of the material and corresponds to defects (including oxygen vacancies and inhomogeneity) in the material. $\Gamma_{1,2,3,4}$ refer to the four energy dissipation (loss) mechanisms considered in the original model (see Appendix, Section A.1 for a detailed discussion



on ξ_s and $\Gamma_{1,2,3,4}$ in the ESI†). Therefore, the dielectric loss, *i.e.* the loss tangent of a ferroelectric material, could be expressed as

$$\tan \delta = \frac{\text{Im}[\varepsilon(E, T, f, x, \xi_s)]}{\text{Re}[\varepsilon(E, T, f, x, \xi_s)]} \quad (2)$$

All model constants in the current model are set to be identical with those of the original Vendik model⁸ and the constants referring to BST material are tabulated in Appendix, Section A.3 in the ESI.†

2.2 Improved Vendik model for high frequencies

By comparing simulations from the original model with our measurement data for BST at high radio frequencies (> 1 GHz), it was observed that the original Vendik model needs to be refined to apply for high frequencies. For example, as can be seen from Fig. 2, between 8 GHz and 12 GHz, BST64 displays a permittivity around 50 depending on different synthesis conditions whereas the original model simulations (with defect parameter $\xi_s = 0.8$) yield values in the order of thousands. This inaccuracy can be attributed to the fact that the contribution of the reduced polarization at high frequencies in ferroelectrics was not considered in the original model.

For a typical dielectric placed in an electric field E between two flat electrodes, the relationship between the internal polarization P_{int} and dielectric constant ε_r is described as

$$\varepsilon_r = 1 + \frac{P_{\text{int}}}{\varepsilon_0 E} \quad (3)$$

where ε_0 represents the permittivity of free space. This basic equation tells that, under such condition, ε_r is positively proportional to internal polarization inside the dielectric. Therefore, at high frequencies, the permittivity of BST will be much reduced as dipoles are unable to respond to the changing directions of the alternating field, whereas at intermediate frequencies, the dipoles can partially reorient with the change of the alternating field direction, but will increasingly lag behind as the frequency increases.³² Specifically at the ferroelectric phase, polarization in

BST material will be reduced with increasing frequency of biasing voltage as domain wall motion cannot follow the alternating field³³ and/or at cryogenic temperatures where domain wall motion is thermally frozen.³³

Having studied measurement data published in ref. 34 and 35 and our own data on BST64, we expect, for pure BST material, a steady high permittivity at low frequencies (1 kHz–1 MHz) and a dramatic drop near 1 GHz. By considering the firm dependence of the permittivity on frequency, we introduce a modified Vendik model with

$$\varepsilon(E, T, f, x, \xi_s) = \frac{\varepsilon_{00}(x)}{[G(E, T, x, \xi_s)K(f)]^{-1} + \sum_{q=1}^4 \Gamma_q(E, T, f, x, \xi_s)} \quad (4)$$

where $K(f)$ is written as

$$K(f) = k_a \tanh[k_b \ln(f) + k_c] + k_d \quad (5)$$

The constants k_a , k_b , k_c , k_d in eqn (5) relating to the modified model were found by a curve fitting process between our measurement data on BST64. The model constants were found to be -0.442 , 0.490 , -3.2 and 0.453 respectively and the details of curve fitting can be found in Appendix, Section B.2 and Fig. B7 in the ESI.† After the introduction of frequency-dependent factor $K(f)$, calculation results are now comparable with those from measurements, as can be observed in Fig. 2.

2.3 Ferroelectric–dielectric composite modelling

MgO doping is one of the most prevalent methods being used to alter the dielectric properties of BST ceramics, especially because the formed composition still maintains the ABO₃ perovskite structure and the macroscopic ferroelectric behaviour at low concentration of MgO doping. For instance, a single-phase solid solution can be achieved at MgO doping levels up to 5 mol% whereas with increasing MgO concentration, multi-phase composites are obtained at 20 mol% MgO, confirming the emergence of BST–MgO interfaces in the material.³⁶ This will give rise to charged defects at the interfaces corresponding to higher extrinsic losses in the composite especially at low frequencies.³⁷ Based on the above improved model and taking MgO-doped BST as an example, we develop a semi-empirical model for ferroelectric–dielectric ‘composites’, for the sake of convenience, which includes the single-phase BST–MgO composition with low doping levels of MgO as well.

With MgO doping, initially the smaller B-site Ti⁴⁺ ions get replaced by larger Mg²⁺ ions and the oxygen vacancies will increase causing more defects resulting in a reduced dielectric constant of the material. A further increase in MgO content results in A-site Ba²⁺ ions being replaced by Mg²⁺ due to the high oxygen vacancies and this further reduces the dielectric constant.³⁸ Moreover, even at the paraelectric phase, polar regions are still identified in BST material and the movement of these regions will also be restricted by new complexes brought by MgO doping which in turn decreases the dielectric constant in the paraelectric phase. As mentioned, in multi-phase

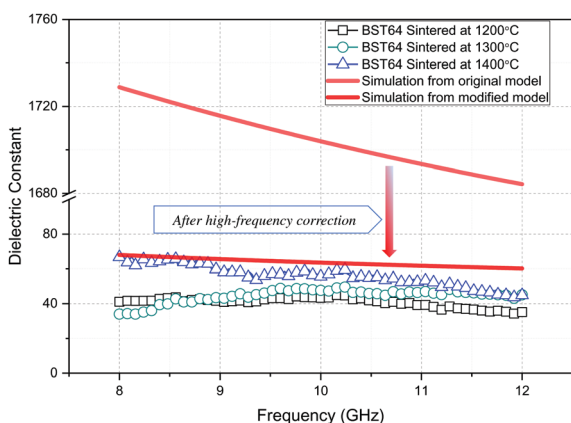


Fig. 2 Experimentally obtained permittivity of BST64 ceramics between 8 and 12 GHz synthesized at different sintering temperatures, measured at room temperature, compared to simulation data from both original and modified models, where ξ_s is set to 0.8.



