Electronic Supplementary Information (ESI)

Importance of Raw Material Features for Prediction of Flux Growth of Al_2O_3 Crystals Using Machine Learning

Tetsuya Yamada,[a,b] Takanori Watanabe,[c] Kazuaki Hatsusaka,[c] Jianjun Yuan,[c] Michihisa Koyama,[a] Katsuya Teshima[a,b]

[a]*Research Initiative for Supra-Materials, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan*
[b]*Department of Materials Chemistry, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan*
[c]*Central Research Laboratories, DIC Corporation, 631, Sakado, Sakura, Chiba 285-8668, Japan*

*Contact address of corresponding author**
teshima@shinshu-u.ac.jp

Contents

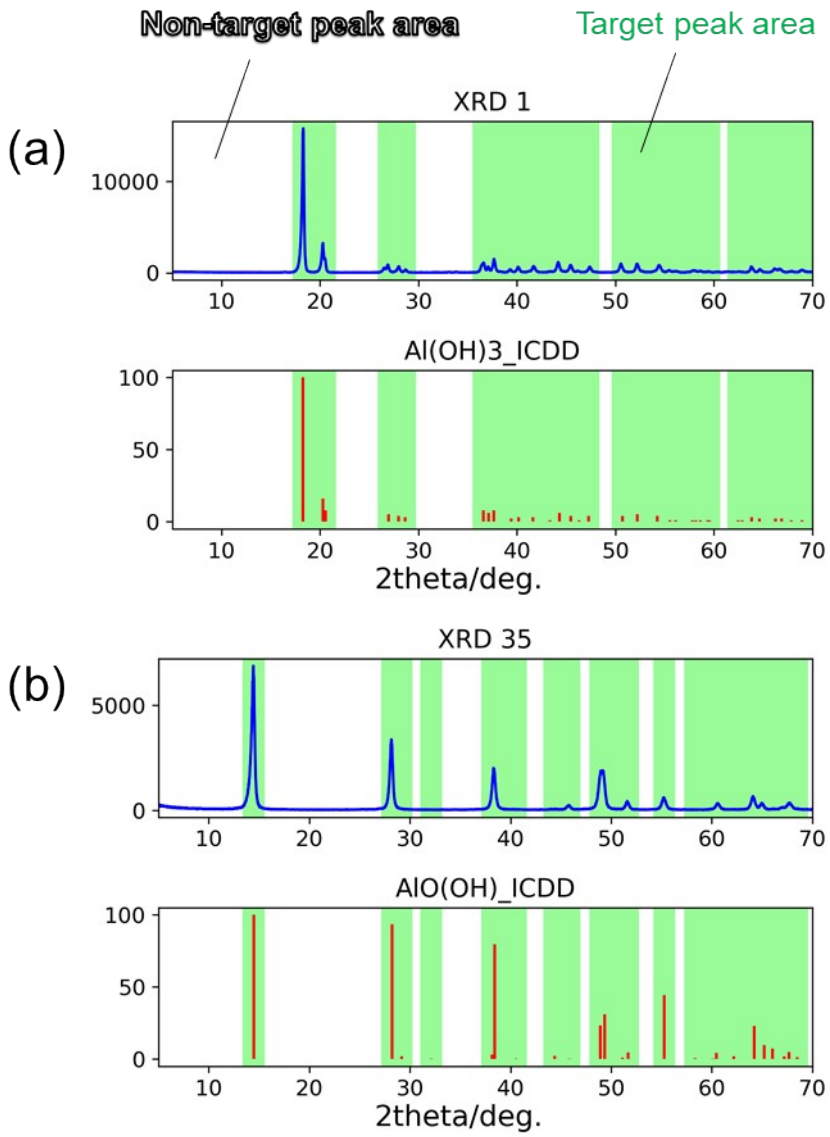

Figure S1 Definition of target and non-target peak areas in XRD profiles for (a) $Al(OH)_3$, (b) AlO(OH). Pale green area indicates target peak area, and other area does non-target peak area.

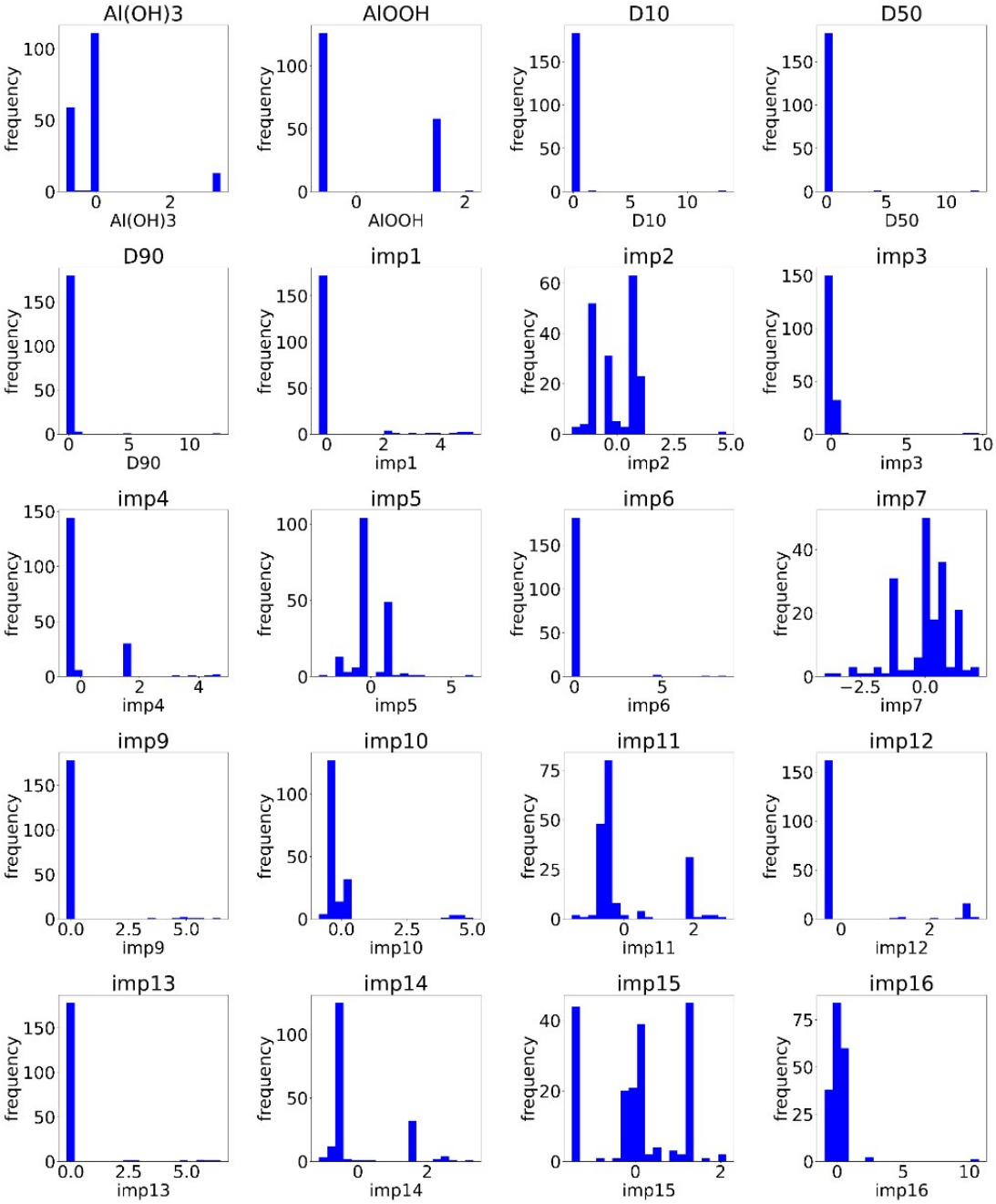

Figure S2-1 Histograms for all descriptors in input data.

