Supplementary Information for A Physics-Regularized Machine Learning Framework for Predicting Time–Temperature–Transformation Curves in Alloys: Application to Uranium-Based Alloys

Sunidhi Garg[1], **Jishnu Bhattacharyya**[1], **Vineet V. Joshi**[2], **Sean R. Agnew**[1], **and Prasanna V. Balachandran**[1,*]

[1]University of Virginia, Department of Materials Science and Engineering, Charlottesville, 22903, Virginia, USA
[2]Pacific Northwest National Laboratory, Richland, 99352, Washington, USA
*pvb5e@virginia.edu

1 Methods

1.1 Dataset Curation

The 39 distinct alloys TTT curves collected for different U-Mo-X alloys are listed below. 38 alloys are used and studied in this work (U-12Mo is excluded).

Table 1. Alloy compositions for the experimentally calculated TTT curves data used in this study. * TTT curves calculated by Repas, Steiner and Peterson[1–3]. ** TTT curves calculated by Repas and Peterson. *** TTT curve calculated by Peterson et al. and is excluded from this study for training the ML models. The U-10Mo TTT curve calculated by Peterson et al. is also excluded from this study for training the ML models.

U-3Mo	U-6.3Mo-1.8Nb	U-5.3Mo-1.9Ru
U-5Mo	U-7.2Mo-0.088Nb	U-6.3Mo-1.9Ru
U-7Mo	U-7.3Mo-1.1Nb	U-7.2Mo-0.95Ru
U-7.13Mo	U-5.3Mo-0.48Cr	U-5Mo-Zr
U-8.13Mo	U-6.3Mo-0.49Cr	U-5Mo-3Zr
U-8Mo*	U-7.2Mo-0.29Cr	U-7Mo-Zr
U-10Mo**	U-7.3Mo-0.49Cr	U-7Mo-2Zr
U-12Mo***	U-6.2Mo-1.7Re	U-7Mo-3Zr
U-5Mo-5Nb	U-6.2Mo-3.4Re	U-7.2Mo-0.86Zr
U-4.4Mo-1.7Nb	U-7.2Mo-1.7Re	U-10Mo-Zr
U-5.3Mo-0.85Nb	U-5.3Mo-0.93Ru	U-10Mo-2Zr
U-5.4Mo-1.7Nb	U-4.4Mo-1.9Ru	U-10Mo-4Zr
U-6.2Mo-0.86Nb	U-6.2Mo-0.49Ru	U-10Mo-6Zr

1.2 Feature Extraction

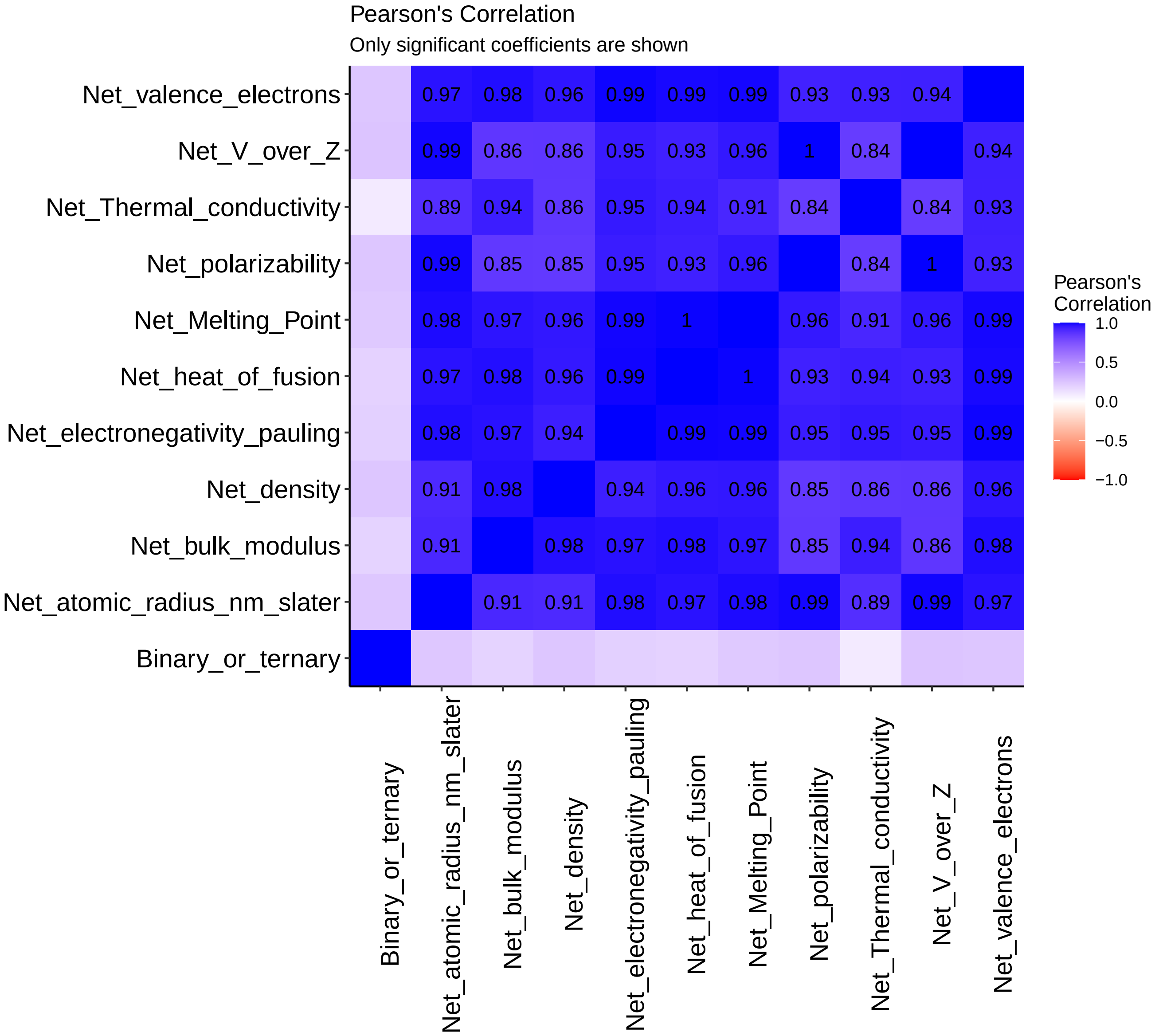

Figure 1. Pearson's correlation coefficient between the 11 descriptors used in this work. Each elemental descriptor was multiplied by its respective weight percentage in the alloy, excluding uranium, and averaged over the total composition to get the net descriptor value. All the elemental descriptor excluding N_C are highly correlated with each having a PCC value > 0.8.

