Supporting Information for:

Sequential Design of Adsorption Simulations in Metal-Organic Frameworks

Krishnendu Mukherjee,[1] Alexander Dowling,[1] and Yamil J. Colón*[1]

[1]Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN, 46556, USA
*Corresponding author: ycolon@nd.edu

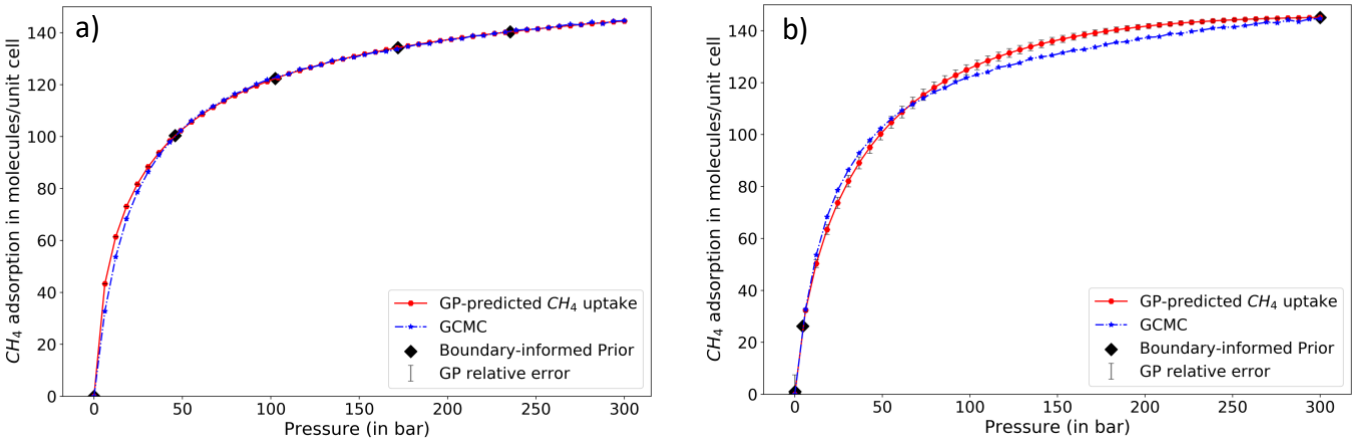

Figure S1. Final GP fit methane isotherm comparison with GCMC simulations (ground-truth) for a) linear-spaced LHS prior, and b) log-spaced LHS prior. We find the linear spaced prior has more deviations in the low-pressure region and has good accuracy for high pressure. For log-spaced prior there is large deviation at high pressure and moderate agreement with the GCMC simulation at low pressure.