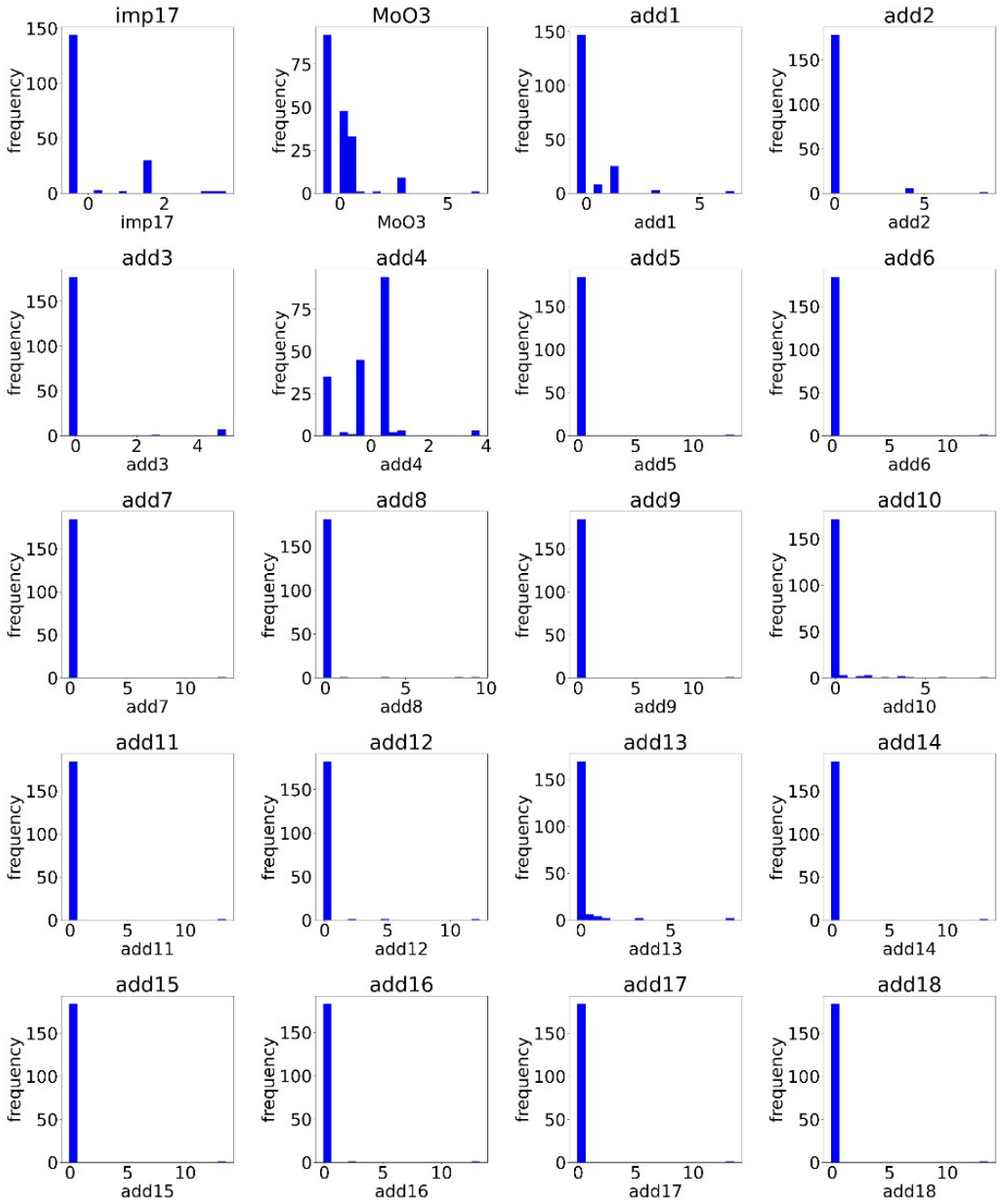

Figure S2-2 Histograms for all descriptors in input data.

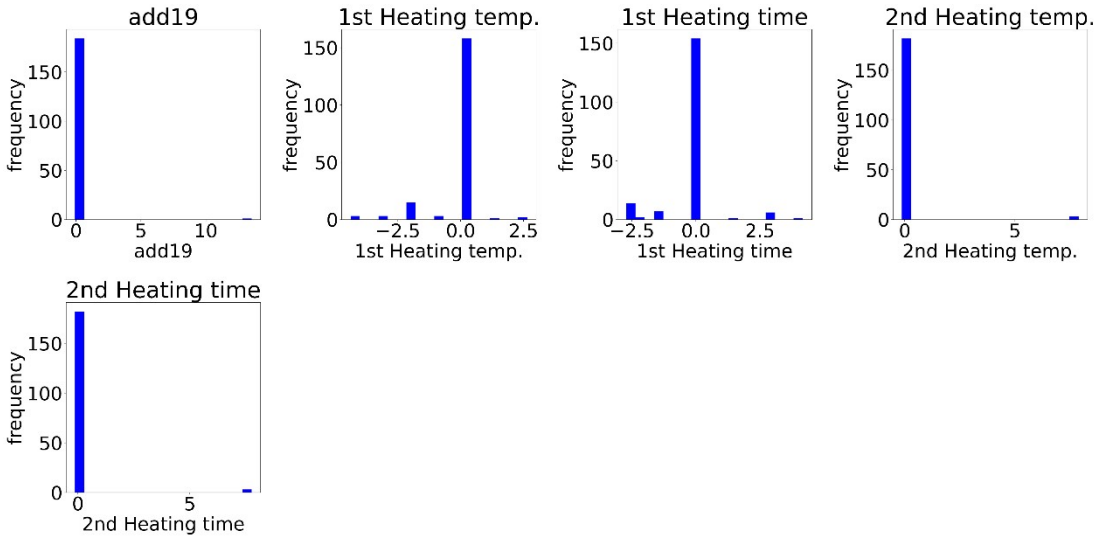

Figure S2-3 Histograms for all descriptors in input data.

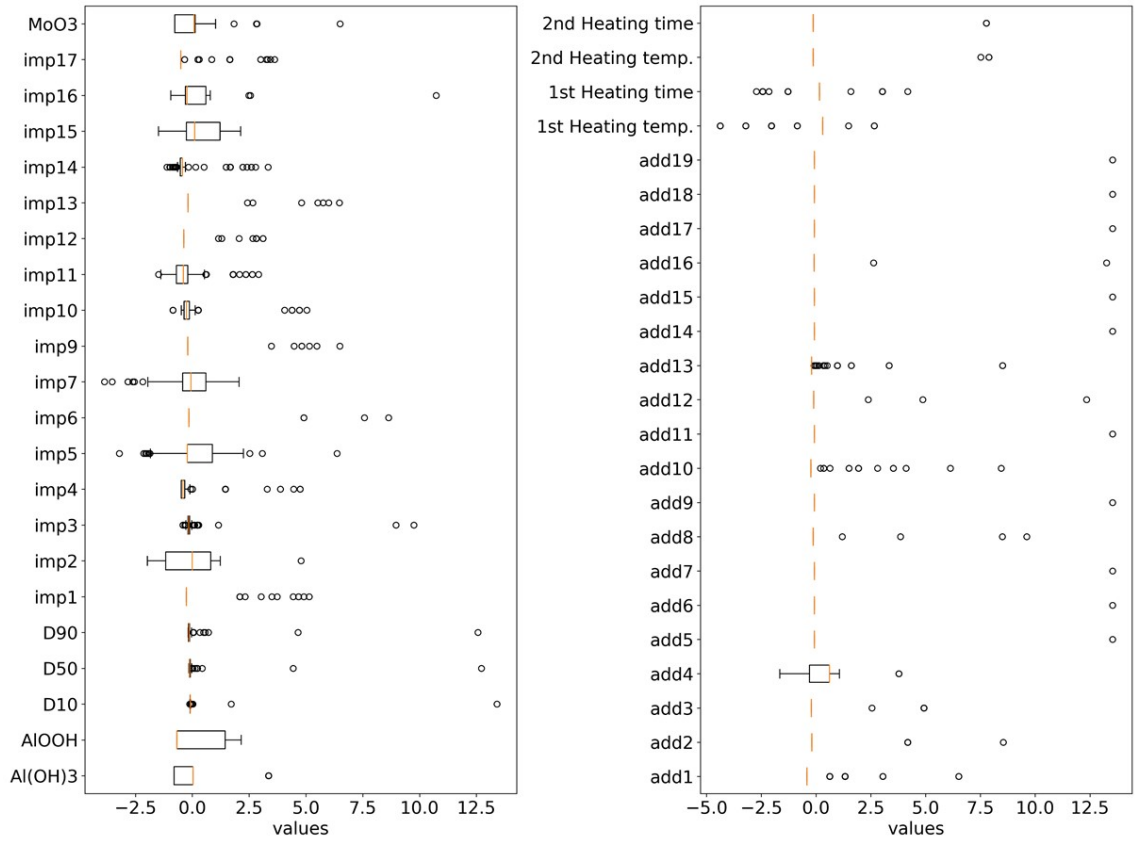

Figure S3 Box plot for all descriptors in input data.

