IDG-DREAM Round 1 Bootstrap Analysis

Robert Allaway

12/03/2018

This script takes the Round 1 IDG-DREAM results and performs the following steps:

-Bootstrap evaluation of the predictions - paired resampling with replacement - to assess stability of results for all six metrics.
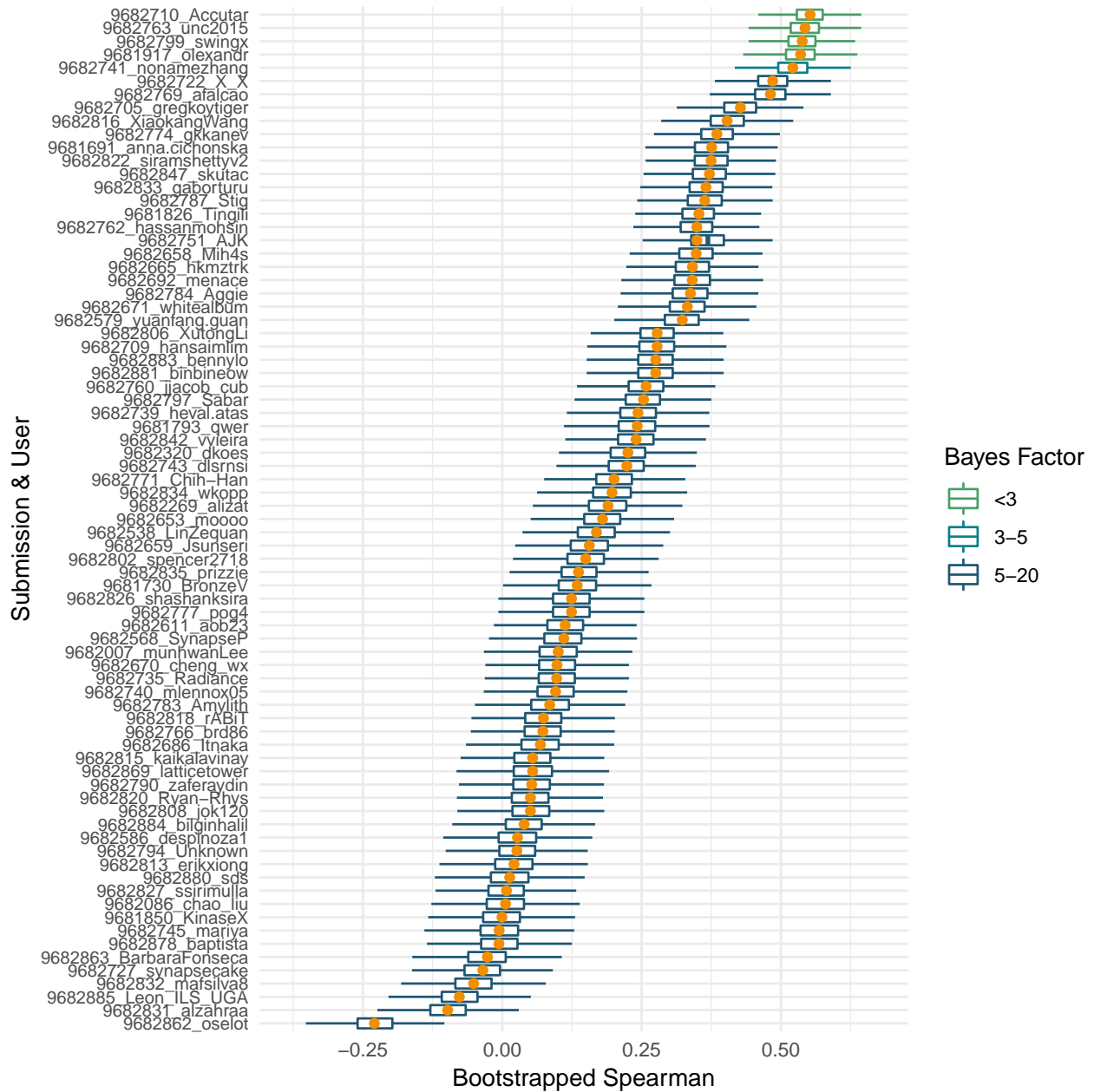
The bootstrapping takes a while, so if rerunning this Rmd please be aware of that.
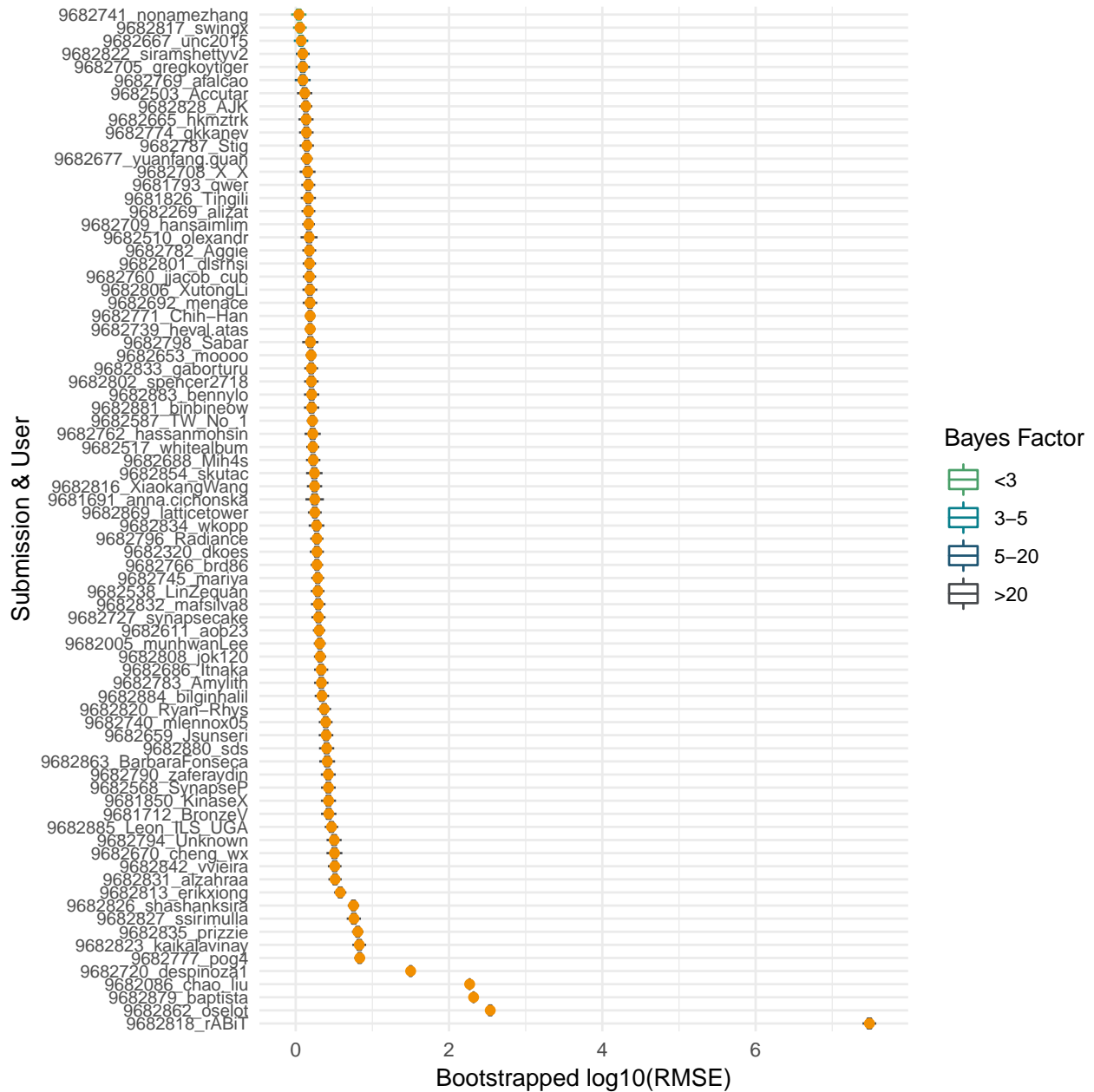
Bootstrapping

