Electronic Supporting Information for:

Machine learning the quantum topology of chemical bonds

Michal Michalski [1)], Slawomir Berski [2)]

[1)] Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology, 08028, Barcelona, Spain.

[2)] Faculty of Chemistry, University of Wroclaw, 50383, Wroclaw, Poland.

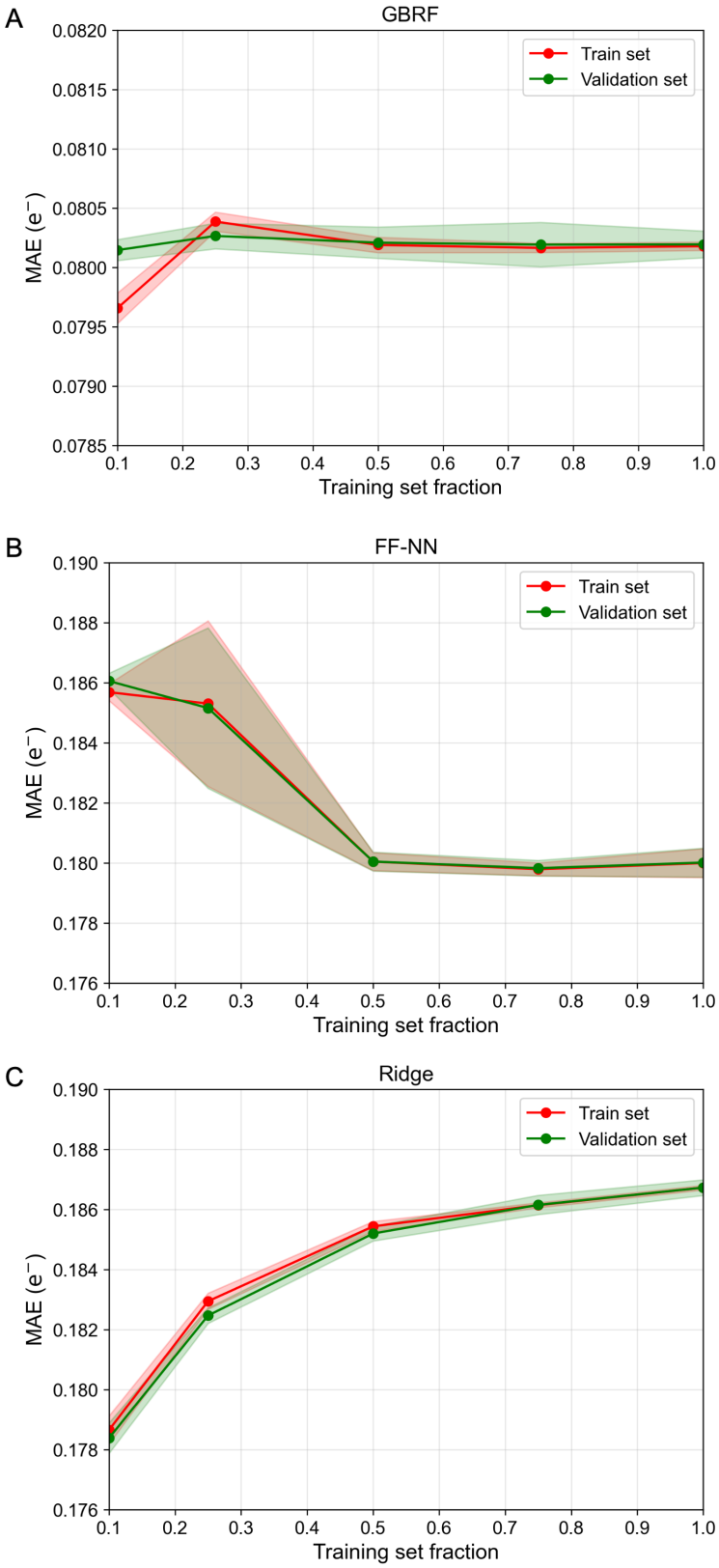

Figure S1. Learning curves showing mean absolute error (MAE) as a function of training set fraction for (a) Gradient Boosting Random Forest regression (GBRF), (b) Feed-Forward Neural Network (FF-NN), and (c) Ridge regression (Ridge) models. Red and green lines correspond to training and validation sets, respectively. Shaded areas indicate standard deviation across five-fold cross-validation.