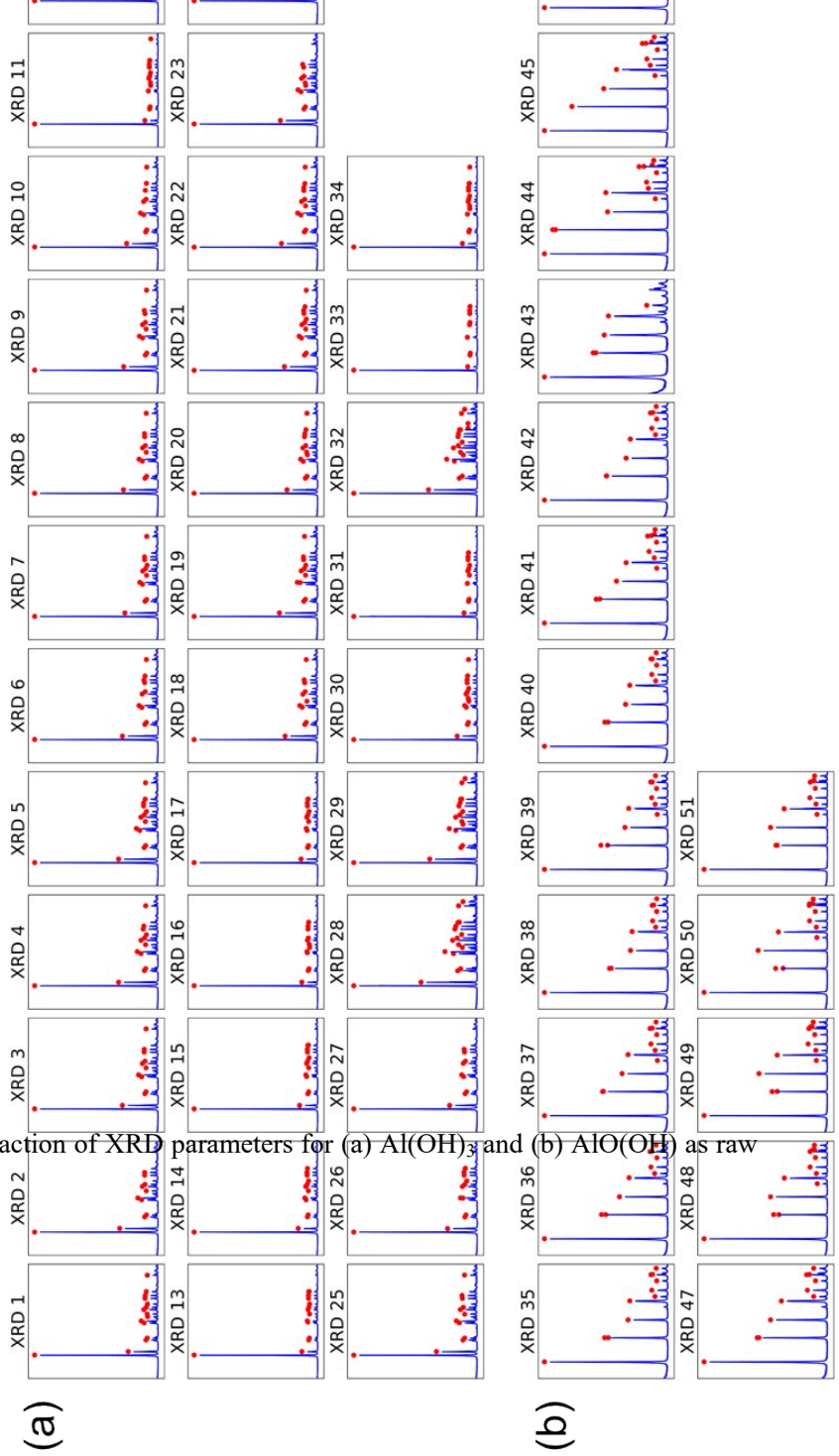

Figure S4 Identification and extraction of XRD parameters for (a) $Al(OH)_3$ and (b) AlO(OH) as raw materials.

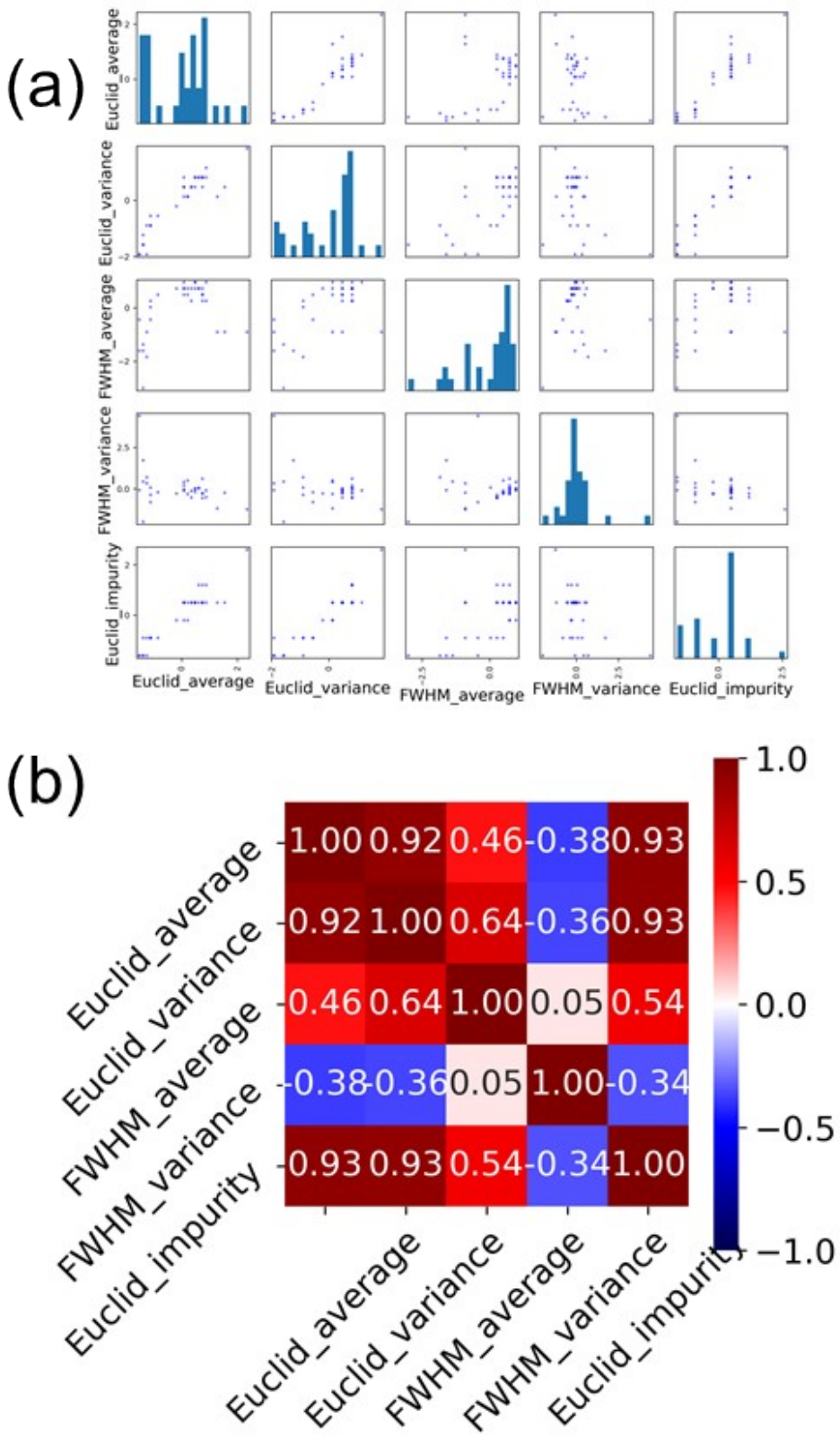

Figure S5 Scatter plots and heat maps of x-ray diffraction (XRD) descriptors for $Al(OH)_3$ (a,b).

