Investigation of Opto-Electronic Properties and Stability of Mixed-Cation Mixed-Halide Perovskite Materials with Machine-Learning Implementation

Nicolae Filipoiu 1, Tudor Luca Mitran 1, Dragos Victor Anghel 1,2, Mihaela Florea 3, Ioana Pintilie 3, Andrei Manolescu 4 and George Alexandru Nemnes 1,2,5,*

1 Horia Hulubei National Institute for Physics and Nuclear Engineering, 077126 Magurele, Ilfov, Romania; nicolae.filipoiu@nipne.ro (N.F.); tudor.mitran@nipne.ro (T.L.M.); dragos@theory.nipne.ro (D.V.A.)
2 Research Institute of the University of Bucharest (ICUB), Mihail Kogalniceanu Blvd 36-46, 050107 Bucharest, Romania
3 National Institute of Materials Physics, 077125 Magurele, Ilfov, Romania; mihaela.florea@infim.ro (M.F.); ioana@infim.ro (I.P.)
4 Department of Engineering, Reykjavik University, Menntavegur 1, IS-102 Reykjavik, Iceland; manoles@ru.is
5 Faculty of Physics, Materials and Devices for Electronics and Optoelectronics Research Center, University of Bucharest, 077125 Magurele, Ilfov, Romania
6 Correspondence: nemnes@solid.fizica.unibuc.ro

Abstract: The feasibility of mixed-cation mixed-halogen perovskites of formula $A_xA'_1-x\text{Pb}_yX_{x-y}X'_{3-y-z}$ is analyzed from the perspective of structural stability, opto-electronic properties and possible degradation mechanisms. Using density functional theory (DFT) calculations aided by machine-learning (ML) methods, the structurally stable compositions are further evaluated for the highest absorption and optimal stability. Here, the role of the halogen mixtures is demonstrated in tuning the contrasting trends of optical absorption and stability. Similarly, binary organic cation mixtures are found to significantly influence the degradation, while they have a lesser, but still visible effect on the opto-electronic properties. The combined framework of high-throughput calculations and ML techniques such as the linear regression methods, random forests and artificial neural networks offers the necessary grounds for an efficient exploration of multi-dimensional compositional spaces.

Keywords: mixed-cation; mixed-halide; perovskite; optical absorption; degradation mechanisms; machine-learning techniques

1. Introduction

Halide perovskites are currently regarded as one of the most prominent materials for photovoltaic technology, following the rapid advancement in photoconversion efficiency (PCE) reaching up to 25.2% and due to the low fabrication costs, which makes perovskite solar cells (PSCs) a strong competing candidate for the standard silicon technology [1,2]. Although the rather high PCEs reflect the excellent opto-electronic properties of halide perovskites, there are still stability issues that hamper commercialization. Therefore, a lot of effort is currently dedicated to mitigating the intrinsic and extrinsic degradation processes, without compromising too much the absorption and photoconversion quality of the optically active materials.

The intrinsic degradation mechanisms concern pristine perovskite materials subjected to temperature [3] and illumination [4], while extrinsic degradation mechanisms are determined by foreign species and molecules such as oxygen [5,6] and water [7]. These are typically addressed by chemical engineering, most notably, using mixtures of cations [8] and halogens [9], doping and passivation [10]. The multitude of new organic cations renders a wide variety of low-dimensional perovskite structures, where ion migration is inhibited by the relatively large cations and their embedded hydrophobic moieties can...
reduce the detrimental effect of water molecules. These encouraging results must be weighted with the typical increase in the bandgap, either due to lighter halogens or due to confinement effects in low-dimensional perovskites [11,12], thus reducing the total optical absorption. Therefore, a careful optimization needs to be undertaken, particularly with respect to cation and halogen mixtures, to find the proper compositions, often setting a trade-off between stability and opto-electronic performance.

From the theoretical perspective, the opto-electronic properties have been investigated extensively in numerous ab initio studies [13–16], based on density functional theory (DFT), which routinely provides absorption spectra, defect formation energies and activation energies for ion migration. Despite certain well-known deficiencies of DFT with respect to bandgap estimation, depending on the halide perovskite system, hybrid functionals and other post-DFT methods [17] can provide reasonably accurate results, except with an additional, quite significant, numerical effort. Far fewer investigations concern stability properties, most of them being focused on individual methylammonium lead halides (MAPbX₃), with the most prominent case being the methylammonium lead iodide (MAPI). A decomposition pathway in the case of water exposure was proposed by Frost et al. [18] and the instability in moist environments was also investigated by Tenuta et al. [19]. Thermodynamic evaluation of stability matched by solution calorimetry was performed by Ivanov et al. [20]. An intrinsic instability of MAPI was pointed out in the theoretical studies [21,22] by acid–base theory and DFT calculations. Senocrate et al. considered four degradation pathways for MAPbX₃ perovskites [23], which include both intrinsic and extrinsic mechanisms. The quite noticeable discrepancies between different stability trends found in the literature, particularly with respect to the halogen contribution, were discussed by Cioccioli et al. [24].

Due to the large number of possible compositions, many potentially interesting configurations remain unexplored. More recently, high-throughput calculations performed by Li et al. in connection with machine-learning (ML) techniques have been used to assess and predict stability properties of double-halide perovskites, focusing on combinations of inorganic and some organic small-sized cations [25]. Additionally, Park et al. showed a data-driven enhancement of mixed-cation perovskites of formula AₓA’₁₋ₓPbX₃ obtained by ML models [26]. Wu et al. employed a gradient-boosting regression algorithm to predict energy gaps in hybrid organic–inorganic perovskites [27]. Corre et al. demonstrated a ML algorithm targeting the recombination losses based on large-scale drift-diffusion simulations to construct the dataset [28]. Several ML techniques were analyzed for the prediction of opto-electronic properties as detailed by Workman et al. [29]. ML methods and strategies to select potential candidates for opto-electronic applications are presented in a recent review [30].

Here we perform a systematic investigation of both opto-electronic properties and stability for a selection of mixed-cation mixed-halogen perovskites. Of particular interest are double-cation and triple-halogen perovskites of formula AₓA’₁₋ₓPbX₉X’2X”₃₋₂₋₂. Using DFT calculations we determine for each composition the absorption spectra and formation energies corresponding to two potential degradation mechanisms: one intrinsic mechanism triggered by rising temperature and one extrinsic mechanism involving the reaction with molecular oxygen. Employing ML techniques, with the aim of optimizing both the optical absorption and stability, we establish an efficient procedure for the selection of optimal candidates to be used as active materials for the development of stable PSCs.

2. Mixed-Cation Mixed-Halide Perovskites

Mixed-cation perovskites have been seen to offer good prospects for enhancing the stability of PSCs [31]. Despite numerous experimental reports concerning the mitigation of the degradation effects, there is still a lack of predictive approaches that can guide the synthesis process. Here, we consider binary mixtures of the following organic cations: methylammonium (MA), formamidinium (FA), ethylammonium (EA), guanidinium (GA), formamidium (FM), hydroxylammonium (HA) and hydrazinium (HZ), as indicated in Table 1 and shown
in Figure 1a. According to their effective radii and the resulting Goldschmidt tolerance factors, these cations can easily fit into a 3D perovskite structure. However, not all of them have been proved so far to acquire a 3D structure, as supplementary interactions with the inorganic Pb-X cage, particularly by hydrogen bonds and van der Waals forces, can destabilize the 3D structure. Moreover, in the case of cation mixtures it is rather unclear how far the criteria based on Goldschmidt factor and octahedral ratio can provide a definite answer regarding the realization of a 3D structure. Larger cations typically give rise to 2D–3D or columnar 1D structures that can be adjusted by the proportions of large/small cations [32]. A more detailed analysis can take into account structural properties of the cations and alternative definitions, e.g., the globularity factor [33] may be introduced.

Table 1. A-site cations considered here used to form 3D mixed-halide perovskites. The Goldschmidt tolerance factor is calculated for each halogen type. The listed cation effective radii (except FM) were adopted from Kieslich et al. [34]. The ionic radii of the constituent elements are $R_{\text{Pb}}^{\text{eff}} = 119$, $R_{\text{I}}^{\text{eff}} = 220$, $R_{\text{Br}}^{\text{eff}} = 196$, $R_{\text{Cl}}^{\text{eff}} = 188$, $R_{\text{N}}^{\text{eff}} = 146$, $R_{\text{O}}^{\text{eff}} = 135$, adopted from Shannon [35]. All ionic radii are given in [pm] units.

| A-Cations | MA | FA | EA | GA | FM | HA | HZ |
|-----------|----|----|----|----|----|----|----|
| Formula   | CH$_3$NH$_3$ | CH(NH$_2$)$_2$ | (CH$_3$CH$_2$)NH$_2$ | (NH$_2$)$_2$C | NH$_2$COH | NH$_2$OH | NH$_3$NH$_2$ |
| 3D?       | yes | yes | yes | yes | no $^1$ | no $^2$ | no $^3$ |
| $R_{\text{eff}}^\text{I}$ | 217 | 253 | 274 | 278 | 254 $^*$ | 216 | 217 |
| $T_f$ (I)  | 0.91 | 0.99 | 1.03 | 1.04 | 0.99 $^*$ | 0.91 | 0.91 |
| $T_f$ (Br) | 0.93 | 1.01 | 1.06 | 1.06 | 1.01 $^*$ | 0.92 | 0.93 |
| $T_f$ (Cl) | 0.94 | 1.02 | 1.07 | 1.08 | 1.03 $^*$ | 0.94 | 0.94 |

(1) used as additive [36]. (2) 3D/1D phases are observed [37]. (3) 3D for $x < 0.3$ in MA$_{1-x}$HZ$_x$ mixtures [38]. (*) our estimate based on DFT calculations.

