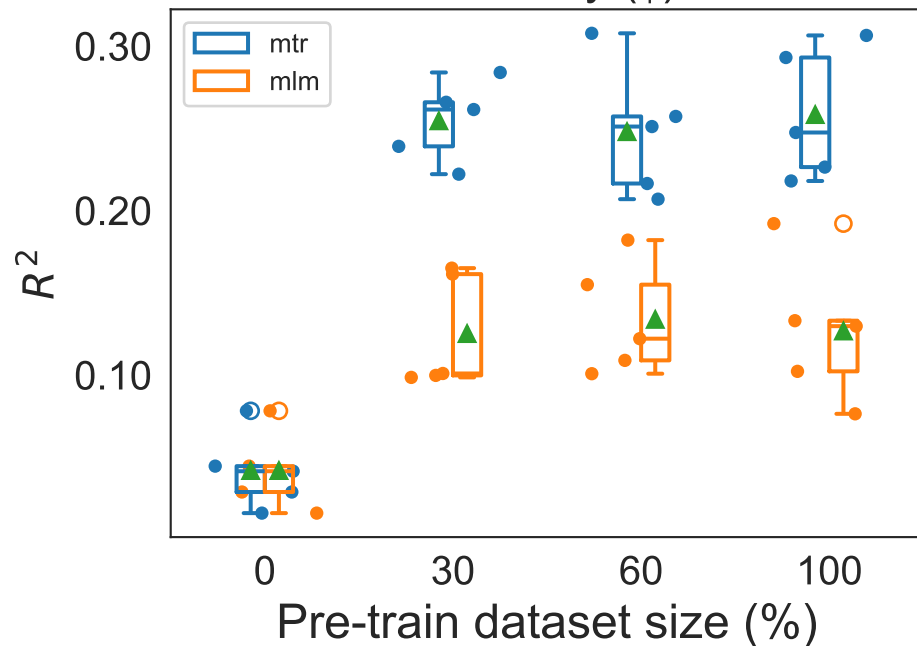
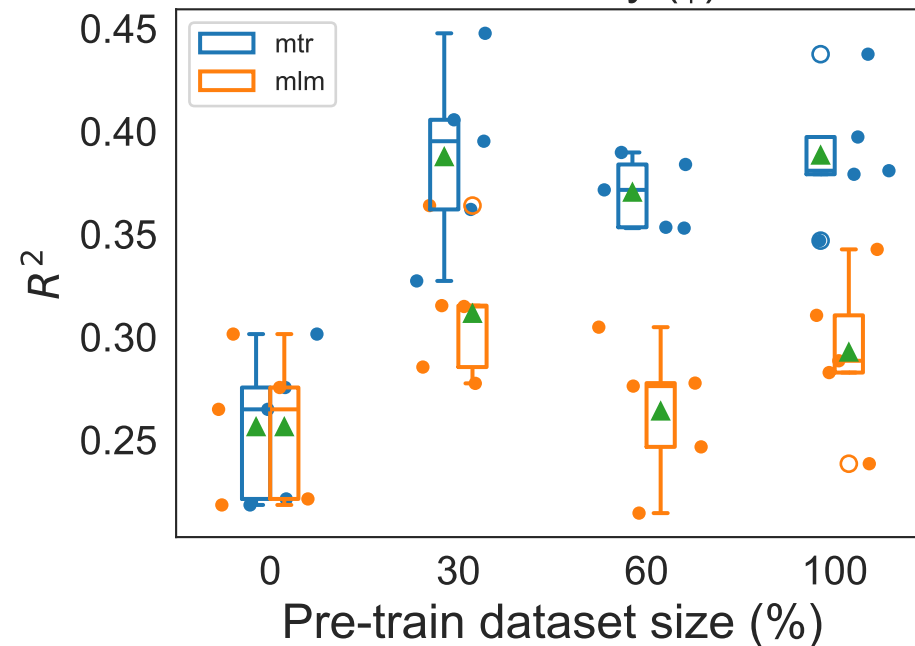


Comparing pre-training dataset sizes with respect to pre-training objective / split = datasail

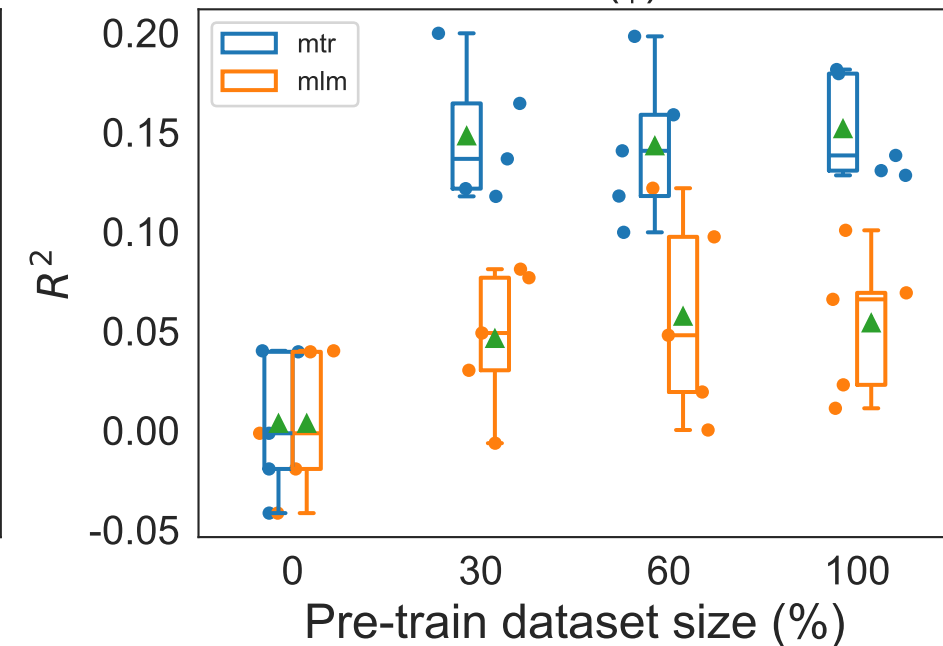
Solubility (\uparrow)