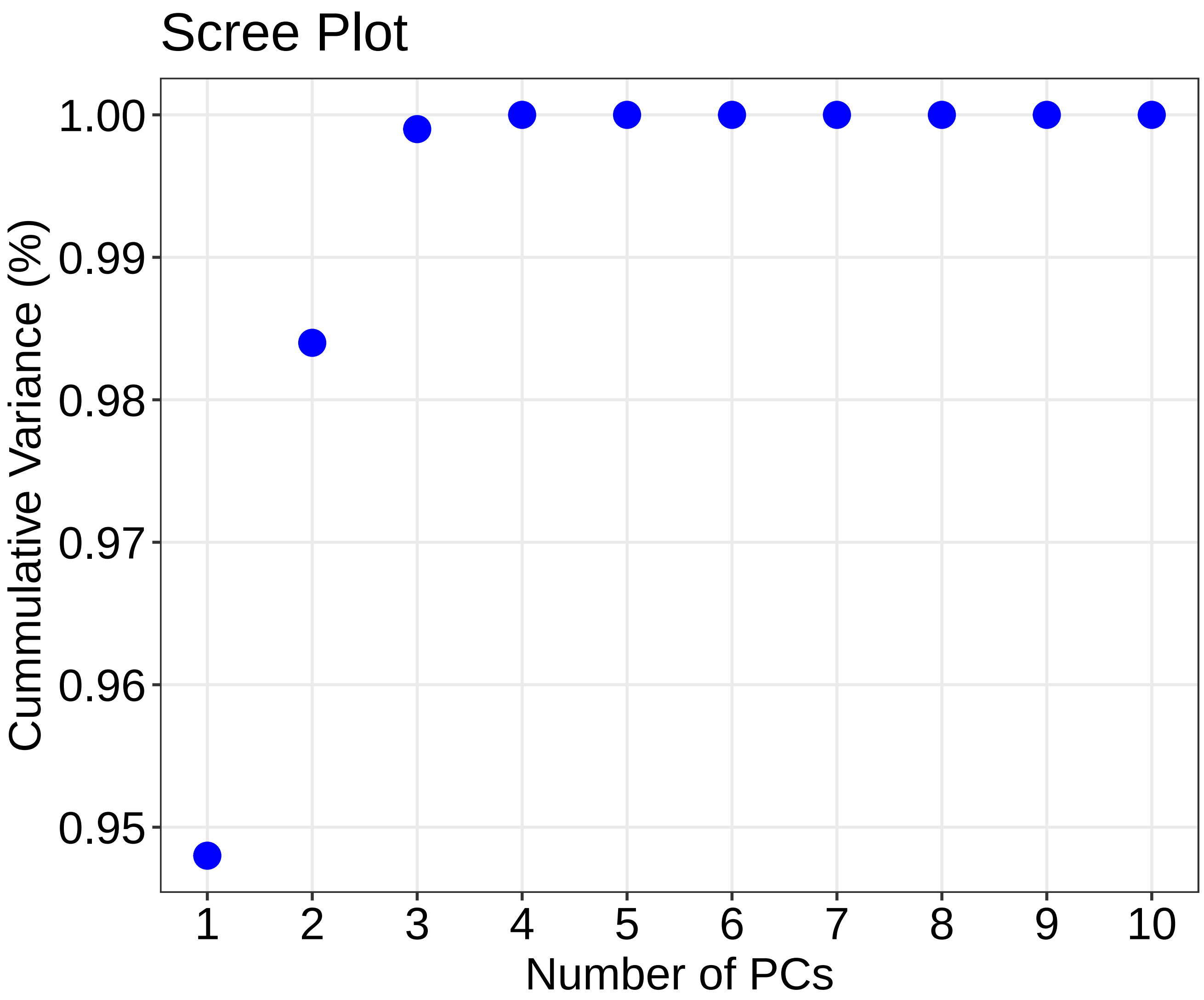

Figure 2. The cumulative variance captured by the 10 PCs which represent the highly correlated 10 elemental descriptors above. The cumulative variance increases as the number of PCs increase and saturates at one representing 100% captured variance after three PCs.

Table 2. Loadings which represent the weights of the elemental descriptors in PC1 (94.8%), PC2 (3.6%) and PC3 (1.5%). All the 10 elemental descriptors are of similar importance in PC1 with comparable weights for all. Net thermal conductivity, net polarizability and net V/Z have loadings > 0.4 and dominate PC2 with net polarizability and net V/Z being inversely proportional. Net thermal conductivity and net density are inversely proportional and dominate PC3 with loading values > 0.4.

Elemental Descriptor	**PC1 Loadings**	**PC2 Loadings**	**PC3 Loadings**
Net atomic radius (Slater)	-0.320	0.283	-0.053
Net valence electron	-0.323	-0.088	0.086
Net electronegativity (Pauling)	-0.324	-0.066	-0.141
Net melting point	-0.324	0.025	0.116
Net thermal conductivity	-0.302	**-0.425**	**-0.685**
Net bulk modulus	-0.315	-0.391	0.142
Net density	-0.311	-0.230	**0.673**
Net polarizability	-0.309	**-0.504**	-0.108
Net heat of fusion	-0.323	-0.116	-0.016
Net V/Z	-0.309	**0.504**	-0.035

1.3 Machine learning with Custom Loss Function

1.3.1 eLASSO

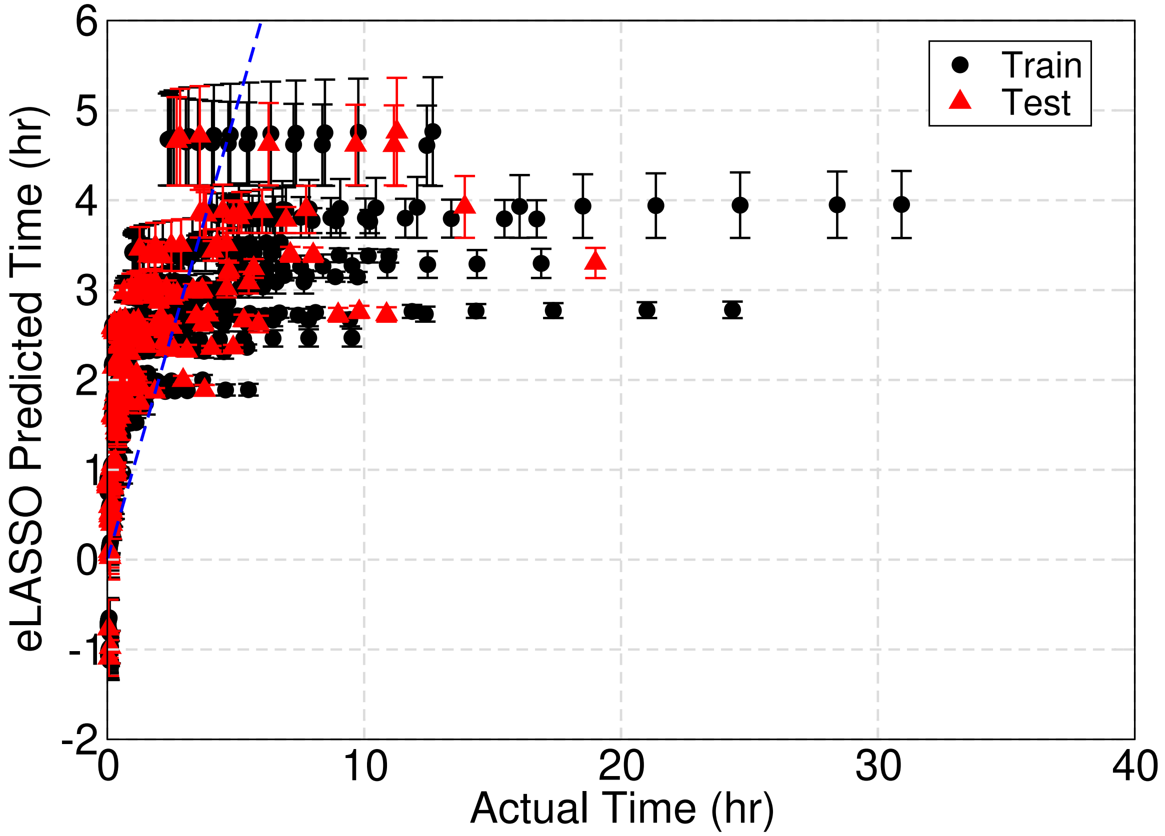

Figure 3. Performance plots for eLASSO model on train and test set. The R^2 values for train and test set here are 0.27 and 0.35, respectively. The blue dashed line represents the line on which the predicted and actual time coincide. Points from both train and test set deviate significantly from this blue dashed line indicating the poor model performance.

