A new approach to design multicomponent metallic glasses using the mendeleev number

Anurag Bajpai\textsuperscript{a}, Jatin Bhatt\textsuperscript{b}, Nilesh P. Gurao\textsuperscript{a} and Krishanu Biswas\textsuperscript{a}

\textsuperscript{a}Department of Materials Science and Engineering, Indian Institute of Technology, Kanpur, Uttar Pradesh, India; \textsuperscript{b}Metallurgical and Materials Engineering Department, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India

ABSTRACT

Designing novel Multicomponent Metallic Glasses (MMGs) based on empirical parameters such as enthalpy of mixing ($\Delta H_{\text{mix}}$) and configurational entropy ($\Delta S_{\text{mix}}$) is a time-consuming exercise that requires various assumptions, limiting the capability to predict new MMG compositions. The current study involves constructing a modified Mendeleev Number ($MNP$) element scale based on many important elemental properties that impact the glass forming phenomena. Machine learning (ML) was used to assess the competence of the proposed $MNP$ to predict MMGs. The ML findings demonstrate that proposed $MNP$ can be utilised as a salient attribute to predict MMGs with 87.8\% cross-validation accuracy. Further, the mean square variation in the $MNP$ of the alloy constituents ($\Delta MNP$) provides a delineated zone of glass forming multicomponent alloys. In summary, the research work presents a novel phenomenological coordinate system that can effectively predict new MMGs while avoiding the limitations of empirical parameters based design strategies.

1. Introduction

Immense scientific curiosity and potential technological interest in new engineering materials have guided the development of complex concentrated materials (CCMs), popularly known as high entropy materials [1–4]. A derivative of the CCMs, multicomponent metallic glasses (MMGs) [5–7], have also emerged as versatile and high-performance materials due to their excellent physical, mechanical, and functional properties [8, 9]. However, the research on MMGs has been slow, primarily due to the scarcity of effective design methodologies to develop and expand the MMG compositional space with new alloy systems. MMGs are mainly designed based on empirical thermodynamic and structural parameters; namely mixing enthalpy ($\Delta H_{\text{mix}}$) [10], relative atomic
size difference ($\delta$) [11] and valence electron concentration (VEC) [12]. Such design approaches require various assumptions to evaluate the empirical parameters, thus limiting their effectiveness in predicting phase formation in multicomponent alloys. Besides, they do not often include information about other important physical factors that play a vital role in understanding the amorphous phase formation in metallic alloys. Therefore, it is crucial to probe new scientific strategies for alloy design to decipher the phenomenon of amorphous phase formation in the multicomponent alloy systems from a fresh perspective.

A coherent chemical space, essentially a coordinate system, in which materials with comparable attributes are tightly connected has not been extensively studied to effectively design and develop new materials. This kind of approach can potentially allow the designing of new MMG compositions with comparable, if not better, properties to the existing ones. Pettifor investigated a similar concept using ‘structure maps’ in 1984, postulating a well-structured chemical space by rearranging the elements in the periodic table [13]. He created a chemical scale based on the assigned Mendeleev number (MN): an integer indicating the position of an element on the scale. According to Pettifor, binary compounds with a similar structural system occupy the same region in a two-dimensional map developed using the MNs. Later, the technique worked successfully for other $A_xB_y$ type compounds as well [14]. Although Pettifor developed the chemical scale and Mendeleev number based on series of experiments conducted on a few hundred binary compounds, the study offered a successful way of ordering the elements, which was later validated by subsequent research [15, 16]. In a similar way, Villars et al. [17] suggested an alternative element enumeration method (called the periodic number, PN) in 2008, stressing more on the relevance of valence electrons in compound formation. Similarly, Allahyari et al. [18] proposed a new Mendeleev Number based on the atomic size and electronegativity that work as a Universal sequence of elements (USE). These studies indicate that such phenomenological coordinates can be utilised to effectively predict the nature of new materials in a computationally inexpensive framework.

In this context, the present investigation proposes a new Mendeleev Number ($MN_P$) based coordinate system as a straightforward, scientifically meaningful, and completely non-empirical method for the design of MMGs. It is anticipated that a non-empirical Mendeleev Number based strategy would outperform the existing empirical parameters and models in predicting the new glass forming alloy compositions. The phase prediction capability of the proposed Mendeleev Number ($MN_P$) scale was investigated using the machine learning (ML) approach. Several ML classification algorithms were trained using a comprehensive and compositionally diverse input dataset of multicomponent alloys. The results validated the hypothesis while providing an excellent estimate of design space using $MN_P$ to propose new MMGs.
2. Method

2.1. The new mendeleev number

In contrast to Pettifor, who developed $MN$ experimentally [13], we designed a non-empirical coordinate system using key non-empirical elemental attributes. In the context of MMGs, glass formation in multicomponent alloys is a complex phenomenon, sensitive to various factors such as material thermodynamics, structural inhomogeneity, and chemical makeup, along with the kinetics involved. Therefore, several critical elemental attributes, namely, atomic radius ($r$), electronegativity ($\chi$), valence, melting temperature ($T_m$), work function ($\phi$) and electron density at the boundary Wigner–Seitz cell ($n_{ws}$) were initially considered for design of the $MNP$ coordinate system from different scientific perspectives. In the present study, Pauling’s electronegativity scale ($\chi_P$) and the metallic radius of the elements were used. We also removed the valence, which is not constant for many elements, for the sake of simplicity and accuracy. The values of $r$, $\chi_P$, $T_m$, $\phi$ and $n_{ws}$ as well as the $MNP$ for the 61 elements are provided in Table 1. The elemental attributes were converted into dimensionless parameters with a value between 0 and 1 through normalisation, using Equation 1. The linear scaling technique was used as the values of the attributes were uniformly distributed, and no segregation of data was observed.

$$x_{\text{norm}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$

where, $x$ is the original value, $x_{\text{norm}}$ is the normalised value, $x_{\text{max}}$ and $x_{\text{min}}$ are the maximum and minimum values of the attribute $x$, respectively. The linear scaling technique allows for similar weights for all elemental attributes, resulting in no bias when considering them together.