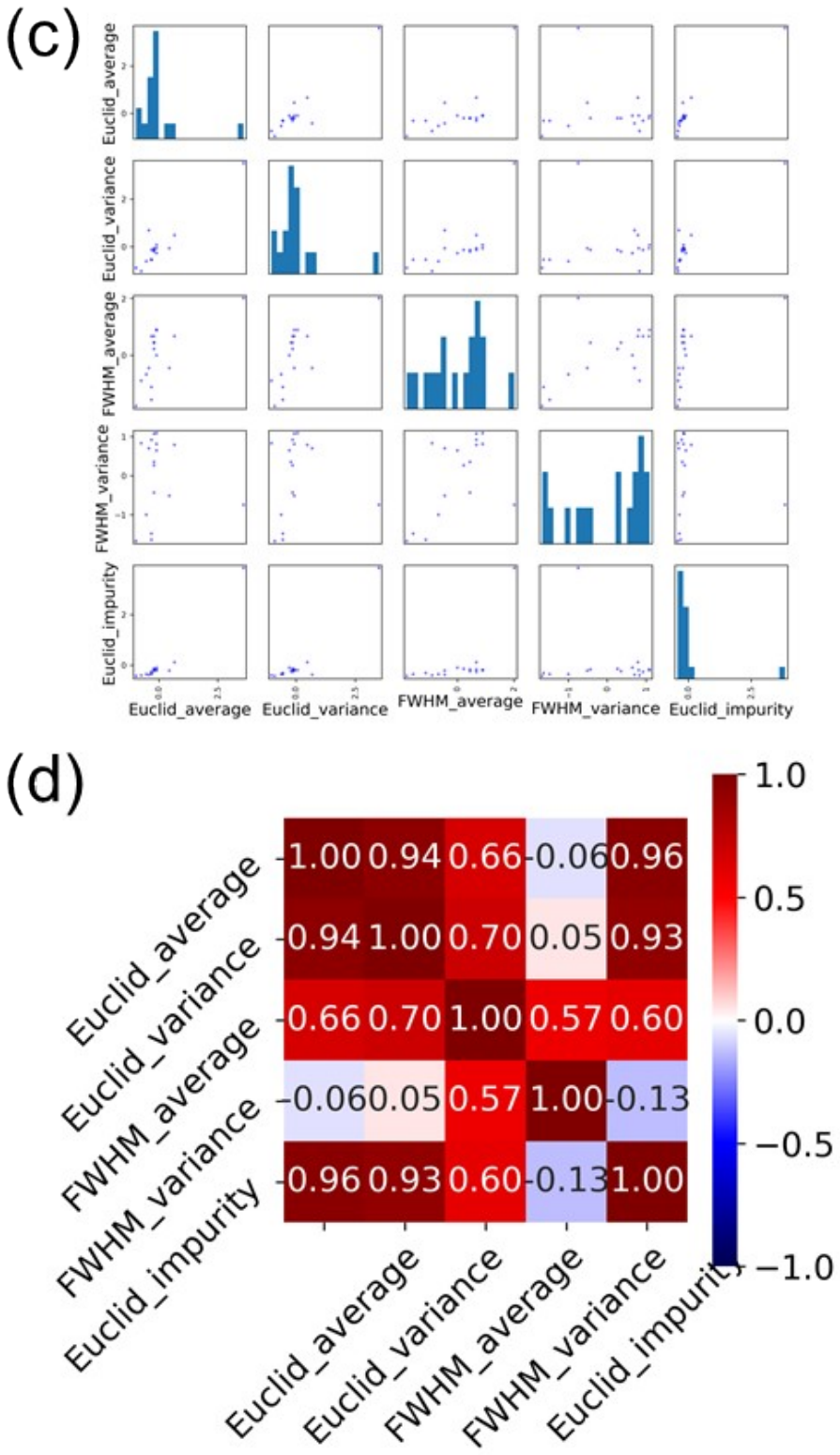

Figure S5 Scatter plots and heat maps of x-ray diffraction (XRD) descriptors for AlO(OH) (c,d).

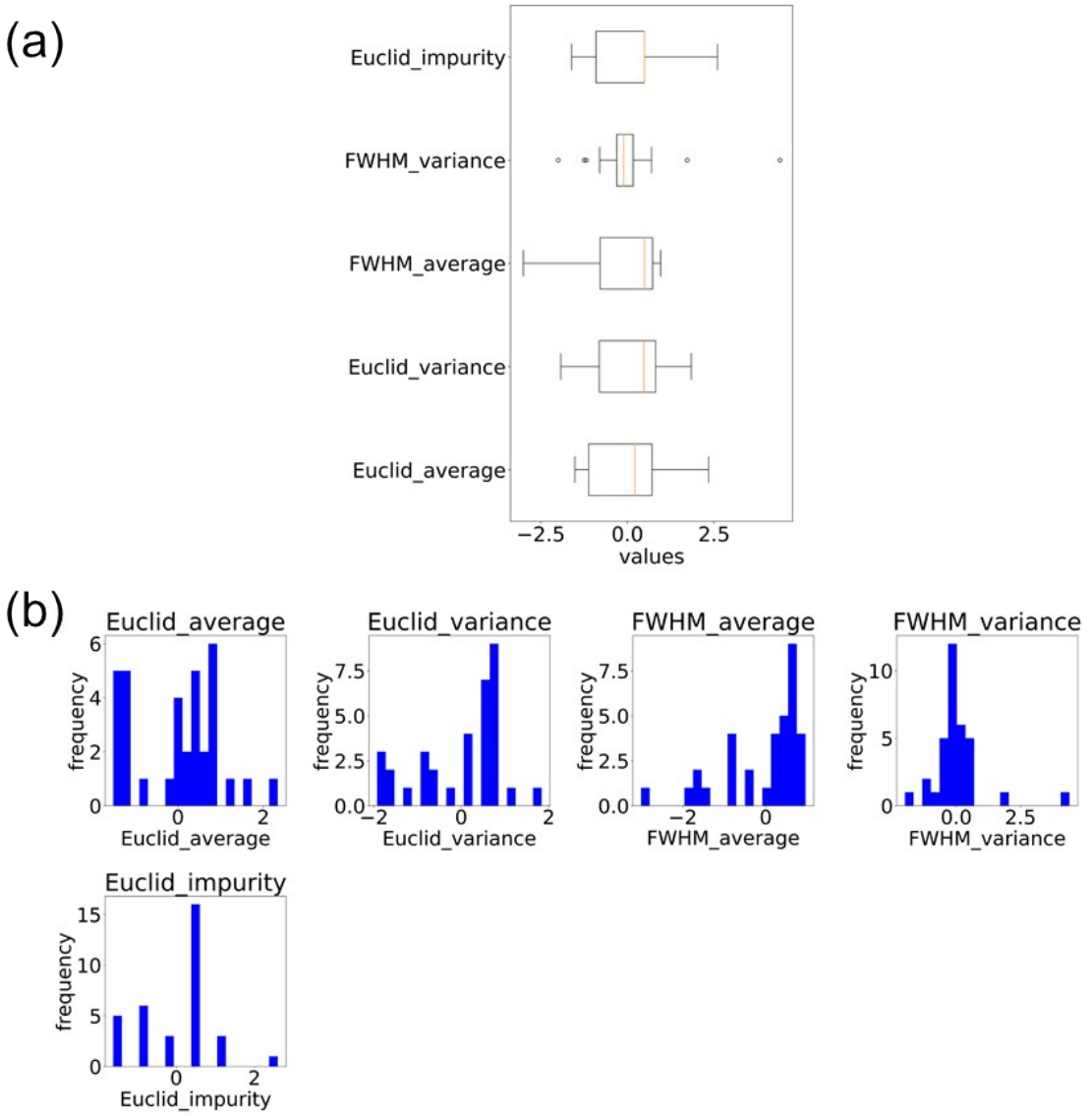

Figure S6 Box plots and histograms of XRD descriptors for $Al(OH)_3$ (a,b).

