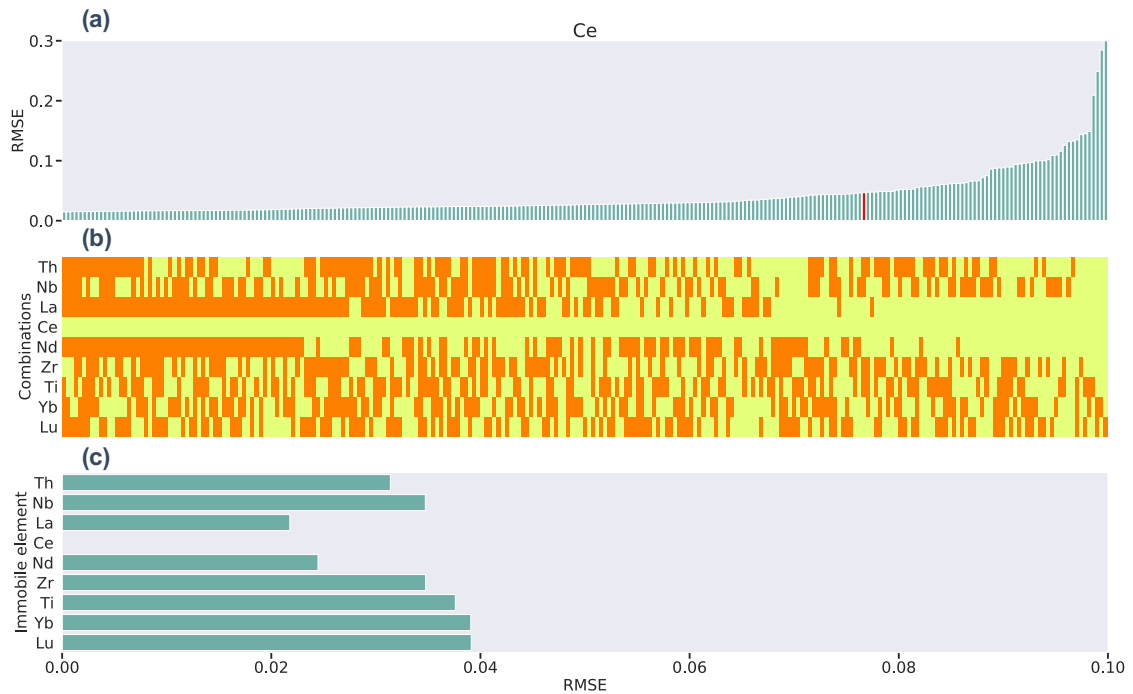


1 **Supplementary Information**

2



3

4 **Supplementary Figure S1. (a)** Average RMSE scores for Ce for each combination of

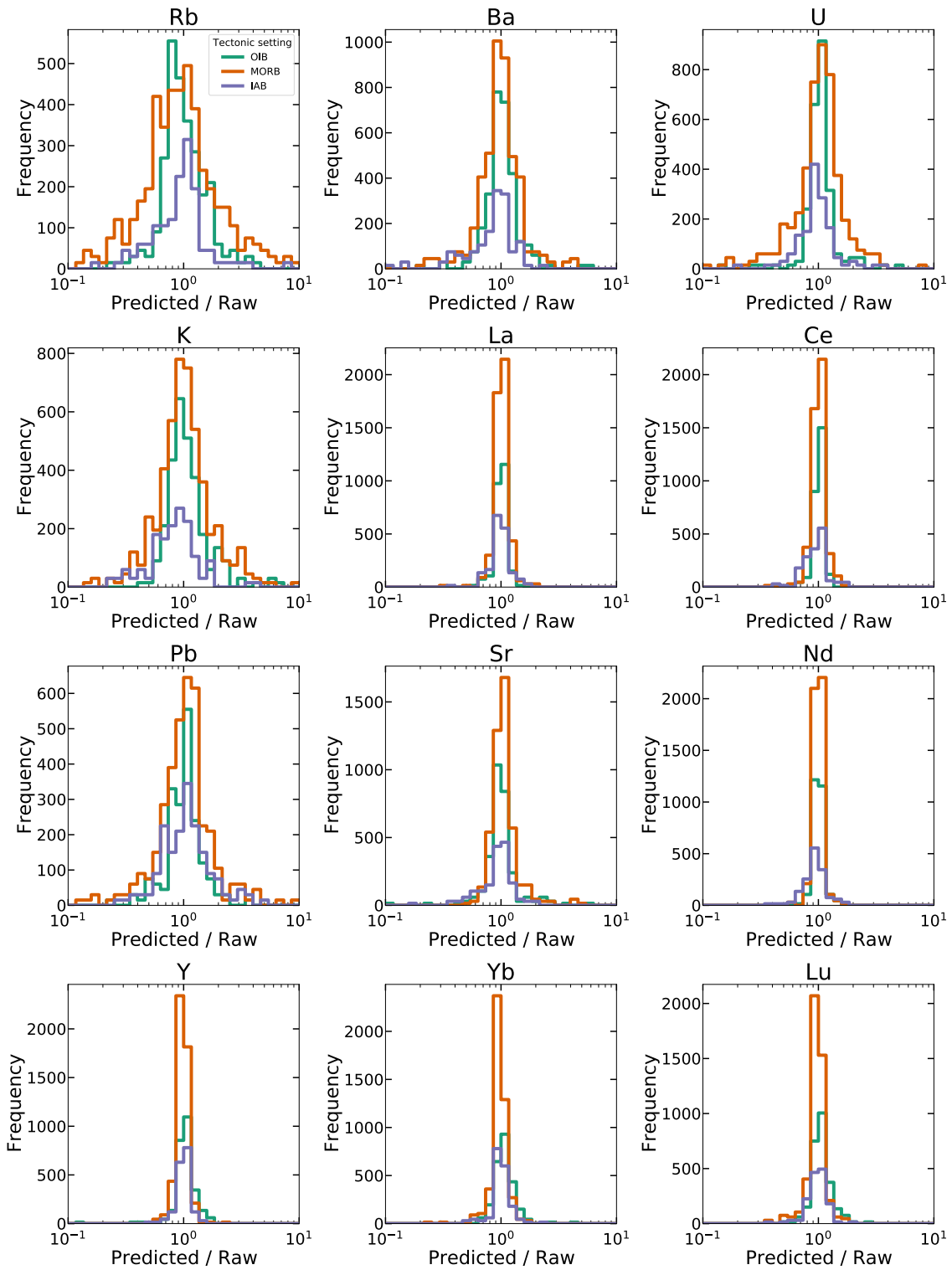
5 input elements (511 cases), red one indicate the input combination of Th, Nb, Zr, and Ti.

