Analyzing Tony Gitter's chemical data set

Available on https://zenodo.org/record/1411506#.X7w4gs1KhPY

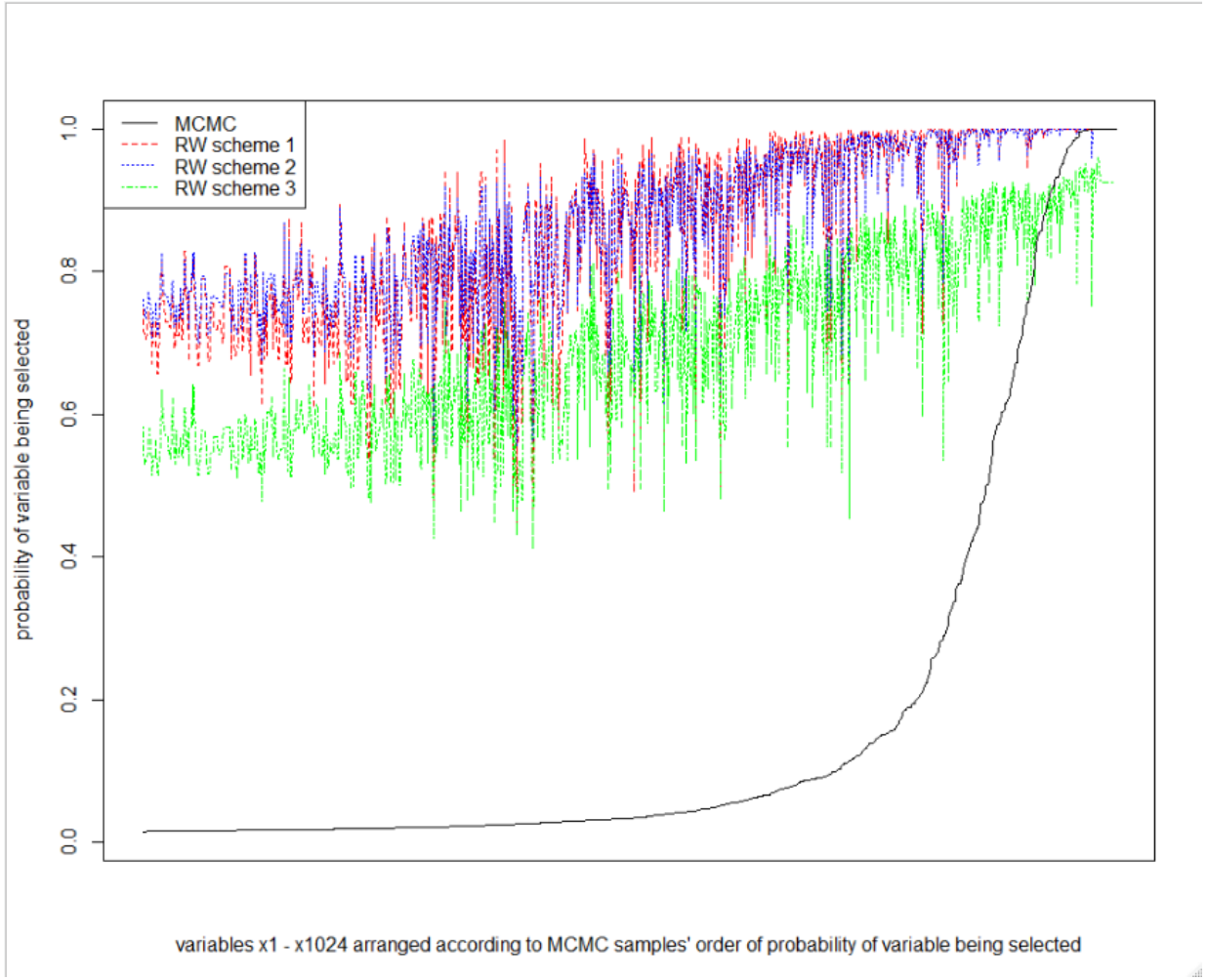
Training

- *basad*
  - Use default package setting
  - We use MCMC β samples exactly as they are for predictions, even though MCMC does provide probability of variables selected.