composites, the BST–MgO interfaces will introduce more charged defects in the material. As a consequence, the value of ζ_s^c related to the defects in the material needs to be elevated. Therefore, another term ζ_{Mg}^c , which is positively correlated with MgO doping content, is added to the previous parameter ζ_s , where the new defect parameter ζ_s^c could be presented as

$$\zeta_s^c = \zeta_s + \zeta_{Mg}^c \quad (6)$$

The replacement of Ti^{4+} ions by Mg^{2+} suppresses the domain-wall motion which in turn reduces the dielectric loss because the loss originated by domain-wall motion is quite significant, especially near T_c . In contrast, the defects caused by oxygen vacancies result in a slight increase of the loss at high temperatures, generally above 350 K for BST–MgO composites.³⁸ Therefore, we introduce a factor K_{Mg} which is dependent on MgO content ζ_{Mg} as

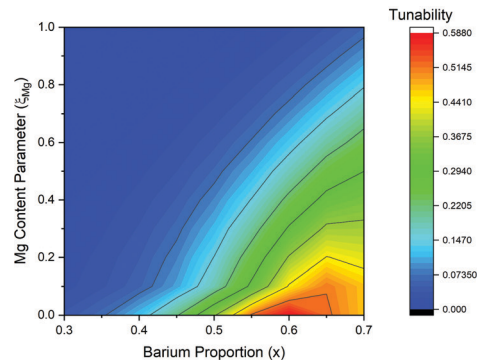
$$K_{Mg} = \exp(-c\zeta_{Mg}^c) \quad (7)$$

In the developed ferroelectric–dielectric composite model, the equations relating to the loss tangent of the material are scaled by a factor of K_{Mg} and the previous defect parameter (ζ_s) is substituted with the new parameter ζ_s^c . The value of c is positive and determines the dependence of K_{Mg} on doping content factor ζ_{Mg}^c and for the sake of convenience, we assume $c = 1$ for the present model. More detailed explanation on analytic equations reflecting the changes brought by MgO doping can be found in Appendix, Section B.2 in the ESI.[†]

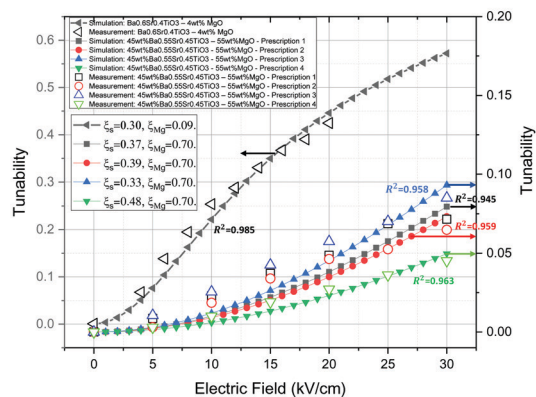
The contour plot in Fig. 3a presents the simulation results of the tunability of BST–MgO composites with various barium proportions and MgO contents as obtained by the developed model. In the plot, with increasing MgO content, we can observe a lower tunability from the composite. Moreover, we fit the measured tunability data of BST–MgO materials (lightly doped BST64 and heavily doped composites of 4 different prescriptions) extracted from ref. 37,39 with our modified theoretical model simulations. For the curve fitting process, the best fitting was obtained with optimum values of defect parameter ζ_s and Mg content parameter ζ_{Mg}^c , as shown in Fig. 3b. For lightly doped BST64 ($Ba_{0.6}Sr_{0.4}TiO_3$ –4 wt% MgO), we have a very low MgO parameter calculated at 0.09. When heavily doped, the BST–MgO composites (45 wt% $Ba_{0.55}Sr_{0.45}TiO_3$ –55 wt% MgO) become much less tunable, smaller than 10% for all four different prescriptions. We assume that the MgO content parameter does not change ($\zeta_{Mg}^c = 0.7$) as the MgO doping level is the same in all prescriptions. Hence, different defect parameters were obtained for each prescription of BST composites respectively, which agrees with the theoretical definition for ζ_s since lattice parameters vary between different prescription approaches (see Appendix, Section A.2 in the ESI[†]).

2.4 Data collection

We created an experimental database of bulk BST composites by screening the published data. This database contains information such as Curie temperature, grain size, dielectric constant at both T_c and room temperature, tunability and loss



(a)



(b)

Fig. 3 (a) Contour plot of room temperature tunability at 20 $kV\ cm^{-1}$ calculated by the modified model, plotting versus barium proportions from 0.3 to 0.7 and MgO content parameter ζ_{Mg} from 0 to 1; (b) simulated tunability curves fitting with measurement data of different BST–MgO composites (extracted from previous literature:^{37,39} $Ba_{0.6}Sr_{0.4}TiO_3$ –4 wt% MgO and 4 different prescriptions of 45 wt% $Ba_{0.55}Sr_{0.45}TiO_3$ –55 wt% MgO), obtained with the highest R -square values respectively. The direction of the arrows pointing to indicates the referred y-axis for each set of data. For each set of simulated curves, defect parameter ζ_s and Mg content parameter ζ_{Mg}^c were calculated respectively.