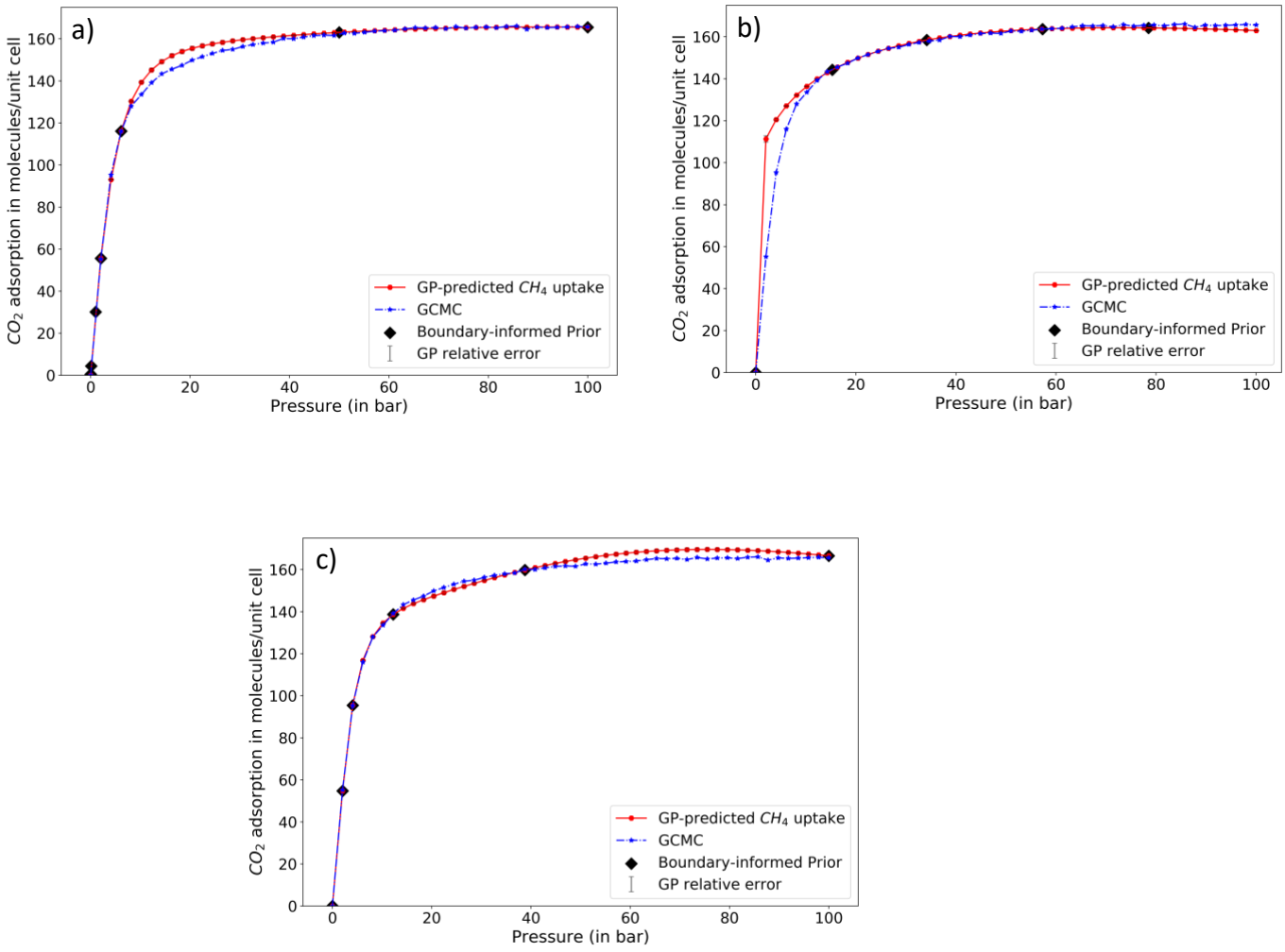

Figure S2. Final GP fit CO_2 isotherm comparison with GCMC simulations (ground-truth) for a) boundary-informed prior, b) linear-spaced LHS prior, and c) log-spaced LHS prior. We observe that boundary-informed prior overpredicts at the adsorption rise zone but performs very well at both the low- and high-pressure region. For the LHS-based priors, as was observed for methane adsorption, linear-spaced prior performs better at high pressure while log-spaced one does significantly better at low pressure.