2 Results

2.1 Post hoc Model Explanation

2.1.1 Global Feature Importance (GFI)

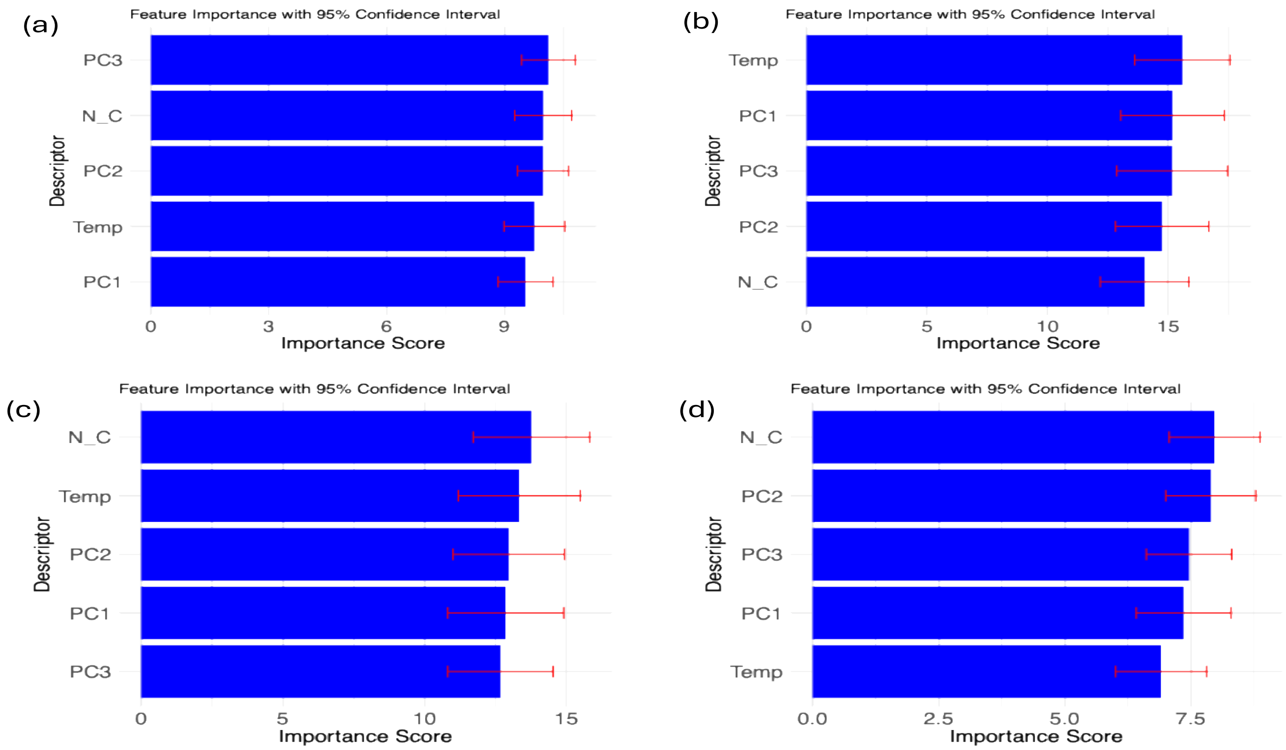

Figure 4. GFI plot for the eSVR model with loss function: (a) $\mathscr{L}_{MSE}$, (b) $\mathscr{L}_{MSE}$+$\mathscr{L}_{R}$, (c) $\mathscr{L}_{MSE}$+$\mathscr{L}_{R}+\mathscr{L}_{C}$ and (d) $\mathscr{L}_{MSE}$+$\mathscr{L}_{C}$. GFI shows nearly equal importance for all the five descriptors although their relative rankings vary with different regularization terms. For the models with the $\mathscr{L}_{C}$ term in the loss function (c) and (d), N_c has the highest importance compared to other descriptors.

2.1.2 Partial Dependence Plots (PDP)

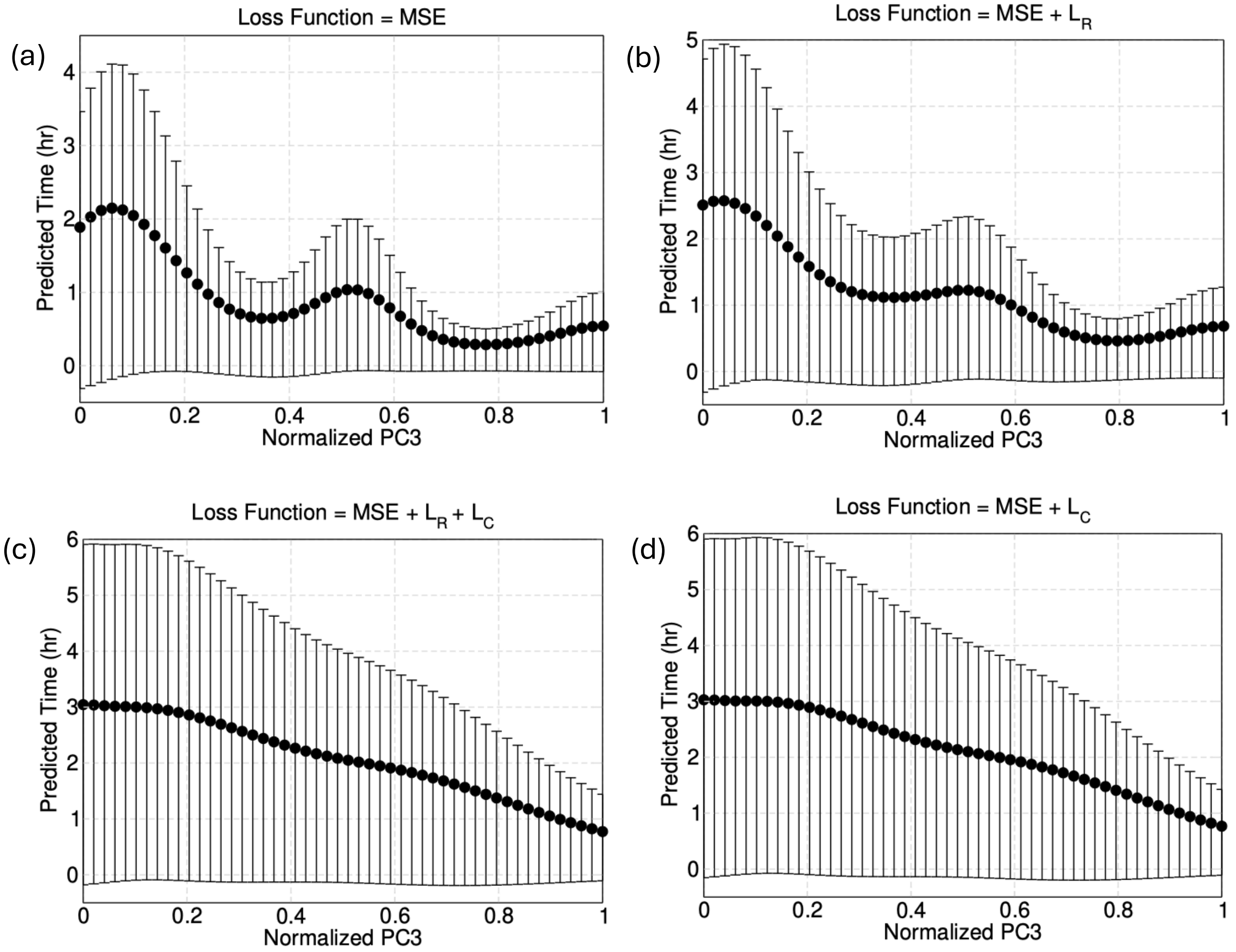

Figure 5. PDP for predicted time versus normalized PC3 (capturing 1.5% variance) from the four eSVR models with different loss function terms, (a) $\mathscr{L}_{MSE}$, (b) $\mathscr{L}_{MSE}$+$\mathscr{L}_R$, (c) $\mathscr{L}_{MSE}$+$\mathscr{L}_R + \mathscr{L}_C$ and (d) $\mathscr{L}_{MSE}$+$\mathscr{L}_C$. There is an overall decrease in the predicted time values as the PC3 values increase in (a) and (b) but the decrease is not linear. Upon adding the $\mathscr{L}_C$ term to the loss function this decrease becomes nearly linear as shown in (c) and (d).