To bootstrap a given prediction file this script: randomly samples 430 times from prediction file. Compute 3 metrics for those random predictions: Spearman correlation (metric 1) RMSE (metric 2) Average AUC - tiebreaker for metrics 1,2

Repeat 10000x per prediction file to generate a distribution of bootstrapped scores per prediction file. I then plotted the top 20 predictions for each metric using the leaderboard value, superimposed on the distribution of the bootstrapped prediction. Bars are ranked best to worst performer (based on single leaderboard value). Orange circles are the actual leaderboard value.
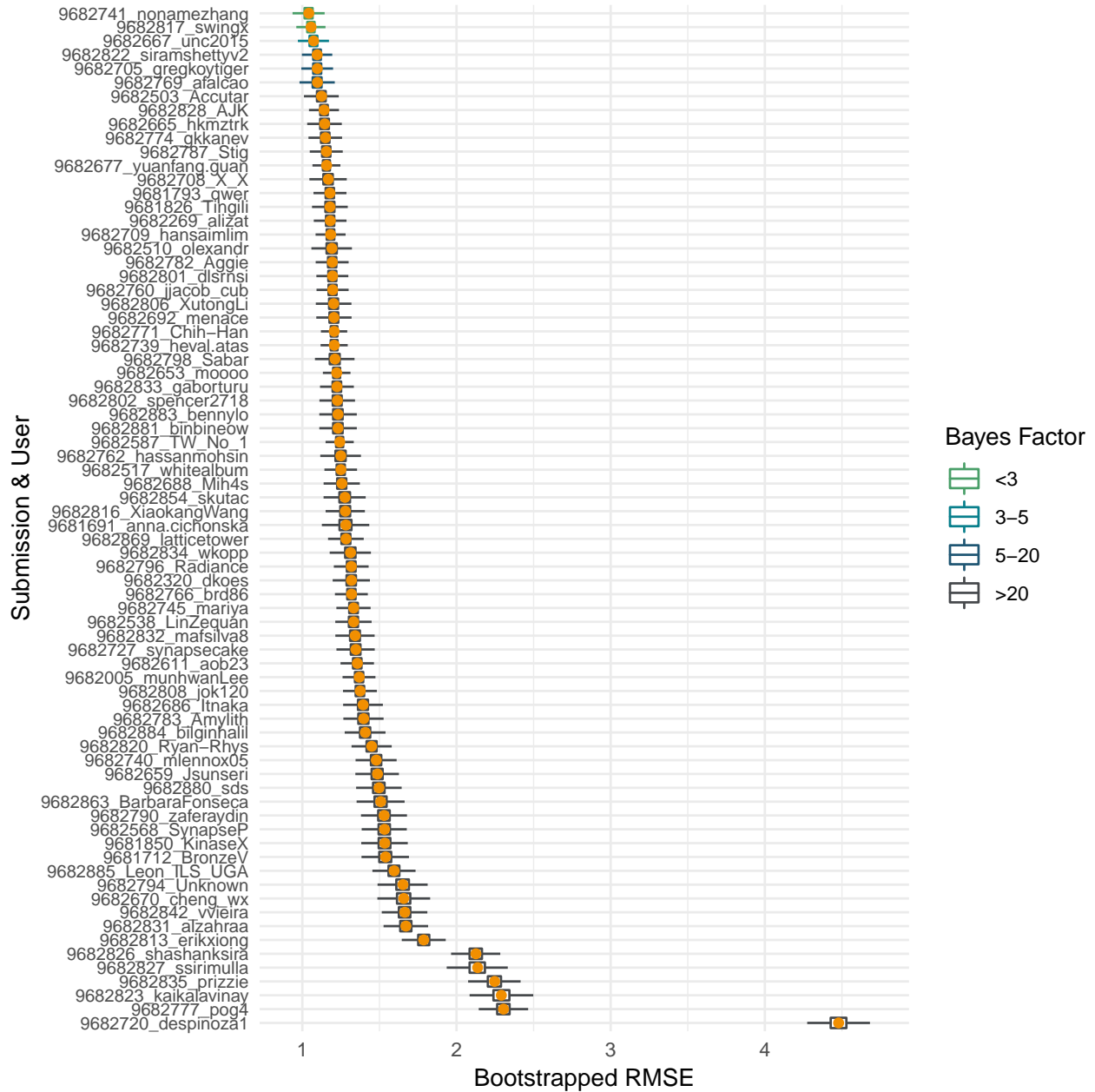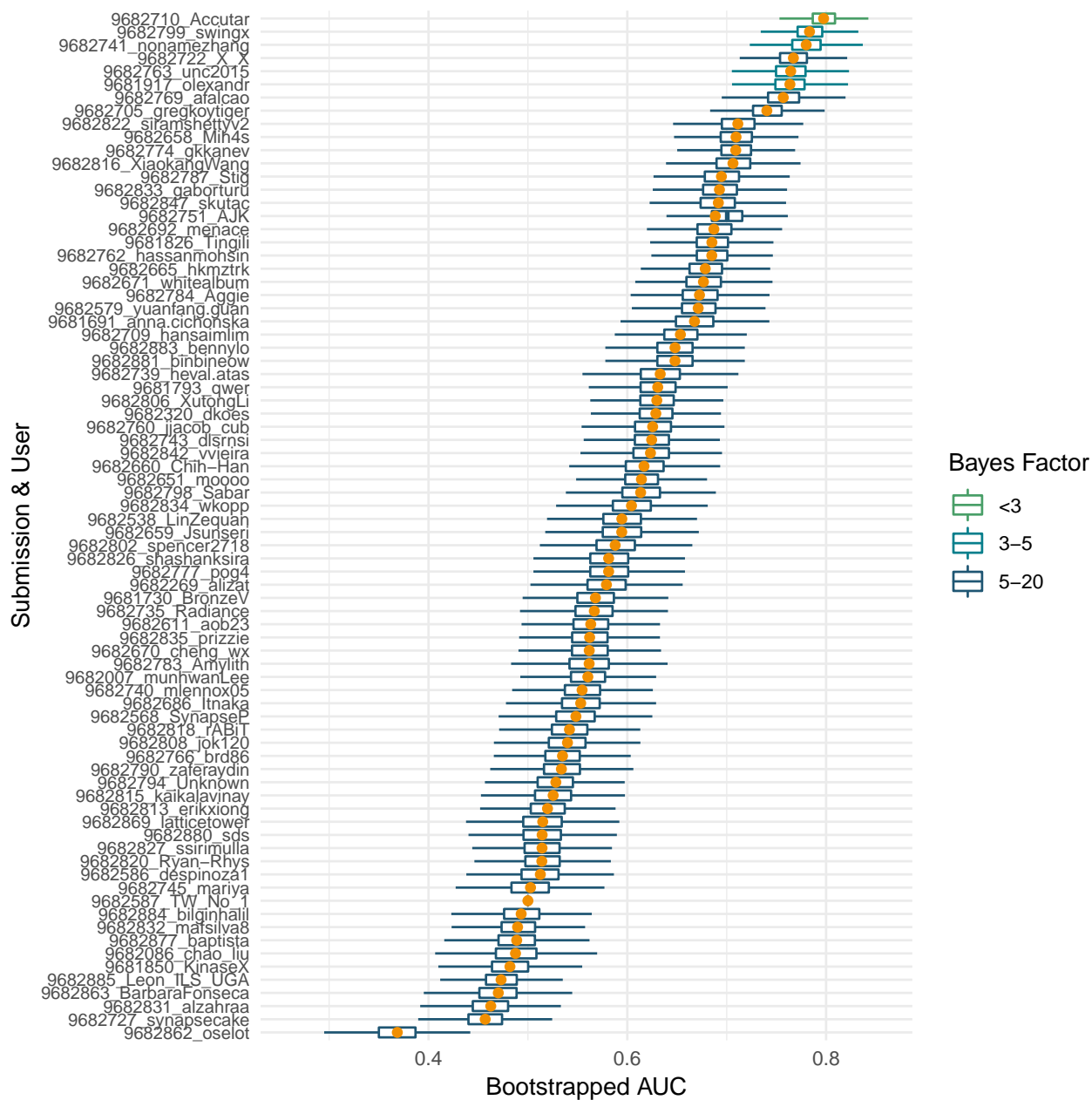
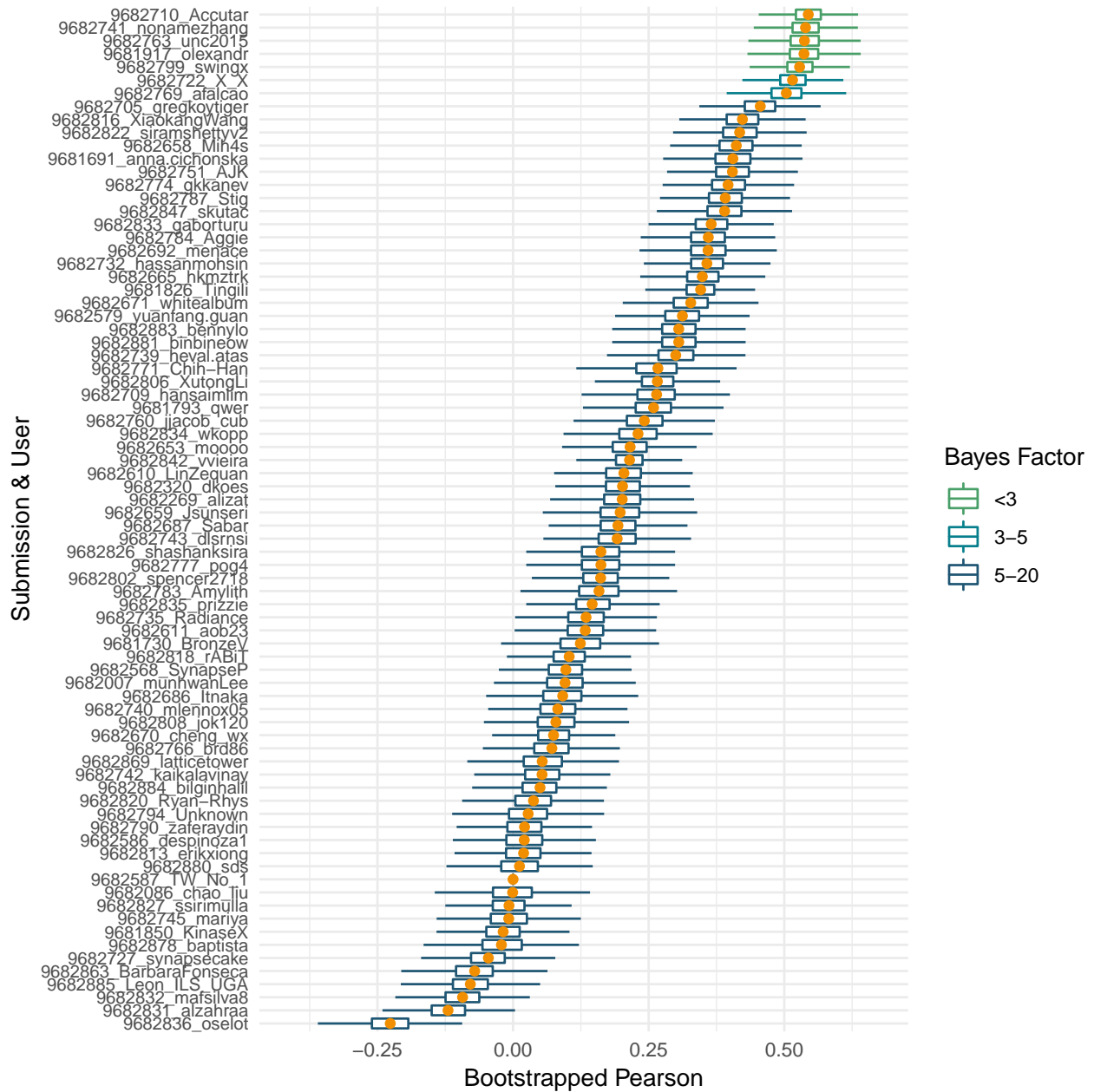# Spearman bootstrap results