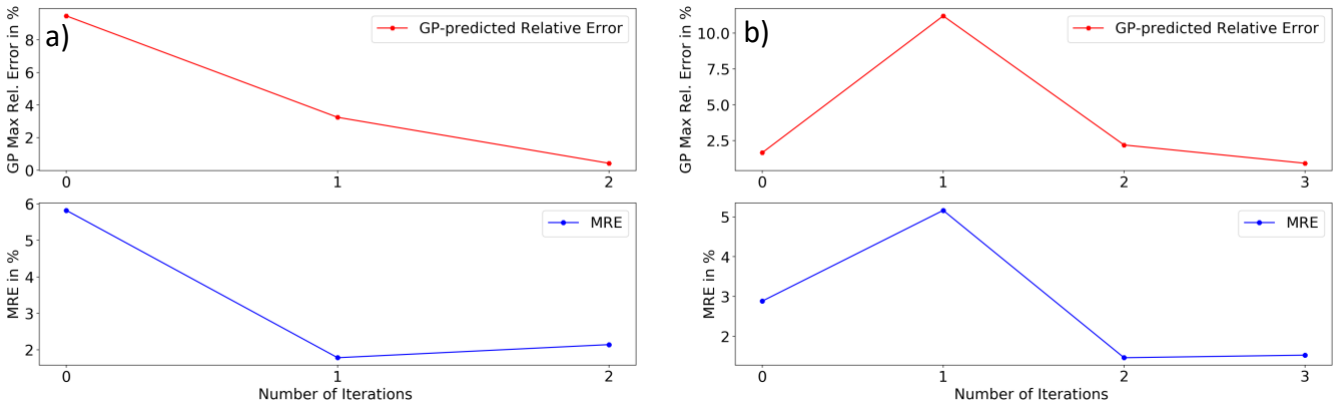

Figure S3. Maximum GP relative error and MRE (in %) with respect to number of AL iterations for a) methane adsorption, b) CO_2 adsorption for boundary-informed prior. This was done for AL for simulating adsorption isotherm in Cu-BTC at a temperature of 300 K (section 3.1 and 3.2). We find here that methane adsorption took only 2 number of AL iteration to converge at a 2 % while CO_2 adsorption took 3 iterations of AL.

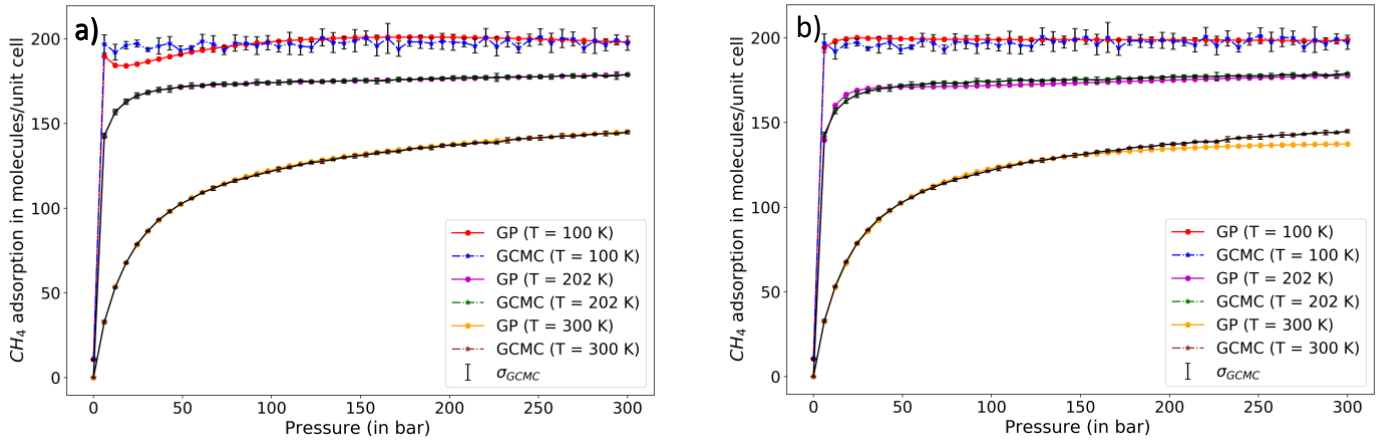

Figure S4. Comparison of GP-predicted methane uptake with GCMC simulation predicted for pressure range of 10^{-6} to 100 bar, at temperature of 100 K, 202 K and 300 K for a) linear-spaced LHS prior, and b) log-spaced LHS prior. We observe here that linear-spaced prior had good agreement with GCMC results at high temperature but for low temperature (at 100 K), there was error at the low-pressure region. For log-spaced prior we observe that the final model has good accuracy at low temperature but for high temperature of 300 K, there is deviation with GCMC simulation at high pressure.