Besides the MAPbI$_3$ compounds, the synthesis of single-cation perovskites has been reported for FAPbI$_3$ [39], EAPbI$_3$ [40], GAPbI$_3$ [41]. In contrast, a recent study indicates that FM was not successfully embedded in the perovskite structure, but rather it was included as an additive [36], despite the similarity with the FA cation. The pure usage of HA cations yields structures with face-sharing octahedra that differ from the standard, corner-sharing perovskites [37]. This is despite the fact that the Goldsmith factor would indicate a typical 3D structure is possible and the anomaly was explained by the formation of hydrogen bonds. Additionally, DFT simulations predict that a perovskite structure can still be obtained in the case of HAPbI$_3$, although significant distortions of the Pb-I cage are present [42].

For binary MA–FA mixtures, a significant opto-electronic performance and stability was indicated [43]. MA–EA mixtures have been recently proposed for the fabrication of efficient light-emitting devices [44] and PSCs [45,46], while Wang et al. shows the detrimental effect of EAPbI$_3$ on the perovskite thin film morphology [47]. Similarly, MA–GA mixtures are associated with enhanced stability [41,48]. MA–HZ and FA–HZ mixtures were used to stabilize the MASnI$_3$ [49] and FAPbI$_3$ [50] perovskites. Moreover, ternary mixtures such as MA–FA–Cs [51] and FA2GA–Cs [52,53] proved enhanced efficiency, defect tolerance, and phase stability.

In addition, mixed halides also play an important role with regards to both efficiency and stability [54,55]. Here, we consider all three halogens with different proportions. By introducing lighter halogens (Br,Cl) the bandgap can be tuned, which is particularly relevant for tandem PSCs. Although the resulting increased bandgaps lower the optical absorption, the stability of bromide-based perovskites is significantly enhanced [56] and this is confirmed by thermal decomposition of MAPbX$_3$ perovskites [57]. Furthermore, the stability seems to be also linked to the A-site cation mixture as observed in the case of halide segregation [58].
Figure 1. (a) Organic cations used in perovskite structures (A and A’); (b) The pattern of presentation in each of the figures below. We will show multiple compositions of mixed-cation and mixed-halide perovskites: three binary mixtures of cations (three composite circles, each with 16 dots, corresponding to the indicated proportions of A and A’), three binary halogen mixtures (dots on the circles) and four ternary halogen mixtures (dots within the circles). The middle dot corresponds to equal halogen proportions. The colors are set to encode the magnitude of the different quantities (measures of optical absorption and stability).

Using binary mixtures of the aforementioned seven A-site cations and three halogens we assemble the perovskite structures of formula $A_xA'_{1-x}PbX_yX'_{1-y}X''_{3-y-2}$. We consider 5 $A_xA'_{1-x}$ compositions, corresponding to $x = 0, 0.25, 0.5, 0.75, 1$ and, for each case, 16 different halogen proportions are taken: 12 of them correspond to binary halogen mixtures of type $X_yX'_1_{-y}$ similar to A-A’ proportions, three proportions of triple cations of type $X_{0.25}X_{0.25}X''_{0.50}$ and the case of equal halogen proportions, $I_{0.33}Br_{0.33}Cl_{0.33}$. This results in a total of 1120 distinct samples to be analyzed. Since the multi-dimensional compositional space poses a difficulty in having a concise representation of the data sets of interest, we adopted the scheme depicted in Figure 1b, where each dot represents a given structure and the colors map the values of different quantities derived for each composition (e.g., integrated optical absorption coefficients, formation energies).

3. Methods
3.1. Density Functional Theory Calculations

The perovskite structures are constructed starting from the orthorhombic MAPbI$_3$ case containing 48 atoms in the unit cell, obtained by substituting the A-site cations and halogens on the original sites of MA and I, respectively. The initial positions of the organic cations are set to match the center of the mass as well as the primary direction of the molecular dipole, set by the C-N bond.

Structural relaxations are performed by DFT calculations implemented in the SIESTA package [59]. The linear scaling of the computational time with the system size, ensured by the strictly localized basis set, is essential for performing efficiently the high-throughput DFT calculations. Norm-conserving Troullier-Martins pseudopotentials are used, and the exchange-correlation potential corresponds to Ceperley–Alder parametrization [60]. We employ a double-$\zeta$ polarized basis set and the real space grid is fixed by a mesh cutoff of 150 Ry. Although reproducing the correct bandgap in halide perovskites may require, in general, the inclusion of hybrid functionals and spin-orbit interaction, which are considerably more computationally demanding, for lead-based perovskites there is a fortuitous cancellation of errors rendering rather accurate band structures. Following relaxations, optical calculations are performed, which yield the complex dielectric function, and the energy dependent absorption coefficient is determined by the polycrystal approximation, i.e., an averaged value with respect to randomly incident optical vectors.
These premises offer a good compromise between accuracy and computational efficiency for these specific set of systems, making a rather extensive study feasible, which offers a broad perspective over the quantities of interest.

3.2. Assessment of 3D Perovskite Structures, Opto-Electronic Properties and Stability

Optimizing PSCs in terms of efficiency and stability is directly connected with enhancing the optical absorption under the standard AM 1.5 G irradiance conditions, while there should also be an effective mitigation of potential degradation effects.

As already noted, according to previous experimental evidence, not all seven compounds based on the cations considered here (see Figure 1a) were observed to form a standard 3D perovskite configuration. Therefore, to identify possible lattice distortions for the relaxed structures, we first perform a Pb-X bond length analysis. We collect all 24 Pb-X bond lengths, corresponding to the four octahedra of the unit cell and monitor the deviations from the reference structures, MAPbX$_3$ (X = I,Br,Cl), which have a relatively narrow distribution. Typically, the de-stabilization of the inorganic cage has, as a primary effect, a wider spread of bond lengths values. As one halogen atom tends to be localized closer to one lead atom, which may be indicative of a potential tendency for bond breaking, a corresponding increase of the distance to the other, nearest neighbor lead atom is observed. We characterize these deviations by $\Delta l_{\text{min}} = l_{\text{min}} - l_{0,\text{min}}^{(0)}$ and $\Delta l_{\text{max}} = l_{\text{max}} - l_{0,\text{max}}^{(0)}$, where $l_{0,\text{min}}^{(0)}$ and $l_{0,\text{max}}^{(0)}$ are the minimum and maximum bond lengths values for Pb-X in the MAPbX$_3$ perovskites, respectively. Since the stoichiometry is imposed for a particular set of DFT calculations, a transition to other structural configurations that require different Pb/I proportions cannot be observed here. However, our aim is to detect possible issues related to the formation of the standard 3D perovskite structures, which are visible in the Pb-X bond lengths statistics, in particular in the deviations obtained for the minimum and maximum bond lengths. Different quantitative criteria may be imposed for selecting feasible structures. Naturally, in the case of mixed compounds, the proportions of different cations of various sizes and chemical behaviors, in connection with the halogen distribution as well, will determine the likelihood for a perovskite to form a three-dimensional structure.

To rank the specific perovskite structures according to their optical efficiency, we determine the flux of photons $\Phi$ at given depth $d$, using the spectral (solar) irradiance $SI(\lambda)$ and the calculated wavelength-dependent absorption coefficient $\alpha(\lambda)$. Applying the Beer-Lambert law we find:

$$
\Phi(d) = \int d\lambda \Phi_0(\lambda) \exp[-\alpha(\lambda)d] = \int d\lambda \frac{SI(\lambda)}{E_\lambda} \exp[-\alpha(\lambda)d],
$$

where the incident distribution of photon flux over the relevant wavelength range is $\Phi_0(\lambda) = SI(\lambda)/E_\lambda$ and $E_\lambda$ is the photon energy. Using Equation (1), one can define the fraction of total absorbed light at depth $d$ as an index $f_a(d) = (\Phi(0) - \Phi(d))/\Phi(0)$. The $f_a$ index shall be further used to evaluate the absorption of perovskite layers, as it is connected to the maximum number of carriers that would be ideally generated for the solar spectrum.