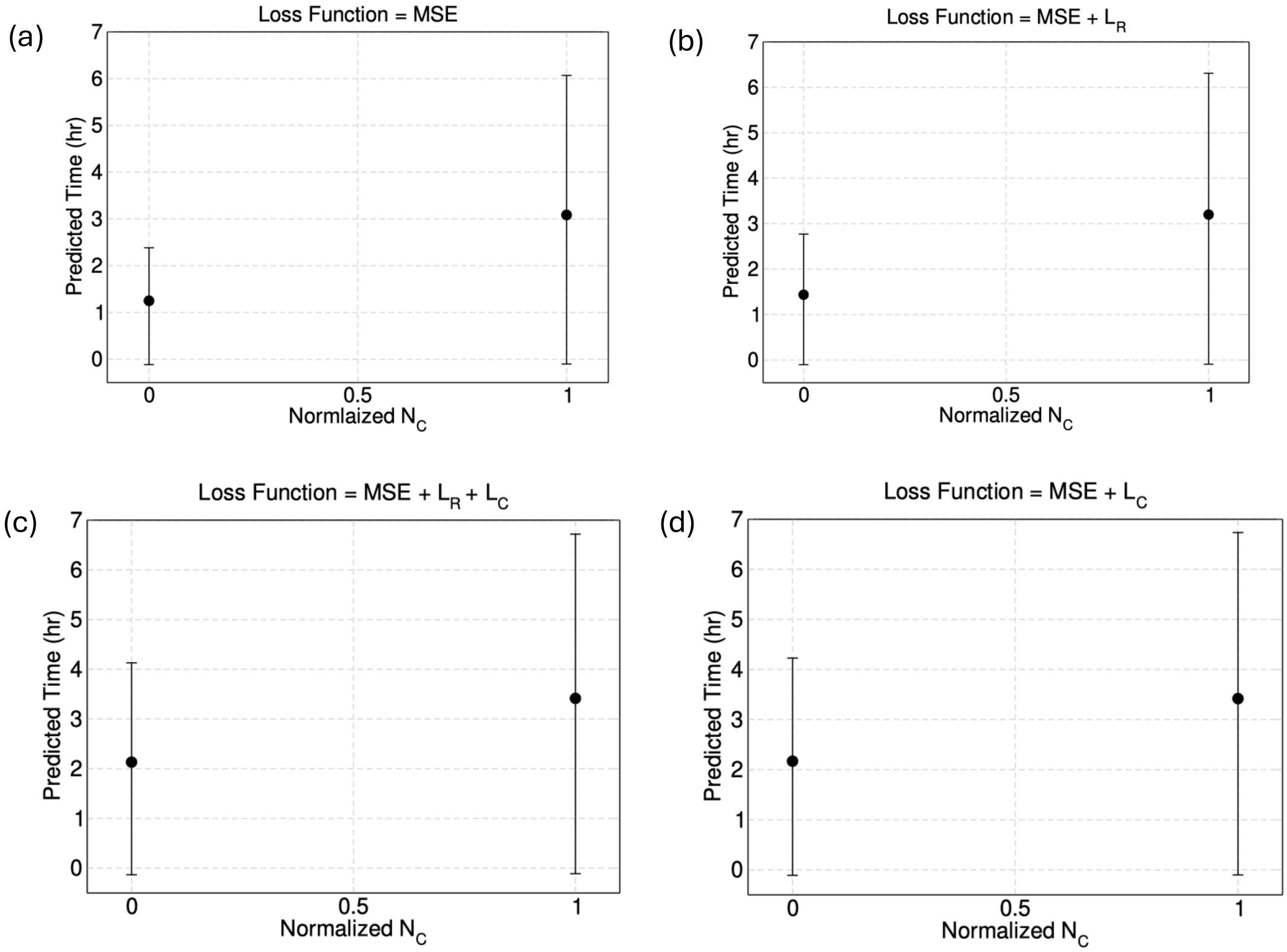

Figure 6. PDP for predicted time versus normalized N_C from the four eSVR models with different loss function terms, (a) $\mathscr{L}_{MSE}$, (b) $\mathscr{L}_{MSE}$+$\mathscr{L}_R$, (c) $\mathscr{L}_{MSE}$+$\mathscr{L}_R + \mathscr{L}_C$ and (d) $\mathscr{L}_{MSE}$+$\mathscr{L}_C$. Here, zero N_C implies binary alloys (U-Mo) and one implies ternary alloys (U-Mo-X). There is an overall increase in the predicted time values as the N_C values increase from two to three implying phase stability upon adding a ternary element X to the U-Mo binary alloy. The PDP trend remains the same for all the four models irrespective of the loss function

2.2 ML Predicted TTT Curves

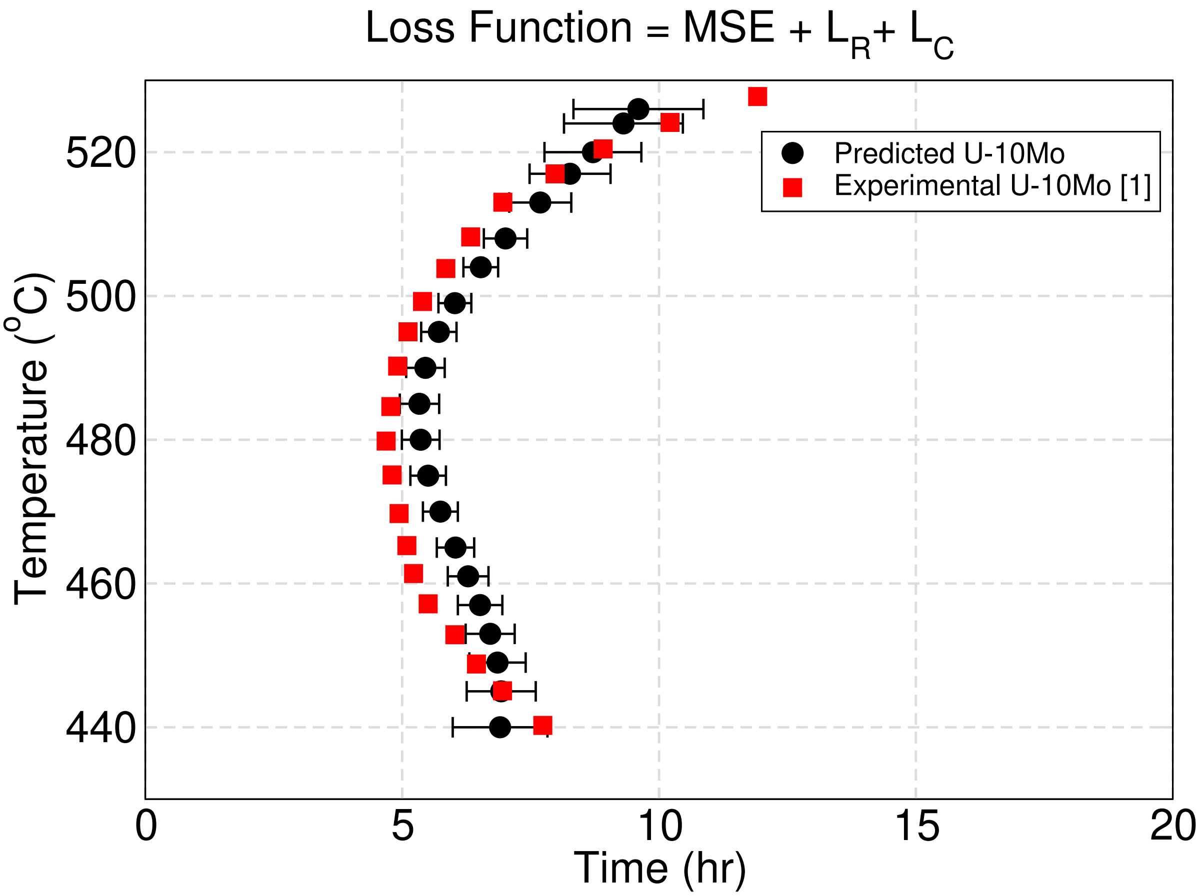

Figure 7. Predicted and experimentally calculated TTT curve for U-10Mo. The predicted data (black circles) from the model employing $\mathscr{L}_{MSE} + \mathscr{L}_R + \mathscr{L}_C$ loss function aligns well with the experimentally calculated data (red squares) in terms of both the predicted time values and the C-shape nature of the TTT curve[1].

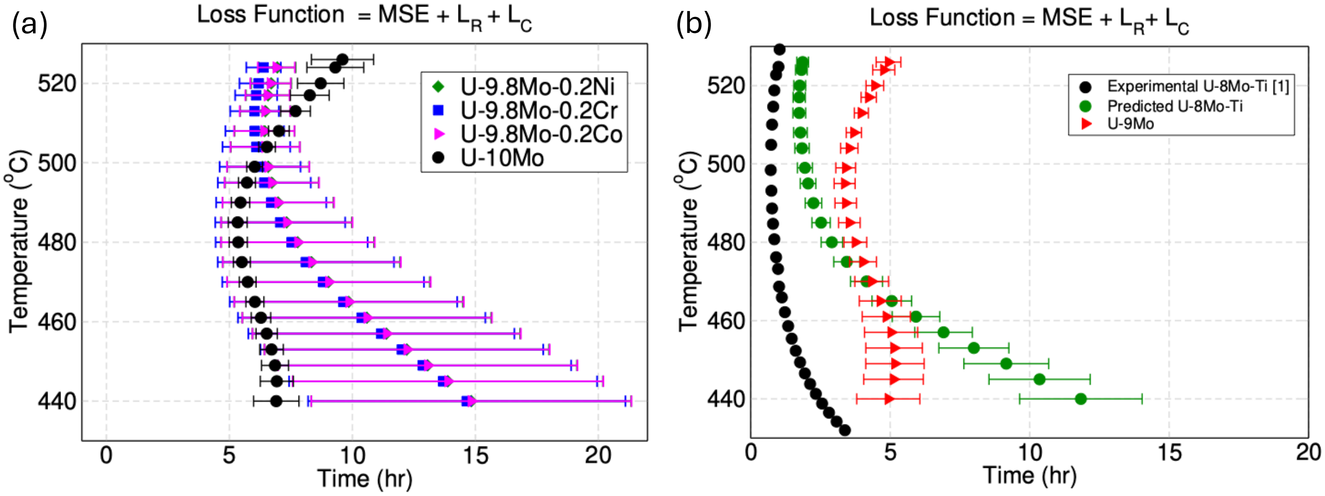

Figure 8. The predicted TTT curve for (a) U-Mo-X (X= Ni, Cr, Co) and U-10Mo alloys, (b) U-8Mo-Ti and U-9Mo alloys from the model employing $\mathscr{L}_{MSE} + \mathscr{L}_R + \mathscr{L}_C$ loss function. For U-Mo-X (X= Ni, Cr, Co) alloys in (a) represented by green diamonds, blue squares, magenta triangles respectively, the predicted TTT curves are shifted to the right compared to the baseline U-10Mo predicted curve (black circles) showing increased γ-phase stability on minor addition of ternary element (Ni, Co and Cr) to U-10Mo. The nose temperature corresponding to the predicted time is also higher than the U-10Mo TTT curve. For U-8Mo-Ti in (b), the predicted nose time (green circles) aligns closely with the experimentally calculated U-8Mo-Ti nose time (black circles)[1]. The predicted TTT curve is however shifted to higher temperatures compared to the experimental U-8Mo-Ti curve. Both the predicted and experimental TTT curve show the same trend of reduction in γ-phase stability compared to the predicted baseline U-9Mo TTT curve (red triangles) as experimentally reported.

