Variable importance-weighted Random Forest

ADDITIONAL SIMULATION STUDIES

Regression

4. $y = 10\,sin(10\pi x_1) + \epsilon$ (same functional form as in 1), with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}$$

and $\epsilon \sim N(0,1)$.

5. $y = 10\,sin(\pi x_1 x_2) + 20(x_3\text{-}0.05)^2 + 10x_4 + 5x_5 + \epsilon$ (same functional form as in 2), with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}$$

and $\epsilon \sim N(0,1)$.

6. $y = f(x) + \epsilon$, with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}$$

and $\epsilon \sim N(0,1)$, and $f(x)$ follows a tree structure as in Figure S1.

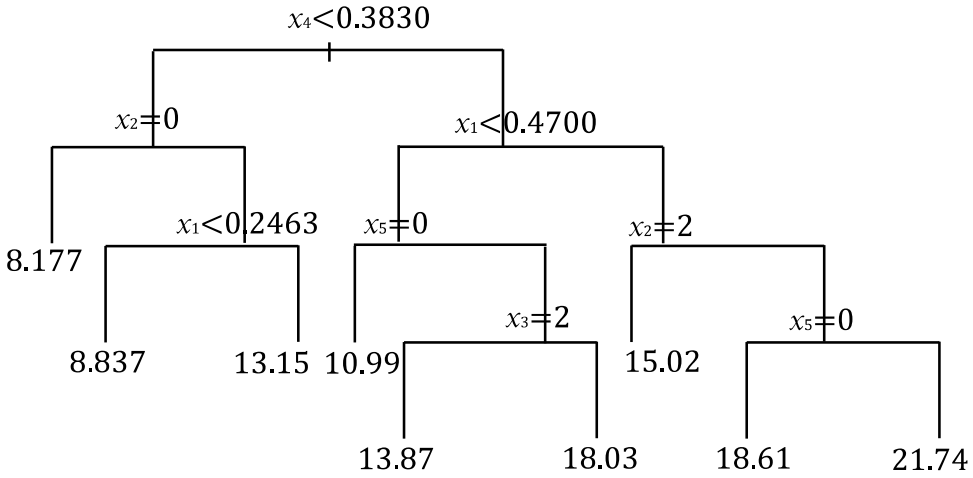

Figure S1. Tree structure used in regression Model 6.

Models 7–9 have the same forms as Models 1–3, respectively, except that though $x_1, x_2, \cdots, x_d$ follows $Unif(0,1)$ marginally, $cor(x_1, x_6) = cor(x_2, x_7) = cor(x_3, x_8) = cor(x_4, x_9) = cor(x_5, x_{10}) = \rho$. We set $\rho = 0.3$ and 0.5.