Consequently, the arithmetic product of the dimensionless parameters ($r$, $\chi_P$, $T_m$, $\phi$ and $n_{ws}$) for the 61 elements chosen from the periodic table was evaluated. Elements were then arranged based on the product value. The element having the highest product value was designated as $MNP = 1$, and the element with the smallest product value was assigned $MNP = 61$ on the coordinate scale. Obtaining the product of the normalised dimensionless attributes allows their contribution in equal proportions. Machine learning (ML) was subsequently used to evaluate the capability of $MNP$ coordinate system in predicting MMGs.

2.2. The dataset and features

A dataset comprising of 1732 reported multicomponent alloys (Table T1, Supplementary Information) was created by a thorough investigation of data available in the literature. The dataset consisted of 908 MMGs and 824 crystalline multicomponent alloys (CMAs). The $MNP$ coordinate system was used to
Table 1. Values for various elemental attributes and scaling of elements based on the proposed Mendeleev Number (MNP) [19–21].

| Element | $r$ (nm) | $\chi_p$ | $T_m$ (K) | $\phi$ (V) | $n_{sw}$ (d.u.) | $MNP$ |
|---------|---------|----------|----------|----------|----------------|-------|
| Ru      | 0.134   | 2.2      | 2607     | 4.8      | 5.64           | 1     |
| Pt      | 0.139   | 2.28     | 2041     | 5.4      | 6.33           | 2     |
| Nb      | 0.143   | 1.6      | 2750     | 4.2      | 5.55           | 3     |
| Rh      | 0.135   | 2.28     | 2237     | 4.98     | 4.66           | 4     |
| Pd      | 0.138   | 2.2      | 1828     | 5.0      | 5.36           | 5     |
| Cr      | 0.125   | 1.66     | 2180     | 4.4      | 5.36           | 6     |
| Gd      | 0.150   | 1.9      | 1687     | 4.85     | 6.43           | 7     |
| Si      | 0.115   | 2.2      | 3306     | 5.93     | 1.77           | 8     |
| Au      | 0.144   | 2.54     | 1337     | 5.32     | 3.87           | 9     |
| U       | 0.142   | 1.38     | 1405     | 3.7      | 6.43           | 10    |
| Lu      | 0.174   | 1.27     | 1925     | 3.01     | 6.03           | 11    |
| Cu      | 0.128   | 1.9      | 1358     | 4.7      | 5.18           | 12    |
| Tm      | 0.156   | 1.25     | 1818     | 2.86     | 6.13           | 13    |
| Ir      | 0.136   | 2.2      | 2719     | 5.6      | 1.60           | 14    |
| Re      | 0.177   | 2.3      | 1941     | 4.72     | 1.52           | 15    |
| Er      | 0.176   | 1.24     | 1802     | 2.84     | 6.03           | 16    |
| Ho      | 0.177   | 1.23     | 1734     | 2.9      | 6.03           | 17    |
| B       | 0.138   | 2.04     | 2349     | 4.75     | 3.75           | 18    |
| Ta      | 0.143   | 1.5      | 3290     | 4.22     | 2.00           | 19    |
| Tb      | 0.178   | 1.21     | 1629     | 3.0      | 6.13           | 20    |
| Mo      | 0.136   | 2.16     | 2896     | 4.3      | 1.60           | 21    |
| Ti      | 0.146   | 1.54     | 1941     | 4.1      | 3.38           | 22    |
| Dy      | 0.177   | 1.22     | 1680     | 2.75     | 6.03           | 23    |
| Fe      | 0.124   | 1.83     | 1811     | 4.65     | 3.18           | 24    |
| Zr      | 0.160   | 1.33     | 2128     | 4.0      | 3.18           | 25    |
| Sm      | 0.181   | 1.17     | 1345     | 2.95     | 6.33           | 26    |
| C       | 0.077   | 2.55     | 3823     | 5.0      | 1.56           | 27    |
| La      | 0.188   | 1.1      | 1193     | 3.4      | 6.13           | 28    |
| Nd      | 0.164   | 1.14     | 1297     | 3.1      | 6.13           | 29    |
| Y       | 0.180   | 1.22     | 1799     | 3.1      | 4.33           | 30    |
| Hf      | 0.158   | 1.3      | 2506     | 3.51     | 2.57           | 31    |
| Pr      | 0.165   | 1.13     | 1208     | 2.7      | 6.33           | 32    |
| V       | 0.132   | 1.63     | 2183     | 4.44     | 1.91           | 33    |
| Tc      | 0.183   | 1.9      | 2430     | 3.0      | 2.05           | 34    |
| Ag      | 0.145   | 1.93     | 1235     | 4.64     | 2.69           | 35    |
| Ce      | 0.183   | 1.12     | 1068     | 2.8      | 5.93           | 36    |
| W       | 0.137   | 2.36     | 3695     | 4.55     | 0.55           | 37    |
| Co      | 0.125   | 1.88     | 1768     | 4.7      | 1.91           | 38    |
| Ca      | 0.198   | 1       | 1115     | 2.71     | 6.86           | 39    |
| Pb      | 0.175   | 2.33     | 601      | 4.18     | 4.25           | 40    |
| Ba      | 0.218   | 0.89     | 1000     | 4.25     | 4.41           | 41    |
| Zn      | 0.140   | 1.65     | 693      | 4.3      | 5.93           | 42    |
| Bi      | 0.143   | 2.02     | 544      | 5.08     | 4.10           | 43    |
| Sr      | 0.215   | 0.95     | 1050     | 2.76     | 6.13           | 44    |
| Ge      | 0.124   | 2.01     | 1211     | 5.0      | 2.25           | 45    |
| Sn      | 0.162   | 1.96     | 505      | 4.35     | 5.45           | 46    |
| Yb      | 0.170   | 1.26     | 1097     | 2.45     | 5.09           | 47    |
| Al      | 0.143   | 1.61     | 933      | 4.19     | 2.69           | 48    |
| Ga      | 0.139   | 1.81     | 303      | 4.25     | 5.55           | 49    |
| In      | 0.166   | 1.78     | 430      | 4.08     | 2.92           | 50    |
| Mn      | 0.135   | 1.55     | 1519     | 3.9      | 0.94           | 51    |
| Ni      | 0.125   | 1.91     | 1728     | 5.15     | 0.55           | 52    |
| Mg      | 0.160   | 1.31     | 923      | 3.66     | 1.30           | 53    |
| P       | 0.106   | 2.19     | 317      | 5.1      | 1.52           | 54    |
| Na      | 0.190   | 0.93     | 371      | 2.36     | 4.17           | 55    |
| Cd      | 0.161   | 1.69     | 594      | 4.08     | 0.75           | 56    |
| Sc      | 0.164   | 1.36     | 1814     | 3.5      | 0.22           | 57    |
| Be      | 0.113   | 1.57     | 1560     | 2.32     | 0.53           | 58    |
| Rb      | 0.265   | 0.82     | 312      | 2.261    | 2.30           | 59    |
| Li      | 0.152   | 0.98     | 454      | 3.1      | 0.27           | 60    |
create the features space consisting of compositionally weighted average $MNP$ (Equation 2) as well as the mean squared difference in $MNP$ (Equation 3) as input parameters. The elemental attributes used to design the $MNP$ coordinate system were not included in the feature space for ML modelling.