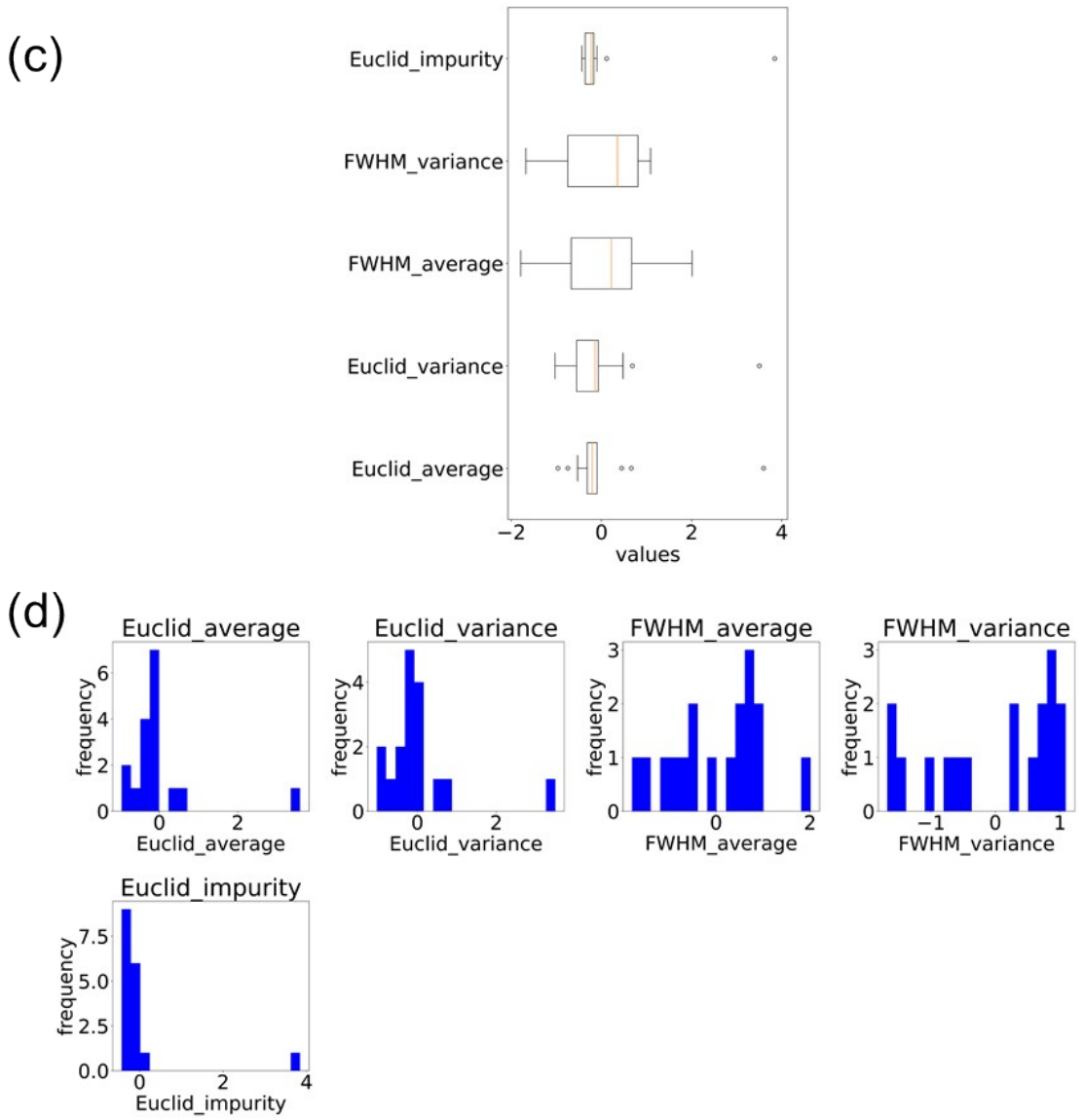

Figure S6 Box plots and histograms of XRD descriptors for AlO(OH) (c,d).

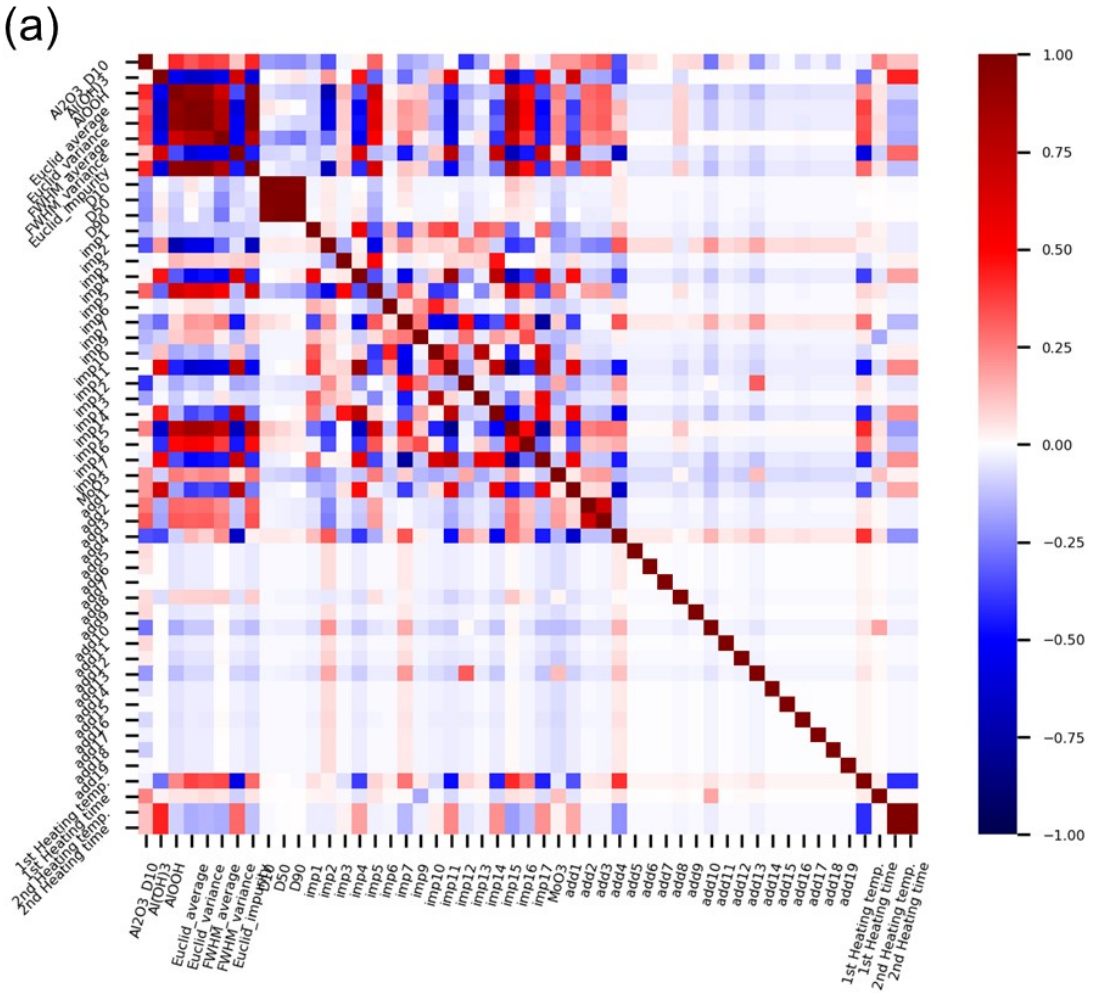

Figure S7 (a) Correlation-coefficients plots for all data.

