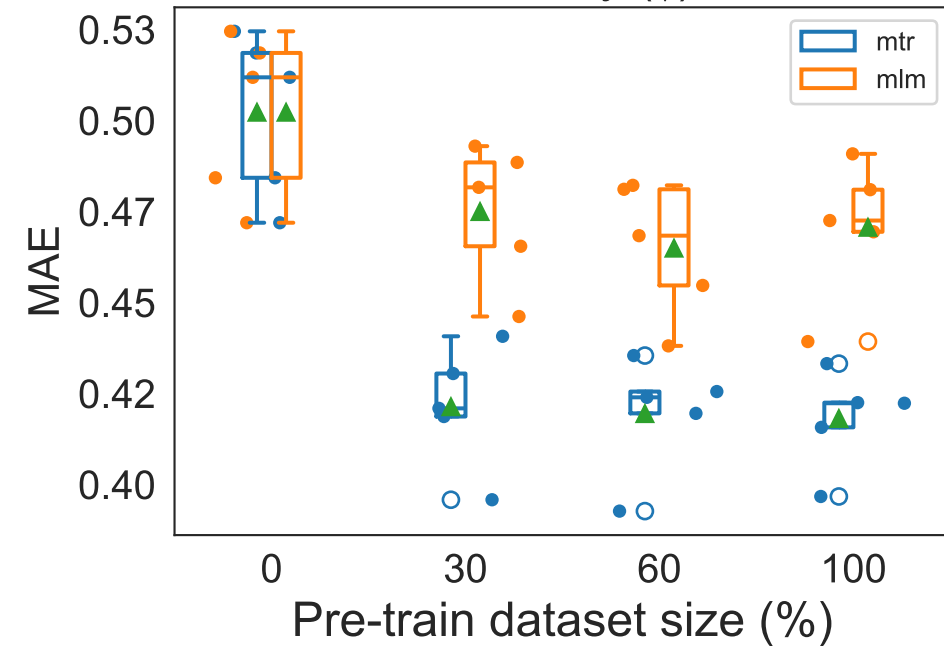
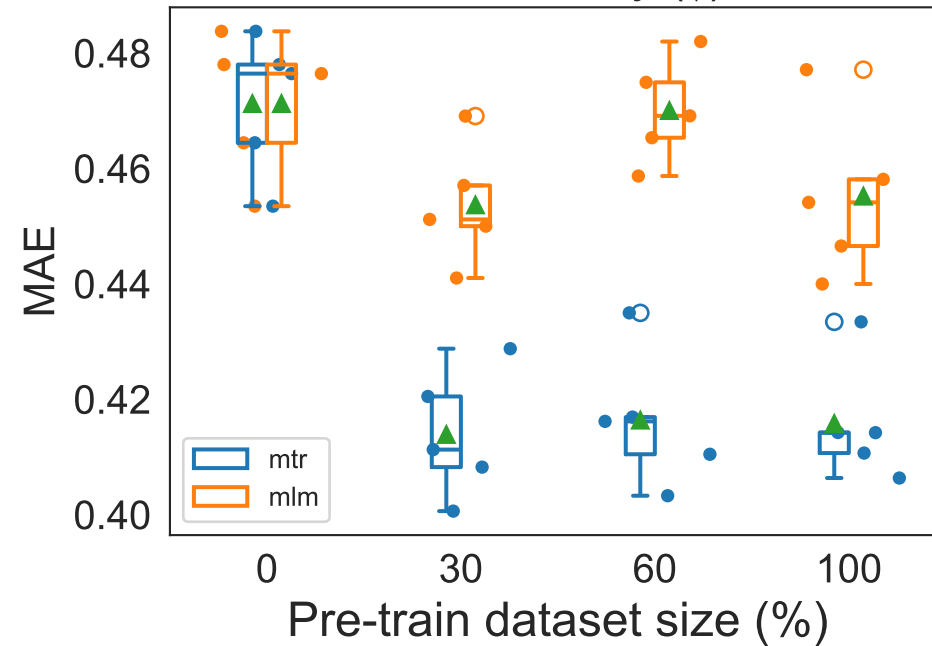


Comparing pre-training dataset sizes with respect to pre-training objective / split = random

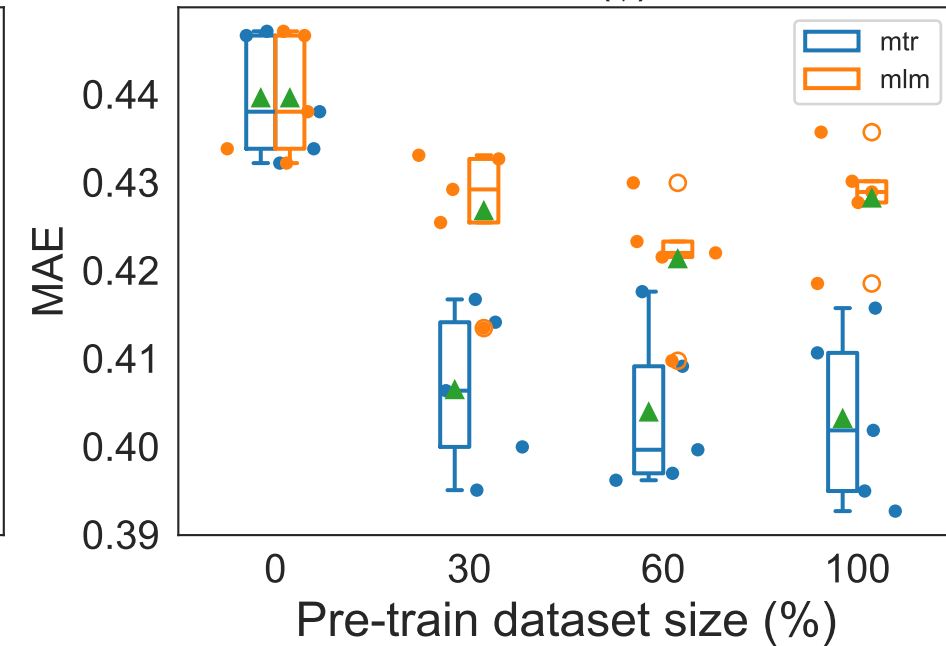
Solubility (\downarrow)