$$\bar{MNP} = \sum_{i=1}^{n} c_i (MNP)_i$$ \hfill (2)

$$\Delta MNP = \sum_{i=1}^{n} c_i [\bar{MNP} - (MNP)_i]^2$$ \hfill (3)

Here, $c_i$ is the atomic composition and $(MNP)_i$ is the $MNP$ coordinate of $i^{th}$ element. The input dataset was randomly divided into ‘training set’ and ‘test set’ in the ratio of 80:20. However, complete randomisation in the distribution of training and test datasets is not always preferable. There can be a case where the model learns from a collection of data points and then applies the trained model to another comparable set of data points. As a result, the evaluation of model performance is skewed [22]. Thus, it is critical to build the test dataset having sufficient dissimilarity to the training dataset. Therefore, following the random distribution of input data into train and test datasets, the alloys with values of both $\bar{MNP}$ and $\Delta MNP$ in the test dataset overlapping or similar to the training dataset were replaced with the same number of alloys from the input dataset that had discrete values for $\bar{MNP}$ and $\Delta MNP$. The training and test datasets were thus designed to get a considerable number of alloys in the test dataset had distinct values of $\bar{MNP}$ and $\Delta MNP$ from the ones in the training dataset. This increased the heterogeneity between the training and testing datasets. The test set was later used to perform an unbiased evaluation of the trained models as well as the $MNP$ coordinate system.

2.3. The ML models

Several ML classifiers, namely Logistic Regression (LR), K-Nearest Neighbours (KNN), Support Vector Machine (SVM) and Random Forest (RF), were considered to develop an efficient ML model. Logistic regression (LR) is a classification algorithm that assigns observations to different classes based on the concept of probability. The LR algorithm uses the sigmoid function

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

as the activation function [23], which provides a probability score between 0 and 1 as output for the given input. The LR classifier produces a set of classifications depending on the likelihood or probability threshold of different classes [23–25]. The cross-entropy loss is used as the cost function, which is used to analyze and minimise the deviation between the predicted and actual output.
K-Nearest Neighbours (KNN) is a supervised machine learning technique used for both classification and regression problems. In the context of classification, the KNN classifier uses the majority voting principle to identify the class of data points [23, 25]. The value of K determines the number of neighbours being looked at while deciding the target data point class. The Euclidean distance is used to determine the nearest neighbours, and the target data point is categorised based on the class of K nearest neighbours [23, 26].