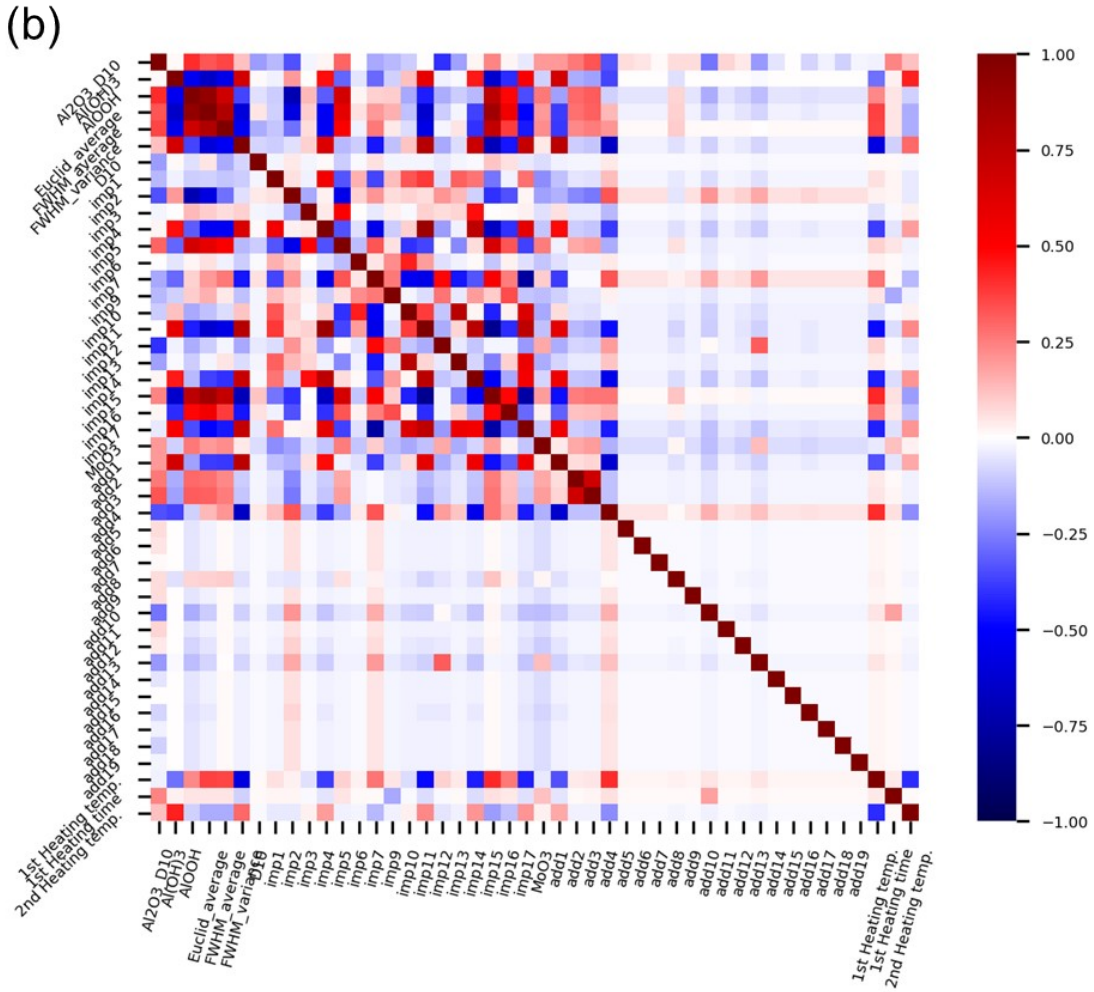

Figure S7 (b) Correlation-coefficients plots for all data after deletion of strong correlation factors.

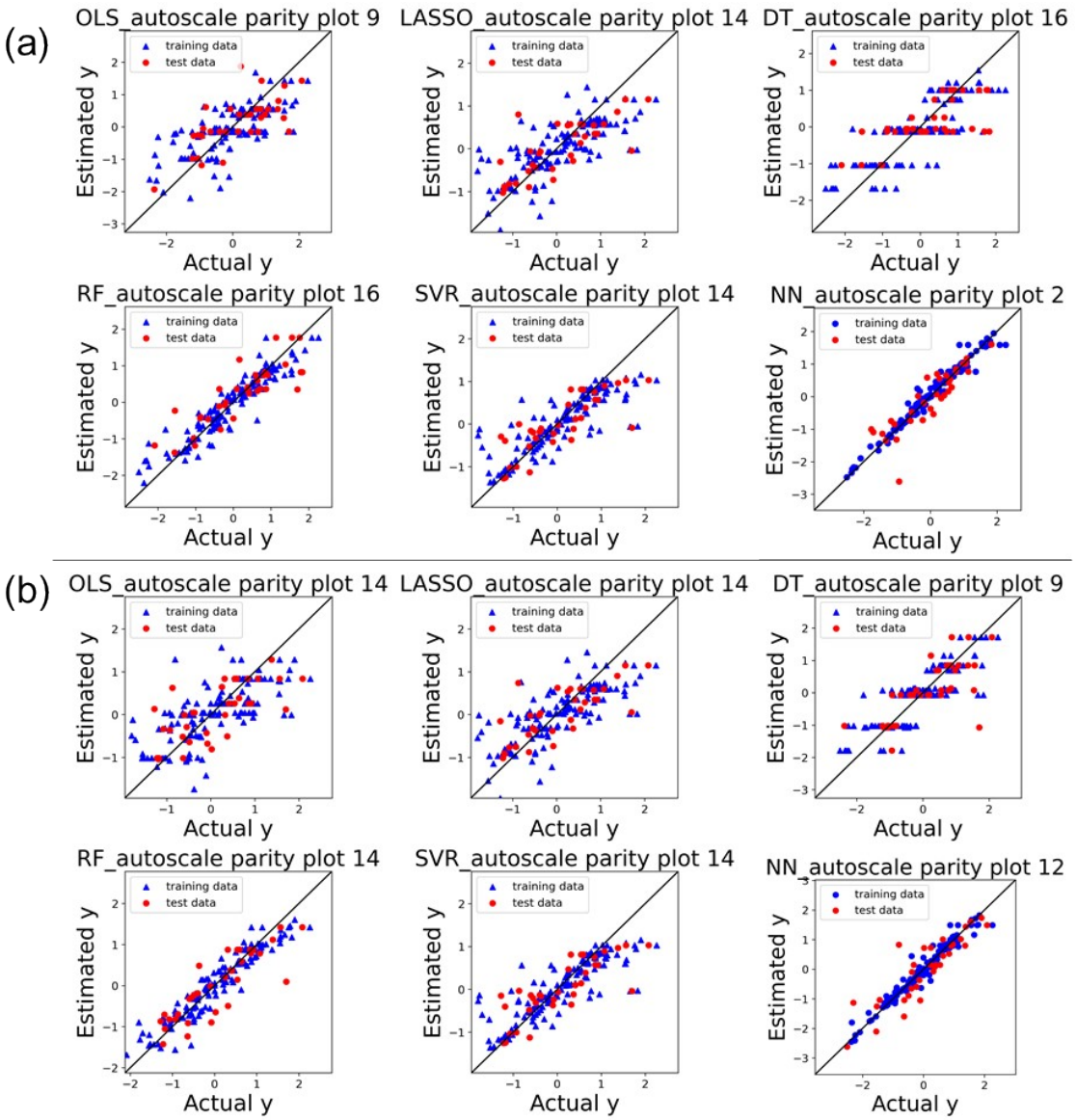

Figure S8 Parity plots of OLS, LASSO, DT, RF, SVR, and NN for dataset (a) without XRD, PD parameters, (b) without XRD parameters.