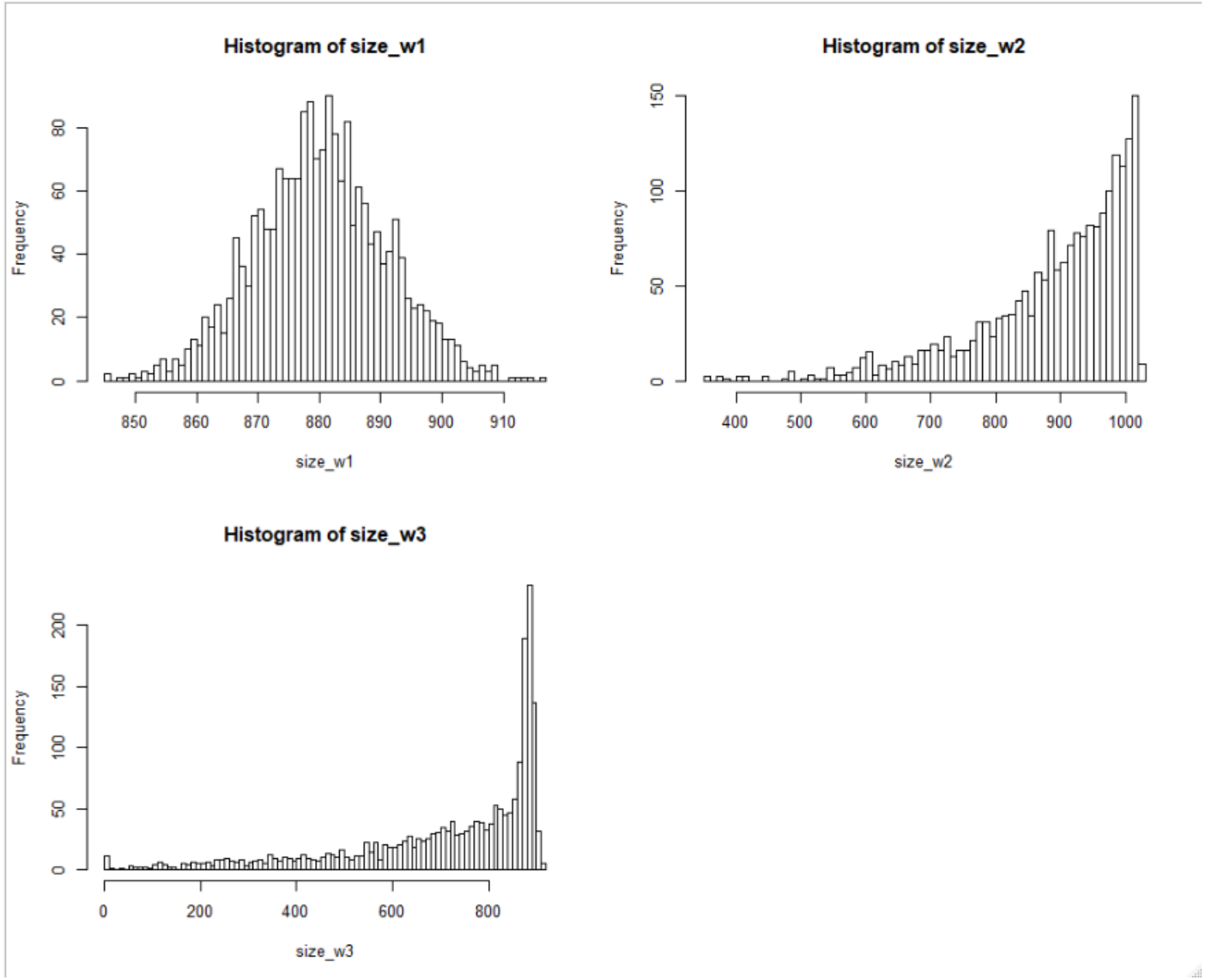
- *glmnet*
  - *cv.glmnet* gives lambda.min = 0.001 and lambda.1se = 0.006
  - Try lambda.min

- Less sparsity constraints on β's for both MCMC and RW.