2.3 Predictions on Previously Unexplored U-Mo-X Alloys

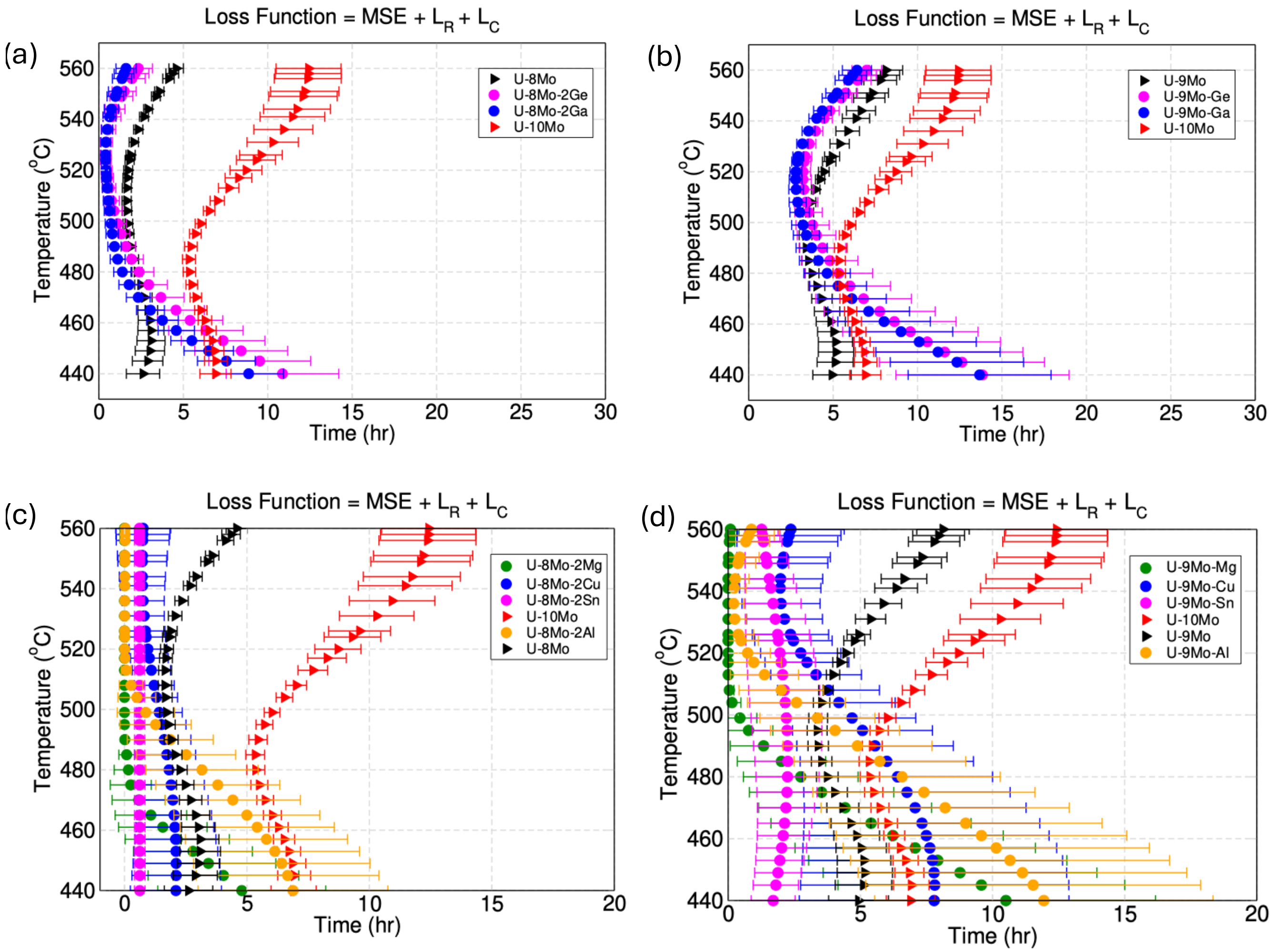

Figure 9. Predicted TTT curves for previously unexplored (a) U-8Mo-2X and (b) U-9Mo-X alloys with X= Ge (magenta circles) and Ga (blue circles). These new alloys are compared with the predicted TTT curves of baseline alloys, U-8Mo (black triangles) and U-10Mo (red triangles) in (a) and U-9Mo (black triangles) and U-10Mo (red triangles) in (b). All the predicted curves for unexplored alloys have higher predicted TTT temperature values compared to the predicted U-10Mo curve. The destabilizing effect of adding Ge and Ga as a ternary element in U-10Mo is reflected in the shift of the TTT curves to lower TTT time values indicating lower nucleation time for α uranium for these alloyed compositions compared to the U-10Mo curve. The α phase nucleation time, represented by the nose of the TTT curves for the unexplored alloys lies near the nose of the binary U-8Mo and U-9Mo alloys in (a) and (b), respectively.

2.4 Benchmarking the Performance of the Physics-Regularized eSVR model against Single-Point Prediction

Table 3. Optimized hyperparameters for the GPR model using the automatic relevance determination (ARD) kernel.

Hyperparameters	**Values**
Lengthscale (σ_{m-N_c})	11.6
Lengthscale (σ_{m-PC1})	36068.31
Lengthscale (σ_{m-PC2})	2.1
Lengthscale (σ_{m-PC3})	5.8
Kernel standard deviation (σ_f)	3.9
Standard deviation of observed noise	0.48
Coefficient of basis function	0.55

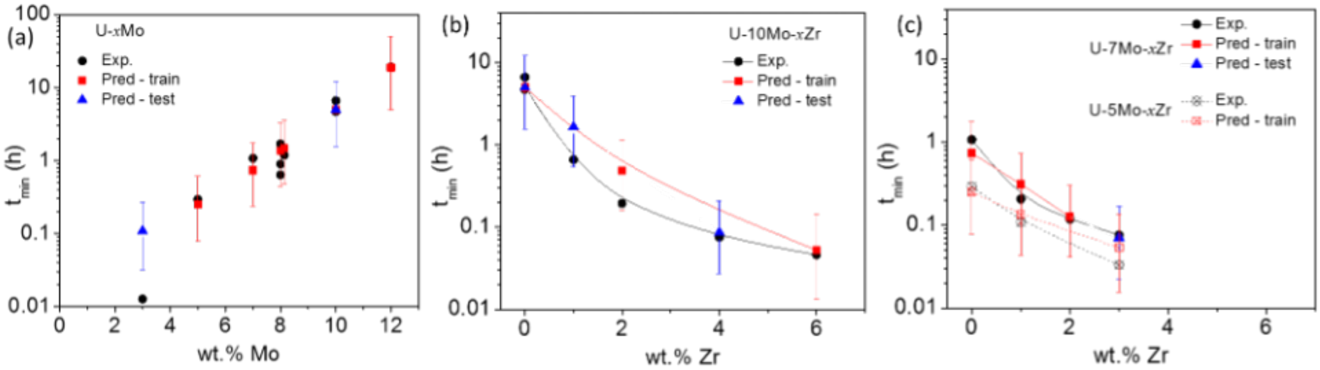

Figure 10. (a) Experimentally observed time of the TTT nose and that predicted by the GPR model, for both training and test data set for (a) binary U-Mo alloys, as a function of Mo content. (b) ternary U-10Mo-xZr alloys as a function of Zr content and (c) ternary U-7Mo-xZr and U-5Mo-xZr alloys as a function of Zr content, showing that the GPR model captures the alloy dependence of the TTT nose

