Accessing the Usefulness of Atomic Adsorption Configurations in Predicting the Adsorption Properties of Molecules with Machine Learning

Walter Malone*,Johnathan van der Hyde† , Abdelkader Kara†

*Department of Physics, Tuskegee University, 1200 W. Montgomery Rd. Tuskegee, AL 36088

†Department of Physics, University of Central Florida, 4000 Central Florida Blvd. Orlando, Florida, 32816

Supplementary Information

Number of XH_n Configurations in Training/Validation Dataset	0	10	50	250	500	750	1250	Maximum
NH Only				0.73	0.55	0.49	0.60	0.27
N + H + NH	0.76	0.5	0.46	0.33	0.28	0.25	0.23	
X2				0.34	0.25	0.25	0.23	
X5				0.32	0.25	0.25	0.20	
OH Only				0.50	0.41	0.46	0.33	0.22
O + H + OH	0.52	0.58	0.43	0.30	0.30	0.27	0.22	
X2				0.31	0.27	0.25	0.19	
X5				0.30	0.27	0.24	0.19	
H_2O Only				0.11	0.09	0.10	0.08	0.07
O + H + H_2O	0.3	0.3	0.21	0.14	0.12	0.11	0.08	
X2				0.15	0.13	0.11	0.09	
X5				0.14	0.10	0.09	0.07	
CH Only				0.70	0.62	0.54	0.35	0.21
C + H + CH	9.16	0.65	0.51	0.36	0.35	0.32	0.24	
X2				0.36	0.32	0.29	0.24	
X5				0.36	0.29	0.30	0.22	
CH_2 Only				0.48	0.41	0.37	0.41	0.25
C + H + CH_2	3.24	3.04	0.61	0.40	0.30	0.26	0.29	
X2				0.37	0.27	0.24	0.25	
X5				0.34	0.27	0.23	0.24	
CH_3 Only				0.34	0.32	0.31	0.31	0.16
C + H + CH_3	5.81	1.13	0.75	0.37	0.25	0.23	0.22	
X2				0.32	0.22	0.24	0.18	

	X5				0.27	0.23	0.19	0.17	

TABLE S1: MAE for HIP-NN models trained with 0, 10, 50, 250, 500, 750, and 1250 XH_n configurations in the training and validation dataset. Rows, in order, are models with XH_n configurations only in the training and validation dataset, models with X, H, and XH_n configurations in the training and validation dataset, and models with X, H, and XH_n configurations in the training and validation dataset with XH_n configurations weighted twofold and fivefold.
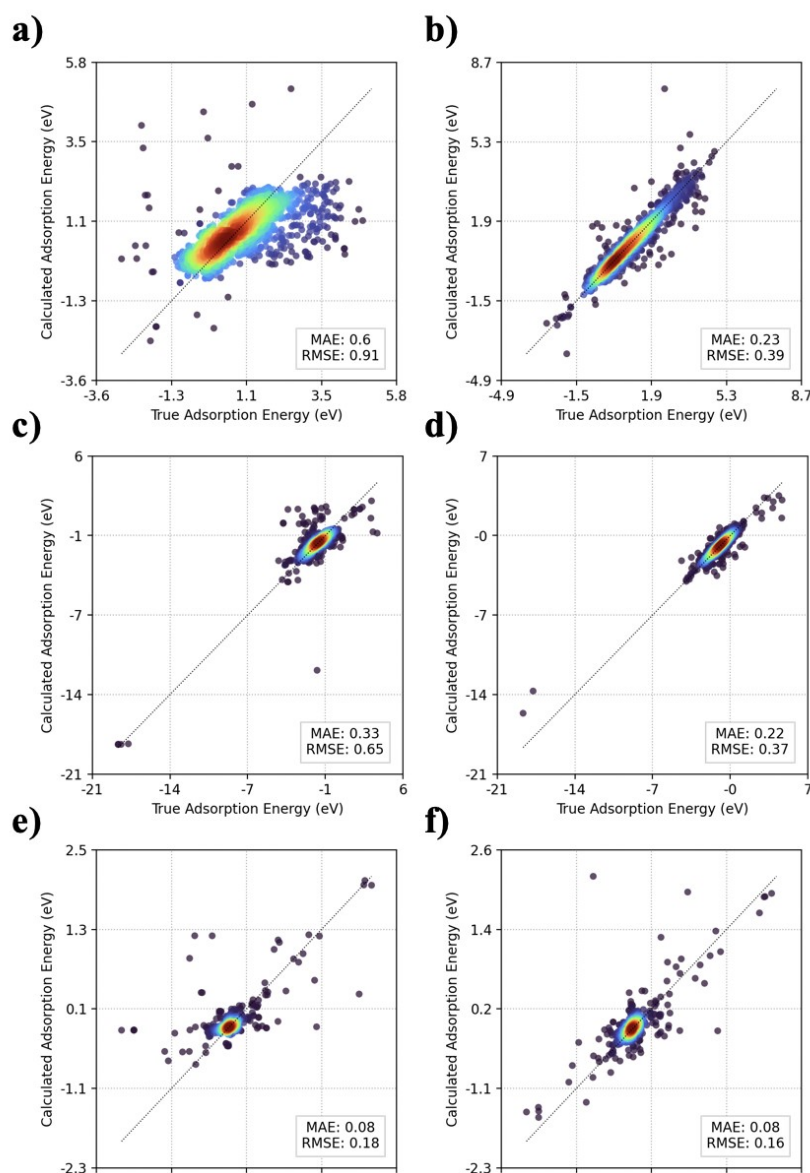

FIGURE S1: Calculated adsorption energy vs true adsorption energy with 1250 XH_n configurations in the training and validation dataset for a) NH without and b) with N and H adsorption configurations in the training and validation dataset, c) OH without and d) with O and H adsorption configurations in the training and validation dataset, and e) H_2O without and f) with O and H adsorption configurations in the training and validation dataset.