Compare probability of variables selected

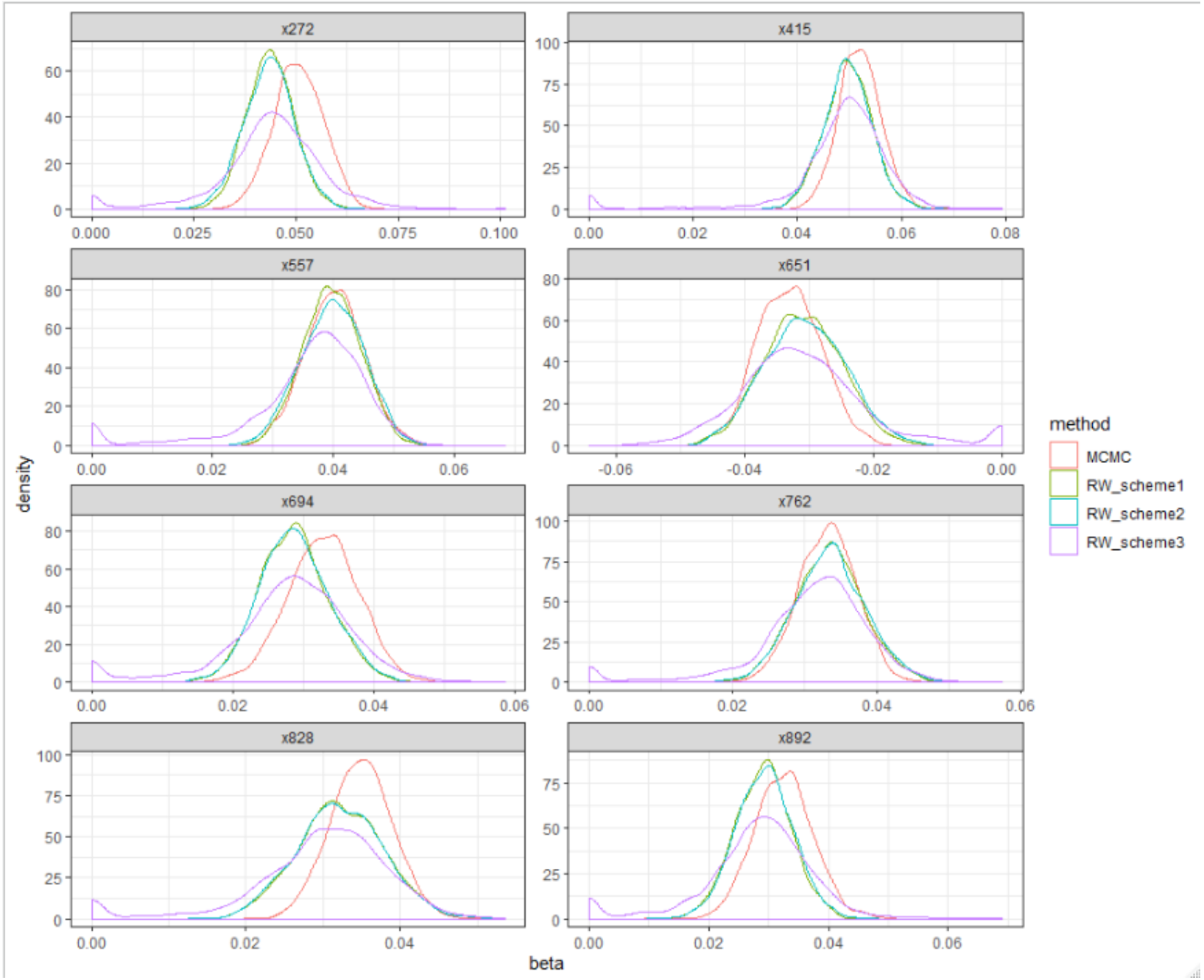


Compare model sizes

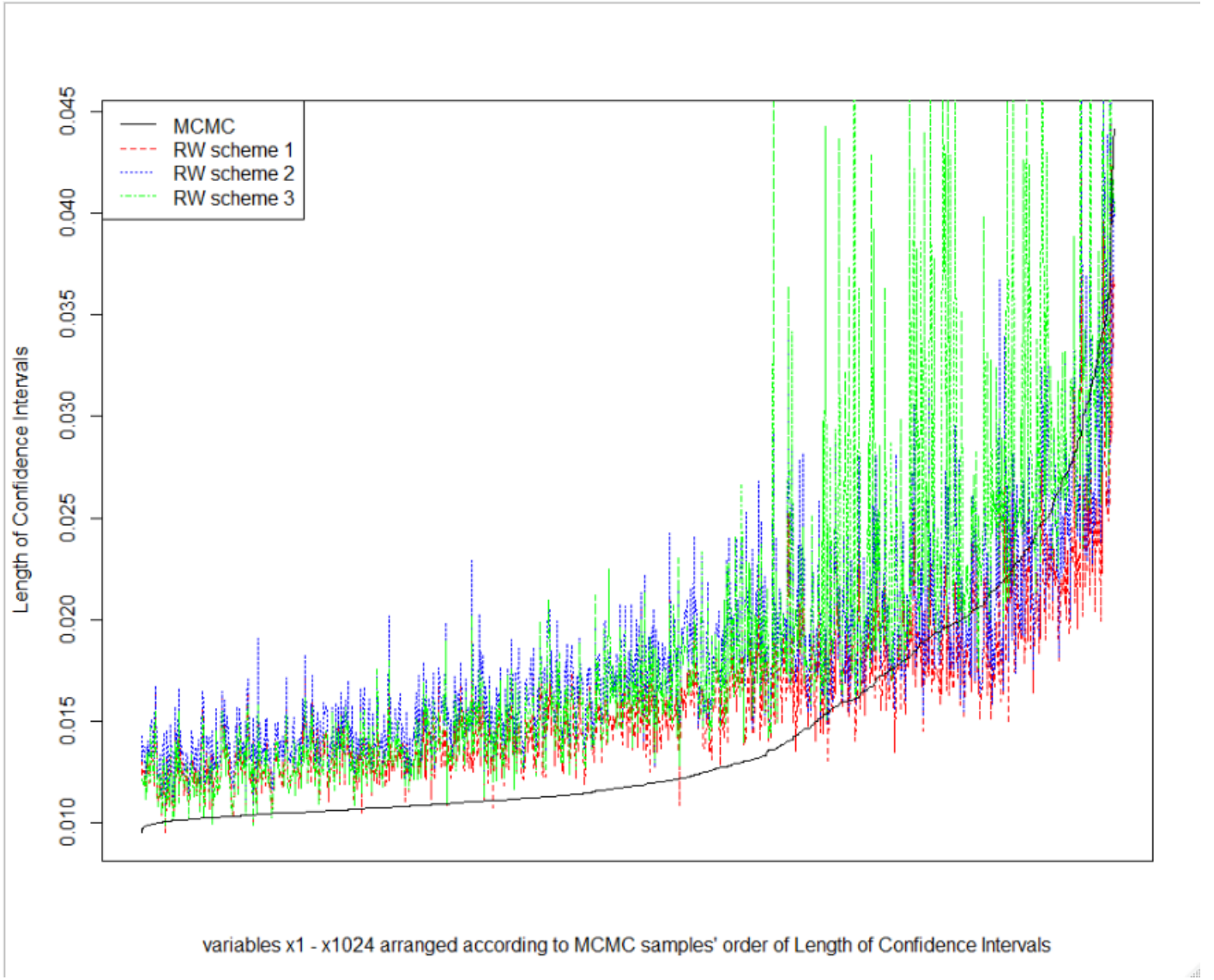


Compare marginal posterior distributions of β's

- Pick variables whose probability of being selected is larger than 0.955 for all RW schemes 1, 2 and 3.
- There are 8 of them.
- Plot marginal distributions of their sampled β's.


x272
x415
x557
x651
x694
x762
x828
x892
density
beta
method
MCMC
RW_scheme1
RW_scheme2
RW_scheme3