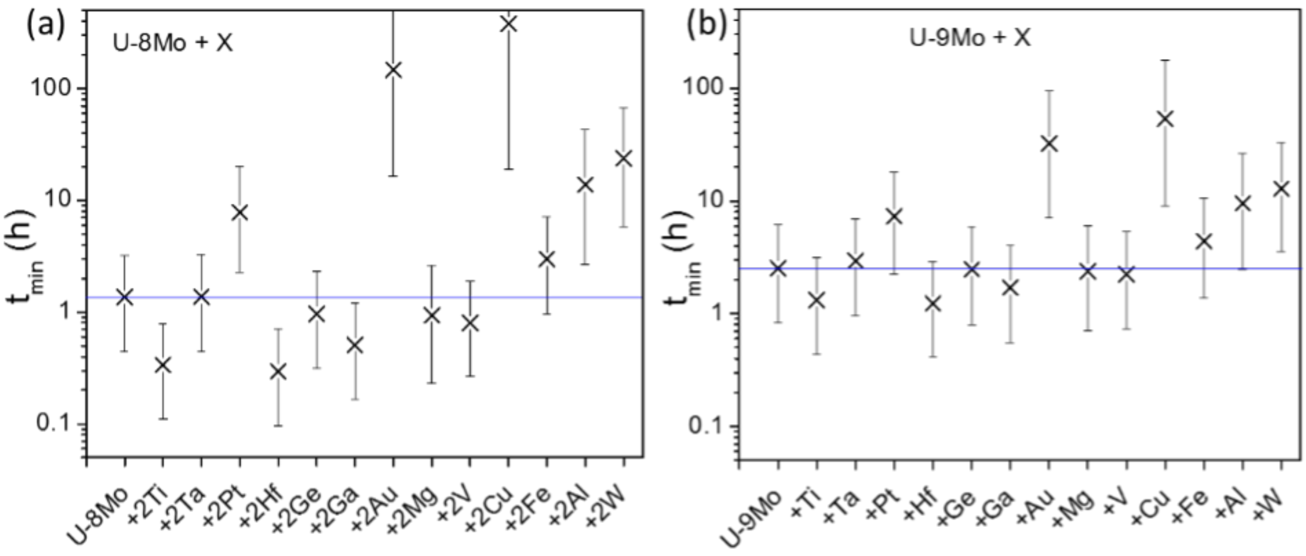

Figure 11. TTT nose time predictions from the GPR model for (a) U-8Mo-2X based alloys with X = Ti, Ta, Pt, Hf, Ge, Ga, Au, Mg, V, Cu, Fe, Al and W, (b) U-9Mo-X based alloys with X = Ti, Ta, Pt, Hf, Ge, Ga, Au, Mg, V, Cu, Fe, Al and W. The blue line represents the TTT nose time for U-8Mo and U-9Mo in (a) and (b), respectively

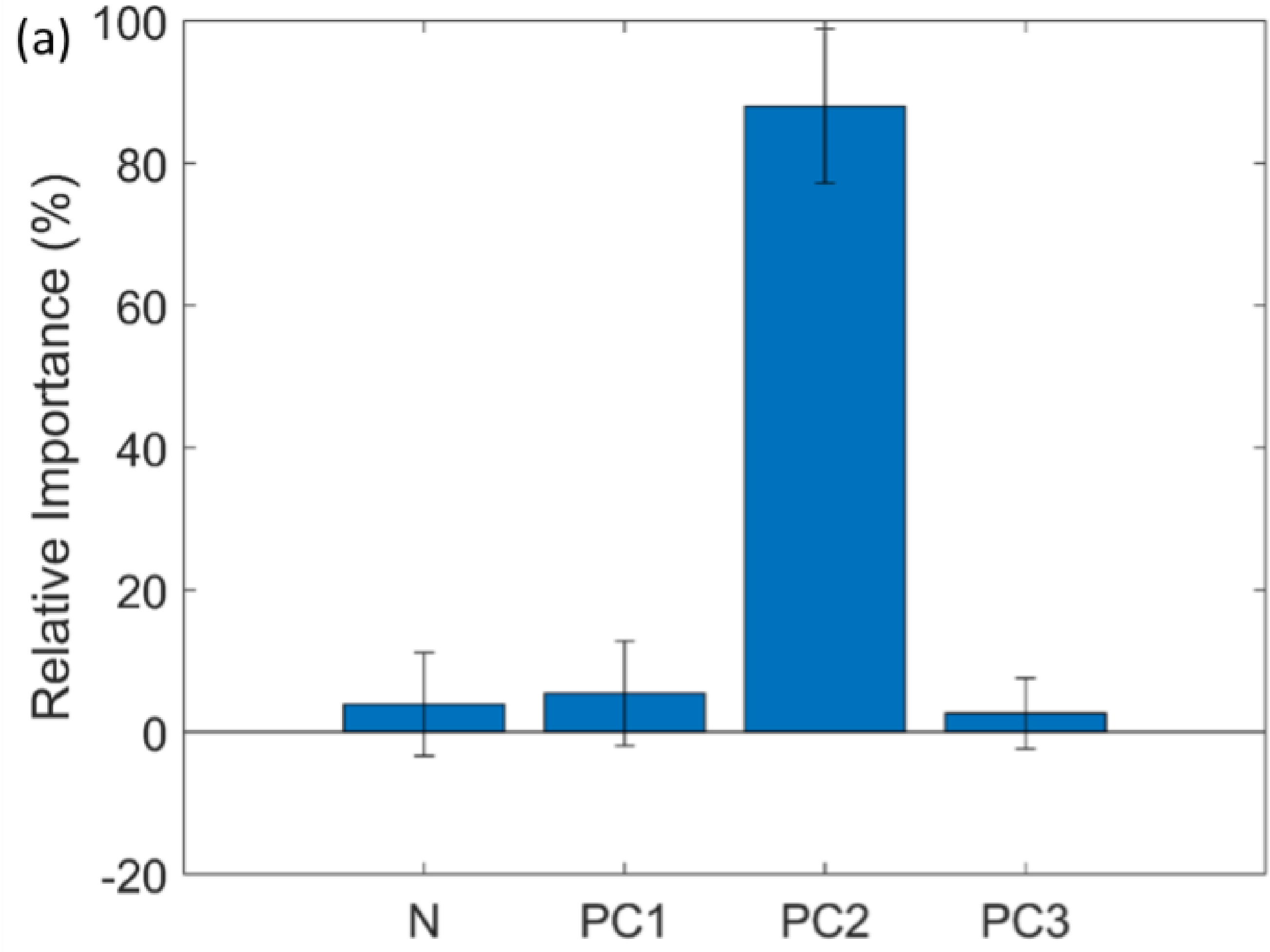

Figure 12. GFI plot for the GPR model showing higher importance for PC2 compared to other descriptors (N, PC1, PC3) in predicting the TTT nose time values.