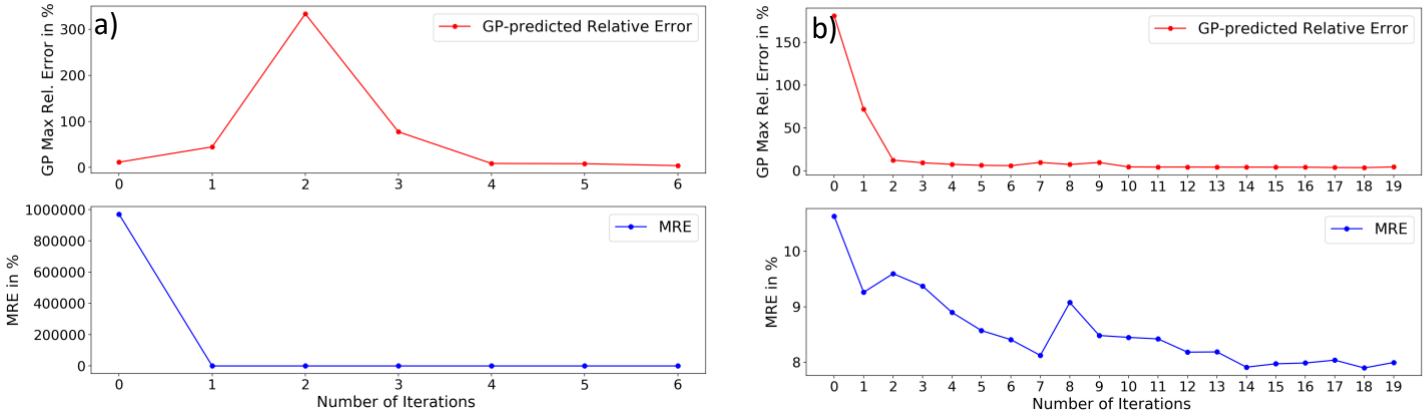

Figure S5. Maximum GP relative error and MRE (in %) with respect to number of AL iterations for methane adsorption in two features. The pressure range here is 10^{-6} to 300 bar, and temperature range of 100 K to 300 K. The plots are for a) linear-spaced LHS, b) log-spaced LHS. We find a very high maximum GP relative error for linear-spaced prior and a jump to higher GP relative error and a coming back to below 2 %. This same error for log-spaced prior goes much smoothly to below 2 % but take a greater number of iterations.