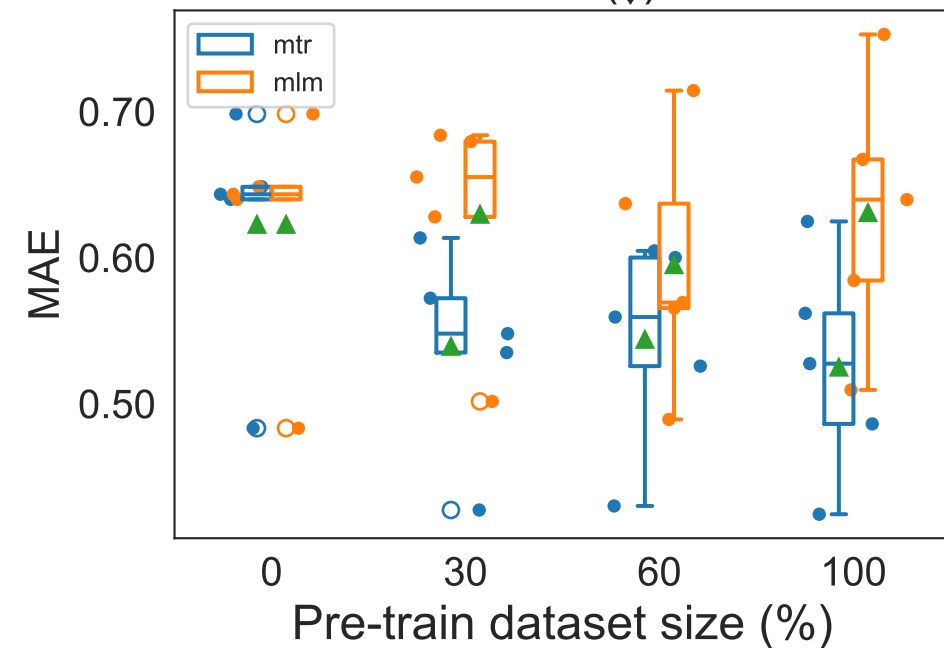
Permeability (\downarrow)



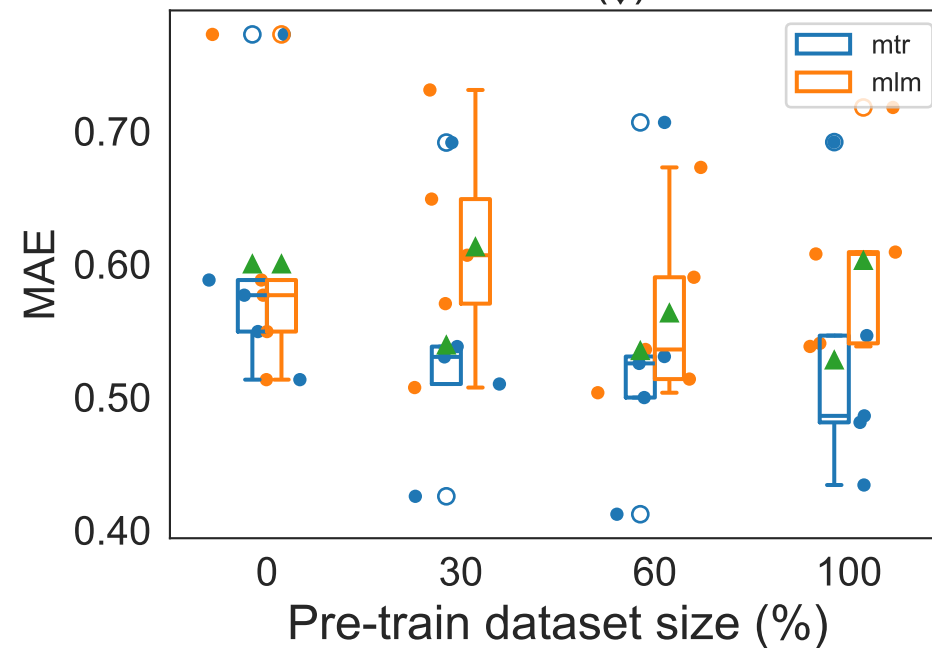
HLM (\downarrow)



hPPB (\downarrow)



rPPB (\downarrow)



RLM (\downarrow)

