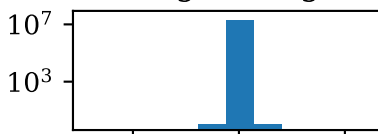
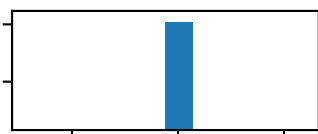


number of model predictions

strengthening24



strengthening24  
baseline



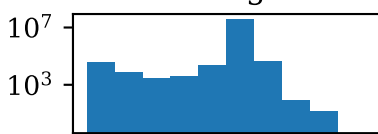
conditional  
strengthening



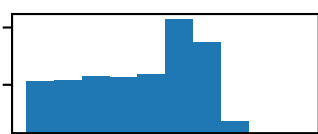
conditional  
strengthening  
baseline



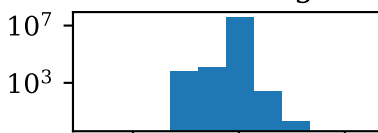
proportional  
change



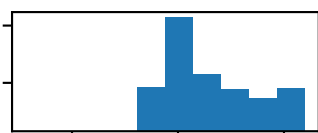
proportional  
change  
baseline



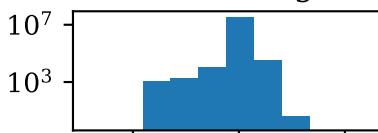
conditional  
weakening



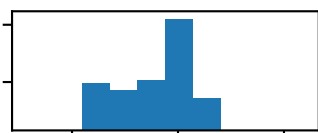
conditional  
weakening  
baseline



weakening



weakening  
baseline



$\log_{\text{scale}}(\text{clean}) / \log_{\text{scale}}(\text{ablated})$