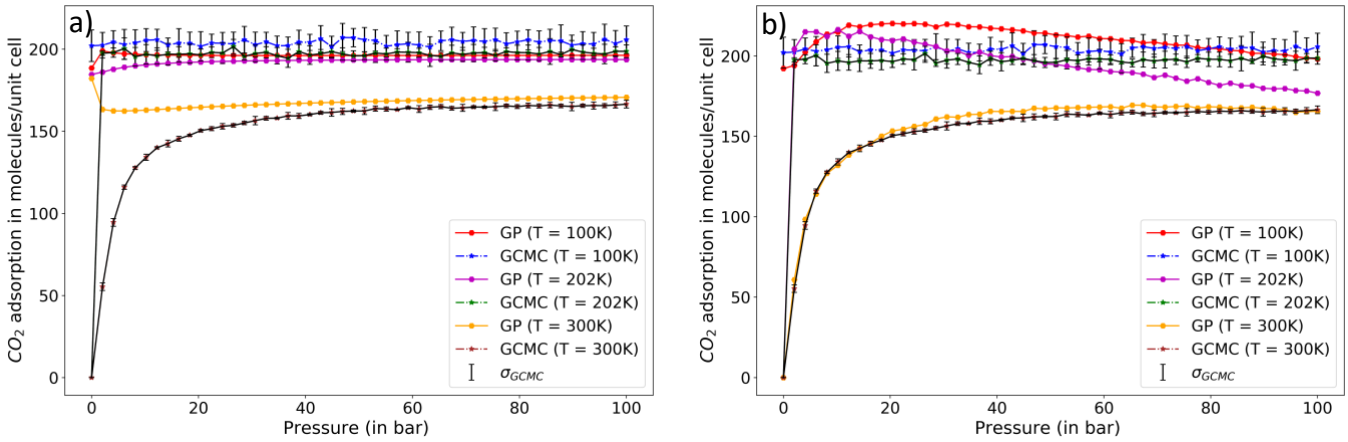

Figure S6. Comparison of GP-predicted CO_2 uptake with GCMC simulation predicted for pressure range of 10^{-6} to 100 bar, at temperature of 100 K, 202 K and 300 K for a) linear-spaced LHS prior, and b) log-spaced LHS prior. We find a high disagreement with GCMC for linear-spaced prior for all the temperature at low-pressure region. For log-spaced there is a high error at low temperature and this error is more pronounced at the high-pressure region.

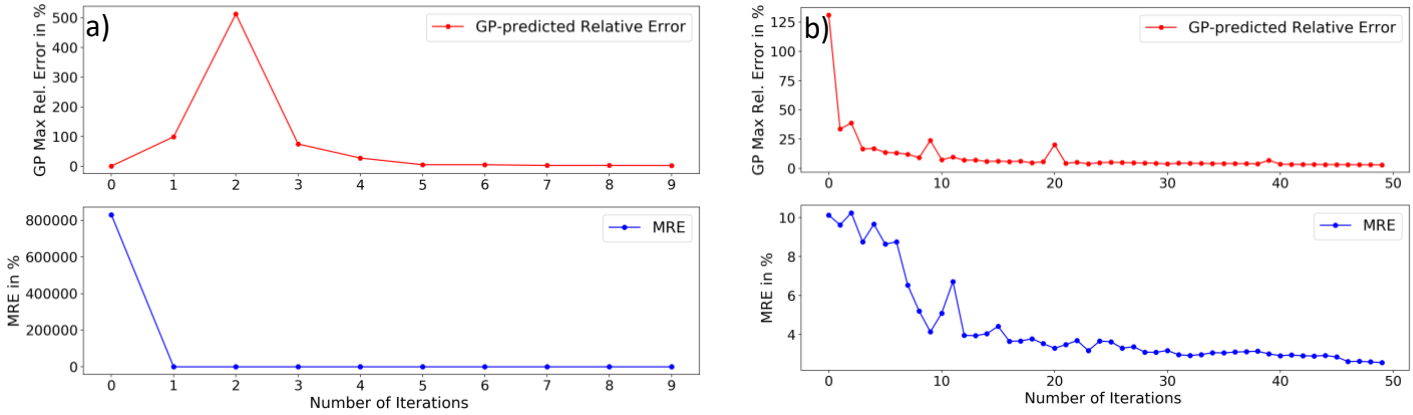

Figure S7. Maximum GP relative error and MRE (in %) with respect to number of AL iterations for 2 features for CO_2 adsorption. The pressure range here is 10^{-6} to 300 bar, and temperature range is 100 K to 300 K. The plots are for a) linear-spaced LHS, b) log-spaced LHS. Please note the convergence limit for CO_2 adsorption for two features was set to 3 % while for methane it was 2 %. Here also we observe a spike in GP maximum relative error for linear-spaced prior and then decreasing to go below 3 % limit. For log-spaced prior we a smooth decline except few points of rise. The log-spaced prior takes considerably a greater number of iterations to converge but has a comparable final MRE of 2.64 % while linear one has 2.79 %.