Support Vector Machine (SVM) classification algorithm is built on the principle of locating a hyperplane that optimally splits the data into distinct domains [23]. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane are attributed to different classes. The algorithm works with the objective of finding a plane with the maximum margin, i.e., the maximum distance between data points of both classes. The hyperplane is decided based on support vectors, which are data points nearest to the hyperplane [23, 25, 27].

Random Forest (RF) is an ensemble of a number of individual decision trees that work together. Each individual tree produces a class prediction using a series of branched conditions, and the class with most votes becomes the final prediction. The RF algorithm is an improvement on decision trees based algorithms [23]. Decision trees are simple and flexible but are considered greedy algorithms since the focus of each node is on its own optimal output. Thus, while each node in a decision tree provides the most optimal solution locally, the algorithm fails to provide the global optimal solution or output [28]. The RF model, on the other hand, consists of a large number of relatively uncorrelated trees operating as a combination that protect each other from their individual errors. This is done using the technique called ‘bagging’ or bootstrap aggregation [23, 29]. Each tree in the RF learns from a random selection of data points during training. The samples are drawn with replacement, implying that certain samples will be used several times in a single tree. The theory is that by training each tree on diverse samples, even if each tree has a large variation with regard to a specific set of training data, the total variance of the forest will be smaller, but not at the expense of increasing bias. Predictions are made on the dataset by averaging the predictions of each decision tree [25, 29].

The ML models were trained using the 6-fold cross-validation method, in which the training set was randomly partitioned into six disjoint subsets, five of which were utilised for training, and the remaining one was used for validation. Grid search algorithm was employed to optimise the hyperparameters for the models. Python 3.8 software was used to perform the calculations as well as training and assessment of the ML models with ‘scikit-learn’ library [23]. The performance of trained ML models was evaluated using several metrics such as accuracy, confusion matrix, precision, recall, F1 score and ROC curve.
3. Results and discussion

3.1. The proposed mendeleev number (MNp)

Figure 1 shows the relation between the elemental attributes \((r, \chi_P, T_m, \phi\) and \(n_{ws}\)) in binary combinations. The data has been linearly fit and \(R^2\) score for the binary correlation between the features is provided in the figure. The \(R^2\) values suggest that there is some, albeit weak, association among the attributes. Therefore, their combination can be employed as a single attribute to predict the properties of an element. From the scientific perspective, the chosen elemental attributes can explain glass forming characteristics of the multicomponent alloys from various viewpoints. Earlier studies by Inoue [30], Egami and Waseda [31] and Miracle [32] have emphasised the importance of topological consideration for designing MMGs by stabilising atomic clusters found in amorphous materials [33]. A large atomic size mismatch constrains the solubility of constituent elements in the competing crystalline phases as redistribution of atoms from liquid to solid is required during the crystallization process. The redistribution process can alter the composition and interfacial energy at the solid/liquid interface. This slows the nucleation process through structural influence and, as a result, can promote glass formation. Further, an adequate atomic size distribution can alter the cluster packing of the undercooled melts by improving their packing densities, lowering the ground-state energy, and stabilising the undercooled liquids, leading to an increase in GFA [31]. Moreover, a high internal strain originating from the difference in atomic size also destabilises the crystalline matrix [31]. Thus, the atomic size difference assessed in terms of the atomic radius of the constituent elements can be an essential feature in explaining the structural modalities in MMGs. According

![Figure 1](image-url)
to the Hume–Rothery principles for solid solution formation, the electronegativity difference among alloy constituents reflects the bonding behaviour of atomic pairs [34]. An appropriate mismatch in electronegativity favours the formation of specific atomic pairs (or atomic clusters), which restricts the solubility among the constituent elements in the competing crystalline phases resulting in better glass-forming ability (GFA) [35].

The atomic attraction between the atoms of the constituent elements in MMG is dependent on their work function ($\phi$). The atomic attraction promotes dense packing and is crucial to glass formation. Thus, a significant difference in the work function between the constituent elements is an ideal condition for generating high viscosity, improving the GFA in multicomponent alloys [36]. Additionally, a negative enthalpy of mixing ($\Delta H_{\text{mix}}$) also promotes attraction among atoms and the formation of dense packed clusters (icosahedral type). According to Miedema’s model, the $\Delta H_{\text{mix}}$ for the interaction between the constituent elements in an alloy is mainly decided by the quantity

$$
\left[ \frac{Q}{P} \times (\Delta n_{\text{ws}})^2 - (\Delta \phi)^2 \right]
$$

where $Q/P$ is a constant independent of alloy composition [37]. Here, $\phi$ is the work function of an element, and $\Delta \phi$ relates to the difference in the chemical potential between the two elements. This is analogous to the electronegativity difference. Therefore, the sign of $\Delta H_{\text{mix}}$ for an alloy is primarily dependent on the relative value of $\Delta n_{\text{ws}}$ and $\Delta \phi$. The alloys having a relatively smaller difference in the electronic density at the boundary of the Wigner–Seitz cell ($\Delta n_{\text{ws}}$) and a considerable difference in their work function results in higher GFA [37]. In addition, the electronic density difference at the boundary of the Wigner–Seitz cell ($\Delta n_{\text{ws}}$) explains the transfer of charges during alloy formation. A large value of $\Delta n_{\text{ws}}$ implies incompatible atomic combinations that repel each other, leading to poor atomic packing and reduced GFA [37, 38]. Furthermore, melting temperature has an important contribution to kinetic constraints that can influence the glass formation phenomenon. During solidification at glass transition, the formation of short-range ordered clusters is also determined by the solvent–solvent atoms bonding among the solute-centered clusters. A lower melting temperature allows for slower cooling rates while bypassing crystallization. It provides sufficient time for the solvent and solute atoms to rearrange themselves in tightly packed short-range ordered icosahedral clusters, enhancing the GFA [39]. Based on the above discussion, the proposed approach includes key elemental attributes that can incorporate thermodynamics, topology, chemical, and kinetics of glass formation, directly or indirectly, into a single phenomenological coordinate to design new MMGs.