To compare the stabilities of the perovskite compounds, we considered one intrinsic mechanism, which can be triggered and influenced by factors such as temperature, illumination, bias conditions, usually found during the operation of the PSC and one extrinsic mechanism, which assumes the presence of a foreign specie, namely oxygen.

One possible intrinsic degradation mechanism corresponds to the decomposition of the perovskite into gaseous phases, composed of organic molecules and HX acids, and a solid precipitate, PbX$_2$, as given by the following reaction:

$$
A_x A_{1-x}' PbX_3 X''_3 \rightarrow xA + (1-x)A' + \frac{y}{3} (HX + PbX_2) + \frac{z}{3} (HX' + PbX_2') + \frac{3 - y - z}{3} (HX'' + PbX_2'')
$$

(2)
Here $\tilde{A}$, $\tilde{A}'$ represent neutral molecules, which result by de-protonation of the respective cations. Equation (2) actually brings into a more general form one of the degradation mechanisms proposed in Ref. [23,61] discussed for MAPb$_3$. Other degradation mechanisms are possible, e.g., with the formation of MAX precursor or splitting of the organic cation to ammonia in the case of MAPbX$_3$ perovskites, typically in the presence of moisture. However, in the subsequent analysis, this reaction pathway also has the advantage of a uniform treatment of the mixed perovskite compounds, irrespective of the nature of the intrinsic and extrinsic mechanisms proposed in Ref. [23,61] discussed for MAPbX$_3$ perovskites, typically in the presence of moisture.

In addition to the intrinsic mechanism, we consider the extrinsic influence of molecular oxygen, which can be described by the following reaction pathway [23]:

$$\text{A}_x\text{A}'_{1-x}\text{Pb}x'y_2z'x''_{3-y-z} + \frac{1}{4}\text{O}_2 \rightarrow x\tilde{A} + (1-x)\tilde{A}' + \frac{1}{2}\text{H}_2\text{O}$$

$$+ \frac{y}{3} \left(\frac{1}{2}x_2 + \text{PbX}_2\right) + \frac{z}{3} \left(\frac{1}{2}x'_{2} + \text{PbX}'_{2}\right)$$

$$+ \frac{3 - y - z}{3} \left(\frac{1}{2}x''_{2} + \text{PbX}''_{2}\right)$$

(3)

Although extrinsic elements such as O$_2$ and moisture can be prevented by a proper encapsulation, an assessment of such degradation mechanisms is important for the fabrication procedure as well as for the overall reliability in time of the PSC modules.

The formation energies, $E_f^{(i)}$ and $E_f^{(e)}$, associated with the reaction pathways are defined in a compact form, for intrinsic and extrinsic degradation mechanisms, respectively, as:

$$E_f^{(i)} = \sum_i x_{p_i}^{(i)} E_{p_i}^{(i)} - E_{\text{prv}}$$

(4)

$$E_f^{(e)} = \sum_i x_{p_i}^{(e)} E_{p_i}^{(e)} - E_{\text{prv}} - \frac{1}{4} E_{\text{O}_2}$$

(5)

where reaction products for the intrinsic and extrinsic mechanisms are, respectively, $p_i^{(i)} = A, A', HX, HX',HX''$, PbX$_2$, PbX'$_2$, PbX''$_2$ and $p_i^{(e)} = A, A', X_2, X'_2, X''_2, \text{PbX}_2, \text{PbX}'_2, \text{PbX}''_2$. The coefficients $x_{p_i}^{(i)}$ and $x_{p_i}^{(e)}$ correspond to molar fractions of the reaction products in Equations (2) and (3), respectively. The total energies of the perovskite compound and molecular oxygen are denoted by $E_{\text{prv}}$ and $E_{\text{O}_2}$, respectively. Based on the formation energies we define properly scaled stability indices $s_f^{(i/e)} = (E_f^{(i/e)} - E_{f_{\text{min}}}^{\text{prv}})/(E_{f_{\text{max}}}^{\text{prv}} - E_{f_{\text{min}}}^{\text{prv}})$, which provide an indication over relative stability of the perovskite compounds, where $E_{f_{\text{min}}}^{\text{prv}}$ and $E_{f_{\text{max}}}^{\text{prv}}$ represent the minimum and maximum values in the set, respectively.

### 3.3. Machine-Learning Models

To select the perovskite structures of interest, a rather extensive exploration of the compositional space is typically required. High-throughput calculations for structural relaxations can meet these goals, but use significant computing resources. Tuning the cation and/or halogen proportions brings a large degree of similitude, overlapping properties and tendencies, so that investigating the entire collection of systems can be more efficiently approached using suitable ML techniques.

The input of a ML approach is typically based on an accessible relevant information which is correlated with the targeted output. Here, we introduce a flexible feature vector that contains the proportions of all entities that define the compositional space under investigation. By entity we refer to groups of atoms, as in the case of organic cations (MA, FA, etc.) or single atoms (e.g., Pb, X). Please note that this entity-based approach can be employed to more complex perovskite structures with arbitrary number of cations, dimensionality and other functional groups as halogen replacements. Instead of considering species proportions as input vectors (i.e., single elements indexed by the atomic numbers), which would seem more general, this approach takes advantage of already identified functional groups...
of atoms (i.e., entities), which are directly correlated with the targeted output (e.g., band gap and optical absorption, formation energy). The predicted quantities are absorption and stability indices, \( f_a \) and \( f_s^{\text{rlt}} \), respectively, and the minimum and maximum bond lengths deviations, \( l_{\text{min}} \) and \( l_{\text{max}} \). We first check the feasibility of linear regression models and, to this end, we employ multivariate least squares (MLS) fitting. However, due to their simplicity the linear regression models have limited applicability and more complex behavior can be accurately captured by non-linear models, such as random forests (RF) and artificial neural networks (ANNs). Therefore, we compare and assess the suitability of different ML methods (MLS, RF, ANN) for the prediction of targeted quantities.

In all three methods we chose \( N_{\text{train}} \) randomly chosen structures as a subset for model training and the same feature vectors based on entity proportions. The rest of the samples, \( N_{\text{test}} \), is used as a test set in the case of MLS and RF methods, while for the ANN model we further split it into two subsets, i.e., for validation and test. The MLS and RF are implemented by Scikit-Learn [62]. Within the RF method a random forest regressor is implemented, which employs classifying decision trees. Here, the entire train subset is used to build the trees. The ANN is employed as an alternative method to compare the performance achieved by the different approaches. The ANN architecture is based on a 3-hidden layer fully connected network, with 100, 50, 25 neurons in each layer. We use ReLU activation function and the weight and biases are initialized from a random uniform distribution. A typical parametric configuration (learning rate of \( 10^{-5} \), \( \beta_1 = 0.9 \), \( \beta_2 = 0.999 \)) was chosen for the Adam optimizer. The output layer contains a single neuron without activation function, which provides the targeted quantity. Our ANN implementation makes use of TensorFlow with Keras frontend [63,64].

4. Results

The 1120 structures are evaluated from the point of view of structural stability, optical absorption and chemical stability with respect to the two degradation mechanisms.

The distributions of the Pb-X minimum and maximum bonds lengths are depicted in Figure 2a,b, respectively, using a red-green-blue encoding to represent different halogen proportions. In the case of \( l_{\text{min}} \) one can clearly see the three groups formed by each halogen, while the mixtures are found in between. By contrast, the distribution of \( l_{\text{max}} \) is broader, with some very large values for Pb-X bond lengths indicating a significant distortion of the inorganic cage.

![Figure 2](image_url)

**Figure 2.** (a) Minimum and (b) maximum bond length distribution for mixtures of halogens, expressed in [Å], using the RGB encoding: iodine (red), bromide (green) and chlorine (blue).

Based on the bond lengths statistics, a binary classification of the perovskite structures is performed, which is shown to depend on the nature of cations, halogens and their proportions, as depicted in Figure 3. If following structural relaxations, the minimum or the maximum Pb-X bond length deviate more than a fraction \( f_{\text{bl}} \) from the corresponding
bond length value in the MAPbX₃ systems, which are taken as reference, i.e., |Δl_{min}/Δl_{max}| > f_{bl} × l_{min}/l_{max}, the systems are considered unlikely to form the 3D perovskite conformation.