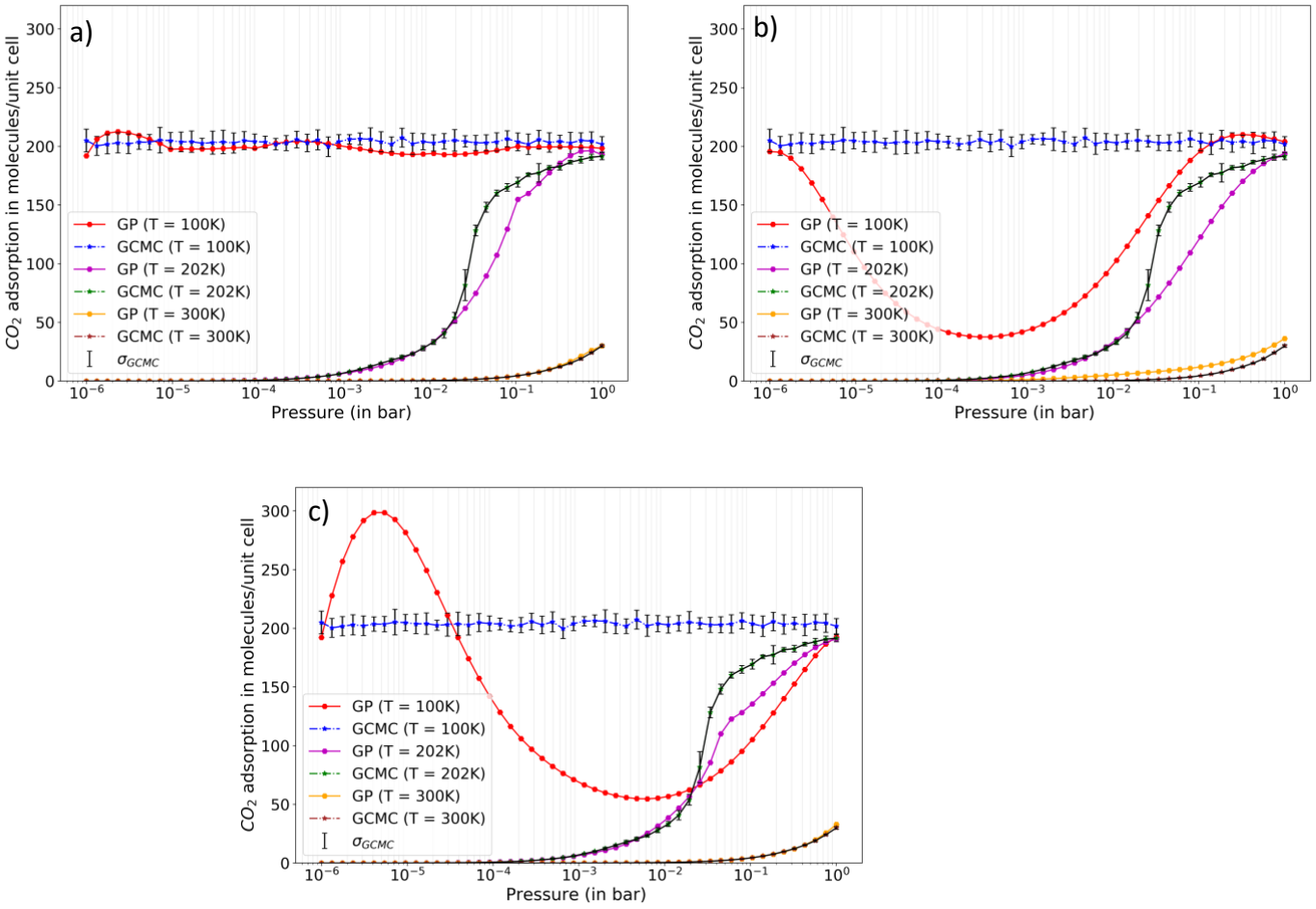

Figure S8. CO_2 uptake comparison between GP and GCMC simulation in Cu-BTC at low pressure range of 10^{-6} to 1 bar for different priors, a) boundary-informed prior, b) linear-spaced LHS prior, and c) log-spaced LHS prior. We find here that boundary performs quite well for at temperature except at the isotherm rise region. Both log-spaced and linear-spaced prior have high error at lowest temperature. Also, at a very low pressure (10^{-6} to 10^{-3} bar) the errors are very high. Since adsorption is basically nil in this zone, the GP has a very uncertainty in prediction. Log-spaced prior performs better than linear-spaced prior at higher temperature. Linear spaced has a poor performance for all the three temperature points.