### 3.2. Evaluation of the ML models

Prior to training and testing any ML model, the statistical information from the input dataset must be extracted to evaluate the input features space [40]. Figure
Figure 2. A $2 \times 2$ scatter matrix describing the quantitative nature of input features.

Figure 2 shows a $2 \times 2$ scatter matrix describing the quantitative nature of input features. The diagonal subplots show the distribution of MMG and CMA phases for only one feature at a time. Off-diagonal subplots show the distribution of two phases between the pair of input features. As can be seen from the diagonal plots, both features, especially $\Delta MN_p$, provide independent regions for MMGs and CMAs along with the overlapping regions where both CMAs and MMGs coexist. The overlapping nature of the diagonal plots shows that a single parameter is insufficient for both MMGs and CMAs and needs both features for developing the ML algorithms. This can be understood by the fact that a single feature ($\overline{MN}_p$ or $\Delta MN_p$) does not have clearly delineated regions for MMGs and CMAs. However, considering both the features ($\overline{MN}_p$ and $\Delta MN_p$) allows for a better prediction since they have different statistical ranges for MMGs and CMAs. The ML models learn from these numerical differences while creating statistical boundaries within the otherwise overlapped regions to categorise MMGs and CMAs. Therefore, the prediction
of MMGs must be carried out using both the features, i.e., $MNP$ and $\Delta MN_p$, together.

The performance of the trained ML models was evaluated using confusion matrix, precision, recall, F1 score and ROC curve. The confusion matrices for the trained models are shown in Figure 3a-d. It is a matrix representing the performance of a machine learning classifier with two or more classes as output. There are four different combinations of predicted and actual classes based on which the training data is classified as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TPs are the number of correct predictions and FNs are the incorrect predictions of the target class. Conversely, TNs are the number of correct predictions, while FPs are the incorrect predictions of the non-target class [25]. The confusion matrices of the models indicate the cross-validation accuracy of 76.9%, 81.8%, 84.9% and 87.8% for the Logistic Regression, Support Vector Machine, K-Nearest Neighbours and Random Forest model, respectively (Figure 3e).

Table 2 provides the hyperparameter grid, best performing hyperparameters, precision, recall and F1 score for the trained models. Precision is defined as the ratio of True Positives to all Positives, while recall is the measure of the ability of a model to accurately predict True Positives. Table 2 shows that the RF model outperforms the other models based on referred metrics. The precision of the RF model in predicting MMGs is 90.2, suggesting that approximately 90% of predictions in MMG class are accurate. On the other hand, recall of the RF

![Figure 3](image-url)

Figure 3. Confusion matrix for (a) Logistic Regression (LR), (b) K-Nearest Neighbours (KNN), (c) Support Vector Machine (SVM), (d) Random Forest (RF) model; (d) cross-validation accuracy; and (e) ROC curve for various ML models.
model for MMG class is 90.5, signifying the accurate identification of $\sim$91% of actual MMG compositions. The F1 score (given as the harmonic mean of precision and recall) for the prediction of MMG class using the RF model is 0.9, indicating the high accuracy of the RF model for MMGs classification among the selected models. A high precision, recall and F1 score for the trained RF classifier (> 90%) substantiate its superior prediction capability over the other models. Differentiation in the performance of various ML models lies in their ability to accurately classify MMGs and CMAs from the overlapping regions in diagonal plots in Figure 2. Training results for all the models suggest that the alloys that are incorrectly predicted come from the overlapped regions. The majority of incorrectly predicted alloys either lie in the central region of overlapped areas where the values of $\overrightarrow{MNN}$ and $\Delta MNN$ are similar to each other or the alloys away from the populated region. The receiver operating characteristic (ROC) curves for the selected models on training dataset are shown in Figure 3f. The ROC curve is obtained by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). For a classifier that picks either class with equal probability, the ROC curve coincides with the line connecting the coordinates (0, 0) and (1, 1) and has an area under the curve (AUC)
The ROC curve of a classifier with extraordinary classification capability arches towards the point (0, 1), indicating that it is the ideal classifier. As evident from Figure 3f, the trained RF model has the largest value of AUC \( = 0.93 \). This analysis clearly suggests that the RF model outperforms the other classifiers.