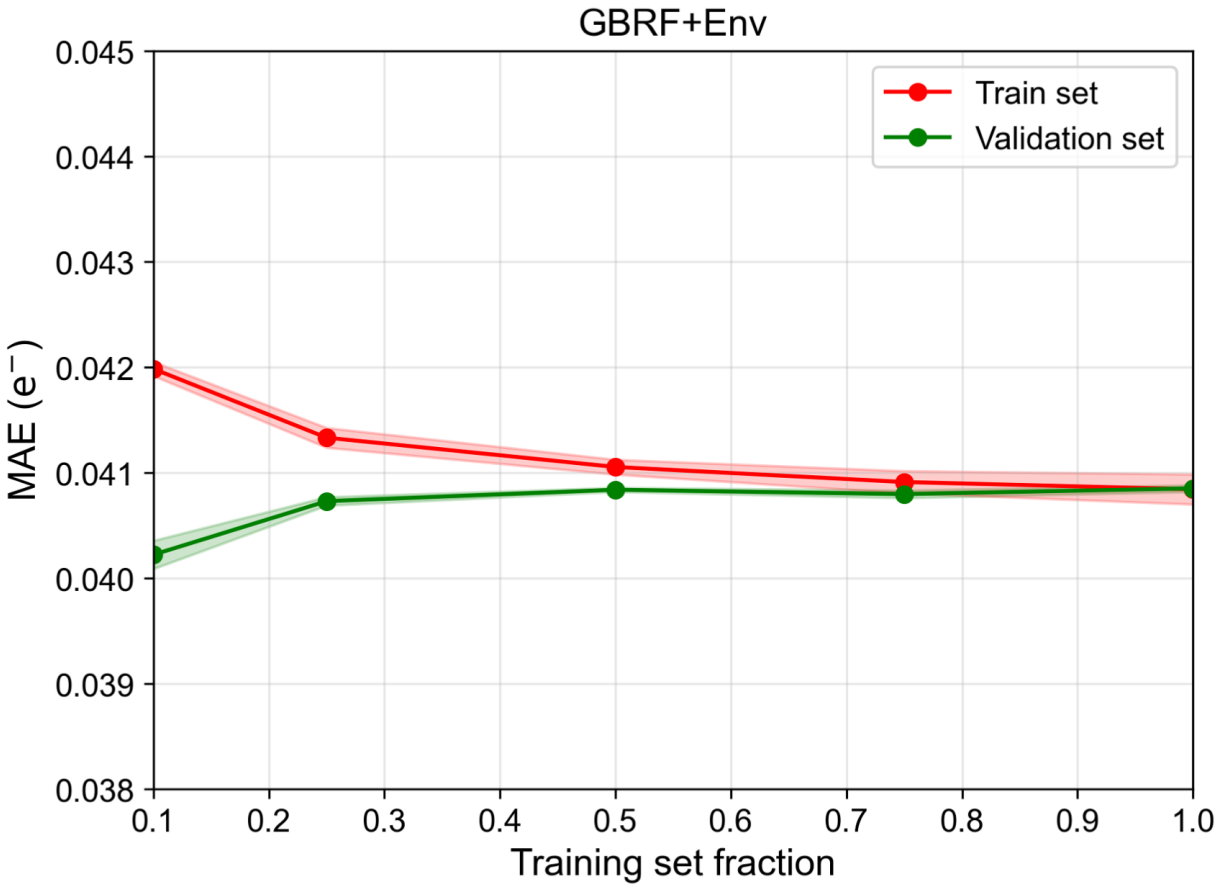

Figure S2. Learning curve for the Gradient Boosting Random Forest model with explicit environmental features (GBRF+Env), showing mean absolute error (MAE) as a function of training set fraction for the training (red) and validation (green) sets. Shaded areas indicate the standard deviation across five-fold cross-validation.