Model Performance Comparison Tables

1 Original Paper Results

Table 1: Original Results from Paper

Models	**MATR1**	**MATR2**	**HUST**	**SNL**	**CLO**	**CRUH**	**CRUSH**	**MIX**
Dummy regressor	398	510	419	466	331	239	576	573
"Variance" model	136	211	398	360	179	118	506	521
"Discharge" model	329	**149**	**322**	267	143	76	> 1000	> 1000
"Full" model	167	> 1000	335	433	**138**	93	> 1000	331
Ridge regression	116	184	> 1000	242	169	65	> 1000	372
PCR	**90**	187	435	**200**	197	68	560	376
PLSR	104	181	431	242	176	**60**	535	383
Gaussian process	154	224	> 1000	251	204	115	> 1000	573
XGBoost	334	799	395	547	215	119	**330**	205
Random forest	168±9	233±7	368±7	532±25	192±2	81±1	416±5	**197±0**
MLP	149±3	275±27	459±9	370±81	146±5	103±4	565±9	451±42
CNN	102±94	228±104	465±75	924±267	> 1000	174±92	545±11	272±101
LSTM	119±11	219±33	443±29	539±40	222±12	105±10	519±39	268±9
Transformer	135±13	364±25	391±11	424±23	187±14	81±8	550±21	271±16

2 Reproduced Results

Table 2: Reproduced Results

Model	**MATR1**	**MATR2**	**HUST**	**SNL**	**CLO**	**CRUH**	**CRUSH**	**MIX**
discharge_model	329.0481	148.5821	321.5733	266.6539	143.7331	76.2897	13527.6553	1735.3196
dummy	398.8229	510.6284	419.5217	466.1380	331.2758	239.8081	576.1726	573.1156
full_model	166.7985	1074.4001	335.0267	944.4740	138.9457	98.0711	526.8979	332.1685
gpr	153.9130	223.7908	54230.9697	254.5108	203.6551	114.4373	140246444603	595769336
pcr	123.1122	197.5088	434.8146	186.0309	197.1068	68.7659	557.3392	376.8111
plsr	103.6526	180.5339	431.2634	236.1880	176.0135	64.6230	533.8002	383.4731
rf	164.3123	231.4127	374.3944	524.7751	192.0838	82.2843	439.2723	199.8695
ridge	115.7892	183.7997	54194.8625	244.3086	169.9931	65.5719	2268509	372.0149
variance_model	136.1296	211.2364	398.1423	360.3418	179.0192	118.3824	506.8378	521.2020
xgb	333.6645	798.8044	394.7010	547.2919	214.5909	119.3553	341.1997	204.7945
mlp	139.8188	312.8503	440.4353	471.4404	148.8035	97.6600	561.1003	406.7105
lstm	105.8686	290.1655	445.6887	494.2307	222.6967	120.4489	525.0753	247.9591
transformer	83.1587	226.6437	406.6988	353.4602	154.6262	80.1458	592.4867	273.7350
cnn	76.8833	192.1275	154.6262	772.6318	353.4602	327.2960	592.4867	–

3 Comparison Table with Color Coding

Table 3: Comparison Results (Green: Within Mean±SD, Red: Outside Mean±SD)

Model	MATR1	MATR2	HUST	SNL	CLO	CRUH	CRUSH	MIX
discharge_model	329.0	148.6	321.6	266.7	143.7	76.3	13527.7	1735.3
dummy	398.8	510.6	419.5	466.1	331.3	239.8	576.2	573.1
full_model	166.8	1074.4	335.0	944.5	138.9	98.1	526.9	332.2
gpr	153.9	223.8	54230.97	254.5	203.7	114.4	1.40E+11	5.96E+08
pcr	123.1	197.5	434.8	186.0	197.1	68.8	557.3	376.8
plsr	103.7	180.5	431.3	236.2	176.0	64.6	533.8	383.5
rf	164.3	231.4	374.4	524.8	192.1	82.3	439.3	199.9
ridge	115.8	183.8	54194.86	244.3	170.0	65.6	2268509	372.0
variance_model	136.1	211.2	398.1	360.3	179.0	118.4	506.8	521.2
xgb	333.7	798.8	394.7	547.3	214.6	119.4	341.2	204.8
mlp	139.8	312.9	440.4	471.4	148.8	97.7	561.1	406.7
lstm	105.9	290.2	445.7	494.2	222.7	120.4	525.1	247.96
transformer	83.2	226.6	406.7	353.5	154.6	80.1	592.5	273.7
cnn	76.9	192.1	154.6	772.6	353.5	327.3	592.5	–

4 Notes

- Green cells indicate reproduced results are within mean ± standard deviation of original results
- Red cells indicate reproduced results are outside the acceptable range
- For models without reported standard deviation in original paper, a standard deviation of 5 was assumed
- Empty cells (–) indicate missing reproduced results
- Values > 1000 in original results were treated as 1000 for comparison purposes
- Scientific notation used for extremely large values in some cases