tangent values at a given biasing field. The data are spanned from 1 kHz to microwave frequencies. We used WebPlotDigitizer⁴⁰ to extract the data from the plots wherever the data are presented in graphical format rather than in numerical format and these data are provided separately. Altogether, this measurement database contains over 1000 data points for over 100 different BST composite materials. As focussed in this work, the majority of data in this database represented BST–MgO composites; however, other compositions such as BST– $MgAl_2O_4$, BST– Mg_2TiO_4 and BST– $MgZrO_3$ were also present. Another database was created from theoretical model simulations. The database comprises tunability and loss tangent values for different barium proportions ($0.3 \leq x \leq 0.9$) of pure BST and MgO-doped BST composites at different ζ_s ($0.2 \leq \zeta_s \leq 0.8$), ζ_{Mg}^c ($0 \leq \zeta_{Mg}^c \leq 0.8$), electric field ($0 \leq E \leq 30\ kV\ cm^{-1}$) and frequency levels ($f \in \{10^i \mid i \in \mathbb{Z}; i \in [3, 10]\}$ Hz). This simulated database contains around 35 000 data points for 35 different BST composites.



3 Results and discussion

3.1 Deep learning model

The underlying requirement for developing a theoretical model for materials discovery is to improve its accuracy by incorporating experimental data. Machine learning stands out as the obvious choice of learning from data and here we propose a fully connected DNN that acts as the interface between the theoretical model and the measurement dataset. As shown in Fig. 4, a deep learning model is firstly developed to learn from the database generated from theoretical calculations. It then learns from the measurement database to reflect the actual behaviour of the BST material. The trained DL model is used to predict new ferroelectric–dielectric composites and the predictions are fed back to the human to make appropriate adjustments to the theoretical model parameters. This section is divided into two parts. In the first phase, we propose a suitable DL architecture that can fully emulate the theoretical model and verify the trained DL model predictions with theoretical simulations. In the second phase, we employ the ML concept of transfer learning to retrain the verified DL model with the small measurement dataset in such a way that it preserves what it had learned from the theoretical model, while learning the experimental features.

3.1.1 Emulation of the theoretical model and verification.

Deep learning architecture is one of the determining factors of the final model performance. It is vital to foresee the ultimate objective of using deep learning in the problem before proposing an architecture. It has become clear that higher tunabilities and lower loss tangents are two of the most preferred characteristics in tunable devices. However, by observing the simulations, we noticed that loss tangent estimation from the theoretical model

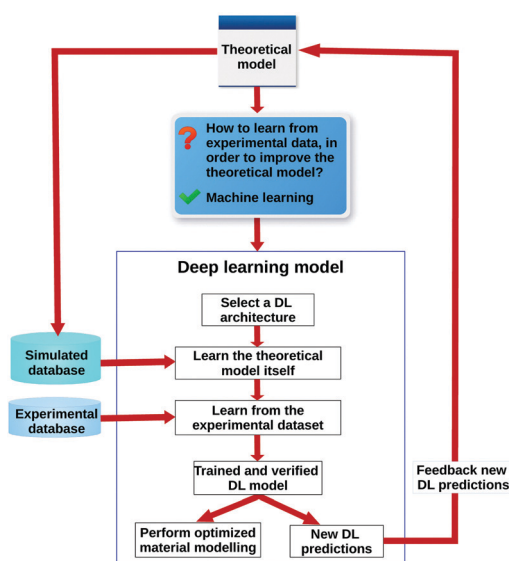


Fig. 4 Deep learning work flow. The DL model first learns the theoretical model itself, followed by the experimental dataset. The accuracy of the trained DL model is quantified and thus it is used to predict new instances that are fed back to the theoretical model where a human could make appropriate adjustments. The DL model is finally used for optimized material modelling.

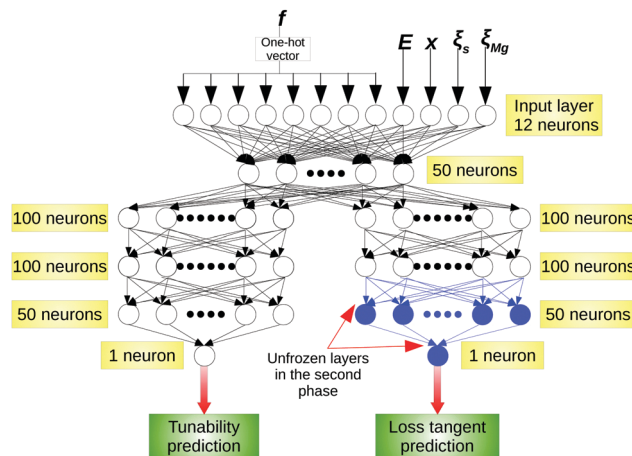


Fig. 5 Deep neural network architecture. The full model is trained on the simulated database in the first phase. Only the blue-shadowed two layers are trained on the measurement database in the second phase to avoid catastrophic forgetting.

could still be improved whereas the tunability estimation matches well with the measurement data. Hence, we envisage the idea of loss tangent improvement and build the DL model upon that.

