Figure S1: The comparison between FLake, PBBM and observations in LST, LE, and HE. **(a-c)** present the comparison of model performance between FLake, PBBM, and observations. **(d-f)** present the scatterplots between FLake and PBBM.
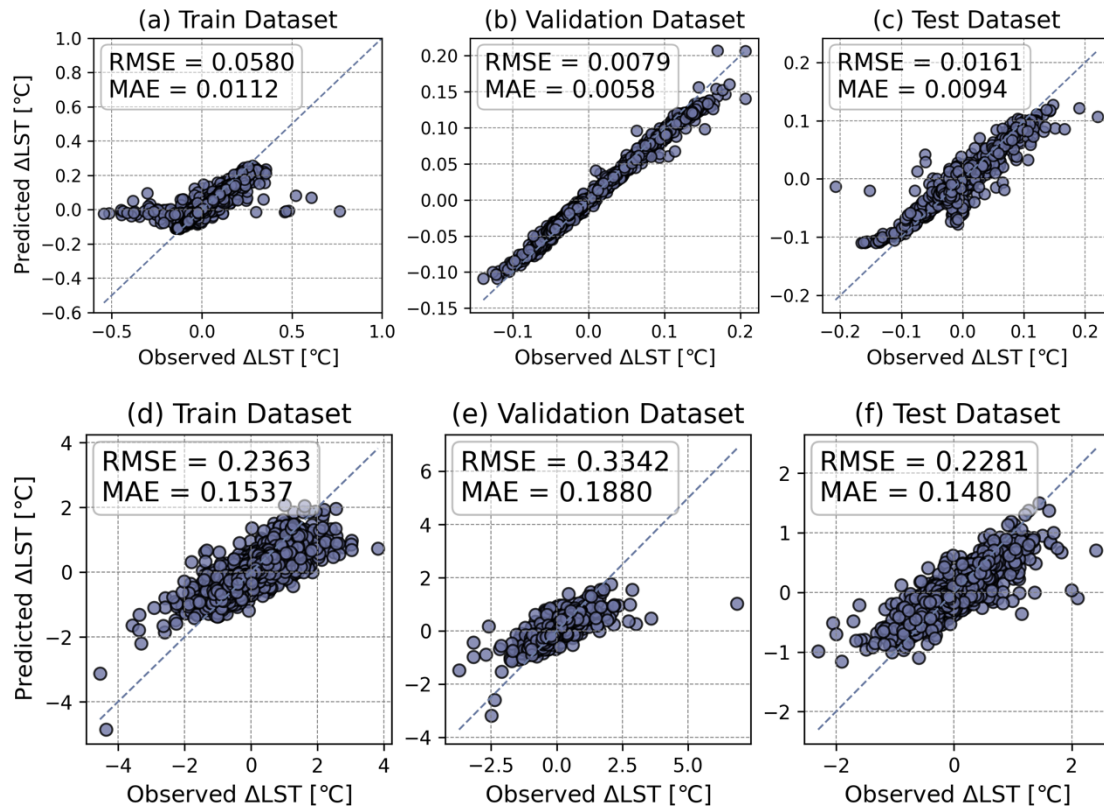
Figure S2: The separate performance of LSTM-based surrogates for Baseline **(a-c)** and TaihuScene **(e-f)** in train, validation, and test datasets, respectively.