The successful prediction of MMGs and CMAs depends on both the chosen features and the model. Different ML models can interpret the input dataset differently. While learning from the data points in the overlapped regions in Figure 2, the LR, KNN, and SVM models fail to perform better than RF because of their inherent characteristics. The major limitation of the LR model is the assumption of linearity between the input features and output. However, this is not true in the present case, and thus it performs the poorest among the chosen models. KNN, on the other hand, uses the ‘similar to neighbours’ approach in classifying unknown data points. In the present case, the overlapped regions of MMGs and CMAs are quite significant. Thus, the KNN algorithm has limited capability to accurately sort such data points within the overlapped region with fewer features. Similarly, the SVM model seeks to generate a hyperplane to categorise data points. However, the SVM model performs relatively poor when target classes overlap significantly because the features might have extreme similarity or overlapping values. Because of the nature of the SVM algorithm, this leads to numerous local optima and, as a result, difficulty in determining the optimal hyperplane for categorisation, as in the present case. The RF algorithm performs the best among the models because of its two inherent characteristics. The first is the ‘bootstrap aggregation’ approach. This involves allowing each individual tree to randomly sample from the input dataset with replacement, resulting in different trees. In the present case, it allows the trees to create a subset from the overlapped and the independent regions, thereby producing a more accurate outcome from each individual tree. The second is the ability of each tree in a random forest to pick a random subset of features. In the case of binary features, it can pick one feature, run the decision tree, and then pick the other, thus producing a more accurate prediction. This can be further visualised using Figure 4, which shows the decision boundaries constructed by various models during training. The area surrounded in pink corresponds to the CMA region, while the area covered in blue corresponds to MMGs. The green markers are representative of MMGs, while the red markers signify CMAs. It is evident from the figures that LR depends on linear correlation and thus forms a linear boundary which is the poorest way to segregate data in the overlapped region. The KNN and SVM models do perform better in designing non-linear classification boundaries but fail when the target classes are significantly overlapping as well as for discrete datapoints that lie far away from the populated region. The figures also show that the decision boundary starts with a linear separator for the LR algorithm and continues to reshape into a more precise form for
KNN and SVM models, respectively. However, in comparison to the KNN and SVM models, the RF model produces a much more precise decision boundary without any outlying regions indicating RF as the most accurately fit model among the four models. The RF model is capable of tuning the decision boundary with smaller subsets of data at a time and thus generates a more accurate prediction boundary with significantly less overfit area.

Overfitting is the most typical problem that any machine learning model encounters, which occurs when an ML model can represent the training set too well and fails to be relevant while testing unknown data. This is because the model learns from both the detail and noise of the training data set, which reduces testing accuracy. We employed $k$-fold cross-validation along with hyperparameter optimisation to solve the issue of overfitting. The number of folds varied from $k=1$ to $k=8$ and the hyperparameters were tuned across the $k$-folds partitioning using the grid search algorithm. Figure 5a shows the learning curves of the training and validation datasets for the RF model across the $k$-fold partitioning prior to hyperparameter optimisation. It can be noted that the cross-validation accuracy, ranging between 0.41–0.61, is considerably poor in comparison to the training accuracy. This indicates that the RF model is overfitted. Figure 5b shows the learning curves of training and validation datasets for the RF model across the $k$-fold partitioning post hyperparameter optimisation. Contrary to Figure 5a, the difference in the training and the cross-validation accuracy in Figure 5b considerably reduces as the number of folds increases from one to
eight and tends to plateau at six folds, where the training and cross-validation accuracies are comparable. This indicates that the RF model is not overfitted post hyperparameter optimisation.

3.3. Evaluation of the MNₚ coordinate system

The trained RF model shows that MNₚ can be used as an independent feature to design MMGs. However, a comprehensive assessment of the ML model as well as the MNₚ coordinate system necessitates their employment on an unknown dataset. This has been achieved using the test set consisting of 318 alloy compositions. The test set was made of 156 reported CMAs and 162 MMGs. These compositions have been provided in Table T2 (Supplementary Information). The trained RF model was employed to predict the possibility of MMG formation for the test dataset. The results suggested a prediction accuracy of 87.7% for MMGs. The confusion matrix of the RF model for the test dataset is shown in Figure S1 (Supplementary Information). The precision, recall and F1 score for the RF model on the test dataset are 0.886, 0.882 and 0.885, respectively. These results indicate that the performance of RF model for test dataset is comparable to training dataset. The reported and the predicted phases for the test dataset have also been provided in Table T2 (Supplementary Information).

Figure 6 shows the ∆MNₚ vs. MNₚ plot overlapped with the phase predictions by the RF model over the reported phase for the test dataset. As evident from the figure, there is no apparent delineation between MMGs and CMAs in the context of MNₚ. The variation of ∆MNₚ, on the other hand, shows distinct zones from MMGs and CMAs. The majority of MMGs (129 of 162) form when ∆MNₚ follows $3 \leq \Delta MN_p \leq 15$. Conversely, $\Delta MN_p > 15$ aids in the formation of CMAs, with 107 of 156 CMAs lying in this region. The accuracy of this segregation for MMGs is 80%, whereas it is 68% for
CMAs. Besides, the RF model accurately classifies 120 out of 132 MMGs in the region $3 \leq \Delta MN_p \leq 15$ and 105 of the 110 CMAs in the region $\Delta MN_p > 15$. This suggests that the RF model has a 91% accuracy in segregating MMGs and CMAs correctly within the proposed regions. This is a significant observation in terms of glass formation. Although the above analysis is based on a few hundred alloys, it provides new insights for visualising the phenomena of glass formation using a property based coordinate system. As previously stated, the design of the $MN_p$ scale is based on the collective interaction of the selected material attributes. Furthermore, the association between the GFA and all elemental features does not follow similar trends. While a high value of $\Delta \phi$ stimulates glass formation, a high $\Delta n_{ws}$ reduces the GFA. Similarly, a larger atomic size mismatch promotes crystalline phase instability. However, beyond a certain point, it results in the formation of intermetallics and metastable phases over the amorphous phase [41]. As a result, while selecting appropriate constituent elements and compositions for MMGs, a balance is required between these attributes. This is precisely what the designed $MN_p$ scale accomplishes by delineating the region for the majority of MMG compositions by finding the appropriate balance between the selected elemental properties, which is crucial in explaining the glass formation phenomenon in multicomponent alloys. Besides, machine learning is merely the tool used here to show the efficacy of $MN_p$ for designing MMGs. The designed $MN_p$, a non-empirical coordinate, can also be used with other computational tools to design MMGs.