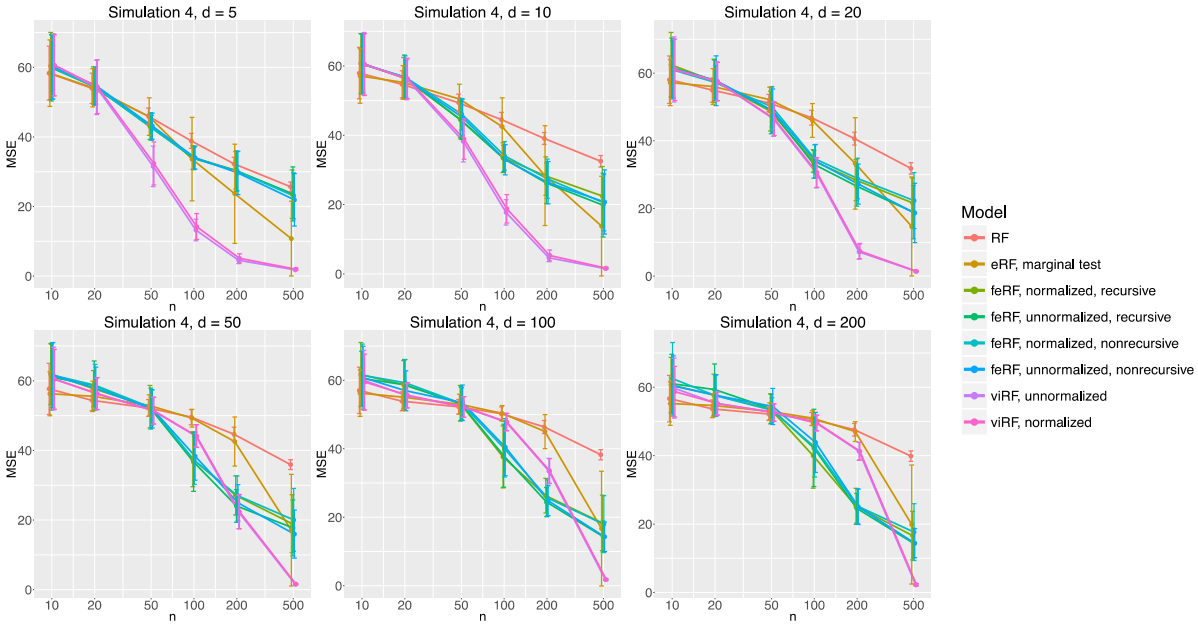

Figure S2. MSE in regression simulation Model 4.

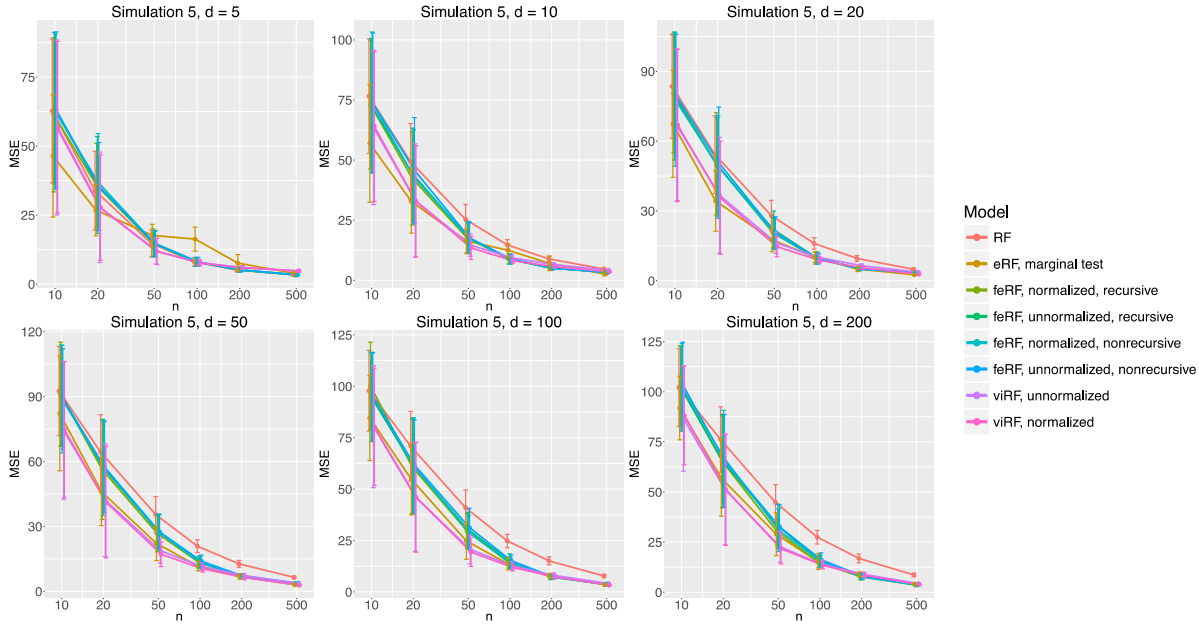

Figure S3. MSE in regression simulation Model 5.

