

average return

140

120

100

80

60

40

20

0

0.5

1.0

1.5

2.0

2.5

3.0

timesteps

1e6

teac  
sac  
ppo  
pcl  
gac