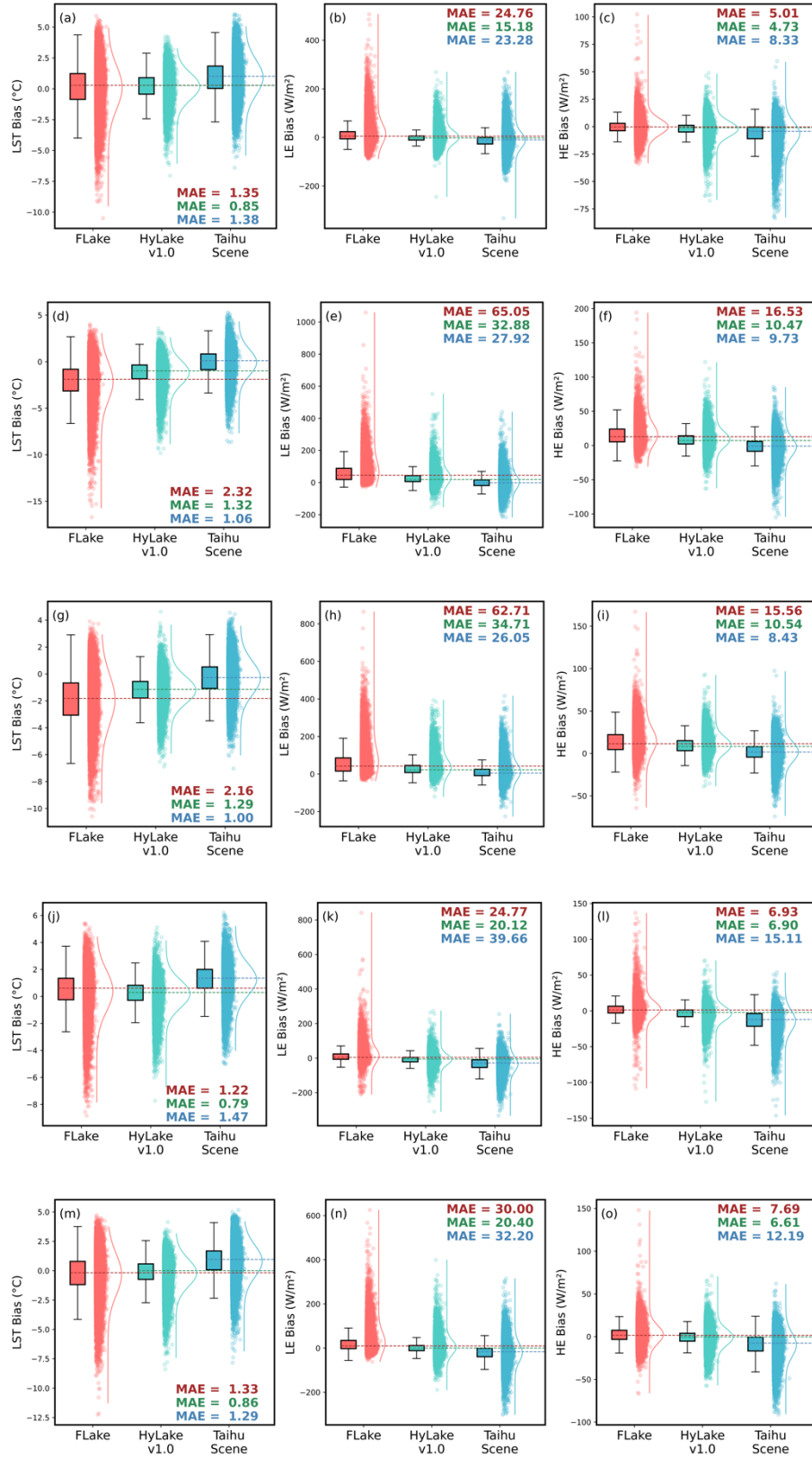
Figure S3: The intercomparison of LST, HE, and LE for each lake site between FLake, HyLake v1.0 and TaihuScene. MLW, BFG, DPK, PTS, and XLS site are **(a-c)**, **(d-f)**, **(g-i)**, **(j-l)**, and **(m-n)** in this Figure, respectively.
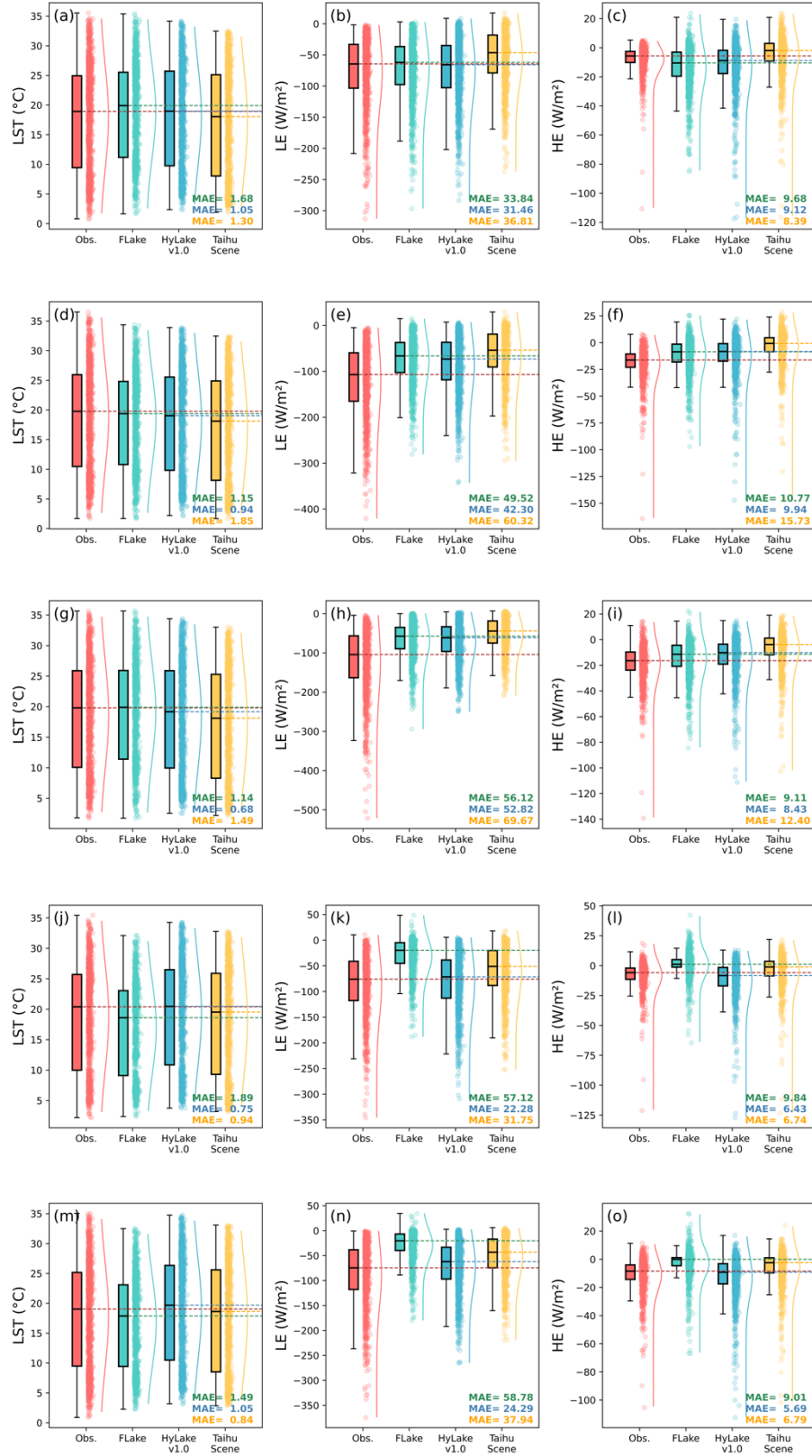
Figure S4: The comparison of LST, HE, and LE for each lake site between observations, FLake, HyLake