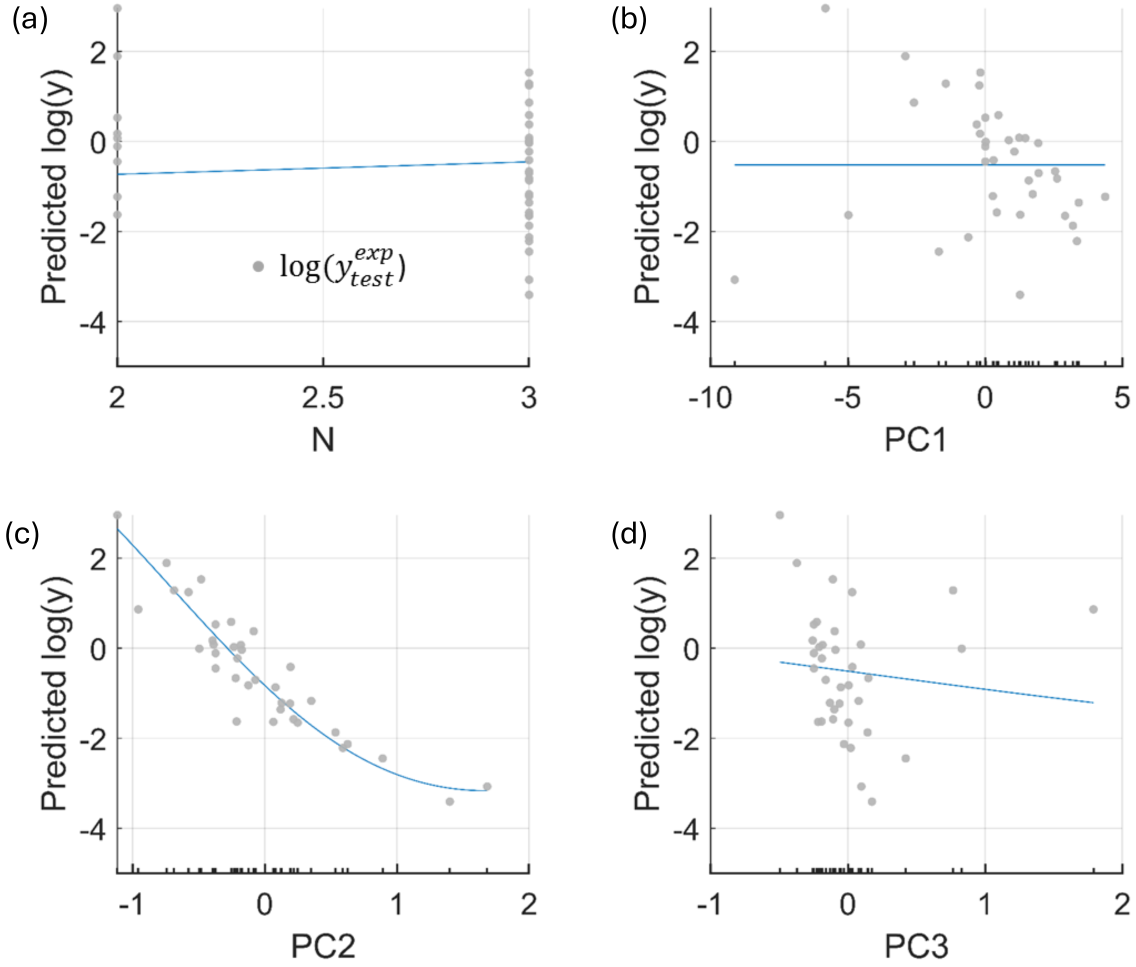

Figure 13. PDP for the four descriptors (a) N, (b) PC1, (c) PC2 and (d) PC3 from the GPR model. The descriptors N, PC1 and PC3 show negligible influence on the predicted TTT nose time values where as PC2 in (c) shows an inverse correlation with the predicted nose time values. The plots here are logarithmic values of predicted nose time versus the descriptor values.

2.5 Predicting TTT curves using Multi-Output Gaussian Process Regression (MO-GPR)

An attempt was made to use MO-GPR to fit the entire TTT curve. The code developed by the Sheffield machine learning group was employed[4]. The advantage of MO-GPR is their capacity to simultaneously learn and infer multiple outputs (two in this case) which have the same set of inputs. This is achieved via coregionalized regression, which involves the use of so-called multiple output kernels connected by a coregionalization matrix, B, forming a latent function:

$$B \otimes K = \begin{pmatrix} B_{1,1} \times K(X_1, X_1) & \cdots & B_{1,D} \times K(X_1, X_D) \\ \vdots & \ddots & \vdots \\ B_{D,1} \times K(X_D, X_1) & \cdots & B_{D,D} \times K(X_D, X_D) \end{pmatrix}$$

Where K is the kernel function, and X_i represents the input corresponding to the i-th output. Note, if $B_{(i,j)} = 0$, then the outputs are independent of each other. In order to ensure that the multiple output kernel is a valid kernel, K has to be a valid kernel and the coregionalization matrix has to be positive definite, i.e., $B = WW^T + \kappa I$, for some matrix W and vector, κ. The hyperparameters involved are thus, the length scale and the variance of the kernels, K and coefficients of the matrix W and vector κ and the noise level. Several different kernels or their combinations, e.g., Radial Basis Function (RBF), Matern52, Matern32, Matern12 (exponential), Bias kernel, and their sums and products were initially explored and, among those evaluated, it was found that a Matern52 kernel provided the best balance among the number of hyperparameters involved, capturing trends in the training data and generalizability. Default hyperparameter optimization was performed by maximizing the log marginal likelihood using the built-in L-BFGS-B algorithm within the GPy code.

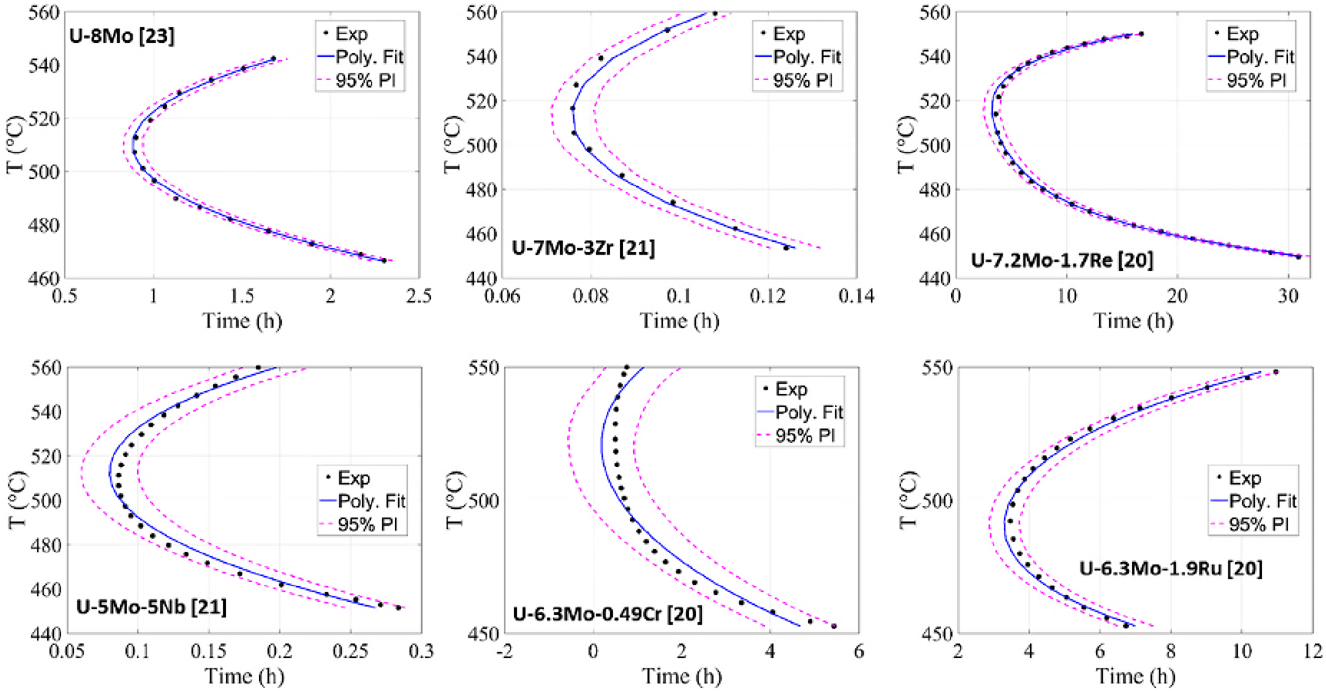

Figure 14. Select examples showing the fit of the empirical polynomial equation to the experimental data for binary and ternary U-Mo-X alloys.

The starting point for this involves fitting the above mentioned dataset of experimentally determined TTT curves for U-based alloys to an empirical expression:

$$\tau = p_4 T^4 + p_3 T^2 + p_2 T + p_1$$

where tau is the transformation time in hours, and T is the temperature in °C, and p_i are the 4 fitting coefficients, which form the output of the MO-GPR model. Figure S14 shows selected examples of the empirical fit to the experimental data, revealing that the chosen expression adequately captures the shape of the TTT curves. The process of hyperparameter selection and optimization was employed, noting that there are now 4 outputs to be described. Figure S15 shows the predicted time values versus the experimentally observed time values along the whole TTT curves, for both the training (blue open circles) and the test (red filled circles) data. The error bars are calculated from the predictions intervals. It is found that the model predictions of the training data lie quite close to the y=x line, however, the predictions on the test set is poor. This suggests that the model is not generalizable and is overfitting the training data without learning the underlying trends.

