Statistical Significance (p-values)
Green = Significant difference

	standard	tong_neuralode	time_indexed_mlp	time_indexed_ssm
standard	—	0.0617	0.0000	0.0082
tong_neuralode	0.0617	—	0.0001	0.0166
time_indexed_mlp	0.0000	0.0001	—	0.2223
time_indexed_ssm	0.0082	0.0166	0.2223	—
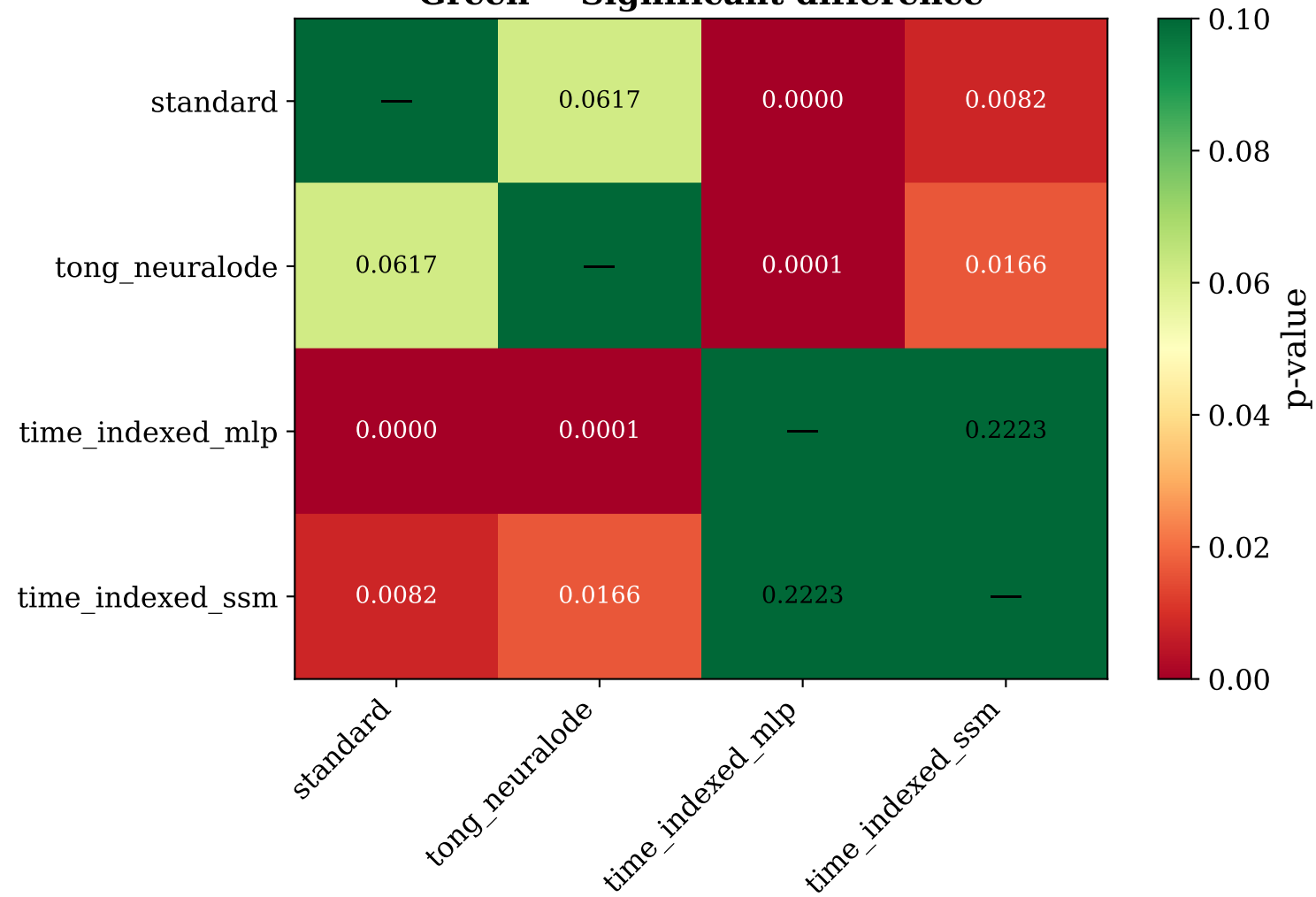
Effect Size (Cohen's d)
Red = Model1 better, Blue = Model2 better

	standard	tong_neuralode	time_indexed_mlp	time_indexed_ssm
standard	—	1.54	5.83	2.47
tong_neuralode	-1.54	—	4.90	2.13
time_indexed_mlp	-5.83	-4.90	—	0.94
time_indexed_ssm	-2.47	-2.13	-0.94	—