We propose a fully connected (dense) deep neural network split into two parts that outputs tunability and loss tangent separately, given frequency, electric field, ξ_s , ξ_{Mg} and barium proportion as input features. This architecture makes it possible to use the experimental database to retrain the layers that are associated with loss tangent, without affecting the tunability. The selection criteria of number of layers and neurons in each layer are described in Appendix, Section D.1 (ESI[†]).⁴¹ Fig. 5 shows the fully connected DNN architecture which includes four hidden layers. Tunability prediction and loss tangent prediction share the input layer and the immediate dense layer. Thenceforth, the network branches off into two parts. The exponential linear unit (ELU) is chosen over the rectified linear unit (ReLU) as the activation function to avoid the vanishing gradient problem and speed up the training process.⁴² Linear activation is used at the output layer. As the simulated dataset was generated for discrete frequencies from 1 kHz to 10 GHz, we convert each of these frequencies to a one-hot vector (Appendix, Section D.3 displays the one-hot vector table in the ESI[†]). This process is known as one-hot encoding where categorical data are converted into a group of bits with a single high (1) bit and all others low (0).

The DNN was implemented in python using keras-2.2.4 library with tensorflow-1.13 backend and the training was done on a RTX 2080 Ti GPU with 11GB memory. L_2 regularization was introduced in all layers except the last two layers of the network to prevent overfitting. K -Fold cross validation ($K = 2$) was performed to evaluate the models by having 3 separate databases for training, validation and testing (see Appendix, Section D.2 in the ESI[†]). Total validation loss settled around 1.48×10^{-5} after about 6 hours of training. Table 1 shows training and validation mean squared errors (MSEs) and coefficient of determination (R^2) of tunability and loss tangent predictions separately.



Table 1 DNN training performance

	Training MSE	Validation MSE	R^2 value
Tunability	2.5×10^{-6}	2.8×10^{-6}	0.998
Loss tangent	9.5×10^{-6}	1.18×10^{-5}	0.997
Total	1.2×10^{-5}	1.48×10^{-5}	> 0.99

The R^2 value is a statistical measure that represents the goodness of a fit of a regression model. In order to validate the DL predictions, we predict the tunability and loss tangent of BST composites that are not present in the database and compare them with the simulations. A composition depends on x , ξ_s and ξ_{Mg} and thus, out-of-database predictions consist of compositions with different values for the above quantities that are unseen by the DL model. Fig. 6 demonstrates theoretical model simulation results and DL predictions for different BST–MgO composites at different frequencies. The accuracy of this DL regression model can be numerically expressed with high R^2 values of both tunability and loss tangent predictions and it becomes quite evident that deep learning can perfectly emulate the theoretical model.

3.1.2 Transfer learning with the measurement data. In the second phase, we enable the pre-trained DL model to learn from the measurement dataset. In order for the database to be

compatible for training, we obtained the equivalent ξ_s and ξ_{Mg} parameters for each of the BST composites present in the dataset by comparing measured tunability values with the theoretical model calculations. Table 2 shows the calculated ξ_s and ξ_{Mg} values for selected BST composites. This completed database is then utilized to learn and improve the loss tangent prediction.

The phenomenon termed “catastrophic forgetting” specifically occurs when a pre-trained neural network is trained with another dataset using the gradient descent algorithm as the new weight updates may not reflect previously learned features. In order to address this issue, we freeze all layers except the last two layers of loss tangent prediction as shown in Fig. 5. Once a layer is frozen, it becomes non-trainable and the weights do not update upon training. Therefore, the weights of the layers associated with tunability do not update and hence the neural network will completely remember tunability characteristics learned from the theoretical model. The first three layers associated with loss tangent will remember the behaviour of loss tangent and will enable learning experimental features by training the last two layers on the measurement data.

We first filter out pure BST and only MgO-doped BST composites from the experimental database. The resulting database contains 170 data on 38 materials out of which 11 materials were selected for the test set. Due to the limitation of data, we retrain two unfrozen layers for low number of epochs, following the early stopping method, in order to prevent overfitting. An overfitted neural network performs well on the training set but has a very poor generalization accuracy. Fig. 7 shows the loss tangent prediction results on the test set. It can be observed that in all the cases, DL predictions are closer to the measurement data rather than the simulated values. For quantification purposes, we introduce a similarity score $s(p, q)$, between two sets p, q . We calculate the mean Euclidean distance $d(p, q)$ between set p and set q each having n elements with shape m as

$$d(p, q) = \frac{\sum_{i=1}^n \sqrt{\sum_{j=1}^m (p_{ij} - q_{ij})^2}}{n} \quad (8)$$

Thus, the similarity score, $s(p, q)$, can be introduced as the inverse of the mean Euclidean distance:

$$s(p, q) = \frac{1}{d(p, q)} \quad (9)$$

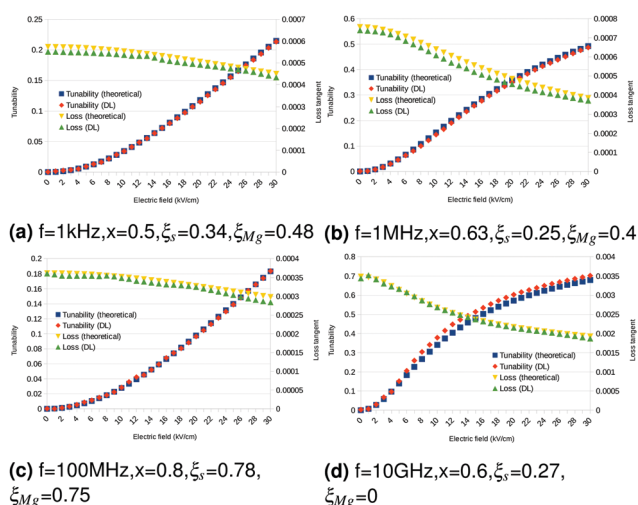


Fig. 6 Theoretical model simulations and deep learning predictions for different BST + MgO composites ((a)–(c)) at different frequencies. Panel (d) represents pure BST64 material.