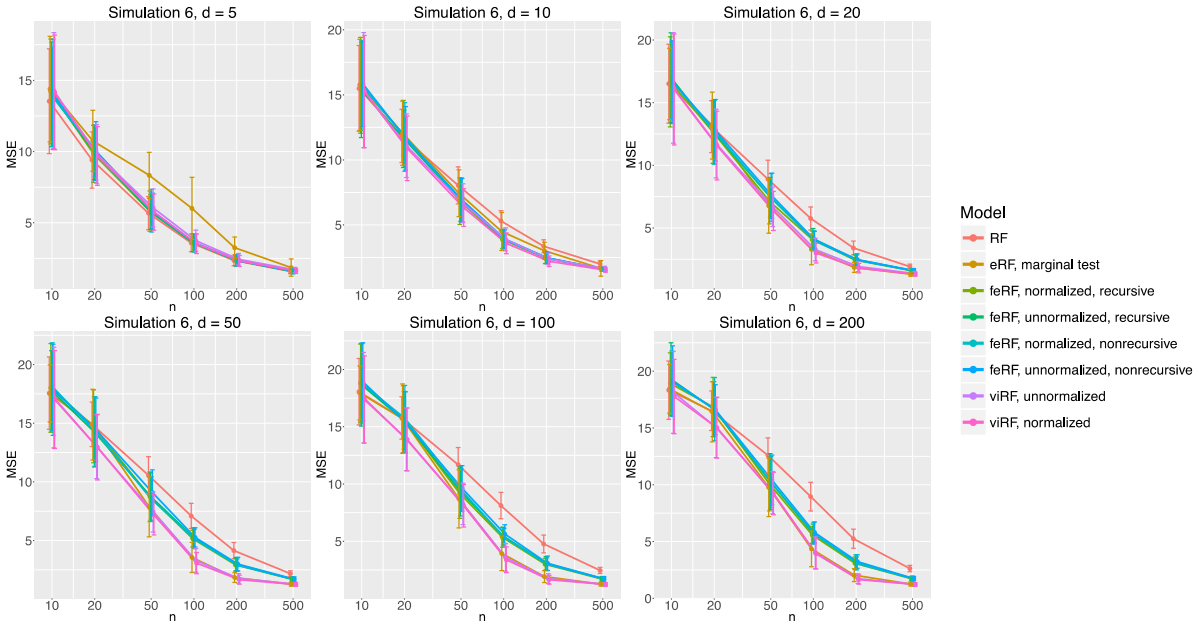

Figure S4. MSE in regression simulation Model 6.

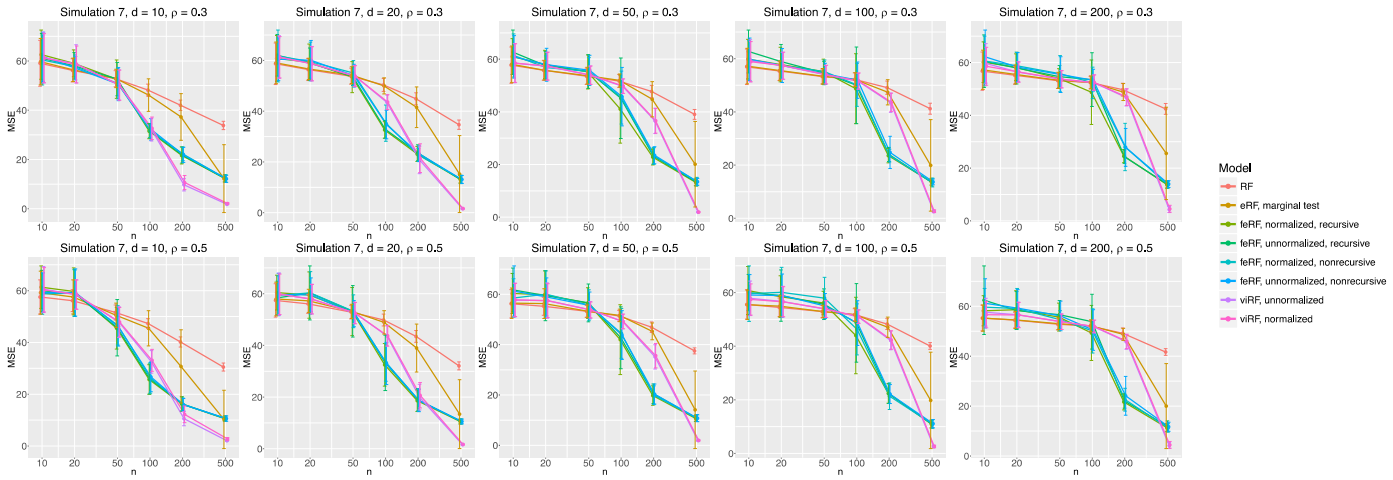

Figure S5. MSE in regression simulation Model 7.