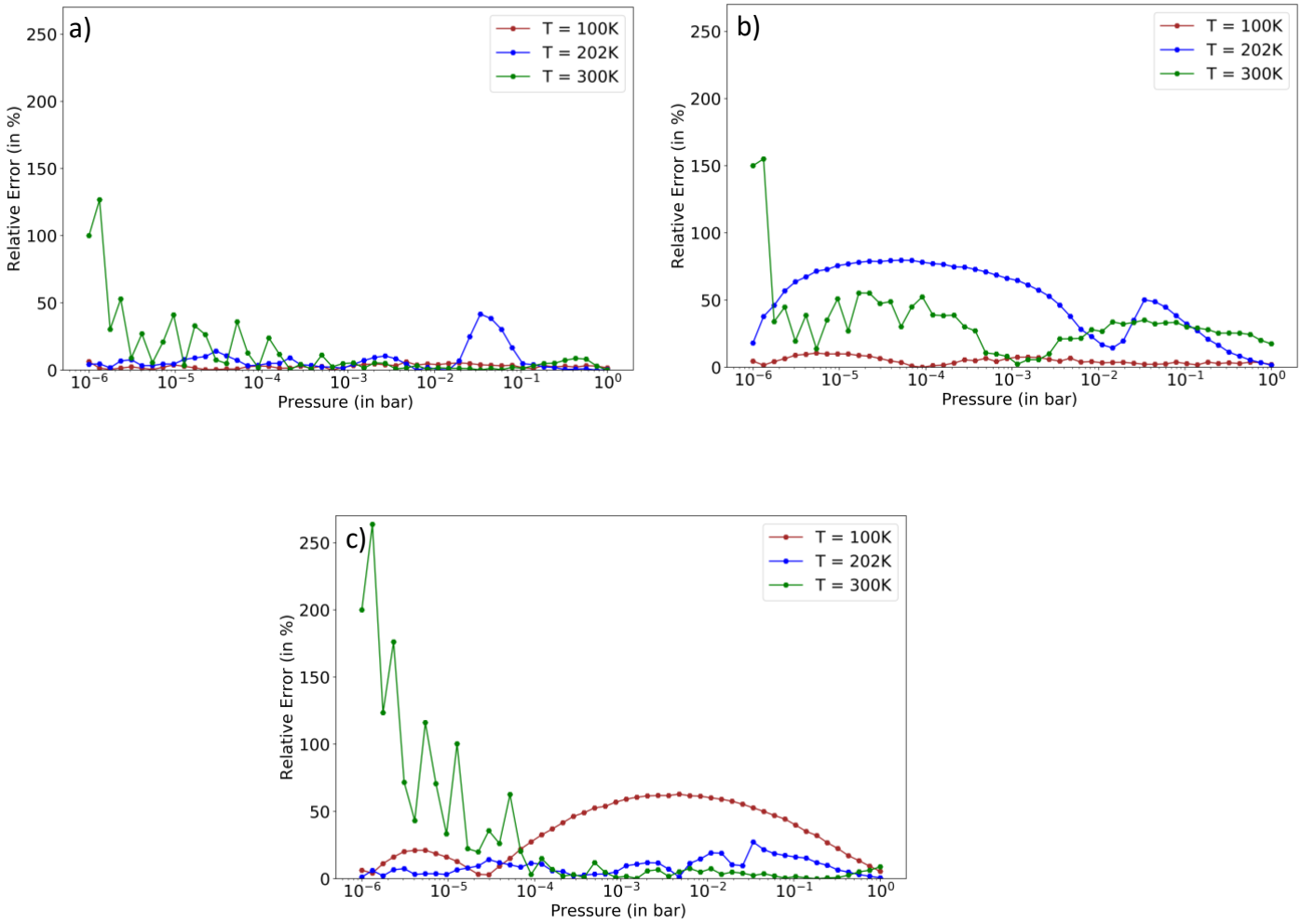

Figure S9. Relative error (in %) comparison between GP and GCMC simulation in Cu-BTC at low pressure range of 10^{-6} to 1 bar for different priors for CO_2 adsorption, a) boundary-informed prior, b) linear-spaced LHS prior, and c) log-spaced LHS prior. We find boundary-informed prior had a lower relative error throughout all the temperature point. Boundary-informed prior had a few points close to and over 50 % relative error at the highest temperature of 300 K. The error decreases for lower temperature. This error might be because adsorption at this pressure range and at high temperature is almost zero and there is a high fluctuation even at the GCMC simulation (refer table 5). Linear-spaced prior had considerably higher relative error, but it does decreases with temperature. Log-spaced prior had a very high relative error at 300 K in the further low-pressure change but the relative error decreases as we decrease the temperature.

