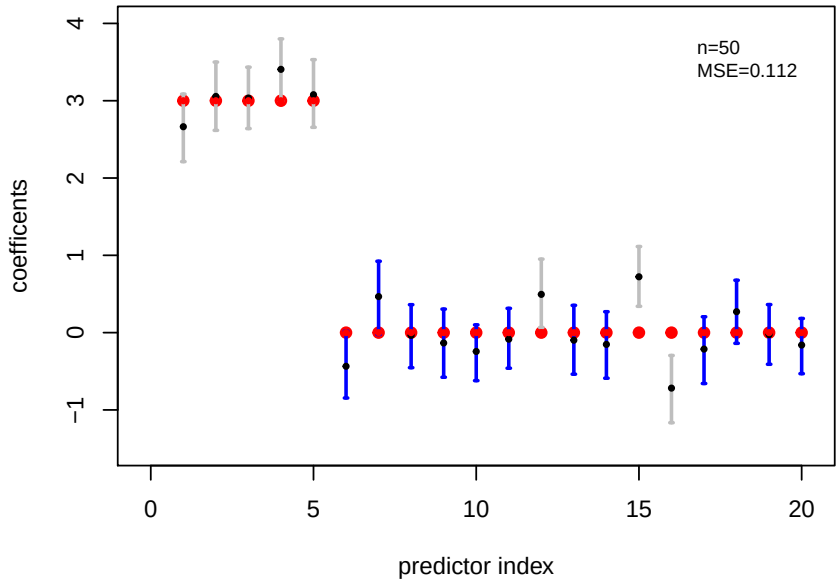
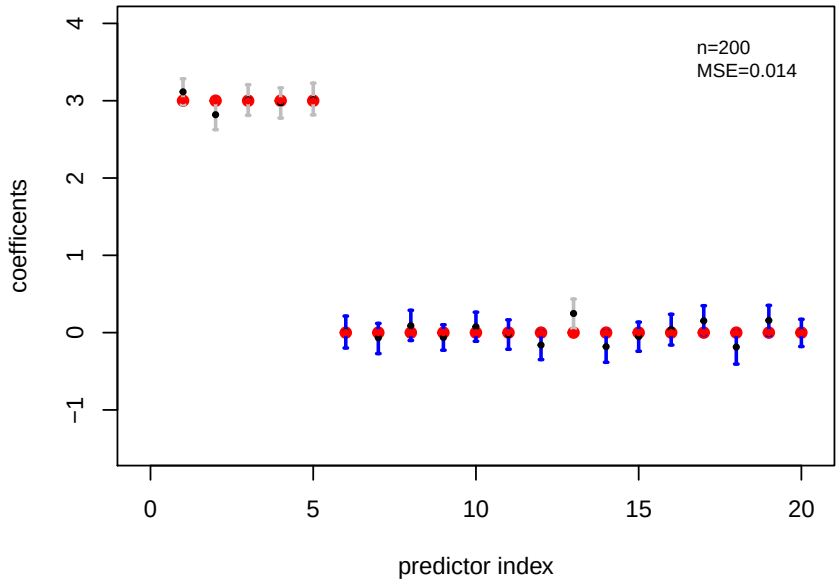
AMS 268 - HW #2