6 **(b)** Combinations of input elements for each model shown in **(a)**. Orange elements are

7 used in combinations, and yellow elements are not used. **(c)** Average RMSE scores for

8 Ce of all the models containing a particular element as an input.

9

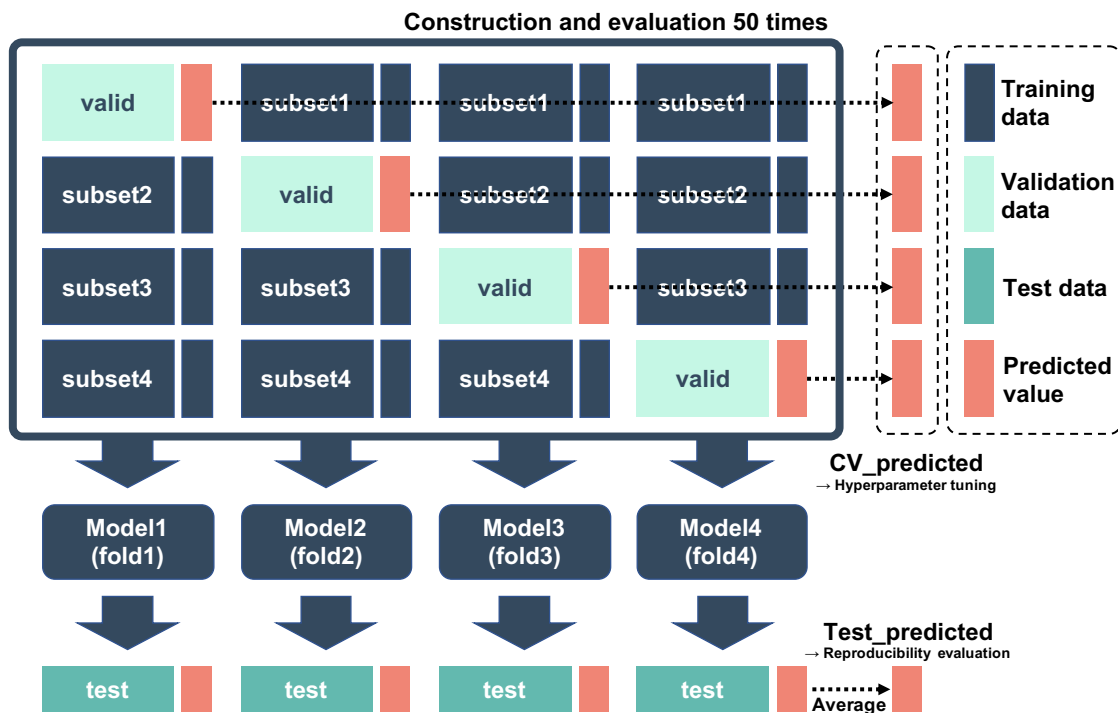


10

11 **Supplementary Figure S2.** Accuracy distribution of test data for each output element.

12 The horizontal axis is the ratio of predicted content to raw (measured) content in \log_{10}

13 scale.



14

15 **Supplementary Figure S3.** Schematic overview of cross-validation. The training data
 16 were divided into K subsets. For each fold, $K - 1$ subsets were used to construct machine
 17 learning models, and the remaining subsets were used to evaluate the constructed models.
 18 The average of the RMSE values obtained from all folds was used for hyperparameter
 19 tuning by Bayesian optimization. (Modified after Ref⁵¹).

20

21

22