Compare length of 95% CI for β's



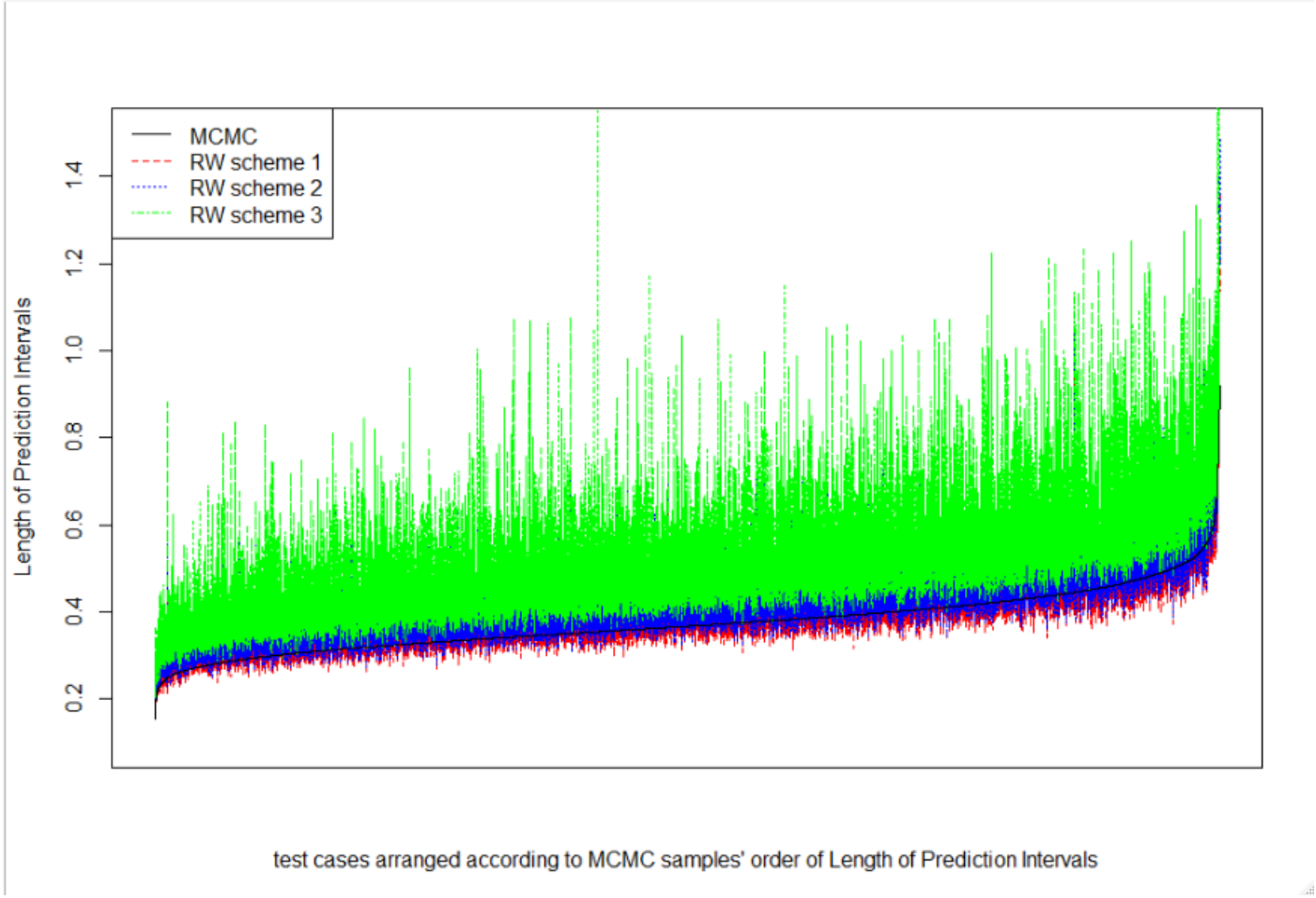
Predictions

- n_test = 22,423
- For predictions, take sampled β's from MCMC or RW, and multiply with test cases' predictors.
- Each test case has a distribution of predictions.