Figure 3 shows a matrix type representation for the 3D classification, where each element corresponds to the description laid out in Figure 1b. The rows correspond to A cations, while the columns are index by A'. This convention is kept the same in the subsequent analyses (e.g., absorption and stability indices). Two criteria are used, set by the fractions f_{bl} of 10% and 5% (a more strict criterium), corresponding to the bond lengths deviations, which indicates the same trend: the MA, FA, EA cations do not induce substantial deformations of the Pb-X cage, as opposed to FM, HA, HZ and, in part, GA, which is the largest cation. This is somewhat opposed to what one would expect, based on the ionic radii alone, since we notice significant differences, as R_{eff}^{MA} < R_{eff}^{FA} < R_{eff}^{EA} < R_{eff}^{GA}, whereas FM, HA, HZ are all comparable in size with MA and FA. However, the interactions with the Pb-X cage seem to have significant destabilization effect, which is also reflected in the lack of reports concerning the attainability of these particular perovskites. Concerning the group pf FM, HA, HZ, our calculations indicate that FM mixtures are statistically the worst candidates, already signaled by the absence of the 3D phase in [36], followed by HA and HZ. However, in the case of HZ, surprisingly, mixtures with large proportions of MA or EA yield rather undistorted structures. A similar situation was experimentally observed in Ref. [38], where MA₁-xHZₓ mixtures were obtained for x < 0.3. We should note that although these results are statistical in nature, the large number of structures provides a consistent overview over the relatively broad class of mixed-cation mixed-halide perovskites.

**Figure 3.** Classification of 3D perovskite structures of formula AₓA’₁₋ₓPbXₙX’ₓX”₃₋ₙ₋ₓ based on PbX bond length statistics: feasible (yellow dots) and likely infeasible (black dots) 3D structures. A and A’ index the rows and columns, respectively. Each matrix element, except the ones on the secondary diagonal, has the representation given in Figure 1b. Two criteria have been considered (marked by shaded areas): f_{bl} = 10% deviations (upper left triangle) and f_{bl} = 5% deviations (lower right triangle) from the smallest or the largest Pb-X bond lengths as established for MAPbX₃ perovskites.
Further on, the optimization of absorption spectra is essential for highly efficient PSCs. Although relatively well positioned for the solar spectrum, the bandgap of MAPI of \( \sim 1.6 \text{ eV} \) is significantly larger than in the case of silicon and a decrease of the bandgap would be therefore desired. However, only a few options are available to achieve smaller gaps, e.g., by changing the species or functional groups in the three possible sites, A, B, X. In fact, by including lighter halogens, we observe the gap increases in correspondence with shorter Pb-X lengths. The band structures for the reference MAPb\( \text{X}_3 \) structures are shown in Figure 4a–c. The absorption spectra indicated in Figure 4d have different overlap with spectral solar irradiance \( S_I \), depending on the halogen type. The substitution of Pb by other group four elements such as Sn, Ge reportedly did not result in similar efficiencies. On the other hand, the influence of the organic cations on the bandgap is typically rather small, but still noticeable in some cases. For example, slightly larger cations such as FA and EA tend to increase the volume of the unit cell from 819.78 Å\(^3\) (MAPI), to 823.47 Å\(^3\) (FAPI) and 884.45 Å\(^3\) (EAPI), which can be correlated with the bandgap decrease. GAPI is a somewhat special case which shows a larger gap and an intermediate volume of 843.81 Å\(^3\). However, the absorption coefficient has a different distribution over wavelengths, with larger values at higher photon energies (4-5 eV) which results in an overall comparable optical performance with FA and EA-based perovskites. The absorption spectra of these mono-cation base perovskites are comparatively presented in Figure 4d.

![Figure 4](image-url)

**Figure 4.** Representative opto-electronic properties. The band structures are indicated for (a) MAPbI\(_3\), (b) MAPbBr\(_3\), (c) MAPbCl\(_3\). The absorption coefficients are depicted comparatively for the different halogens and cations in (d,e), respectively. The spectral solar irradiance \( S_I \) is also represented.

The picture of binary-cation and triple-cation mixtures is more complex; however, it follows the same general trends as depicted in Figure 5. As expected, the halogen proportions have a strong impact over the absorption index \( f_a \), the fraction of total absorbed light at depth \( d \) in a perovskite layer. The bandgap sequence corresponding to different I, Br, Cl proportions is clearly seen in all binary-cation mixtures, as it has direct implications in the absorption spectra. In particular, Br-Cl mixtures indicate the lowest efficiency as absorber materials. However, lighter halogens play an active role for enhancing the stability, therefore the potential detrimental effect with regards to opto-electronic properties should be put in a broader context. A close inspection of data reveals that there are indeed
variations due to the cation mixtures as well, even though at a smaller scale, and the most prominent candidate mixtures rely on EA, GA, FA cations, which prove to be better than the standard MAPbX$_3$ compounds: higher x proportions in EA$_x$-MA$_{1-x}$, GA$_x$-MA$_{1-x}$, FA$_x$-MA$_{1-x}$ mixtures show a rather steady trend of improvement.

Figure 5. Distribution of solar spectrum integrated absorption, measured by $f_a$ index, for two different values of the absorber layer width, $d_1 = 300$ nm (upper left triangle) and $d_2 = 600$ nm (lower right triangle). Significant variations due to halogen proportions are observed, while smaller contributions due to cation mixtures are visible, with particular enhancements for FA, EA and GA binary mixtures.

The primary goal of enhancing the opto-electronic properties is presently increasingly put in the balance with the mitigation of possible degradation mechanisms, which have a greater impact on the overall performance. Here, we consider the intrinsic mechanism defined by Equation (2) and an extrinsic mechanism, which involves the perovskite interaction with molecular oxygen, described by Equation (3). The results concerning the stability index $f_s^{(i/e)}$ are shown in Figure 6, scaled within the reference intervals of [2.9, 5.9] eV and [11.0, 14.8] eV for the intrinsic and extrinsic degradation mechanism, respectively, providing a compact overview regarding the relative stability of the perovskite mixtures.

Concerning the intrinsic degradation mechanism, the Br > I > Cl stability sequence was found as a quite general trend. This is consistent with the experimental study of Pistor et al. showing in situ XRD patterns during heating and decomposition of MAPbX$_3$ perovskites [57], indicating the highest decomposition temperature for MAPbBr$_3$ and the lowest for MAPbCl$_3$. In addition, Brunetti et al. describes MAPbCl$_3$ as the most unstable compound among the same three perovskites [65], while MAPbI$_3$ and MABr$_3$ are reported to be more stable. Ciccioli et al. also finds highest stability for bromide-based compounds with respect to the volatilization reaction [24] and McGovern et al. further argues for the enhanced stability of MAPbBr$_3$ over MAPbI$_3$ by a reduced mobility of MA cations due to shorter Br-X bond lengths [56].
With respect to the organic cation mixtures, the perovskites incorporating GA, FA and EA show the higher stability. This is in line with reported improvements in the overall stability of the PSCs for GA-MA mixtures \[48\], EA-MA mixtures as revealed by first-principles calculations \[46\], while FA-MA mixtures with additional Cs also demonstrate stability enhancement \[66\]. Triple cations systems have been also evaluated starting with the base composition of $\text{MA}_{0.5}\text{FA}_{0.25}\text{Cs}_{0.25}\text{PbI}_3$ and subsequently the highest proportion among the cations is varied along with halogen composition. The intrinsic mechanism described by Equation (2) is slightly modified by the appearance of CsX salts as degradation products. With respect to the reference triple-cation system, $\text{MA}_{0.25}\text{FA}_{0.5}\text{Cs}_{0.25}\text{PbI}_3$ (i.e., larger proportion of FA) has slightly larger formation energy of $\sim 4\%$, followed by a significant decrease by $37\%$ for $\text{MA}_{0.25}\text{FA}_{0.25}\text{Cs}_{0.5}\text{PbI}_3$, where the Cs proportion is the largest. The same trend is quite accurately confirmed by replacing I with Br and Cl, albeit for slightly different values: a small improvement in the stability of $+2\%$ for $\text{MA}_{0.5}\text{FA}_{0.25}\text{Cs}_{0.25}\text{PbBr}_3$ and a decrease in the stability of $\sim 3\%$ for $\text{MA}_{0.5}\text{FA}_{0.25}\text{Cs}_{0.25}\text{PbCl}_3$, as compared to the base composition, $\text{MA}_{0.5}\text{FA}_{0.25}\text{Cs}_{0.25}\text{PbI}_3$.

Our investigation shows that GA-FA and GA-EA are the best candidates. This is despite the fact that the proposed combinations have, on average, larger cation sizes, compared to MA, but also to FM, HA, HZ, which are shown to yield a smaller formation energy for the decomposition into reaction products. Again, this can be correlated with the missing 3D perovskite compounds that can be, at least in theory, produced. Our calculations also suggest that the chemical stability can be enhanced for FM, HA, HZ by mixing them with GA and FA, as long as the structural analysis indicates a 3D structure. In addition, a degradation mechanism resulting in the formation of precursors was also suggested \[23\], although the experimental evidence of cation-halide precursor (A-X) is lacking \[24\], in contrast to PbI$_2$, while the prevalence of the gaseous phases (A, HX) is pointed out. Nevertheless, we able to confirm the intrinsic instability of MAPbI$_3$ with a
formation energy of $-50 \text{ meV}$, similar to Ref. [22]. The extrinsic degradation mechanism conveys a similar picture, with the difference that chlorine-based perovskites are the most stable with respect to oxygen action. The same cations, GA and FA, can provide stability improvements, while least stability is demonstrated for the FM, HA and HZ cations.