Table 2 Calculated ξ_s and ξ_{Mg} parameter values for different BST composites. All measurements and simulations are done at 20 kV cm⁻¹ biasing field. ξ_{Mg} increases with MgO concentration for a particular composite

Material	Frequency	Tunability (measurement)	Tunability (theoretical)	Calculated ξ_s	Calculated ξ_{Mg}
Ba _{0.7} Sr _{0.3} TiO ₃ + 2.5 mol% MgO ⁴³	10 kHz	0.34	0.342	0.2	0.48
Ba _{0.7} Sr _{0.3} TiO ₃ + 7.5 mol% MgO ⁴³	10 kHz	0.26	0.26	0.2	0.69
Ba _{0.7} Sr _{0.3} TiO ₃ + 10 mol% MgO ⁴³	10 kHz	0.22	0.224	0.2	0.8
Ba _{0.6} Sr _{0.4} TiO ₃ + 10 wt% MgO ³⁵	1 MHz	0.166	0.165	0.28	0.74
Ba _{0.6} Sr _{0.4} TiO ₃ + 30 wt% MgO ³⁵	1 MHz	0.148	0.146	0.28	0.81
Ba _{0.6} Sr _{0.4} TiO ₃ + 60 wt% MgO ³⁵	1 MHz	0.099	0.097	0.28	1
Ba _{0.45} Sr _{0.55} TiO ₃ ³⁵	10 GHz	0.152	0.145	0.13	0
Ba _{0.5} Sr _{0.5} TiO ₃ ³⁵	10 GHz	0.25	0.25	0.2	0



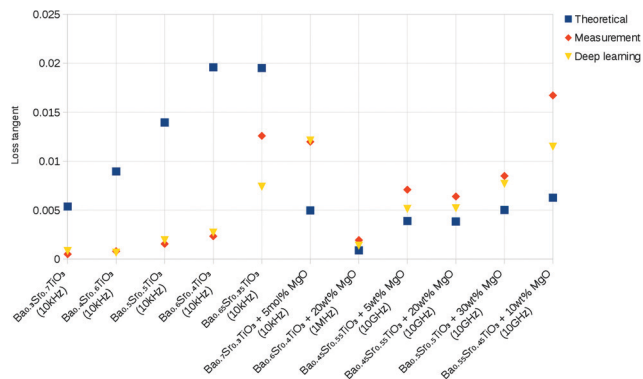


Fig. 7 Comparison of theoretical model simulations, measurements and deep learning predictions on the test set ($E = 20 \text{ kV cm}^{-1}$). It should be noted that the theoretical simulations are carried out after the human-machine interactive learning improvements.

The calculated similarity score between the theoretical simulations and measurements is 142.2, whereas that between the DL predictions and measurements is as high as 675.6. Hence, we can conclude that deep learning offers about 4 times performance improvement in predicting the loss tangent. While we understand that the measurement values can differ significantly depending on the synthesis conditions, it is still essential to develop a model that fits well with the existing data and the proposed DL model shows a better agreement with the experimental measurements.

3.2 Interactive learning framework

New predictions from the ‘transfer-learned’ DL model assist the human to interactively make appropriate adjustments to the theoretical model parameters. As shown in Fig. 8, the interactive learning work flow is a reciprocal process done in two cycles. In step 1, theoretical simulations (tunability and loss tangent) are

carried out to be compared with the experimental data. Then in step 2, a comparison is done by manual inspection and the theoretical model parameters are tweaked heuristically. In the real scenario, by manual comparison with the results from ref. 44 and 45, we found that the calculated loss tangent at low frequencies (1 kHz–100 kHz) is generally much lower estimated. In the original model, the resonance frequency of the low frequency relaxation loss Γ_4 is set to 10 MHz. Therefore, we propose a new low frequency relaxation formula of Γ_5 resonant at $f_5 = 10 \text{ kHz}$ and $\omega_5 = 2\pi f_5$ as

$$\Gamma_5 = A_5 / (1 - i\omega/\omega_5), \quad (10)$$

which corrects the underestimation given by the model. The parameter A_5 is assumed to be equal to the low frequency relaxation parameter of the original model (see Appendix, Section A.3 in the ESI†).

However, the first learning cycle limits us only to the existing data. Therefore, we make use of experimental-data-trained DL model to predict the tunability and loss tangent of new ferroelectric-dielectric composites that are to be compared with the simulations. Theoretical model generated data in step 3 are compared with the DL predictions in step 4 and the model parameters are again tuned to confront with the predictions. In the actual case, we observed that the DL predicted loss tangent at 1 MHz is over 10 times larger than that of the simulated value. This was confirmed by doing further literature review and finding the corresponding experimental value.³⁵ Hence, the subsequent fine-tuning is performed on the theoretical model parameters. In our previous simulations, all model constants were set to be the same as in the original Vendik model in ref. 8 (refer to Appendix, Section A.3 in the ESI†). Since the loss tangent was obviously lower estimated around 1 MHz, we increased the value of coefficient of low-frequency