4. Conclusions

The present study reports a new non-empirical Mendeleev Number ($MN_p$) coordinate system based on key elemental attributes which influence the
glass formation in multi-component alloys. The major conclusions made from the present study are summarised as follows:

- Empirical parameters conventionally used to design multi-component metallic glasses have inherent assumptions associated with them, which limit their universal applicability. In contrast, the proposed $M\text{NP}_p$ scale utilises the non-empirical elemental attributes, circumventing the limitations of the empirical parameters to explain the glass formation.
- The present study also highlights that the electronic and chemical nature of constituents and the kinetics of glass formation significantly impact the phenomenon of glass formation in multi-component alloys.
- Machine Learning modelling results suggest that the mean squared difference in $M\text{NP}_p$ ($\Delta M\text{NP}_p$) is effective in differentiating between MMGs and CMAs, establishing $\Delta M\text{NP}_p$ as a prudent material feature to design multi-component alloys with high throughput.
- The present work provides a new approach which can address the limitations of using empirical parameters to design multi-component metallic glasses and drive their future development.

**Author contributions**

A.B. conceptualised the idea, performed calculations, developed the Machine Learning model, analyzed the results, and wrote the manuscript. J.B., N.P.G., and K.B. supervised and reviewed the manuscript.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**References**

[1] A.S. Sharma, S. Yadav, K. Biswas and B. Basu, *High-entropy alloys and metallic nanocomposites: Processing challenges, microstructure development and property enhancement*. Materials Science and Engineering: R: Reports 131 (2018), pp. 1.
[2] N.P.G. Tazuddin and K. Biswas, *In the quest of single phase multi-component multi-principal high entropy alloys*. J. Alloys Compd. 697 (2017), pp. 434.
[3] N.P. Gurao and K. Biswas, *High Entropy Materials: Critical Review and Way Forward*. Curr. Sci. 118(10) (2020), pp. 1520.
[4] K. Biswas, J.-W. Yeh, P.P. Bhattacharjee and J.T.M. DeHosson, *High entropy alloys: Key issues under passionate debate*. Scr. Mater. 188 (2020), pp. 54.
[5] C.A. Schuh, T.C. Hufnagel and U. Ramamurti, *Mechanical behavior of amorphous alloys*. Acta Mater. 55 (2007), pp. 4067.
[6] M. Chen, *A brief overview of bulk metallic glasses*. NPG Asia Mater. 3 (2011), pp. 82.
[7] A. Bajpai, J. Bhatt, K. Biswas and N.P. Gurao, *A new perspective to thermodynamical designing of high entropy bulk metallic glasses (HE-BMGs)*. Phys. B 595 (2020), pp. 412350.

[8] Y. Wang and J. Xu, *Toughen and harden metallic glass through designing statistical heterogeneity*. Scr. Mater. 113 (2016), pp. 10.

[9] Q. Zhou, Y. Ren, Y. Du, W. Han, D. Hua, H. Zhai, P. Huang, F. Wang and H. Wang, *Identifying the significance of Sn addition on the tribological performance of Ti-based bulk metallic glass composites*. J. Alloys Compd. 780 (2018), pp. 671–679.

[10] A. Takeuchi and A. Inoue, *Classification of Bulk Metallic Glasses by Atomic Size Difference, Heat of Mixing and Period of Constituent Elements and Its Application to Characterization of the Main Alloving Element*. Mater. Trans. 46 (2005), pp. 2817.

[11] Y. Zhang, Y.J. Zhou, J.P. Lin, G.L. Chen and P.K. Liaw, *Solid-Solution Phase Formation Rules for Multi-component Alloys*. Adv. Eng. Mater. 10 (2008), pp. 534.

[12] S. Guo, C. Ng, J. Lu and C.T. Liu, *Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys*. J. Appl. Phys. 109 (2011), pp. 103505.

[13] D.G. Pettifor, *A chemical scale for crystal-structure maps*. Solid State Commun. 51 (1984), pp. 31.

[14] D.G. Pettifor, *Structure maps in alloy design*. J. Chem. Soc., Faraday Trans. 86 (1990), pp. 1209.

[15] P. Villars, K. Brandenburg, M. Berndt, S. LeClair, A. Jackson, Y.H. Pao, B. Igelnik, M. Oxley, B. Bakshi, P. Chen and S. Iwata, *Binary, ternary and quaternary compound former/nonformer prediction via Mendeleev number*. J. Alloys Compd. 317-318 (2001), pp. 26.

[16] H. Glawe, A. Sanna, E.K.U. Gross and M.A.L. Marques, *The optimal one dimensional periodic table: a modified Pettifor chemical scale from data mining*. New J. Phys. 18 (2016), pp. 093011.