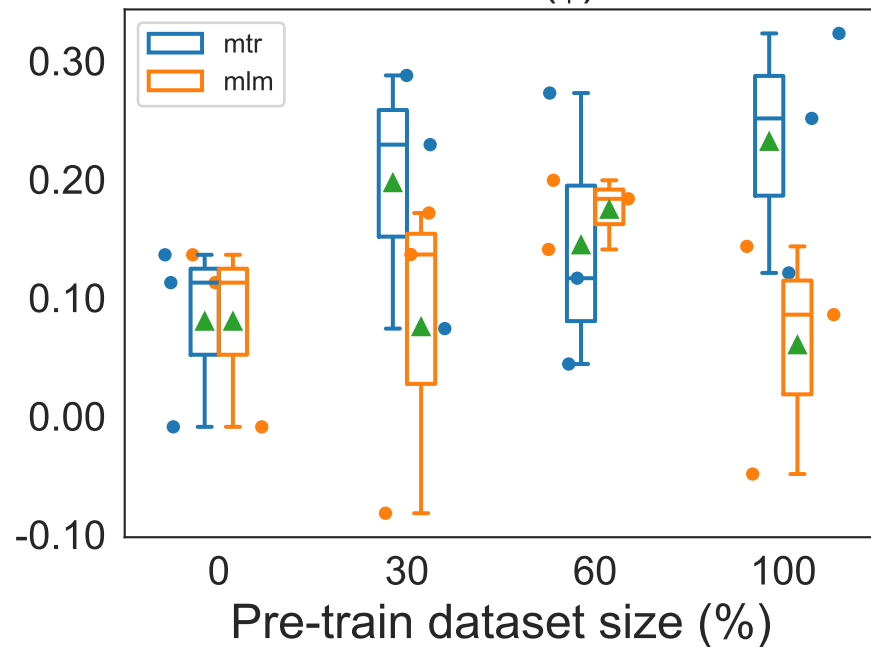
Permeability (\uparrow)



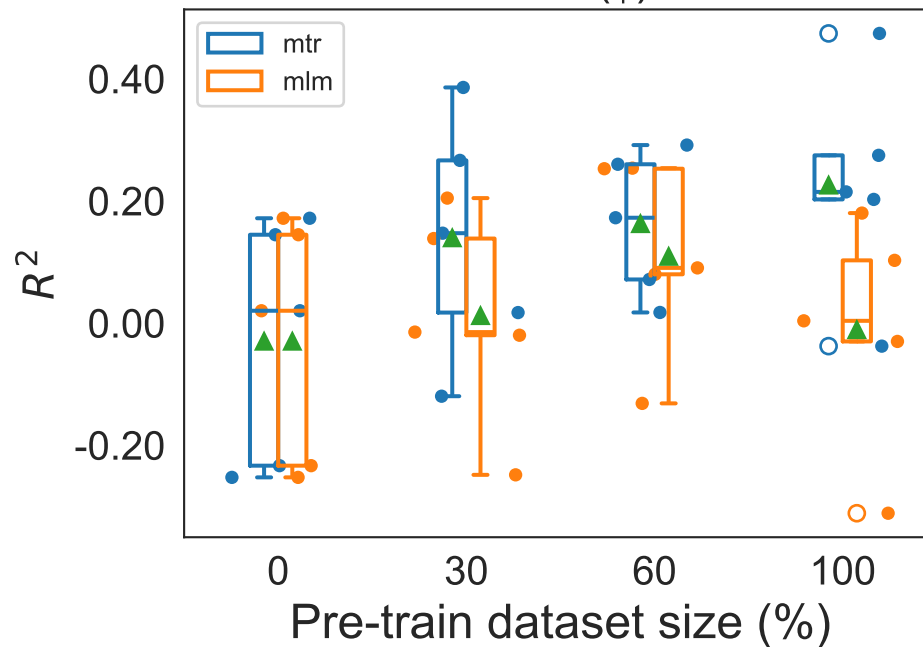
HLM (\uparrow)



hPPB (\uparrow)



rPPB (\uparrow)



RLM (\uparrow)

