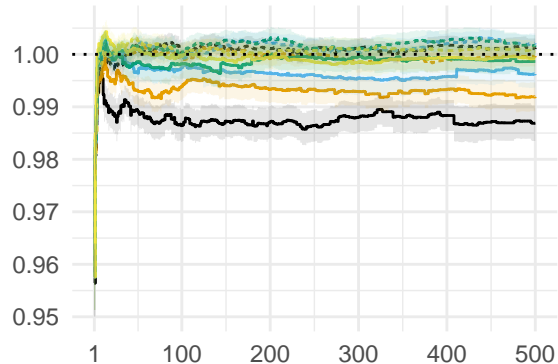
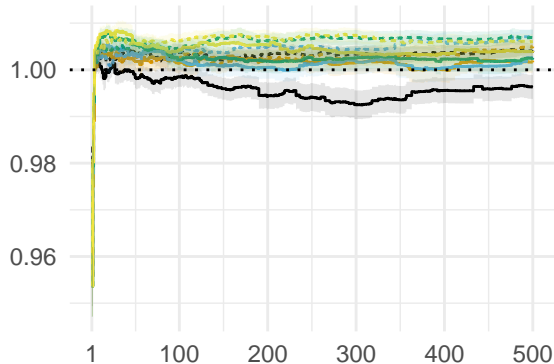


Mean Relative Test Performance

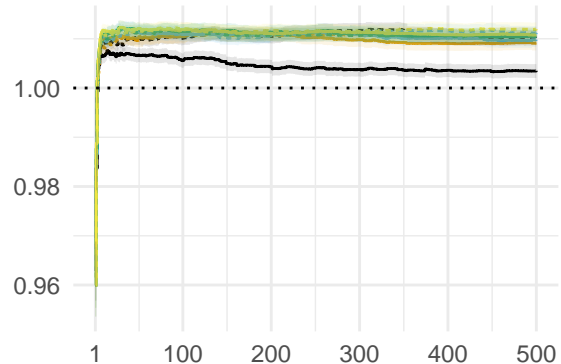
500



1000



5000



No. HPC Evaluations

Holdout



1-fold



2-fold



3-fold



4-fold



5-fold

Reshuffling



FALSE



TRUE