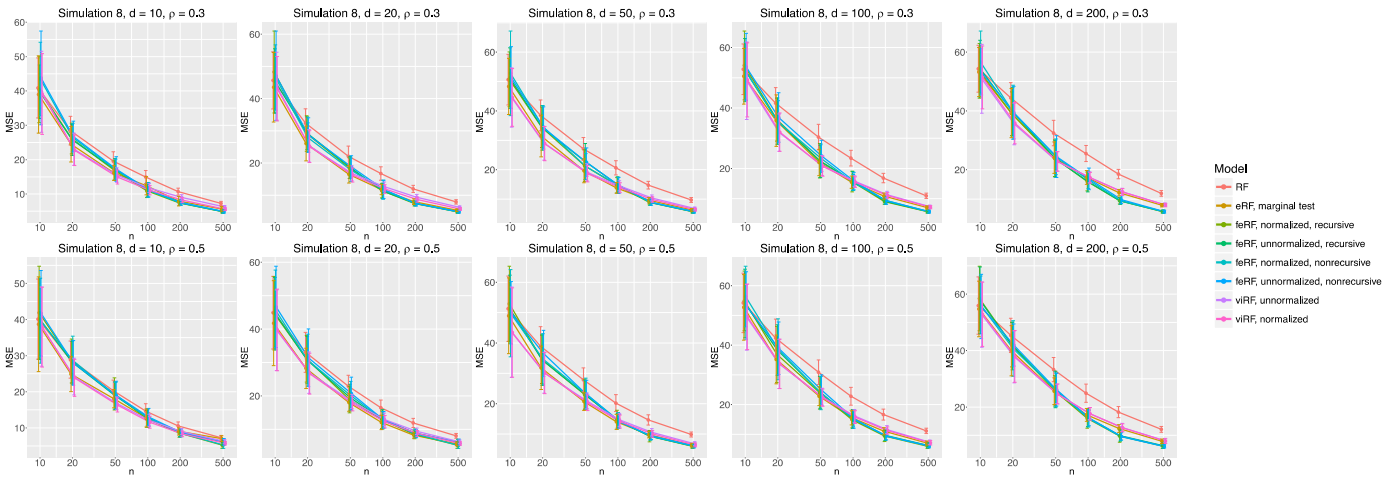

Figure S6. MSE in regression simulation Model 8.

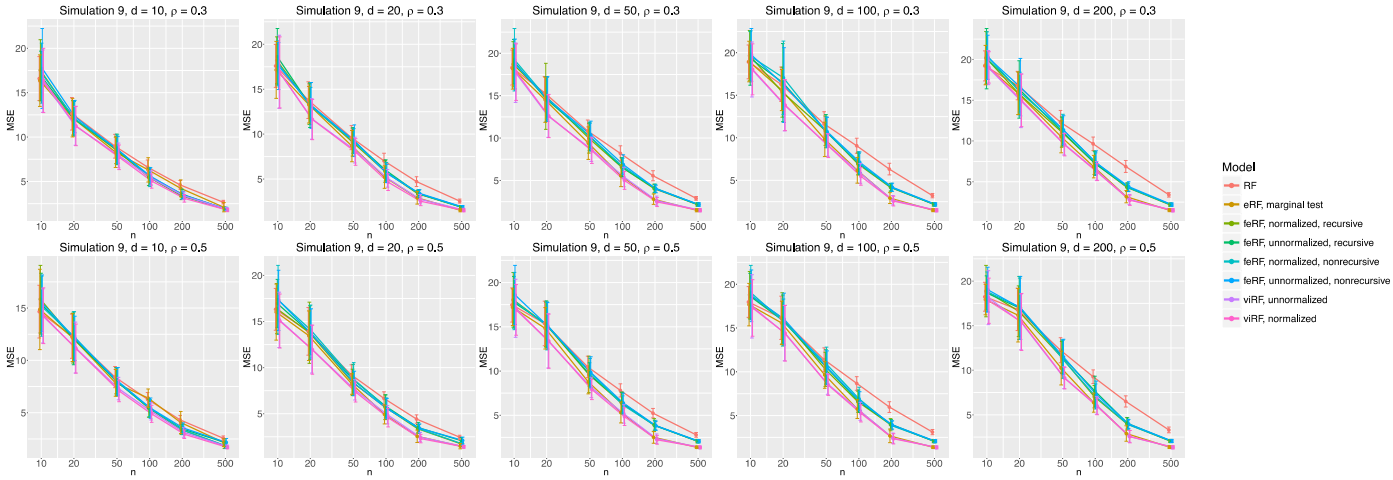

Figure S7. MSE in regression simulation Model 9.

Classification

4. $y \sim Bernoulli\left(\frac{1}{1+exp(1-2x_1)}\right)$, with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}.$$

5. $y \sim Bernoulli\left(\frac{1}{1+exp\left(-\frac{10\,sin(\pi x_1 x_2)+20(x_3-0.05)^2+10x_4+5x_5-20}{3}\right)}\right)$, with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}.$$

6. y follows a tree structure as in Figure S8, with

$$x_i \sim \begin{cases} Unif(0,1), & if\ i = 3k+1 \\ Bernoulli(0.5), & if\ i = 3k+2 \\ Multinomial\big(1,(0.1,0.2,0.7)\big), & if\ i = 3k \end{cases}.$$