95% prediction interval coverage for the ~22,000 test cases is similarly poor for all schemes

- RW scheme 1 = 17.7% (higher than lamb.1se case)
- RW scheme 2 = 19.1% (close to lamb.1se case)
- RW scheme 3 = 23.0% (close to lamb.1se case)
- MCMC = 17.1% (lower than using lamb.1se case)
- Because length of prediction intervals for RW schemes 2 and 3 are relatively wider

Compare 95% prediction interval length



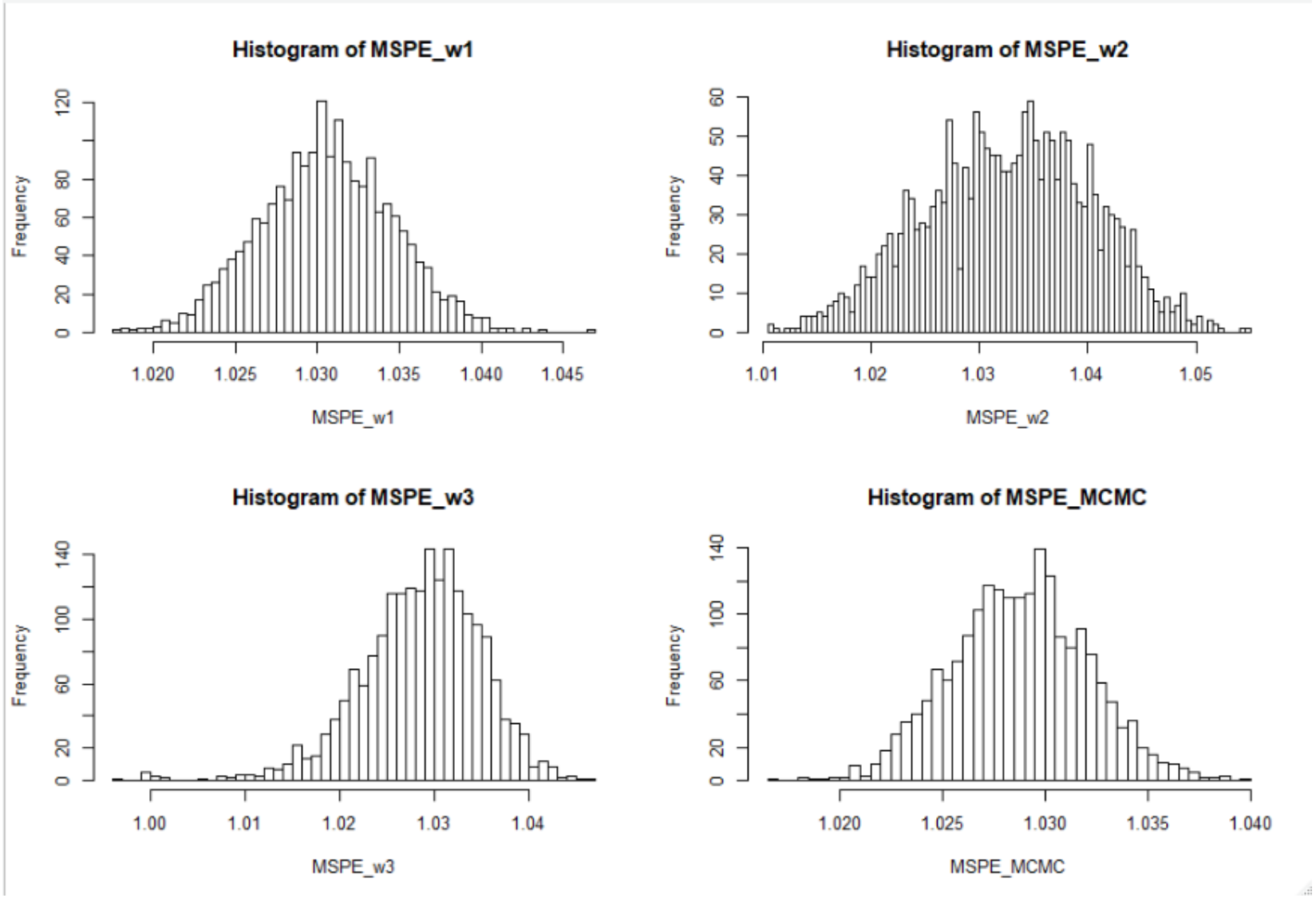
Prediction

- I considered 2 types of MSPE calculations

    1. For each of B=2,000 iterations, compute MSPE (squared prediction errors averaged across all 22,423 test cases)

    2. For each of n_test = 22,423 test cases, compute MSPE (squared prediction errors averaged across all 2,000 predictions)

Distribution of 2,000 MSPEs (averaged across all test cases)



Compare distribution of MSPE for each test case