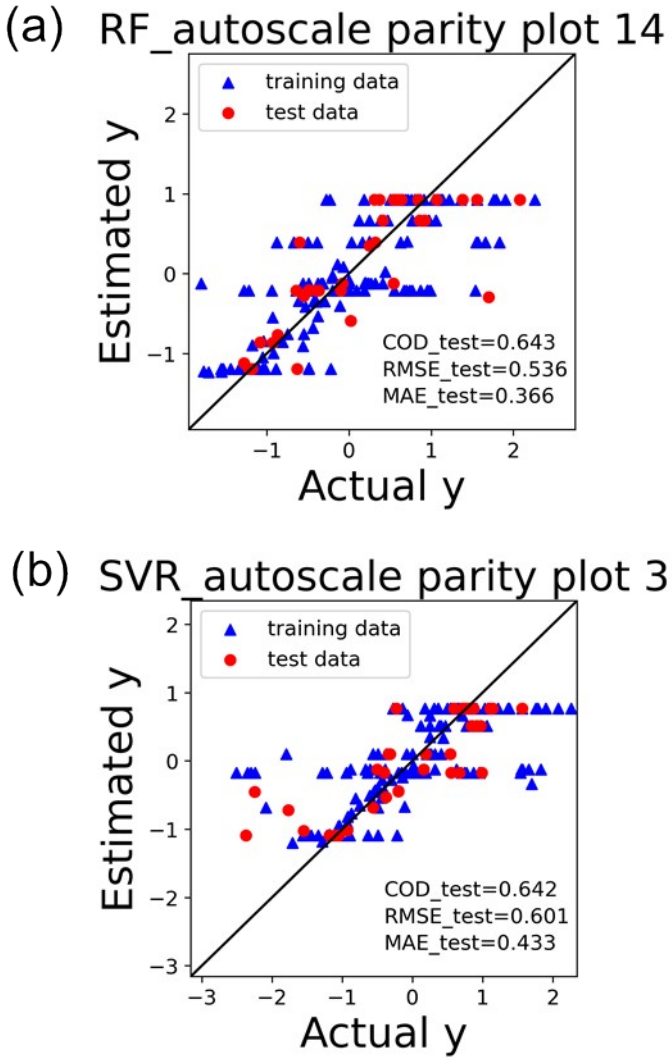

Figure S9 Parity plots of (a) RF, and (b) SVR for dataset with only raw materials, and raw material conditions descriptors.

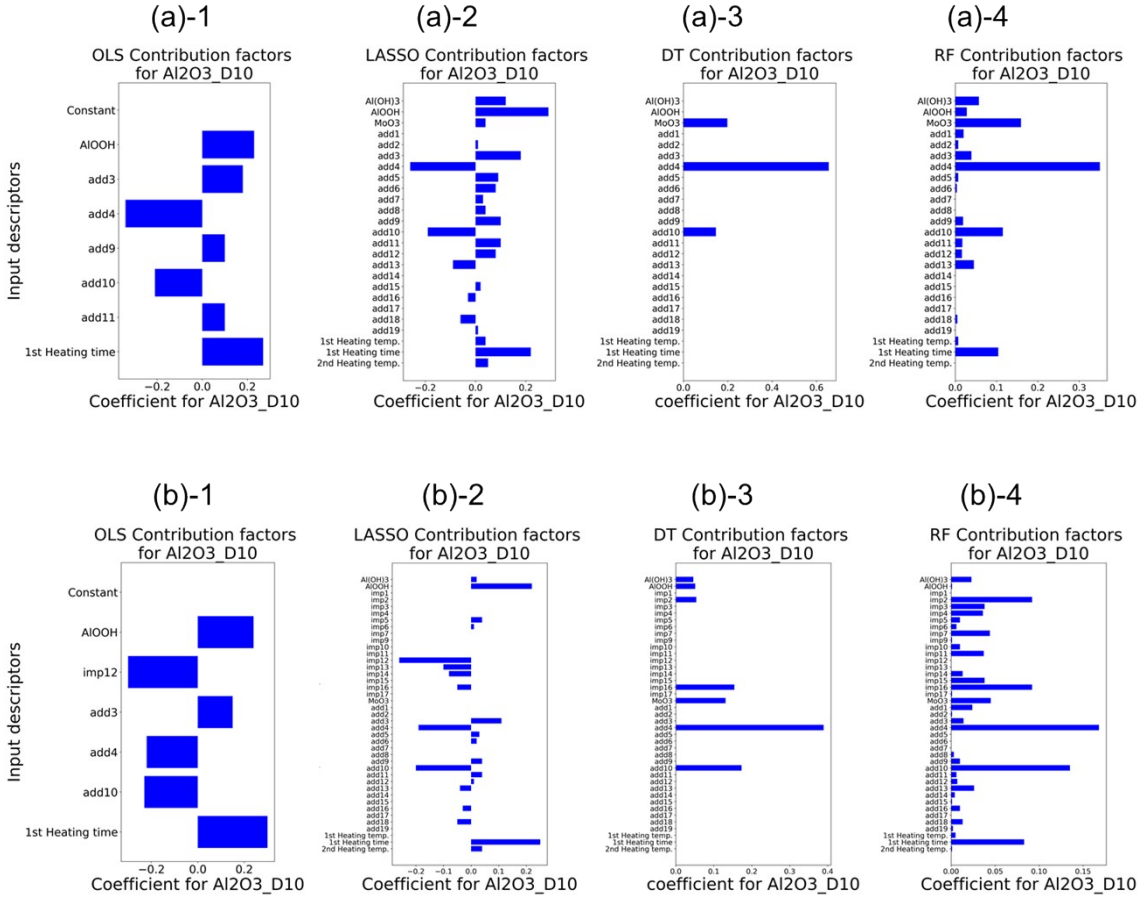

Figure S10 Model coefficients for dataset (a) without XRD, PD, IE parameters (OLS((a)-1), LASSO((a)-2), DT((a)-3), RF((a)-4)), (b) without XRD, PD parameters (OLS((b)-1), LASSO((b)-2), DT((b)-3), RF((b)-4)).

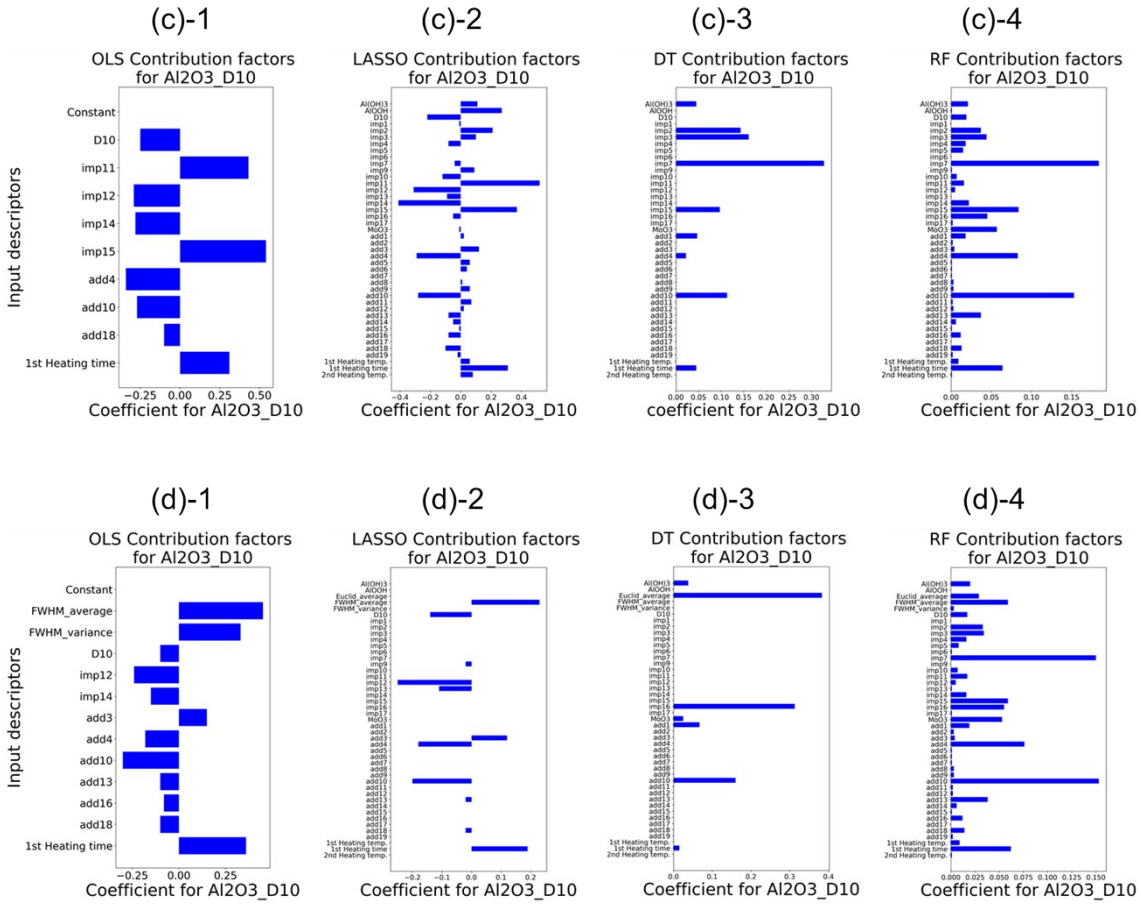

Figure S10 Model coefficients for dataset (c) without XRD parameters (OLS((c)-1), LASSO((c)-2), DT((c)-3), RF((c)-4)), (d) with all parameters (OLS((d)-1), LASSO((d)-2), DT((d)-3), RF((d)-4)).

