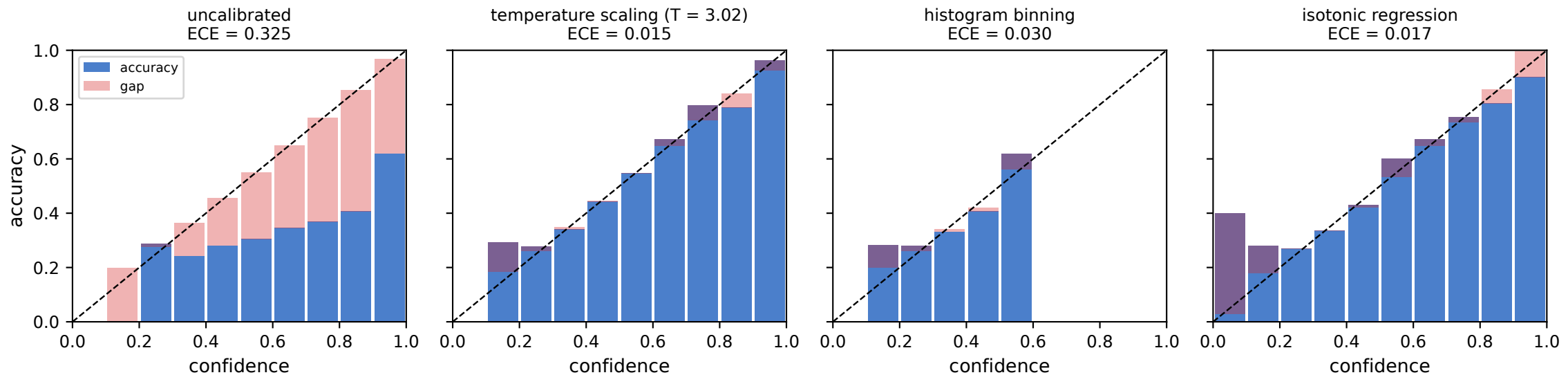

uncalibrated
ECE = 0.325
temperature scaling (T = 3.02)
ECE = 0.015
histogram binning
ECE = 0.030
isotonic regression
ECE = 0.017
accuracy
gap
accuracy
confidence
0.0
0.2
0.4
0.6
0.8
1.0