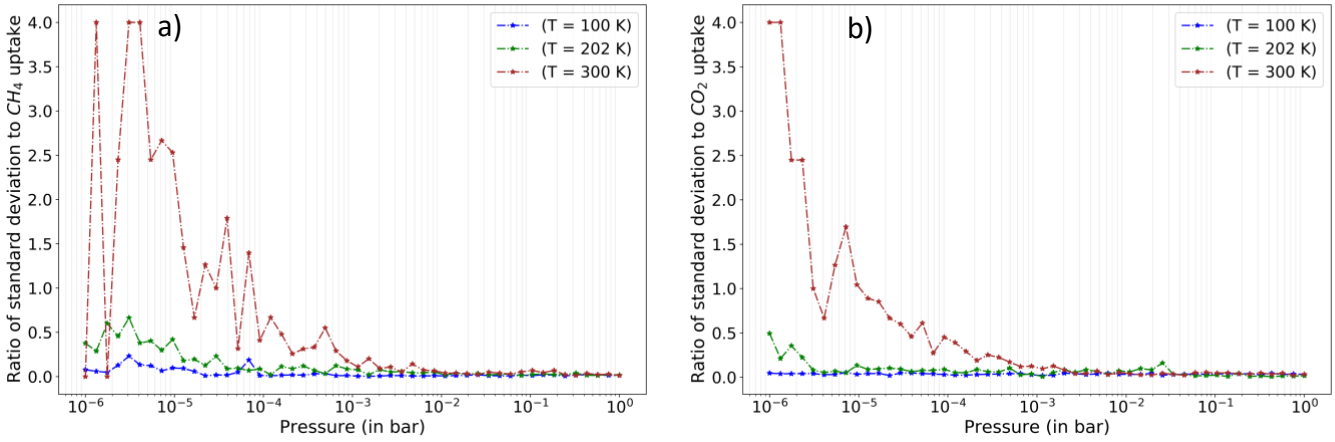

Figure S10. Ratio of standard deviation in GCMC simulation to gas uptake at low pressure range (10^{-6} to 1 bar) in Cu-BTC at different temperatures, a) Methane adsorption, and b) CO_2 adsorption. Adsorption is extremely low (almost nil) for both these gases at high temperature of 300 K, and hence this ratio is very high, especially for the 10^{-6} to 10^{-4} bar range. Since the absolute adsorption is near zero, it is extremely difficult to predict the uptake accurately at this region. Even for temperature of 202 K we see this ratio remains well above 0.5 for a considerable range of pressure. These plots illustrate the high degree of uncertainty in the GCMC simulations and hence the ground truth has a high unreliability in this space which also is manifested in the final GP fit.