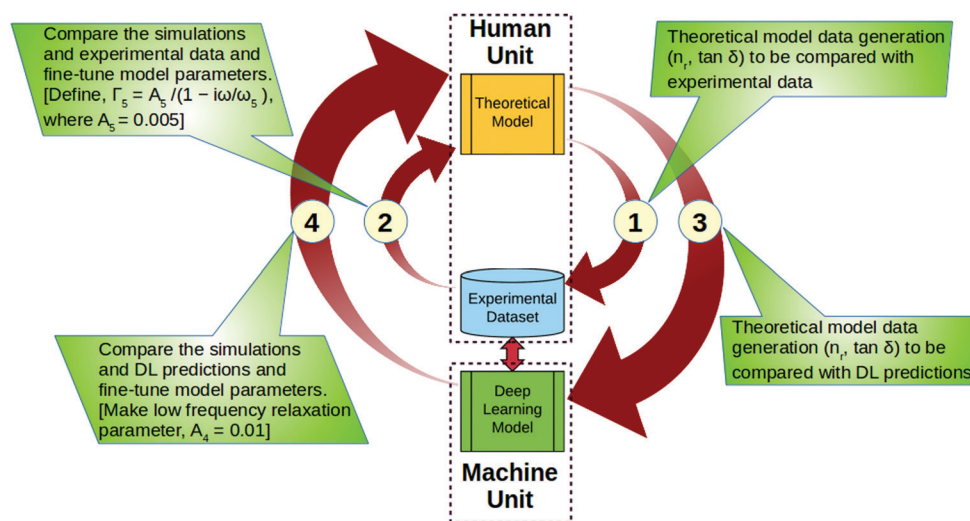


Fig. 8 BST material modelling using human-machine learning interaction. In step 1, theoretical simulations are performed to be compared with the experimental data. In step 2, model parameters are tweaked accordingly by manual inspection and comparison. Step 3 refers to the data generation that are to be compared with machine-generated DL predictions. A human compares these two data and make appropriate refinements to the model parameters in step 4.



loss A_4 . After some heuristic tweaking, here we assign a new value 0.01 for A_4 .

The first learning cycle could be referred to as a manual human learning procedure whereas the second could be identified as a human-machine learning interaction, since the human adjusts model parameters depending on the feedback of a trained DL model. At the end of two learning cycles, the theoretical model has adjusted to the measurement data as well as possible.

3.3 DL optimized materials modelling

We regard an objective function to consider both tunability (n_r) and loss tangent ($\tan \delta$) factors such that optimal material properties can be quantified. Here, we define the figure of merit (FOM/K)^{46,47} factor as

$$K = \frac{n_r}{\tan \delta} \quad (11)$$

Using the trained DL model, we investigate the best FOM values at different frequencies and the corresponding barium proportions and defect parameter values. For the sake of convenience, we perform this for pure BST materials. In the present paper, we investigate FOM at 20 kV cm⁻¹, which is the most frequent value present in our literature data. We range the proportion of barium from 0.5 to 0.7 (assuming the material is at paraelectric state at room temperature when $x \leq 0.7$ ⁴⁸) and the defect factor ζ_s from 0.2 to 0.8. It should be noted that for all practical BST ceramics, there is always some existing defect, and we assume that the initial lowest value of defect factor ζ_s is 0.2.⁴⁹ Several example frequencies were selected from 100 kHz to 10 GHz for the proposed optimization and the temperature is set to be the room temperature at 290 K, at which the dataset was generated. The corresponding combination of x and ζ_s which results in the highest FOM value can theoretically be regarded as the best BST material under the considered frequency and temperature.

However, the predicted loss tangent being too low can result in very high FOM values without revealing much information about the tunability and the overall performance of the material. Hence, while calculating the FOM value from the DL model, we set the minimum threshold of the loss tangent to be 5×10^{-4} , as it is the lowest loss value observed in the experimental dataset.⁵⁰ Table 3 shows the best FOM values and the corresponding x and ζ_s values obtained using the DL predictions at different frequencies for pure BST materials. By observing the DL results, it can be concluded that the best operating frequency for pure BST materials is around 10 MHz as it shows the highest K value among the chosen frequencies. The optimum barium proportion is found to be around 0.63 for most of the frequencies. From Table 3, it can also be noticed

Table 3 Best figure of merit obtained from the DL model at different frequencies at 20 kV cm⁻¹

f	100 kHz	10 MHz	1 GHz	10 GHz
x	0.54	0.63	0.6	0.63
ζ_s	0.68	0.70	0.20	0.20
FOM	273.38	551.49	35.66	14.04

that a low level of defects is preferred in microwave frequencies whereas relatively high defects provide better FOM values in low frequencies.

3.4 Experimental validation

BST64 samples sintered at different temperatures were taken as an example material to be studied (see Appendix, Section C for sample preparation methods in the ESI†). Depending on these different synthesis conditions, the defects may differ for each sample and it is worth investigating whether our model could capture the correct defect parameter ζ_s and the corresponding dielectric properties of these samples. For all the prepared BST64 samples, the dielectric constant was measured from 250 K to 400 K respectively at 100 kHz under a zero external biasing field. Therefore, we perform dielectric constant vs temperature simulations at 100 kHz while keeping $E = 0$ kV cm⁻¹. Fig. 9 shows the best curve fittings between experimental data (circles) and the simulation data (lines) for BST64 synthesized with different sintering methods and temperatures. Through the fitting process, we can obtain unique values of ζ_s for each type of material and the simulation results match quite well with the measurements, especially above the Curie point. We believe that the observed poor fit below the phase transition temperature is due to the simplified calculation of Landau-Ginzburg equations used to derive the Vendik model (see Appendix, Section C in the ESI†). Moreover, microstructures of BST64 pellets under different sintering conditions were investigated by using a FEI Quanta FEG 400 high resolution scanning electron microscope. Table 4 shows the grain sizes of different samples and the calculated ζ_s values. High R^2 values evidence that our model is able to capture the correct dielectric properties of different BST64 materials having different defects. When comparing the values of estimated average grain sizes and defect factors, no clear trend can be concluded as the data size is not enough. However, for samples synthesised by the SPS process, grain sizes are obviously smaller and values of defect factor ζ_s are higher, indicating a greater density of defects in BST material.