These results can be, in part, explained by simple processes. Both intrinsic and extrinsic stability trends suggest that the cation deprotonation resistance is an important factor. To check this, we considered a simplified deprotonation reaction, of type $A_1 \rightarrow \tilde{A} + HI$, where $A$ is the cation and $\tilde{A}$ is the deprotonated molecule, and we found that the reaction energies are in line with the formation energies corresponding to the degradation mechanisms: $MA$ (0.55 eV), $FA$ (1.41 eV), $EA$ (2.09 eV), $FM$ (0.40 eV), $HA$ (0.78 eV) and $HZ$ (1.19 eV). This shows that GA and FA cations have the largest deprotonation energies, while EA is rather close, which translates in higher stability. FM is at the other end, while HA and HZ have intermediate values. Our estimation for MA gives a small value, but still larger than FM. On the other hand, as the intrinsic mechanism has, as final products $HX$, instead of $X_2$ as in the case of the extrinsic degradation mechanism, we obtain that bromide-based perovskites are more stable for the intrinsic mechanism (following the trend $Br > I > Cl$), while chlorine compounds are more stable for the extrinsic one ($Cl > Br > I$). This is in accordance with the fact that HCl is most stable, followed by HBr and HI, with $Cl > Br > I$.

In Table 2, we present the selection of 10 best candidates with the goal of optimizing: (a) absorption index $f_a$; (b) stability index $f_s$; both $f_a$ and $f_s$, with values larger than 0.65. The intrinsic degradation mechanism has been considered. The individual impact of cations and halogens on both optical absorption and intrinsic stability can be seen in Figure 7, as they are highlighted in absorption-stability ($f_a - f_s$) maps. Although in general, one cannot decouple the individual contributions of cations and halogens, some trends become visible at this level too. For instance, high concentration of bromine, translated in a stability enhancement. This is in stark contrast to FM-based mixtures. Regarding the optical absorption, large proportions of EA and GA in mixtures are beneficial, with iodine as the best choice in this respect. Assuming a certain partition of the ($f_a - f_s$) map, one may define selections of candidate perovskites that are more appropriate for certain technological applications. For instance, aiming at both optical and stability performance, we obtain that the best candidates mostly rely on EA-GA and I-Br mixtures as one can see from Table 2, with $EA_{0.25}GA_{0.75}PbI_3$ as the most prominent example. Optimizing absorption only, compounds based on high GA proportions mixed with EA and potentially HZ or FM, with predominantly iodine are found, e.g., $EA_{0.25}GA_{0.75}PbI_3$ or $FA_{0.25}EA_{0.75}PbI_2Pb_2Cl_{0.75}$. On the other hand, for high stability only, perovskites with higher Br contents are ranked higher like $GAPbBr_{2.25}Cl_{0.75}$ or $FA_{0.25}GA_{0.75}PbI_{1.50}Br_{1.50}$, although as a consequence they have smaller absorptions.

| Absorption Index ($f_a$) | Stability Index ($f_s$) | Both $f_a$ and $f_s$ |
|------------------------|------------------------|----------------------|
| $EA_{0.25}GA_{0.75}PbI$ | $GAPbBr_{2.25}Cl_{0.75}$ | $EA_{0.25}GA_{0.75}PbI$ |
| $f_a = 0.81$ and $f_s = 0.77$ | $f_a = 0.26$ and $f_s = 0.97$ | $f_a = 0.81$ and $f_s = 0.77$ |
| $GAPbCl_3$ | $GA_{0.25}Br_{0.75}Cl_{0.25}$ | $GAPb_{2.25}Cl_{0.75}$ |
| $f_a = 0.77$ and $f_s = 0.83$ | $f_a = 0.48$ and $f_s = 0.94$ | $f_a = 0.77$ and $f_s = 0.83$ |
| $EA_{0.25}PbI_3$ | $GAPbCl_3$ | $EA_{0.25}GA_{0.75}Pb_{2.25}Br_{0.75}$ |
| $f_a = 0.77$ and $f_s = 0.54$ | $f_a = 0.13$ and $f_s = 0.94$ | $f_a = 0.75$ and $f_s = 0.80$ |
| $EA_{0.25}GA_{0.75}PbI_3$ | $GAPb_{2.25}Cl_{0.75}$ | $GA_{0.25}HZ_{0.25}PbI_3$ |
| $f_a = 0.76$ and $f_s = 0.63$ | $f_a = 0.58$ and $f_s = 0.93$ | $f_a = 0.73$ and $f_s = 0.73$ |
Table 2. Cont.

| Absorption Index ($f_a$) | Stability Index ($f_s$) | Both $f_a$ and $f_s$ |
|-------------------------|------------------------|----------------------|
| FA$_{0.25}$EA$_{0.50}$PB$_{1.25}$Cl$_{0.25}$ | GAP$_{1.25}$Br$_{0.75}$ | EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$Br$_{0.25}$ |
| $f_a = 0.75$ and $f_s = 0.51$ | $f_a = 0.56$ and $f_s = 0.93$ | $f_a = 0.73$ and $f_s = 0.69$ |
| EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$Br$_{0.25}$ | FA$_{0.25}$GA$_{0.50}$PB$_{1.25}$Br$_{0.25}$ | FA$_{0.25}$GA$_{0.50}$PB$_{1.25}$Br$_{0.25}$ |
| $f_a = 0.75$ and $f_s = 0.80$ | $f_a = 0.59$ and $f_s = 0.91$ | $f_a = 0.72$ and $f_s = 0.75$ |
| EA$_{0.25}$Br$_{0.25}$ | EA$_{0.25}$Br$_{0.25}$ | EA$_{0.25}$Br$_{0.25}$ |
| $f_a = 0.75$ and $f_s = 0.52$ | $f_a = 0.55$ and $f_s = 0.91$ | $f_a = 0.71$ and $f_s = 0.79$ |
| EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ | FA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ | EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ |
| $f_a = 0.75$ and $f_s = 0.41$ | $f_a = 0.66$ and $f_s = 0.90$ | $f_a = 0.71$ and $f_s = 0.66$ |
| EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ | FA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ | FA$_{0.25}$GA$_{0.50}$PB$_{1.25}$ |
| $f_a = 0.73$ and $f_s = 0.27$ | $f_a = 0.43$ and $f_s = 0.89$ | $f_a = 0.68$ and $f_s = 0.70$ |
| EA$_{0.25}$FM$_{0.50}$PB$_{1.25}$ | EAP$_{1.25}$Br$_{0.25}$Cl$_{0.25}$ | EA$_{0.25}$GA$_{0.50}$PB$_{1.25}$Br$_{0.25}$ |
| $f_a = 0.73$ and $f_s = 0.27$ | $f_a = 0.23$ and $f_s = 0.89$ | $f_a = 0.69$ and $f_s = 0.70$ |

Figure 7. Absorption-stability ($f_a - f_s$) maps demonstrating individual contributions from the organic cations and halogens, established for the intrinsic degradation mechanism. Stability enhancements are observed for mixtures containing FA and GA, while large proportions of FM is detrimental. The halogens significantly influence the absorption in accordance with their concentrations ($f_{I^a} > f_{Bra} > f_{Cla}$), while Br brings a shift towards higher formation energies.

One should also mention that compared to realistic conditions, formation energy calculations are idealized for perfect bulk systems, rather than defected perovskite layers with finite grain sizes and interfaces, which can lead to rather different degradation mechanisms. It is expected that these different kinds of defects would lower the formation energies of the resulting compounds. Additionally, the DFT calculations at LDA level typically tend to overestimate the formation energies as well as slight underestimate the band gaps; however, in this context, a reasonable compromise was achieved between computational efficiency and accuracy. This enables a rather extensive overview regarding the mixed perovskite compounds, allowing a direct comparison with respect to both optical efficiency and chemical stability.

Moreover, under certain experimental conditions it is difficult to separate multiple degradation mechanisms. Taking into account the combined action of temperature and oxygen, an enhanced stability should be obtained by including lighter halogens (Br, Cl)
and mixtures containing primarily GA, FA and EA. For these three cations, depending on their sizes and proportions, an abundance of low-dimensional perovskite structures may be formed as in the case of FA-GA mixtures, where 2D layered structures are also observed [67]. Despite their enhanced stability, these low-dimensional perovskites have wider bandgaps which can significantly reduce the light absorption in the case of thin 2D layers. In this respect, 3D perovskites are generally better suited. Therefore, controlling the thickness and chemical composition of the 2D layers remains an active goal for optimizing the overall absorber performance.