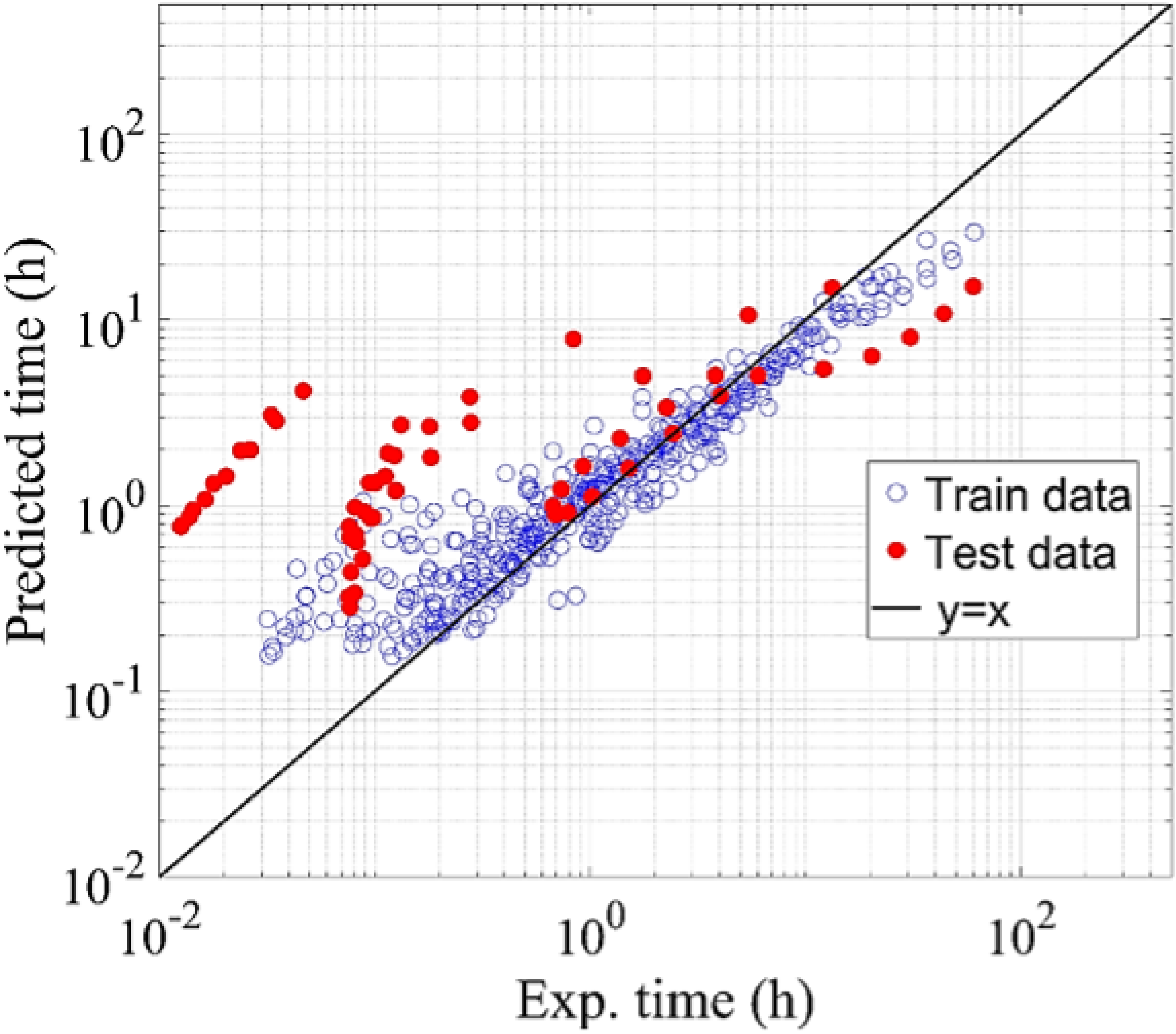

Figure 15. Experimentally observed time values of the full TTT curve vs that predicted by the MO-GP model, for both training and test data set

This is further evident from the examination of the predicted TTT curves of the trained model for a series of binary U-xMo alloys, Figure S16a. The predicted TTT curves are shown as dashed curves for the training set and dotted curves for the test set. As described above for the nose, an increase in the Mo content, the TTT curve shifts to larger times, i.e. the transformation is delayed, and the model does not capture this trend. In particular, the MO-GPR model grossly over predicts the time of transformation for U-3Mo, to be even more than U-5Mo and U-7Mo. A similar observation can also be made for the ternary alloys, U-7Mo-xZr, U-10Mo-xZr and U-5Mo-xZr, where the model predicts the training dataset quite well but performs poorly for the test dataset, reflecting the overfitting and poor generalizability behavior. (Figure S16 (b)-(d)).

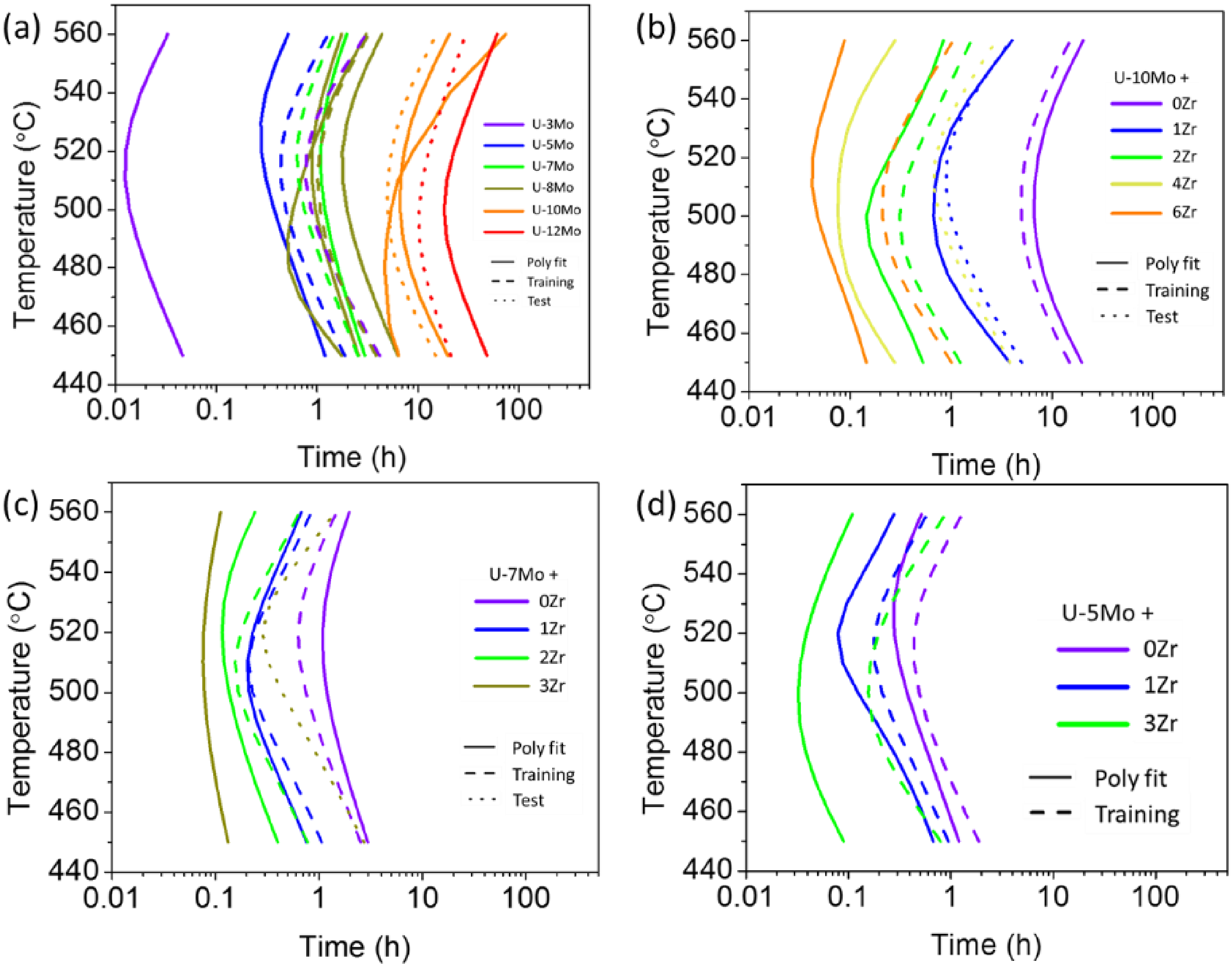

Figure 16. Experimental TTT curves represented by polynomial fit vs that predicted by the MO-GP model for the case of (a) binary U-Mo alloys, as a function of Mo content. (b) ternary U-10Mo-xZr alloys (c) ternary U-7Mo-xZr and (d) U-5Mo-xZr alloys, as a function of Zr content.

Two possible reasons for this are, 1) polynomial coefficients are often highly correlated with each other, and small changes in coefficients lead to large changes in the predictions, and 2) the large number of hyperparameters involved in the MO-GPR, ranging between 10-30.

References

1. Repas, P., Goodenow, R. & Hehemann, R. Transformation characteristics of U-Mo and U-Mo-Ti alloys. *Trans. Am. Soc. Met.* **57** (1964).
2. Steiner, M. *et al.* α-Phase transformation kinetics of U–8 wt% Mo established by in situ neutron diffraction. *J. Nucl. Mater.* **477**, 149–156, DOI: 10.1016/j.jnucmat.2016.05.016 (2016).
3. Peterson, C. *Isothermal transformation study of some uranium-base alloys* (University of California Lawrence Radiation Laboratory, 1964).
4. GPy. GPy: A Gaussian process framework in python. http://github.com/SheffieldML/GPy (since 2012).