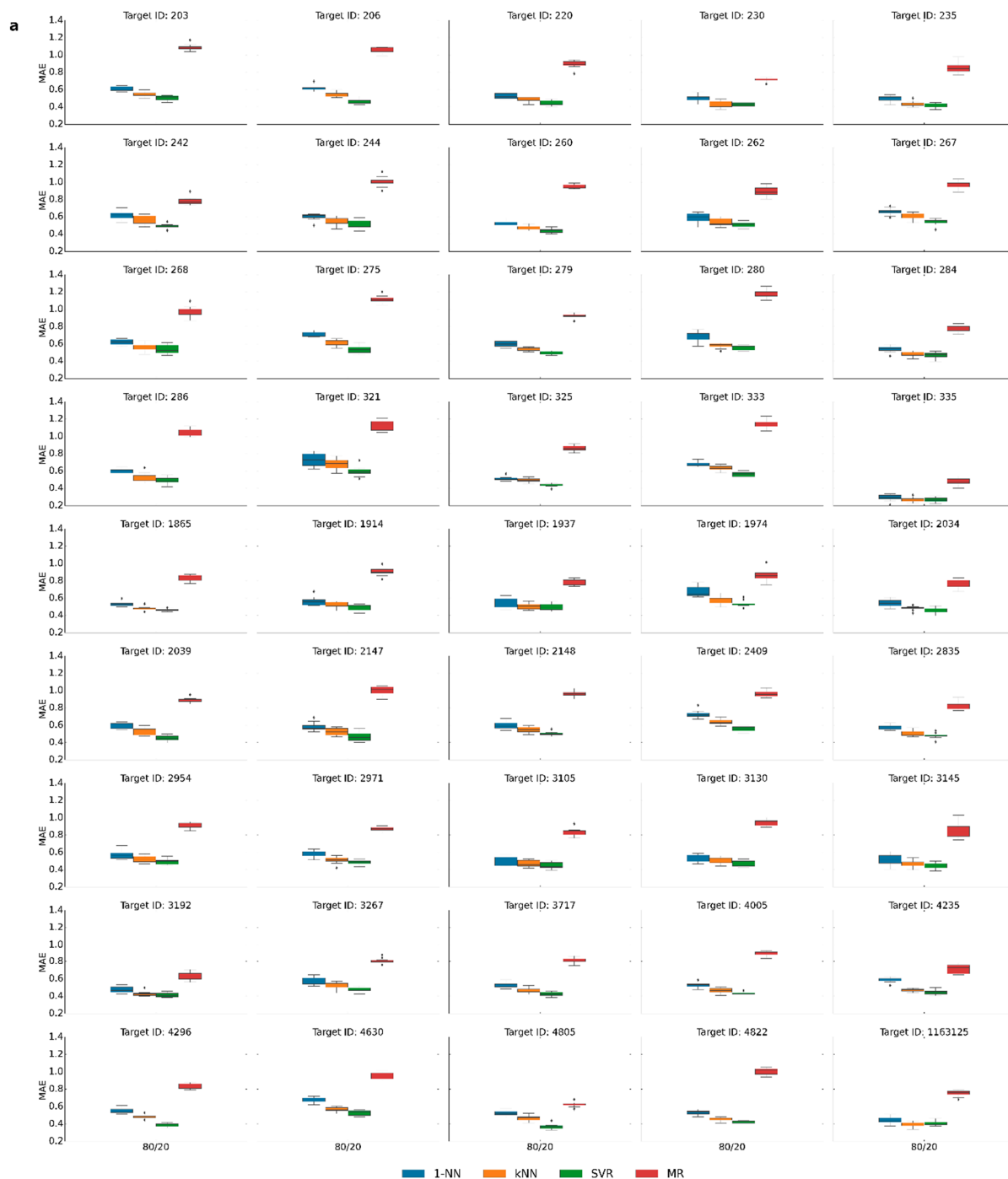


Supplementary Materials



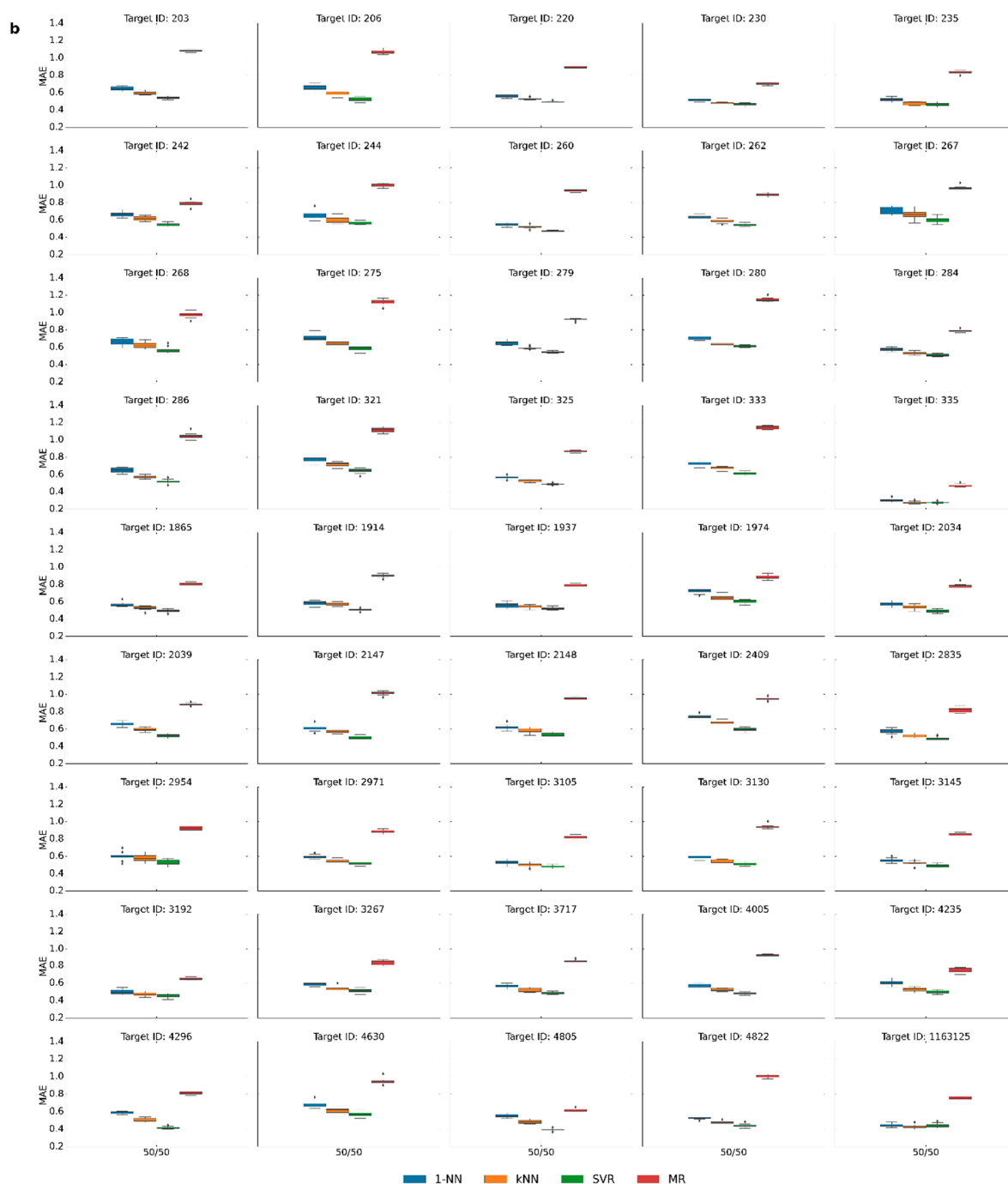


Figure S1. Prediction accuracy. Boxplots report the distribution of MAE values for potency predictions over 10 independent trials on 45 activity classes using 1NN, kNN, SVR, and MR models for training/test set compound splits of (a) 80:20% and (b) 50:50%.