Raquel Barata

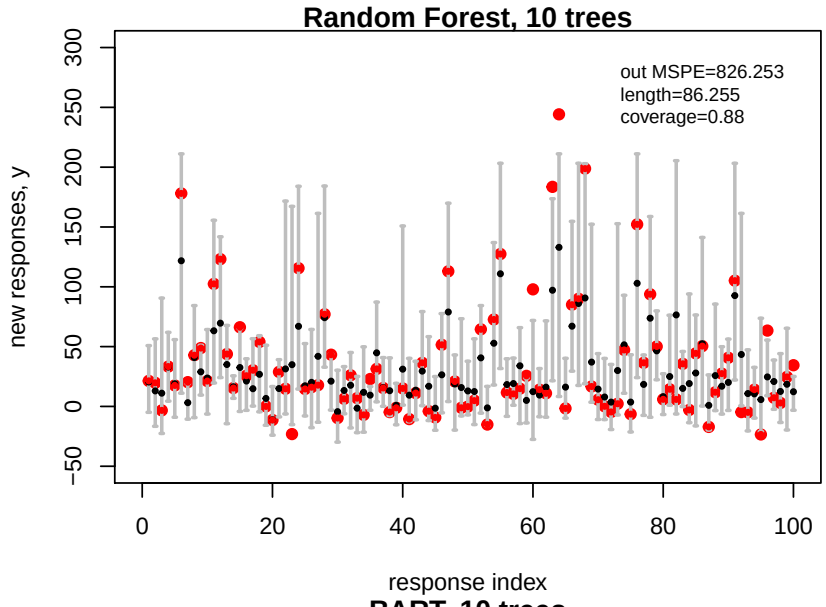
g-prior

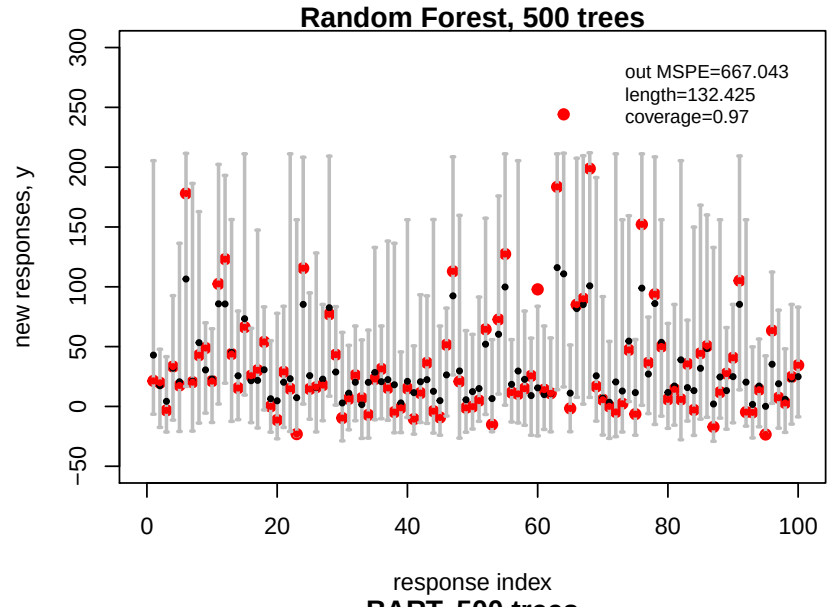


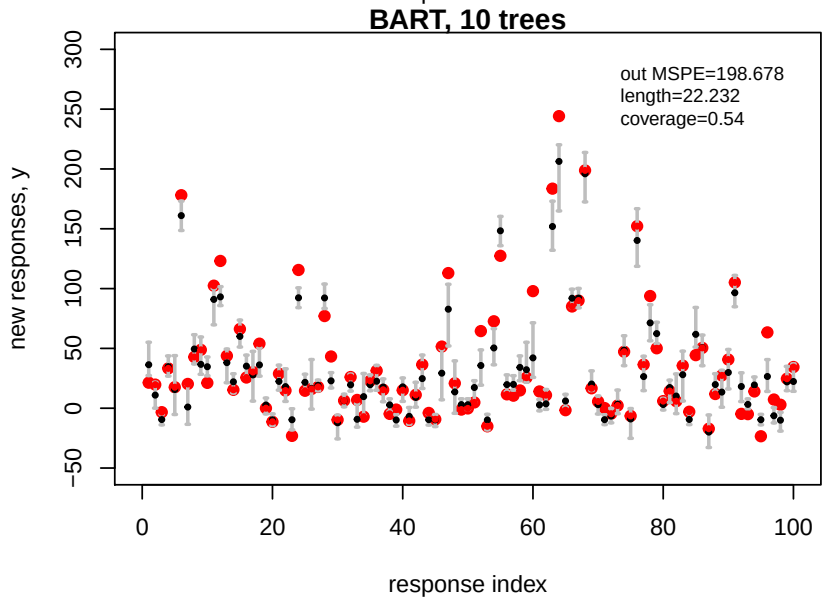


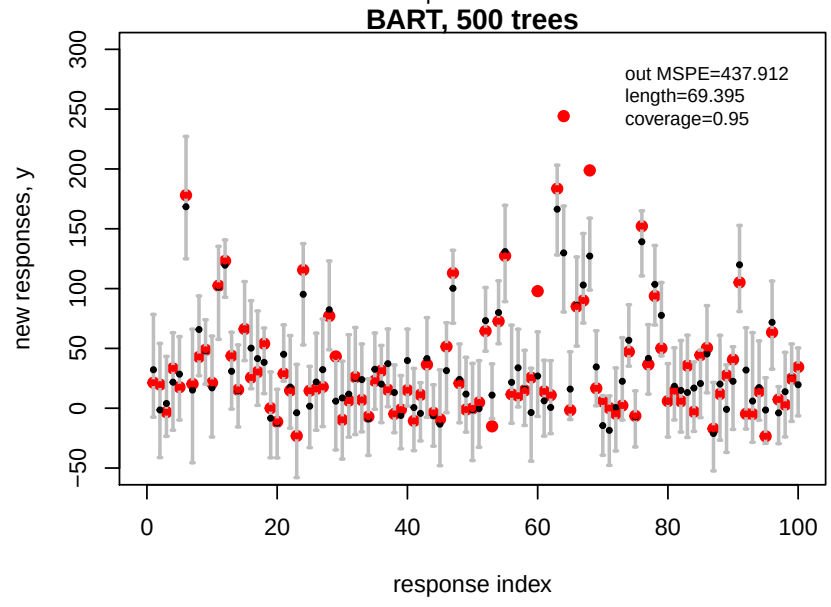
- g fixed as $\max\{n, p^2\}$ respectively for each model.
- True β_j values in red. 95% posterior intervals in grey if interval does not include zero, blue if interval does include zero.
- Test $H_0 : \beta_1 = 0$ -> Reject H_0 since 95% PI does not include 0.
- Test $H_0 : \beta_{10} = 0$ -> Fail to reject H_0 since 95% PI does include 0.