v1.0 and TaihuScene under ERA5 reanalysis datasets. MLW, BFG, DPK, PTS, and XLS site are **(a-c)**, **(d-f)**, **(g-i)**, **(j-l)**, and **(m-n)** in this Figure, respectively.
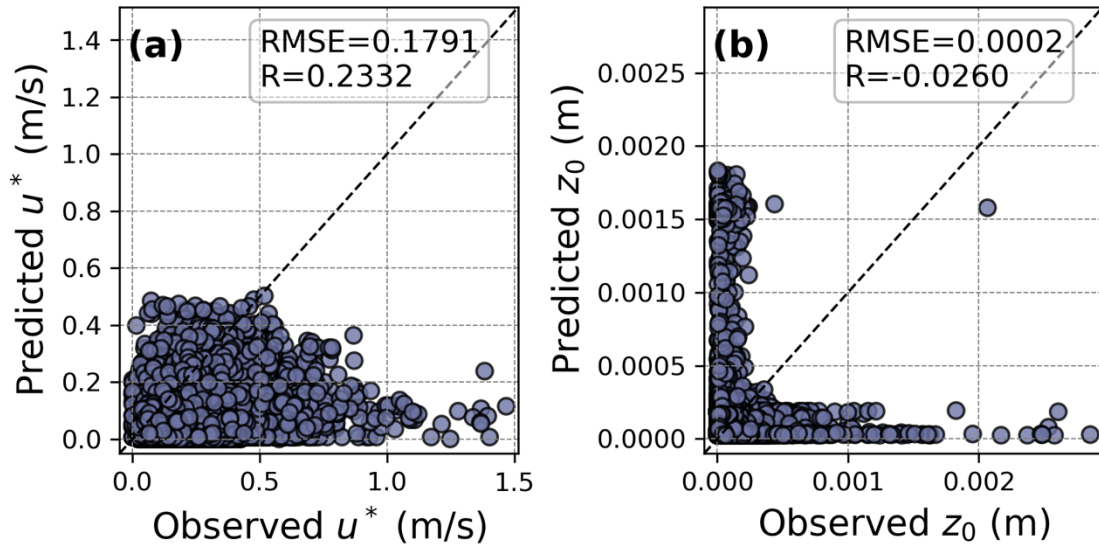
Figure S5: The comparison of friction velocity (u*) and surface roughness length (z_{0m}, m) in MLW lake site between simulation derived from PBBM and HyLake v1.0 and observations.