The multitude of possible compositions, potentially resulting in rather different structures cannot be covered by high-throughput calculations alone. As there is a large degree of overlapping information contained in the analyzed set of structures, regression methods can provide reasonably accurate results using only a relatively small subset. To this end, we investigate in how far relatively complex predictions accounting for structural 3D phase formation, optical absorption and chemical stability altogether can be achieved. The ML results concerning the prediction accuracy, measured by $R^2$ coefficient of determination, are summarized in Table 3. For MSL and RF methods, three different sizes for the train sets were investigated. In this way, we can point out a certain leveling of the prediction accuracies for about 100 random samples in the train set, beyond which a further increase in accuracy is not significant. This is important, as quite accurate results can be achieved using less than 10% from the total of 1120 samples. This is not entirely surprising, since e.g., the action of the halogens on the electronic gap is the same in all the cation mixtures. Linear regression methods such as MLS, using the multi-dimensional input assembled from the entity proportions, generally perform rather well particularly for $f_a$ and $f_s$ indices, with $R^2 > 0.77$ and $R^2 > 92$ on average and small variance, respectively. However, the structural deformations described by $l_{\text{min}}$ and $l_{\text{max}}$ have a more complex behavior and non-linear methods such as RF perform better, e.g., $R^2 = 0.88$ (RF) compared $R^2 = 0.56$ (MLS) for $l_{\text{min}}$. This is particularly true in the case of $l_{\text{min}}$, while the distribution of $l_{\text{max}}$ has a more complicated relation to the structure type distribution and the resulting prediction accuracies are smaller, as one may infer from Figure 2a,b.

The ANN models confirm the prediction accuracies of the previous two methods. This is illustrated in Figure 8, where the reference DFT calculated values and ANN predictions are indicated. Here, $N_{\text{train}} = 100$ random structures were used as before, while the remaining set is equally split into validation and test sets. However, in this case, the overtraining effects is rather small, particularly for $f_a$ and $f_s^{(i/e)}$, so that validation and test sets yield rather similar values for a fixed number of training epochs. The relatively large fluctuations observed for $l_{\text{max}}$, as compared to $l_{\text{min}}$, obtained for the MLS and RF methods, is indicative for a certain limitation of the model’s prediction capacity, which is also found to be the case for the ANN model. Instead, $l_{\text{min}}$ is a better suited parameter that can be strongly correlated with structural deformations for a certain perovskite composition and can be accurately inferred by the different ML techniques.

Table 3. The $R^2$ coefficient of determination, measuring the accuracy of the test sets, in the case of MLS, RF and ANN methods. For MLS and RF, 10 distinct partitions of the train-test sets were employed yielding the minimum, maximum and average $R^2$. The ANN models confirm the performance attained by the previous two methods.
Table 3. Cont.

| Training: | 50 Samples | 100 Samples | 200 Samples |
|-----------|------------|-------------|-------------|
| \(f_a^{(i)}\) | MLS: 0.92 0.93 0.93 | MLS: 0.93 0.94 0.94 | MLS: 0.94 0.94 0.94 |
| | RF : 0.70 0.81 0.77 | RF : 0.81 0.87 0.85 | RF : 0.87 0.91 0.89 |
| | ANN: 0.94 | ANN: 0.95 |  |
| \(l_{\text{min}}\) | MLS: 0.43 0.56 0.48 | MLS: 0.51 0.60 0.56 | MLS: 0.55 0.60 0.58 |
| | RF : 0.81 0.87 0.84 | RF : 0.83 0.86 0.85 | RF : 0.84 0.86 0.85 |
| | ANN: 0.84 | ANN: 0.84 |  |
| \(l_{\text{max}}\) | MLS: 0.12 0.31 0.23 | MLS: 0.11 0.33 0.28 | MLS: 0.27 0.38 0.32 |
| | RF : 0.16 0.37 0.31 | RF : 0.17 0.42 0.33 | RF : 0.20 0.42 0.35 |
| | ANN: 0.43 | ANN: 0.43 |  |

Figure 8. Regression data obtained using ANN models: (a) optical absorption index \(f_a\); (b,c) stability indices for the intrinsic/extrinsic mechanisms, \(f_s^{(i/e)}\); (d,e) minimum/maximum bond lengths.

5. Conclusions

This study encompasses a rather broad class of perovskite materials with mixtures of cations and halogens. The systems are analyzed from both the perspective of optical absorption and the chemical stability, which is essential in the development of high-performance PSCs. The proposed methodology enables an efficient and unified description of many instances with varying proportions of cations and halogens. Some known results are recovered (e.g., based on MA, FA, GA), while some of the less studied compositions (e.g., based on EA, FM, HA, HZ) are integrated into the larger picture. Performing a bond lengths analysis, we first determine which structures are susceptible to acquire 3D perovskite structure or not, followed by an investigation of the chemical stability. One intrinsic degradation mechanism and one extrinsic mechanism subject to addition of molecular oxygen are evaluated for each structure. It was found that the most promising candidates with respect to the intrinsic degradation mechanism contain high proportions of GA and, in particular, FA-GA mixtures, while bromine is highly influential for enhancing stability. The same cation mixtures provide good mitigation of oxygen induced degradation, while here chlorine seems to have a larger impact. The cations which are reportedly less likely to form 3D perovskite structures (FM, HA, HZ), may appear in mixtures (e.g., with GA) to improve their structural and chemical stability. It was also established that the stability can be correlated with the deprotonation resistance of the organic cations. On the other hand, an optimal absorption can be reached by EA-GA and EA-FA mixtures, potentially with small amounts of FA, while iodine is generally the best choice as it induces...
a relatively smaller gap compared to bromine and chlorine. The EA cation, which has not been extensively investigated in previous studies, may prove to be a good candidate for enhanced opto-electronic properties. Somewhat surprisingly, small proportions of FM and HZ added to a GA-based perovskite may also result in a quite high optical performance.

As the high-throughput calculations of this type require extensive numerical resources, we investigated in how far ML techniques can predict the essential quantities. It was established that using random subsets of \( \sim 10\% \) of the calculated structures can still provide accurate results, using feature vectors based on entity proportions, which are readily available. The absorption index \( f_a \) and the stability indices \( f_s^{(i/e)} \) can be predicted using multivariate linear regression methods on a much smaller subset. However, some other properties stemming from bond lengths distribution, notably the minimum and maximum values, require non-linear regression techniques, which were here implemented using the random forests method and artificial neural networks. We identify the variation of the minimum bond length as a suitable, predictable parameter to assess the deformations of the perovskite structures.

In conclusion, we proposed a methodology for an efficient exploration of candidate perovskite materials, to find optimal solutions for both opto-electronic and stability properties as required by the PSC technology. DFT-ML schemes of this type, which have a reasonable compromise between accuracy and efficient exploration of a huge compositional space, are easily scalable and can provide further essential insight into complex perovskite structures, specifically multiple-cation and low-dimensional systems.

Author Contributions: Conceptualization, M.F., I.P., A.M., G.A.N.; Formal analysis, D.V.A.; Investigation, N.F., T.L.M. and G.A.N.; Methodology, A.M. and G.A.N.; Supervision, G.A.N.; Validation, D.V.A., M.F. and I.P.; Writing—original draft, N.F. and G.A.N.; Writing—review and editing, T.L.M., D.V.A., M.F., I.P. and A.M.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant of the Romanian Ministry of Research, Innovation and Digitalization, CCCDI—UEFISCDI, project number PN-III-P2-2.1-PED-2019-1567, within PNCDI III.

Institutional Review Board Statement: None.

Informed Consent Statement: None.