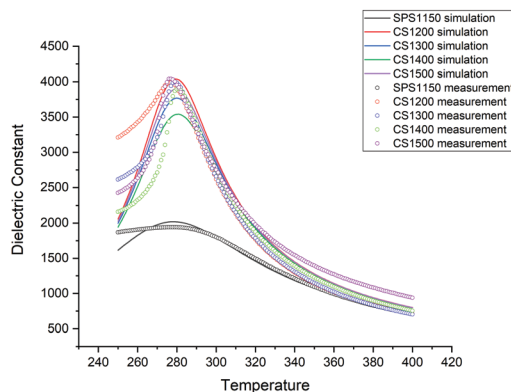


Fig. 9 Best fitting of dielectric constant versus temperature between simulation data (line) and experimental data for pure BST64 materials (at 100 kHz) sintered under different conditions. SPS – spark plasma sintering; CS – conventional sintering.



Table 4 Obtained parameters from simulations for BST64 materials sintered at different temperatures, pressures and time intervals

Synthesis	Calculated ζ_s	Grain size (μm)	R^2 value
SPS (1150/5 m)	0.59	0.5	0.9834
CS (1200/3 h)	0.18	0.8	0.9190
CS (1300/3 h)	0.20	1.2	0.9712
CS (1400/3 h)	0.22	10	0.9728
CS (1500/3 h)	0.19	25	0.9771

4 Conclusions

A new framework of human-machine interactive learning has been developed for accurate modelling of ferroelectric-dielectric composites. By integrating big data generated from a semi-empirical model and the measurement database of sufficient size, we have trained a DL model, which was applied to obtain a refinement of a classical model of ferroelectric materials that can be made to account for multiple unknowns. The model was experimentally validated with BST64 samples synthesised under different sintering conditions and the simulations show a good agreement with the measurements. We believe that this approach has a far reaching implication for application in discovering new material models, especially those analytically unsolvable. As future work, we plan to apply the developed DL model to automate the materials design process as well as perform a thorough analysis on the dependence of less-studied ζ_s and ζ_{Mg} parameters on grain sizes, domain walls and oxygen vacancies in ferroelectric-dielectric composites.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Ms Hanchi Ruan for collecting data. This work is supported in part by the “Software Defined Materials for Dynamic Control of Electromagnetic Waves” (ANIMATE) project (Grant No. EP/R035393/1) and the authors acknowledge Engineering and Physical Sciences Research Council (EPSRC) for providing funding for AOTOMAT (Grant No. EP/P005578/1), TERRA (Grant No. EP/S010009/1), TERALINKS (Grant No. EP/P016421/1) and SYMETA (Grant No. EP/N010493/1). A. I. acknowledges IET AF Harvey Research Prize for funding the PhD studentship.

References

- J. Neugebauer and T. Hickel, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2013, **3**, 438–448.
- J. F. Scott, *npj Comput. Mater.*, 2015, **1**, 15006.
- V. Boddu, F. Endres and P. Steinmann, *Sci. Rep.*, 2017, **7**, 806.
- S. Boyn, J. Grollier, G. Lecerf, B. Xu, N. Locatelli, S. Fusil, S. Girod, C. Carrétero, K. Garcia, S. Xavier, J. Tomas, L. Bellaiche, M. Bibes, A. Barthélémy, S. Saïghi and V. Garcia, *Nat. Commun.*, 2017, **8**, 14736.
- S. Prokhorenko, K. Kalke, Y. Nahas and L. Bellaiche, *npj Comput. Mater.*, 2018, **4**, 80.
- M. Lilienblum, T. Lottermoser, S. Manz, S. M. Selbach, A. Cano and M. Fiebig, *Nat. Phys.*, 2015, **11**, 1070.
- O. G. Vendik and S. P. Zubko, *J. Appl. Phys.*, 1997, **82**, 4475–4483.
- O. G. Vendik, S. P. Zubko and M. A. Nikol'ski, *J. Appl. Phys.*, 2002, **92**, 7448–7452.
- C. Zannoni, R. Mantovani and M. Viceconti, *Med. Eng. Phys.*, 1999, **20**, 735–740.
- J. M. Alison and R. M. Hill, *J. Phys. D: Appl. Phys.*, 1994, **27**, 1291–1299.
- S. R. Phillpot, S. B. Sinnott and A. Asthagiri, *Annu. Rev. Mater. Res.*, 2007, **37**, 239–270.
- Z. Ma, Y. Ma, Z. Chen, F. Zheng, H. Gao, H. Liu and H. Chen, *Ceram. Int.*, 2018, **44**, 4338–4343.
- I. Grinberg, V. R. Cooper and A. M. Rappe, *Nature*, 2002, **419**, 909–911.
- S. Liu, I. Grinberg and A. M. Rappe, *Nature*, 2016, **534**, 360.
- P. S. Bednyakov, B. I. Sturman, T. Sluka, A. K. Tagantsev and P. V. Yudin, *npj Comput. Mater.*, 2018, **4**, 65.
- P. Gao, J. Britson, J. R. Jokisaari, C. T. Nelson, S.-H. Baek, Y. Wang, C.-B. Eom, L.-Q. Chen and X. Pan, *Nat. Commun.*, 2013, **4**, 2791.
- H. W. Park, J. Roh, Y. B. Lee and C. S. Hwang, *Adv. Mater.*, 2019, **31**, 1805266.
- C. Yang, E. Sun, B. Yang and W. Cao, *J. Phys. D: Appl. Phys.*, 2018, **51**, 415303.
- G. H. Gu, J. Noh, I. Kim and Y. Jung, *J. Mater. Chem. A*, 2019, **7**, 17096–17117.
- M. Umehara, H. S. Stein, D. Guevarra, P. F. Newhouse, D. A. Boyd and J. M. Gregoire, *npj Comput. Mater.*, 2019, **5**, 34.
- D. Xue, P. V. Balachandran, J. Hogden, J. Theiler, D. Xue and T. Lookman, *Nat. Commun.*, 2016, **7**, 11241.
- A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi and M. Ceriotti, *Sci. Adv.*, 2017, **3**, e1701816.
- H. S. Stein, D. Guevarra, P. F. Newhouse, E. Soedarmadji and J. M. Gregoire, *Chem. Sci.*, 2019, **10**, 47–55.
- G. Hautier, C. Fischer, V. Ehrlicher, A. Jain and G. Ceder, *Inorg. Chem.*, 2011, **50**, 656–663.
- Y. Liu, T. Zhao, W. Ju and S. Shi, *J. Materiomics*, 2017, **3**, 159–177.
- S. V. Kalinin, B. G. Sumpter and R. K. Archibald, *Nat. Mater.*, 2015, **14**, 973.
- O. G. Vendik and S. P. Zubko, *J. Appl. Phys.*, 2000, **88**, 5343–5350.
- O. G. Vendik, L. T. Ter-Martirosyan and S. P. Zubko, *J. Appl. Phys.*, 1998, **84**, 993–998.
- O. G. Vendik, S. P. Zubko and L. T. Ter-Martirosyan, *Appl. Phys. Lett.*, 1998, **73**, 37–39.
- M. Voigts, W. Menesklou and E. Ivers-Tiffée, *Integr. Ferroelectr.*, 2001, **39**, 383–392.
- M. McCloskey and N. J. Cohen, *Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem*, Psychology of Learning and Motivation, Academic Press, 1989, vol. 24, pp. 109–165.