Table S1 Model coefficients for dataset in table: (a) without XRD, PD, IE parameters.

Descriptor	OLS	LASSO	DT	RF
Al(OH)3	0	0.115239	0	0.057
AlOOH	0.228632	0.292157	0	0.028
MoO3	0	0.034068	0.198	0.159
add1	0	0	0	0.02
add2	0	0.013041	0	0.007
add3	0.180792	0.176013	0	0.039
add4	-0.34131	-0.26007	0.655	0.35
add5	0	0.093305	0	0.007
add6	0	0.080268	0	0.004
add7	0	0.033489	0	0
add8	0	0.045366	0	0
add9	0.100708	0.09944	0	0.019
add10	-0.21342	-0.19206	0.147	0.115
add11	0.103775	0.102507	0	0.017
add12	0	0.078764	0	0.016
add13	0	-0.09341	0	0.045
add14	0	0	0	0
add15	0	0.025053	0	0
add16	0	-0.03525	0	0
add17	0	0	0	0
add18	0	-0.05682	0	0.005
add19	0	0.01125	0	0
1st Heating temp.	0	0.044574	0	0.007
1st Heating time	0.271277	0.22372	0	0.104
2nd Heating temp.	0	0.053213	0	0

Table S1 Model coefficients for dataset in table: (b) without XRD, PD parameters.

Descriptor	OLS	LASSO	DT	RF
Al(OH)3	0	0.016878	0.046	0.023
AlOOH	0.307669	0.215449	0.051	0.001
imp1	0	0	0	0
imp2	0	0	0.054	0.092
imp3	0	0	0	0.038
imp4	0	0	0	0.036
imp5	0	0.038155	0	0.01
imp6	0	0.010093	0	0.006
imp7	0	0	0	0.044
imp9	0	0	0	0.001
imp10	0	0	0	0.01
imp11	0	0	0	0.037
imp12	-0.26473	-0.25601	0	0
imp13	0	-0.09865	0	0
imp14	0	-0.07584	0	0.013
imp15	0	0	0	0.038
imp16	-0.15933	-0.04795	0.154	0.092
imp17	0	0	0	0.001
MoO3	0	0	0.131	0.045
add1	0	0	0	0.024
add2	0	0	0	0.001
add3	0.178204	0.112414	0	0.014
add4	-0.19426	-0.19495	0.389	0.168
add5	0	0.03252	0	0
add6	0	0.019483	0	0
add7	0	0	0	0
add8	0	0	0	0.003
add9	0	0.038655	0	0.01
add10	-0.24624	-0.20022	0.173	0.135
add11	0	0.041722	0	0.006
add12	0	0.006789	0	0.007
add13	0	-0.04274	0	0.026

add14	0	-0.0021	0	0.004
add15	0	0	0	0.001
add16	0	-0.02952	0	0.01
add17	0	0	0	0
add18	0	-0.04658	0	0.013
add19	0	0	0	0.002
1st Heating temp.	0	0	0	0.005
1st Heating time	0.345129	0.248224	0	0.083
2nd Heating temp.	0	0.039911	0	0.001

Table S1 Model coefficients for dataset in table: (c) without XRD parameters.

Descriptor	OLS	LASSO	DT	RF
Al(OH)3	0	0.107784	0.045	0.021
AlOOH	0	0.272957	0	0.001
D10	-0.24999	-0.21816	0	0.019
imp1	0	-0.01263	0	0
imp2	0	0.217685	0.143	0.037
imp3	0	0.099373	0.161	0.044
imp4	0	-0.06652	0	0.018
imp5	0	0	0	0.015
imp6	0	0	0	0
imp7	0	-0.03902	0.328	0.184
imp9	0	0.090353	0	0.001
imp10	0	-0.10648	0	0.007
imp11	0.429305	0.503679	0	0.016
imp12	-0.28634	-0.30561	0	0.005
imp13	0	-0.09246	0	0
imp14	-0.28504	-0.39904	0	0.022
imp15	0.542295	0.369144	0.097	0.084
imp16	0	-0.05661	0	0.045
imp17	0	0	0	0.002
MoO3	0	-0.01635	0	0.057
add1	0	0.018436	0.047	0.018
add2	0	-0.00294	0	0.002
add3	0	0.123136	0	0.004
add4	-0.33834	-0.28743	0.022	0.083
add5	0	0.05581	0	0.001
add6	0	0.042773	0	0.001
add7	0	0	0	0.001
add8	0	0.005769	0	0.003
add9	0	0.061945	0	0.003
add10	-0.26749	-0.28337	0.113	0.153
add11	0	0.065012	0	0.002
add12	0	0.02311	0	0.003

add13	0	-0.07494	0	0.037
add14	0	-0.05491	0	0.006
add15	0	-0.00737	0	0.001
add16	0	-0.08483	0	0.012
add17	0	0	0	0
add18	-0.09675	-0.09939	0	0.013
add19	0	-0.02117	0	0.002
1st Heating temp.	0	0.062647	0	0.009
1st Heating time	0.305291	0.309376	0.045	0.064
2nd Heating temp.	0	0.07661	0	0.001

Table S1 Model coefficients for dataset in table: (d) with all parameters.

Descriptor	OLS	LASSO	DT	RF
Al(OH)3	0	0	0.038	0.02
AlOOH	0	0.002	0	0
Euclid_average	0.297	0	0.383	0.029
FWHM_average	0	0.228	0	0.059
FWHM_variance	0.26	0	0	0.003
D10	-0.166	-0.148	0	0.017
imp1	0	0	0	0
imp2	0	0	0	0.033
imp3	0	0	0	0.034
imp4	0	0	0	0.016
imp5	0	0	0	0.008
imp6	0	0	0	0.001
imp7	0	0	0	0.15
imp9	0	-0.023	0	0
imp10	0	0	0	0.007
imp11	0	0	0	0.017
imp12	-0.229	-0.249	0	0.005
imp13	0	-0.112	0	0.001
imp14	-0.125	0	0	0.016
imp15	0	0	0	0.059
imp16	0	0	0.313	0.055
imp17	0	0	0	0.001
MoO3	0	0	0.025	0.053
add1	0	0	0.067	0.019
add2	0	0	0	0.003
add3	0.151	0.116	0	0.004
add4	-0.19	-0.184	0	0.076
add5	0	0	0	0.001
add6	0	0	0	0.001
add7	0	0	0	0.001
add8	0	0	0	0.003
add9	0	0	0	0.003

add10	-0.277	-0.196	0.16	0.153
add11	0	0	0	0.002
add12	0	0	0	0.002
add13	-0.094	-0.023	0	0.038
add14	0	0	0	0.006
add15	0	0	0	0.001
add16	0	-0.004	0	0.012
add17	0	0	0	0
add18	-0.091	-0.019	0	0.014
add19	0	0	0	0.002
1st Heating temp.	0	0	0	0.009
1st Heating time	0.351	0.193	0.015	0.062
2nd Heating temp.	0	0	0	0.001