Data Availability Statement: None.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Qiu, L.; Ono, L.K.; Qi, Y. Advances and challenges to the commercialization of organic–inorganic halide perovskite solar cell technology. *Mater. Today Energy* 2018, 7, 169–189. [CrossRef]
2. Aranda, C.A.; Caliò, L.; Salado, M. Toward Commercialization of Stable Devices: An Overview on Encapsulation of Hybrid Organic-Inorganic Perovskite Solar Cells. *Crystals* 2021, 11, 519. [CrossRef]
3. Kundu, S.; Kelly, T.L. In situ studies of the degradation mechanisms of perovskite solar cells. *EcoMat* 2020, 2, e12025. [CrossRef]
4. Bryant, D.; Aristidou, N.; Pont, S.; Sanchez-Molina, I.; Chotchunangatchaval, T.; Wheeler, S.; Durrant, J.R.; Haque, S.A. Light and oxygen induced degradation limits the operational stability of methylammonium lead triiodide perovskite solar cells. *Energy Environ. Sci.* 2016, 9, 1655–1660. [CrossRef]
5. Aristidou, N.; Eames, C.; Sanchez-Molina, I.; Bu, X.; Kosco, J.; Islam, M.S.; Haque, S.A. Fast oxygen diffusion and iodide defects mediate oxygen-induced degradation of perovskite solar cells. *Nat. Commun.* 2017, 8, 15218. [CrossRef]
6. Yang, J.; Yuan, Z.; Liu, X.; Braun, S.; Li, Y.; Tang, J.; Gao, F.; Duan, C.; Fahlman, M.; Bao, Q. Oxygen- and Water-Induced Energics Degradation in Organometal Halide Perovskites. *ACS Appl. Mater. Interfaces* 2018, 10, 16225–16230. [CrossRef] [PubMed]
7. Li, M.; Li, H.; Fu, J.; Liang, T.; Ma, W. Recent Progress on the Stability of Perovskite Solar Cells in a Humid Environment. *J. Phys. Chem. C* 2020, 124, 27251–27266. [CrossRef]
8. Xu, F.; Zhang, T.; Li, G.; Zhao, Y. Mixed cation hybrid lead halide perovskites with enhanced performance and stability. *J. Mater. Chem. A* 2017, 5, 11450–11461. [CrossRef]
9. Li, Q.; Zhao, Y.; Zhou, W.; Han, Z.; Fu, R.; Lin, F.; Yu, D.; Zhao, Q. Halogen Engineering for Operationally Stable Perovskite Solar Cells via Sequential Deposition. *Adv. Energy Mater.* 2019, 9, 1902239. [CrossRef]
10. Yang, L.; Xiao, Y.; Han, G.; Chang, Y.; Zhang, Y.; Hou, W.; Lin, J.Y.; Li, H. Enhanced stability and efficiency of perovskite solar cells via bifunctional group passivation with thioalicylic acid. Org. Electron. 2020, 81, 105681. [CrossRef]

11. Stoumpos, C.C.; Cao, D.H.; Clark, D.J.; Young, J.; Rondinelli, J.M.; Jang, J.I.; Hupp, J.T.; Kanatzidis, M.G. Ruddlesden–Popper Hybrid Lead Iodide Perovskite 2D Homologous Semiconductors. Chem. Mater. 2016, 28, 2852–2867. [CrossRef]

12. Ortiz-Cervantes, C.; Carmona-Monroy, P.; Solis-Ibarra, D. Two-Dimensional Halide Perovskites in Solar Cells: 2D or not 2D? ChemSusChem 2019, 12, 1560–1575. [CrossRef] [PubMed]

13. Meggiolaro, D.; De Angelis, F. First-Principles Modeling of Defects in Lead Halide Perovskites: Best Practices and Open Issues. ACS Energy Lett. 2018, 3, 2206–2222. [CrossRef]

14. Brakkee, R.; Williams, R.M. Minimizing Defect States in Lead Halide Perovskite Solar Cell Materials. Appl. Sci. 2020, 10, 3061. [CrossRef]

15. Roknuzzaman, M.; Zhang, C.; Ostrikov, K.K.; Du, A.; Wang, H.; Wang, L.; Tesfamichael, T. Electronic and optical properties of lead-free hybrid double perovskites for photovoltaic and optoelectronic applications. Sci. Rep. 2019, 9, 718. [CrossRef]

16. Maiti, A.; Chatterjee, S.; Peedikakkandy, L.; Pal, A.J. Defects and Their Passivation in Hybrid Halide Perovskites toward Solar Cell Applications. Sol. RRL 2020, 4, 2000505. [CrossRef]

17. Tao, S.X.; Cao, X.; Bobbert, P.A. Accurate and efficient band gap predictions of metal halide perovskites using the DFT-1/2 method: GW accuracy with DFT expense. Sci. Rep. 2017, 7, 14386. [CrossRef]

18. Frost, J.M.; Butler, K.T.; Brivio, F.; Hendon, C.H.; vanSchilfgaarde, M.; Walsh, A. Atomistic Origins of High-Performance in Hybrid Halide Perovskite Solar Cells. Nano Lett. 2014, 14, 2584–2590. [CrossRef]

19. Tenuta, E.; Zheng, C.; Rubel, O. Thermodynamic origin of instability in hybrid halide perovskites. Sci. Rep. 2016, 6, 37654. [CrossRef]

20. Ivanov, I.; Steparuk, A.; Bolyachkina, M.; Tsvetkov, D.; Safronov, A.; Zuev, A. Thermodynamics of formation of hybrid perovskite-type methylammonium lead halides. J. Chem. Thermodyn. 2018, 116, 253–258. [CrossRef]

21. Juarez-Perez, E.J.; Ono, L.K.; Uriarte, I.; Cocinero, E.J.; Qi, Y. Degradation Mechanism and Relative Stability of Methylammonium Halide Based Perovskites Analyzed on the Basis of Acid–Base Theory. ACS Appl. Mater. Interfaces 2019, 11, 12586–12593. [CrossRef]

22. Zhang, Y.Y.; Chen, S.; Xu, P.; Xiang, H.; Gong, X.G.; Walsh, A.; Wei, S.H. Intrinsic Instability of the Hybrid Halide Perovskite Semiconductor CH3NH3PbI3. Chin. Phys. Lett. 2018, 35, 036104. [CrossRef]

23. Senocate, A.; Kim, G.Y.; Grätzel, M.; Maier, J. Thermochemical Stability of Hybrid Halide Perovskites. ACS Energy Lett. 2019, 4, 2899–2870. [CrossRef]

24. Ciccoli, A.; Latini, A. Thermodynamics and the Intrinsic Stability of Lead Halide Perovskites CH3NH3PbX3. J. Phys. Chem. Lett. 2018, 9, 3756–3765. [CrossRef] [PubMed]

25. Li, Z.; Xu, Q.; Sun, Q.; Hou, Z.; Yin, W.J. Thermodynamic Stability Landscape of Halide Double Perovskites via High-Throughput Computing and Machine Learning. Adv. Funct. Mater. 2019, 29, 1807280. [CrossRef]

26. Park, H.; Ali, A.; Mall, R.; Bansmail, H.; Sanvito, S.; El-Mellouhi, F. Data-driven enhancement of cubic phase stability in mixed-cation perovskites. Mach. Learn. Sci. Technol. 2021, 2, 025030. [CrossRef]

27. Wu, T.; Wang, J. Deep Mining Stable and Nontoxic Hybrid Organic–Inorganic Perovskites for Photovoltaics via Progressive Machine Learning. ACS Appl. Materials Interfaces 2020, 12, 57821–57831. [CrossRef]

28. Le Corre, V.M.; Scherker, T.S.; Koopmans, M.; Koster, L.J.A. Identification of the dominant recombination process for perovskite solar cells based on machine learning. Cell Phys. Sci. 2021, 2, 100346. [CrossRef]

29. Workman, M.; Zhi Chen, D.; Musa, S.M. Machine Learning for Predicting Perovskite Solar Cell Opto-Electronic Properties. In Proceedings of the 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 26–27 October 2020; pp. 1–5. [CrossRef]

30. Tao, Q.; Xu, P.; Li, M.; Lu, W. Machine learning for perovskite materials design and discovery. Npj Comput. Mater. 2021, 7, 23. [CrossRef]

31. Bidikoudi, M.; Kymakis, E. Novel approaches and scalability prospects of copper based hole transporting materials for planar perovskite solar cells. J. Mater. Chem. C 2017, 5, 13680–13708. [CrossRef]

32. Kim, E.B.; Akhtar, M.S.; Shin, H.S.; Ameen, S.; Nazeeruddin, M.K. A review on two-dimensional (2D) and 2D-3D multidimensional perovskite solar cells: Perovskites structures, stability, and photovoltaic performances. J. Photochem. Photobiol. C Photochem. Rev. 2021, 68, 100405. [CrossRef]

33. Gholidour, S.; Ali, A.M.; Correa-Baena, J.P.; Turren-Cruz, S.H.; Tajabadi, F.; Tress, W.; Taghavinia, N.; Grätzel, M.; Abate, A.; De Angelis, F.; et al. Globularity-Selected Large Molecules for a New Generation of Multication Perovskites. Adv. Mater. 2017, 29, 1702005. [CrossRef]

34. Kieslich, G.; Sun, S.; Cheetham, A.K. Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog. Chem. Sci. 2014, 5, 4712–4715. [CrossRef]

35. Shannon, R.D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallogr. Sect. A 1976, 32, 751–767. [CrossRef]

36. Raifuku, I.; Chiang, Y.H.; Hou, C.H.; Li, M.H.; Lin, C.F.; Lin, P.Y.; Shyue, J.J.; Chen, P. Formamide iodide: A new cation additive for inhibiting δ-phase formation of formamidinium lead iodide perovskite. Mater. Adv. 2021, 2, 2272–2277. [CrossRef]
37. D’Annibale, A.; Panetta, R.; Tarquini, O.; Colapietro, M.; Quaranta, S.; Cassetta, A.; Barba, L.; Chita, G.; Latini, A. Synthesis, physico-chemical characterization and structure of the elusive hydroxylammonium lead iodide perovskite NH3OHPbI3. *Dalton Trans.* 2019, 48, 5397–5407. [CrossRef]