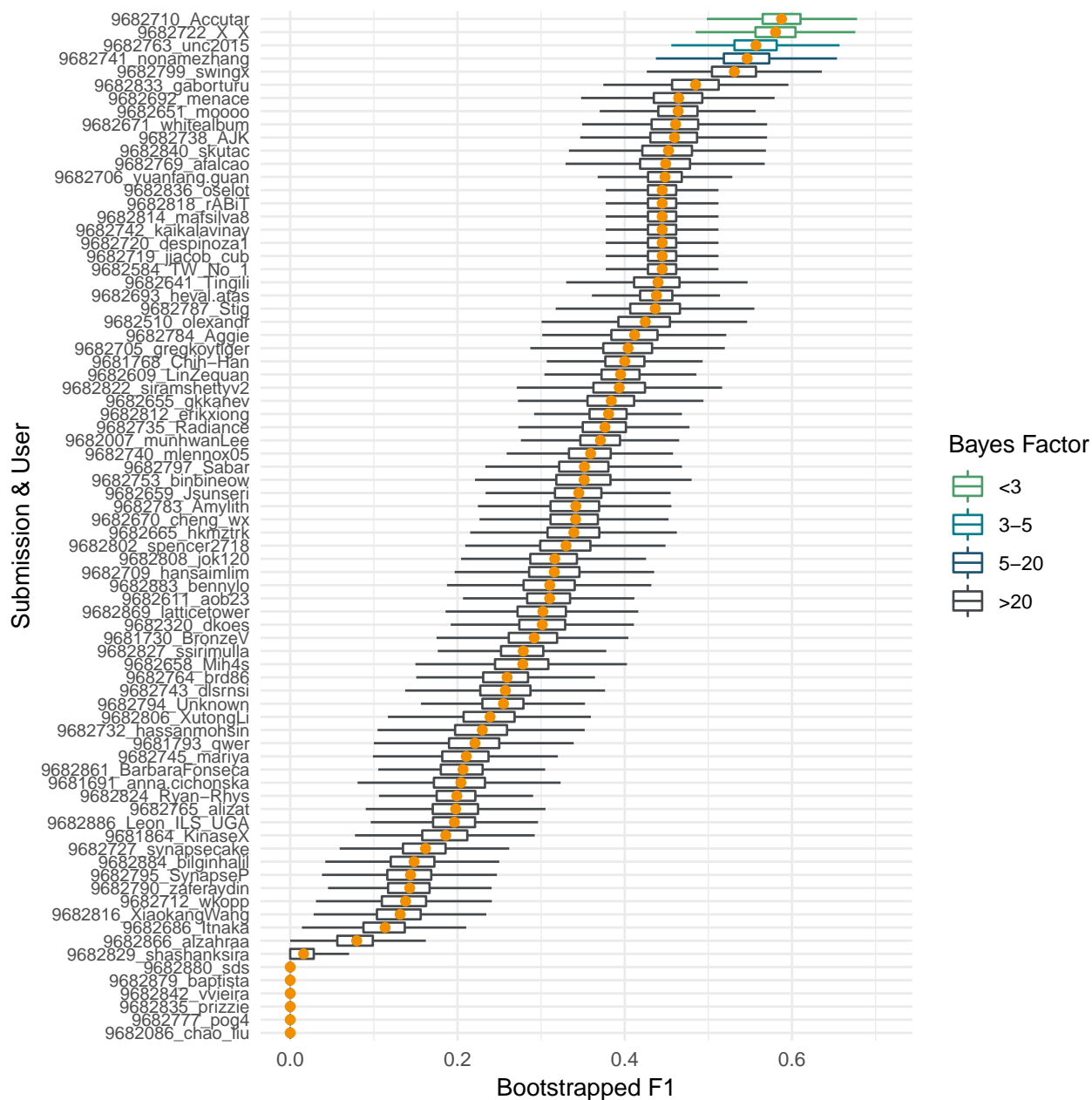**RMSE bootstrap results**

RMSE bootstrap results

Average AUC bootstrap results (standalone)

Pearson bootstrap results

f1 bootstrap results

ci bootstrap results