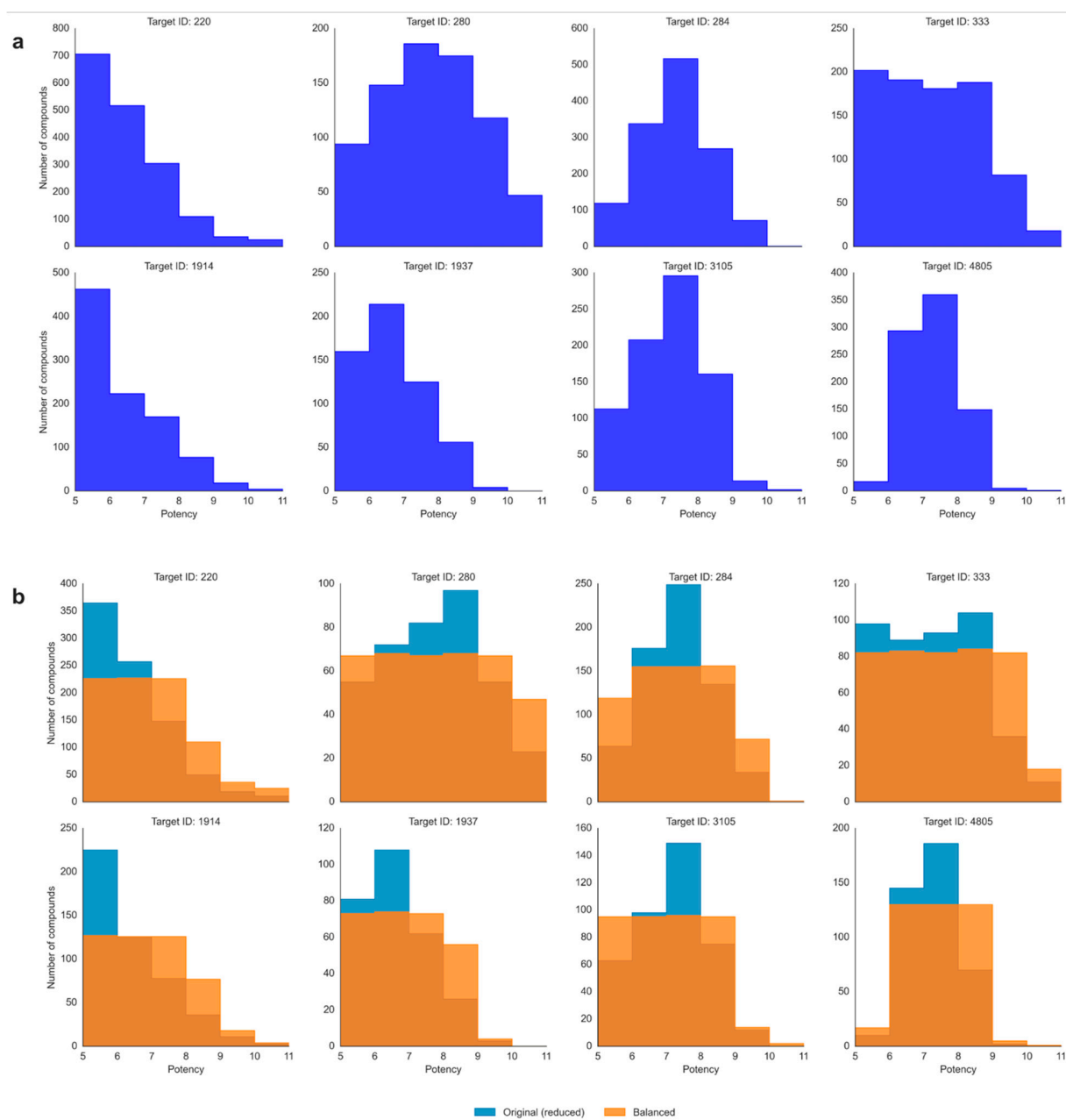


Figure S2. Potency value distributions. For the eight activity classes in Figure 1, binned potency value distributions are shown (a) before and (b) after balancing.