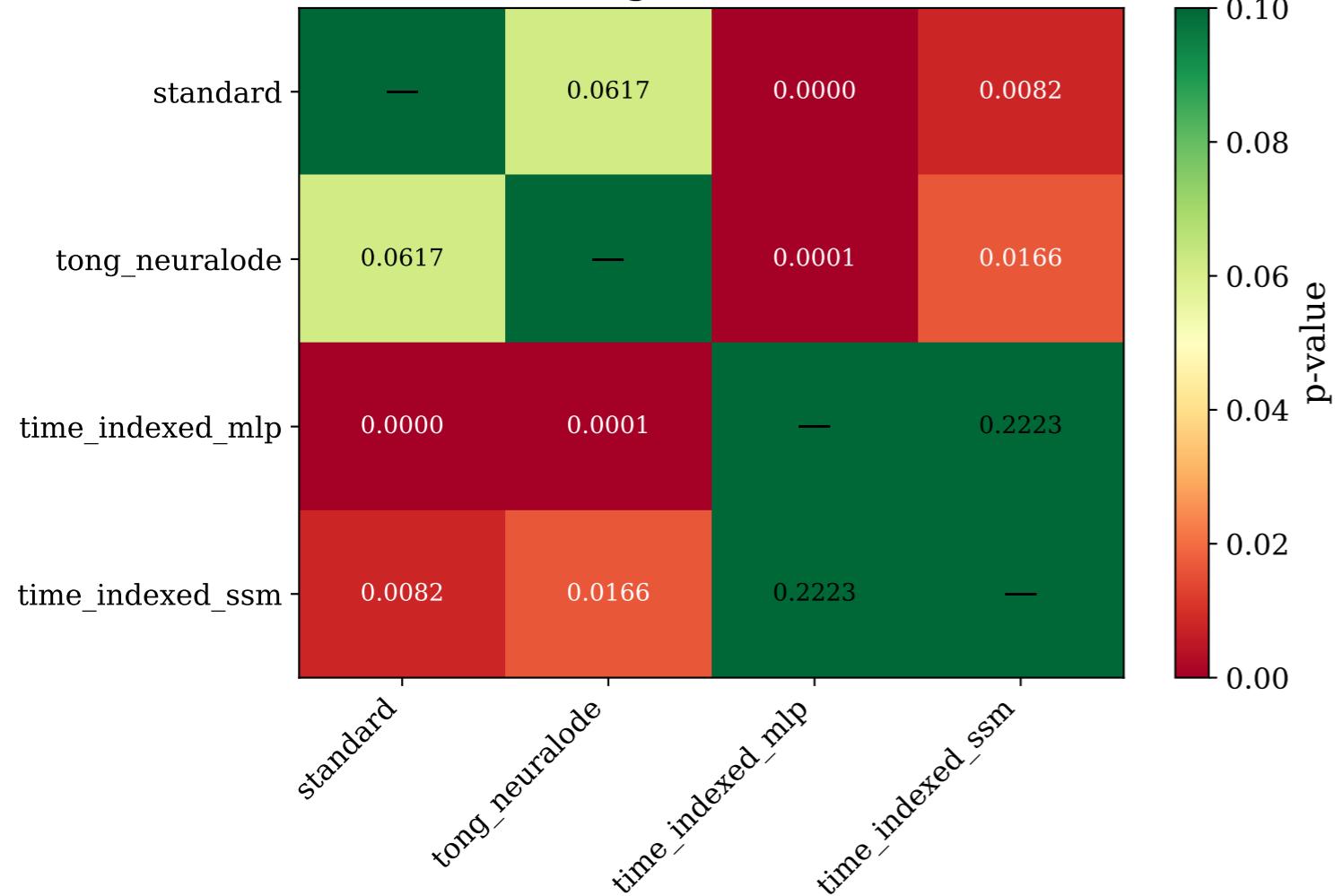


Statistical Significance (p-values)
Green = Significant difference



Effect Size (Cohen's d)
Red = Model1 better, Blue = Model2 better