38. Akbukalov, A.F.; Frolova, L.A.; Anokhin, D.V.; Gerasimov, K.L.; Dremova, N.N.; Troshin, P.A. Hydrazinium-loaded perovskite solar cells with enhanced performance and stability. *J. Mater. Chem. A* 2016, 4, 18378–18382. [CrossRef]

39. Han, Q.; Bae, S.H.; Sun, P.; Hsieh, Y.T.; Yang, Y.M.; Rim, Y.S.; Zhao, H.; Chen, Q.; Shi, W.; Li, G.; et al. Single Crystal Formamidinium Lead Iodide (FAPbI3): Insight into the Structural, Optical, and Electrical Properties. *Adv. Mater.* 2016, 28, 2253–2258. [CrossRef] [PubMed]

40. Ma, D.; Dai, N.; Lan, Y. Solution Route to Single-Crystalline Ethylammonium Lead Halide Microstructures. *ChemistrySelect* 2019, 4, 2174–2180. [CrossRef]

41. Kim, Y.; Bae, C.; Jung, H.S.; Shin, H. Enhanced stability of guanidinium-based organic-inorganic hybrid lead triiodides in resistance switching. *APL Mater.* 2019, 7, 081107. [CrossRef]

42. Becker, M.; Wark, M. Organic Cation Substitution in Hybrid Perovskite CH3NH3PbI3 with Hydroxylammonium (NH3OH+): A First-Principles Study. *J. Phys. Chem. C* 2018, 122, 3548–3557. [CrossRef]

43. Li, W.G.; Rao, H.S.; Chen, B.X.; Wang, X.D.; Kuang, D.B. A formamidinium–methylammonium lead iodide perovskite single crystal exhibiting exceptional optoelectronic properties and long-term stability. *J. Mater. Chem. A* 2017, 5, 19431–19438. [CrossRef]

44. Xiao, Z.; Wang, Q.; Wu, X.; Wu, Y.; Ren, J.; Xiong, Z.; Yang, X. Efficient light-emitting devices based on mixed-cation lead halide perovskites. *Org. Electron.* 2020, 77, 105546. [CrossRef]

45. Wang, Y.; Zhang, T.; Li, G.; Xu, F.; Wang, T.; Li, Y.; Yang, Y.; Zhao, Y. A mixed-cation lead iodide MA1−xEAxPbI3 absorber for perovskite solar cells. *J. Energy Chem.* 2018, 27, 215–218. [CrossRef]

46. Liu, D.; Li, Q.; Wu, K. Ethylammonium as an alternative cation for efficient perovskite solar cells from first-principles calculations. *RSC Adv.* 2019, 9, 7356–7361. [CrossRef]

47. Wang, R.T.; Liu, E.E.; Xu, A.F.; Yang, L.W.; Chen, J.Y.; Xu, G. Ethylammonium Lead Iodide Formation in MAPbI3 Precursor Solutions by DMF Decomposition and Organic Cation Exchange Reaction. *Crystals* 2020, 10, 162. [CrossRef]

48. Jodlowski, A.D.; Roldán-Carmona, C.; Grancini, G.; Salado, M.; Ahmad, S.; Koch, N.; Camacho, L.; de Miguel, G.; Nazeeruddin, M.K.; Zakeeruddin, S.M. Large guanidinium cation mixed with methylammonium in lead iodide perovskites for 19% efficient solar cells. *Nat. Energy* 2017, 2, 972–979. [CrossRef]

49. Tsarev, S.; Boldyreva, A.G.; Luchkin, S.Y.; Elshobaki, M.; Afanasov, M.I.; Stevenson, K.J.; Troshin, P.A. Hydrazinium-assisted stabilisation of methylammonium tin iodide for lead-free perovskite solar cells. *J. Mater. Chem. A* 2018, 6, 21389–21395. [CrossRef]

50. Yu, S.; Liu, H.; Wang, S.; Zhu, H.; Dong, X.; Li, X. Hyridrazinium cation mixed FAPbI3-based perovskite with 1/D3 hybrid dimension structure for efficient and stable solar cells. *Chem. Eng. J.* 2021, 403, 125724. [CrossRef]

51. Saliba, M.; Matsui, T.; Sato, Y.; Domanski, K.; Correa-Baena, J.P.; Nazeeruddin, M.K.; Zakeeruddin, S.M.; Tress, W.; Abate, A.; Hagfeldt, A.; et al. Cesium-containing triple cation perovskite solar cells: Improved stability, reproducibility and high efficiency. *Energy Environ. Sci.* 2016, 9, 1989–1997. [CrossRef]

52. Stoddard, R.J.; Rajagopal, A.; Palmer, R.L.; Braly, I.L.; Jen, A.K.Y.; Hillhouse, H.W. Enhancing Defect Tolerance and Phase Stability of High-Bandgap Perovskites via Guanidinium Alloying. *ACS Energy Lett.* 2018, 3, 1261–1268. [CrossRef]

53. Boziki, A.; Mladenović, M.; Grätzel, M.; Rothlisberger, U. Why choosing the right partner is important: Stabilization of ternary Cs_{y}Gu_{x}A_{(1-y-x)}PbI_{3} perovskites. *Phys. Chem. Chem. Phys.* 2020, 22, 20880–20890. [CrossRef]

54. Ono, L.K.; Juarez-Perez, E.J.; Qi, Y. Progress on Perovskite Materials and Solar Cells with Mixed A-Site Cations and Halide Anions. *ACS Appl. Mater. Interfaces* 2017, 9, 30197–30246. [CrossRef]

55. Zarick, H.F.; Soetan, N.; Erwin, W.R.; Bardhan, R. Mixed halide hybrid perovskites: A paradigm shift in photovoltaics. *J. Mater. Chem. A* 2018, 6, 5507–5537. [CrossRef]

56. McGovern, L.; Futschler, M.H.; Ehrler, B. Understanding the Stability of MAPbBr3 versus MAPbI3: Suppression of Methylammonium Migration and Reduction of Halide Migration. *J. Phys. Chem. Lett.* 2020, 11, 7127–7132. [CrossRef] [PubMed]

57. Pistor, P.; Burwig, T.; Brzuska, C.; Weber, B.; Fränzel, W. Thermal stability and miscibility of co-evaporated methyl ammonium lead halide (MAPbX3, X = I, Br, Cl) thin films analysed by in situ X-ray diffraction. *J. Mater. Chem. A* 2016, 4, 11496–11506. [CrossRef]

58. Knight, A.J.; Borchert, J.; Oliver, R.D.J.; Patel, J.B.; Radaelli, P.G.; Snaith, H.J.; Johnston, M.B.; Herz, L.M. Halide Segregation in Mixed-Halide Perovskites: Influence of A-Site Cations. *ACS Energy Lett.* 2021, 6, 799–808. [CrossRef] [PubMed]

59. Soler, J.M.; Artacho, E.; Gale, J.D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D. The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter* 2002, 14, 2745. [CrossRef]

60. Ceperley, D.M.; Alder, B.J. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.* 1980, 45, 566–569. [CrossRef]

61. Latini, A.; Gigli, G.; Ciccioli, A. A study on the nature of the thermal decomposition of methylammonium lead iodide perovskite, CH3NH3PbI3: An attempt to rationalise contradictory experimental results. *Sustain. Energy Fuels* 2017, 1, 1351–1357. [CrossRef]

62. Scikit-Learn. Available online: https://scikit-learn.org (accessed on 15 July 2021).

63. TensorFlow. Available online: https://www.tensorflow.org (accessed on 15 July 2021).

64. Keras. Available online: https://keras.io/ (accessed on 15 July 2021).
65. Brunetti, B.; Cavallo, C.; Ciccioli, A.; Gigli, G.; Latini, A. On the Thermal and Thermodynamic (In)Stability of Methylammonium Lead Halide Perovskites. Sci. Rep. 2016, 6, 31896. [CrossRef] [PubMed]

66. Valenzano, V.; Cesari, A.; Balzano, F.; Milella, A.; Fracassi, F.; Listorti, A.; Gigli, G.; Rizzo, A.; Uccello-Barretta, G.; Colella, S. Methylammonium-formamidinium reactivity in aged organometal halide perovskite inks. Cell Rep. Phys. Sci. 2021, 2, 100432. [CrossRef]

67. Nazarenko, O.; Kotyrba, M.R.; Yakunin, S.; Aebli, M.; Rainò, G.; Benin, B.M.; Wörle, M.; Kovalenko, M.V. Guanidinium-Formamidinium Lead Iodide: A Layered Perovskite-Related Compound with Red Luminescence at Room Temperature. J. Am. Chem. Soc. 2018, 140, 3850–3853. [CrossRef] [PubMed]