Random Forest vs. BART: $n = 200$, $p = 200$

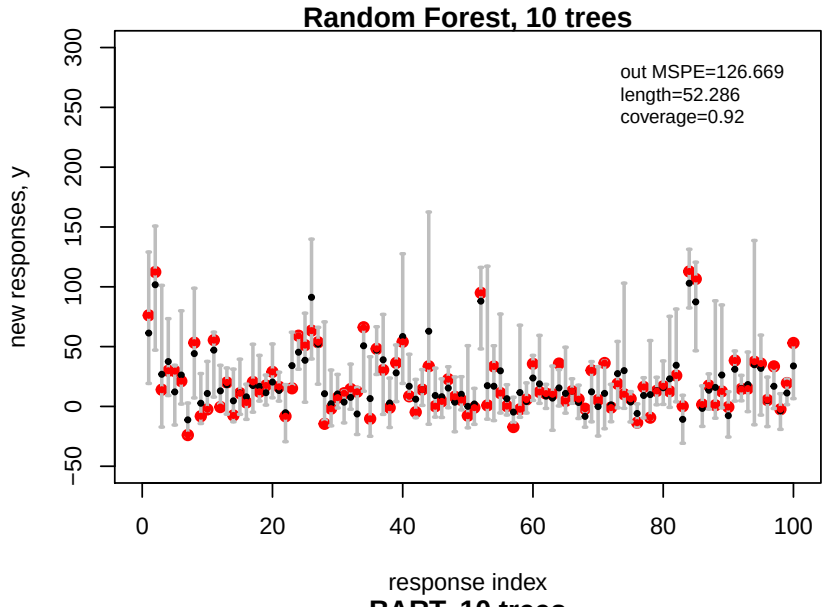


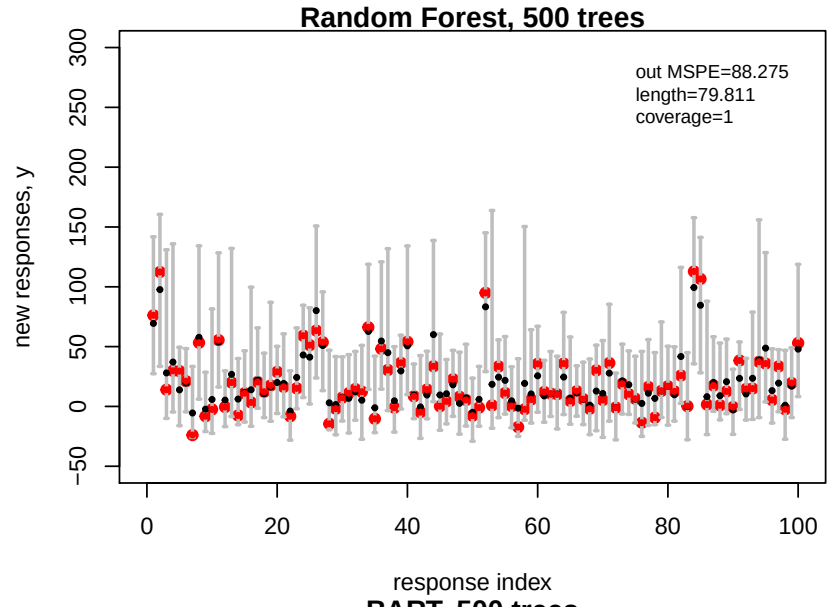


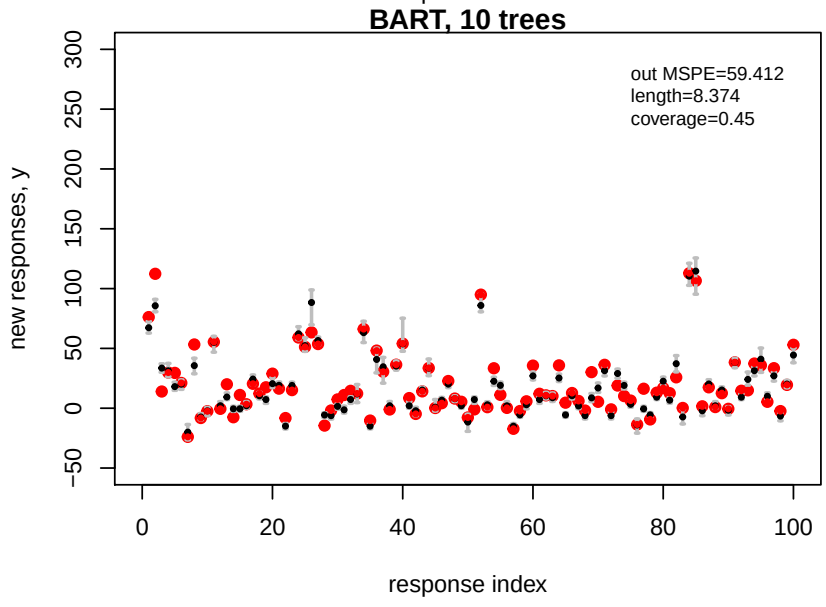


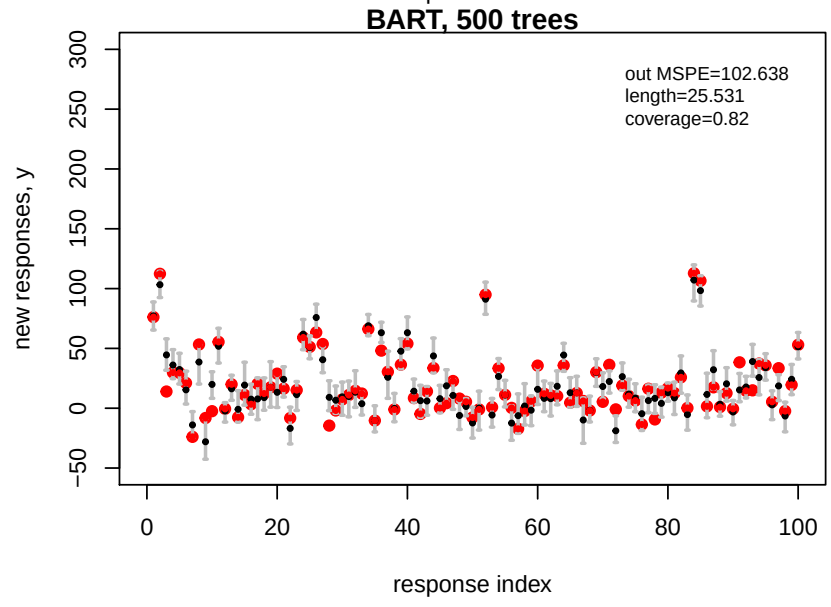


Random Forest vs. BART: $p = 100$, $n = 500$

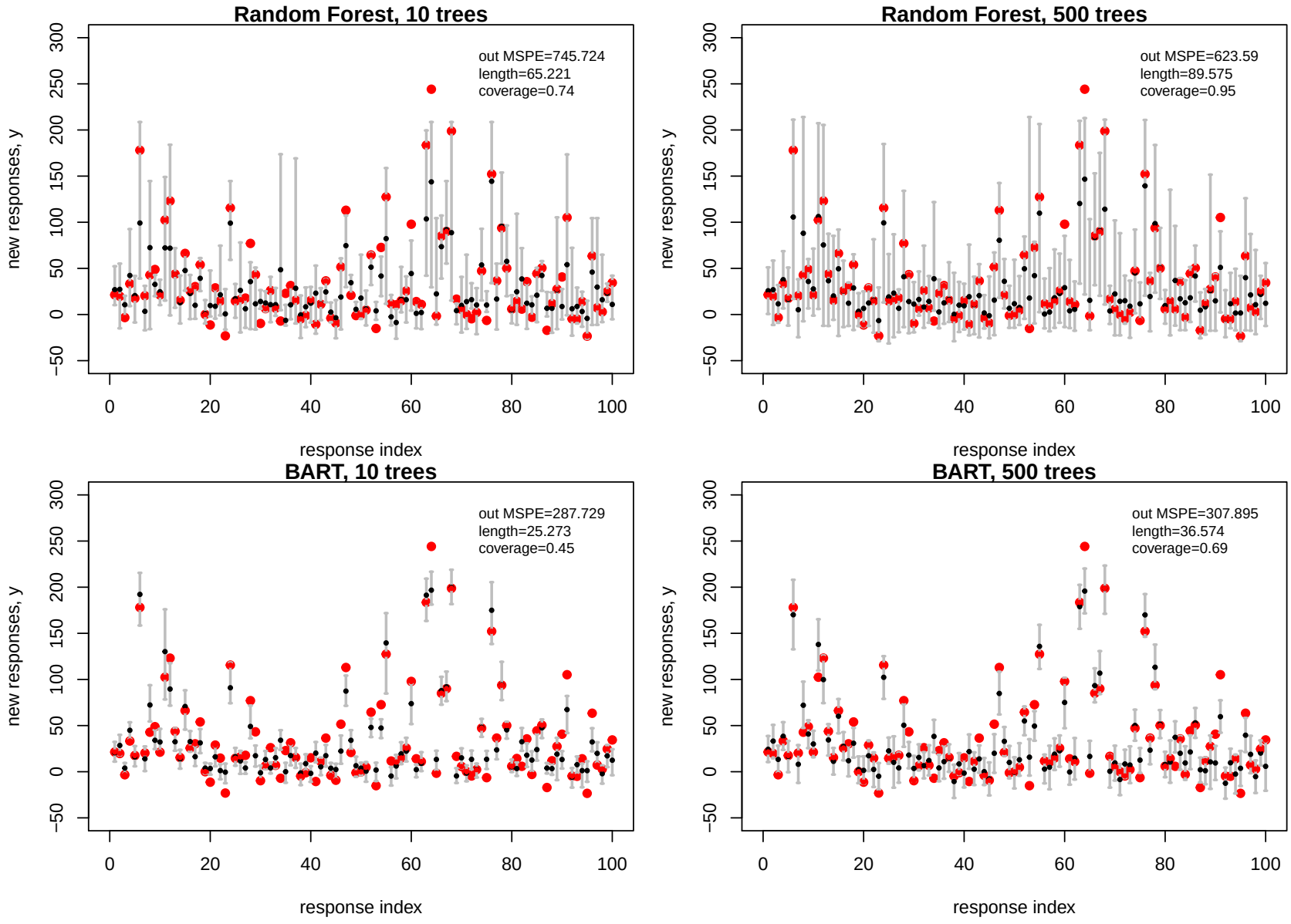




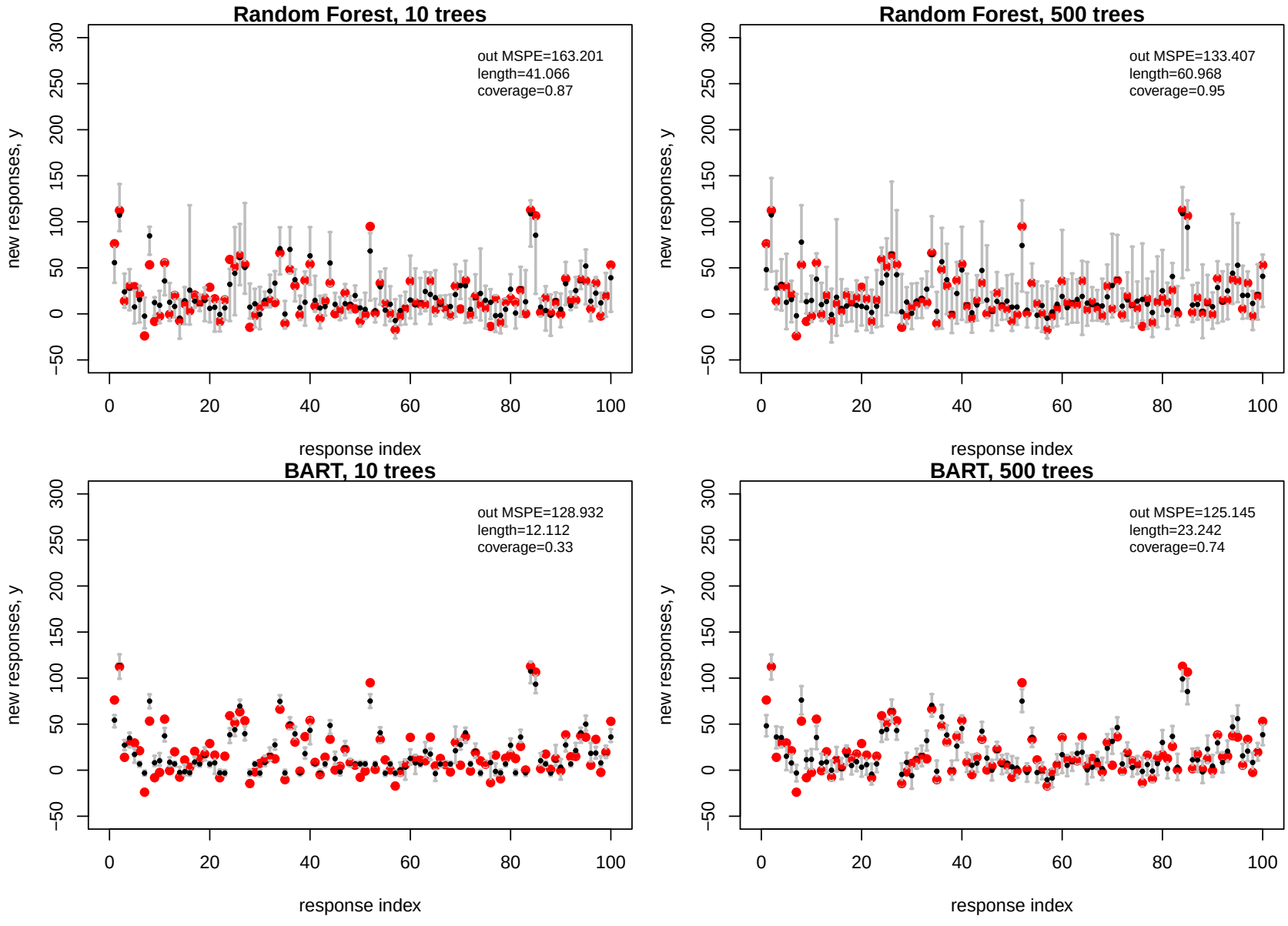




Random Forest vs. BART with added noise to x_1, x_2, and x_9: $p = 200$, $n = 200$



Random Forest vs. BART with added noise to x_1, x_2, and x_9: $p = 100$, $n = 500$



Random Forest vs. BART

- All BART models fit with post-burn-in samples of 3000 and burn-in of 1000
- 95% posterior predicive intervals for 100 new predictors in grey, $y_{est,i}$ in black, and true values $y_{pred,i}$ in red.
- out-of-sample MSPE $= \frac{1}{100}\sum_{i=1}^{50}(y_{pred,i} - y_{est,i})^2$
- "length" is the average length of the 95% predictive intervals
- "coverage" is the proportion of 95% predictive intervals which contain the true values $y_{pred,i}$.