[17] P. Villars, J. Daams, Y. Shikata, K. Rajan and S. Iwata, *A new approach to describe elemental-property parameters*. Chemistry of metals alloys (2008), pp. 1–3.

[18] Z. Allahyari and A.R. Oganov, *Nonempirical Definition of the Mendeleev Numbers: Organizing the Chemical Space*. J. Phys. Chem. C 124 (2020), pp. 23867.

[19] F.R.d. Boer, W.C.M. Mattens, R. Boom, A.R. Miedema and A.K. Niessen, *Cohesion in metals*, North-Holland, Netherlands, 1988.

[20] O. Senkov, D. Miracle, O.N. Senkov and D.B. Miracle, *Effect of the atomic size distribution on glass forming ability of amorphous metallic alloys*. Mater. Res. Bull 36 (2001), pp. 2183.

[21] WebElements. The periodic table on the web (URL: http://www.webelements.com).

[22] L. Ward, S.C. O’Keeffe, J. Stevick, G.R. Jelbert, M. Aykol and C. Wolverton, *A machine learning approach for engineering bulk metallic glass alloys*. Acta Mater. 159 (2018), pp. 102.

[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *Scikit-learn: Machine Learning in {P}ython*. J. Mach. Learn. Res. 12 (2011), pp. 2825.

[24] Y.V. Krishna, U.K. Jaiswal and R.M. R, *Machine learning approach to predict new multiphase high entropy alloys*. Scr. Mater. 197 (2021), pp. 113804.

[25] R. Choudhary and H.K. Gianey, *Comprehensive Review On Supervised Machine Learning Algorithms*, 2017 International Conference on Machine Learning and Data Science (MLDS), 2017. 37.

[26] R. Mitra, A. Bajpai and K. Biswas, *Machine learning based approach for phase prediction in high entropy borides*. Ceram. Int. (2022), pp. 16695–16706.
[27] Z. Li, Z. Long, S. Lei, T. Zhang, X. Liu and D. Kuang, Predicting the glass formation of metallic glasses using machine learning approaches. Comput. Mater. Sci. 197 (2021), pp. 110656.

[28] B. Akhil, A. Bajpai, N.P. Gurao and K. Biswas, Designing hexagonal close packed high entropy alloys using machine learning. Modell. Simul. Mater. Sci. Eng. 29 (2021), pp. 085005.

[29] A. Mishra, L. Kompella, L.M. Sanagavarapu and S. Varam, Ensemble-based machine learning models for phase prediction in high entropy alloys. Comput. Mater. Sci. 1-9 (2021), pp. 111025.

[30] A. Inoue, Stabilization of metallic supercooled liquid and bulk amorphous alloys. Acta Mater. 48 (2000), pp. 279.

[31] T. Egami and Y. Waseda, Atomic size effect on the formability of metallic glasses. J. Non-Cryst. Solids 64 (1984), pp. 113.

[32] D.B. Miracle, A structural model for metallic glasses. Nat. Mater. 3 (2004), pp. 697.

[33] T. Fujita, K. Konno, W. Zhang, V. Kumar, M. Matsuura, A. Inoue, T. Sakurai and M.W. Chen, Atomic-Scale Heterogeneity of a Multicomponent Bulk Metallic Glass with Excellent Glass Forming Ability. Phys. Rev. Lett. 103 (2009), pp. 075502.

[34] Z.P. Lu, C.T. Liu and Y.D. Dong, Effects of atomic bonding nature and size mismatch on thermal stability and glass-forming ability of bulk metallic glasses. J. Non-Cryst. Solids 341 (2004), pp. 93.

[35] W.Y. Liu, H.F. Zhang, A.M. Wang, H. Li and Z.Q. Hu, New criteria of glass forming ability, thermal stability and characteristic temperatures for various bulk metallic glass systems. Mater. Sci. Eng. A 459 (2007), pp. 196.

[36] A.R. Miedema, F.R. De Boer and R. Boom, Predicting heat effects in alloys. Physica B +C 103 (1981), pp. 67.

[37] A.K. Niessen, F.R. de Boer, R. Boom, P.F. de Châtel, W.C.M. Mattens and A.R. Miedema, Model predictions for the enthalpy of formation of transition metal alloys II. Calphad 7 (1983), pp. 51.

[38] L.P. Déo, M.A.B. Mendes, A.M.S. Costa, N.D. Campos Neto and M.F. de Oliveira, Applying a new criterion to predict glass forming alloys in the Zr–Ni–Cu ternary system. J. Alloys Compd. 553 (2013), pp. 212.

[39] C.R. Cao, D.W. Ding, D.Q. Zhao, E. Axinte, H.Y. Bai and W.H. Wang, Correlation between glass transition temperature and melting temperature in metallic glasses. Mater. Des. 60 (2014), pp. 576.

[40] N. Islam, W. Huang and H.L. Zhuang, Machine learning for phase selection in multi-principal element alloys. Comput. Mater. Sci. 150 (2018), pp. 230.

[41] Q. Hu, S. Guo, J.M. Wang, Y.H. Yan, S.S. Chen, D.P. Lu, K.M. Liu, J.Z. Zou and X.R. Zeng, Parametric Study of Amorphous High-Entropy Alloys formation from two New Perspectives: Atomic Radius Modification and Crystalline Structure of Alloying Elements. Sci. Rep. 7 (2017), pp. 39917.