- 32 S. H. Cha and Y. H. Han, *Jpn. J. Appl. Phys.*, 2006, **45**, 7797–7800.
- 33 M. Acosta, N. Novak, V. Rojas, S. Patel, R. Vaish, J. Koruza, G. A. Rossetti and J. Rödel, *Appl. Phys. Rev.*, 2017, **4**, 041305.
- 34 J. Liu, S. Chen, N. Xi, J. Zhang, L. Wang and Z. Wang, *IEICE Electron. Express*, 2016, **13**, DOI: 10.1587/elex.13.20160713.
- 35 L. C. Sengupta and S. Sengupta, *Mater. Res. Innovations*, 1999, **2**, 278–282.
- 36 M. Cole, P. Joshi, M. Ervin, M. Wood and R. Pfeffer, *Thin Solid Films*, 2000, **374**, 34–41.
- 37 U.-C. Chung, C. Elissalde, M. Maglione, C. Estournès, M. Paté and J. P. Ganne, *Appl. Phys. Lett.*, 2008, **92**, 042902.
- 38 M. Zhang, J. Zhai, B. Shen and X. Yao, *J. Am. Ceram. Soc.*, 2011, **94**, 3883–3888.
- 39 H. Jiang, X.-H. Wang, H.-B. Wang, J. Tao and W.-Z. Lu, *Integr. Ferroelectr.*, 2016, **176**, 275–282.
- 40 A. Rohatgi, *WebPlotDigitizer – Extract data from plots, images, and maps*, 2019, <https://automeris.io/WebPlotDigitizer>.
- 41 M. Hagan, H. Demuth, M. Beale and O. De Jesús, *Neural Network Design*, Martin Hagan, 2014.
- 42 A. Shah, E. Kadam, H. Shah, S. Shinde and S. Shingade, Proceedings of the Third International Symposium on Computer Vision and the Internet, New York, NY, USA, 2016, pp. 59–65.
- 43 R. Laishram, K. C. Singh and C. Prakash, *Ceram. Int.*, 2016, **42**, 14970–14975.
- 44 P.-Z. Ge, X.-G. Tang, Q.-X. Liu, Y.-P. Jiang, W.-H. Li and B. Li, *J. Alloys Compd.*, 2018, **731**, 70–77.
- 45 Q. Xu, X.-F. Zhang, Y.-H. Huang, W. Chen, H.-X. Liu, M. Chen and B.-H. Kim, *J. Phys. Chem. Solids*, 2010, **71**, 1550–1556.
- 46 K. B. Chong, L. B. Kong, L. Chen, L. Yan, C. Y. Tan, T. Yang, C. K. Ong and T. Osipowicz, *J. Appl. Phys.*, 2004, **95**, 1416–1419.
- 47 J.-Y. Ha, L. Lin, D.-Y. Jeong, S.-J. Yoon and J.-W. Choi, *Jpn. J. Appl. Phys.*, 2009, **48**, 011402.
- 48 L. Kong, S. Li, T. Zhang, J. Zhai, F. Boey and J. ma, *Prog. Mater. Sci.*, 2010, **55**, 840–893.
- 49 O. G. Vendik and S. P. Zubko, in *The Oxford Handbook of Innovation*, ed. F. Capolino, CRC Press, 6000 Broken Sound Parkway NW, Suite 300, 2009, ch. 33, pp. 266–290.
- 50 J. Zhang, J. Zhai, X. Chou, J. Shao, X. Lu and X. Yao, *Acta Mater.*, 2009, **57**, 4491–4499.