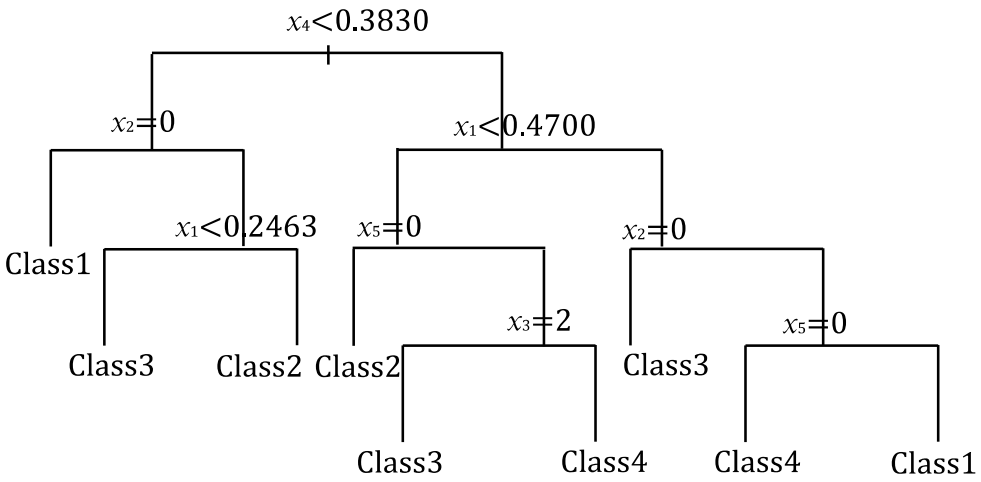

Figure S8. **Tree structure used in classification Model 6.** Noises are added by assigning data point in each terminal node the denoted class with probability 0.9 and any class with probability 0.1/3.

Models 7–9 have the same forms as Models 1–3, respectively, except that though $x_1, x_2, \cdots, x_d$ follows $Unif(0,1)$ marginally, $cor(x_1, x_6) = cor(x_2, x_7) = cor(x_3, x_8) = cor(x_4, x_9) = cor(x_5, x_{10}) = \rho$. We set $\rho = 0.3$ and 0.5.

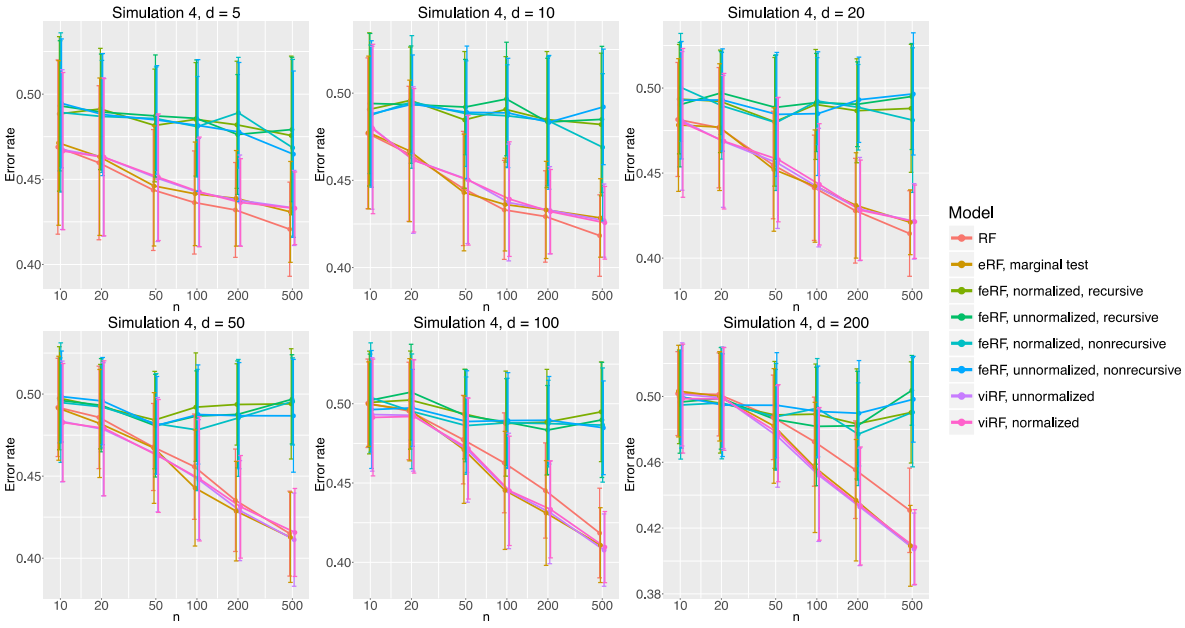

Figure S9. **Error rate in classification simulation Model 4.**

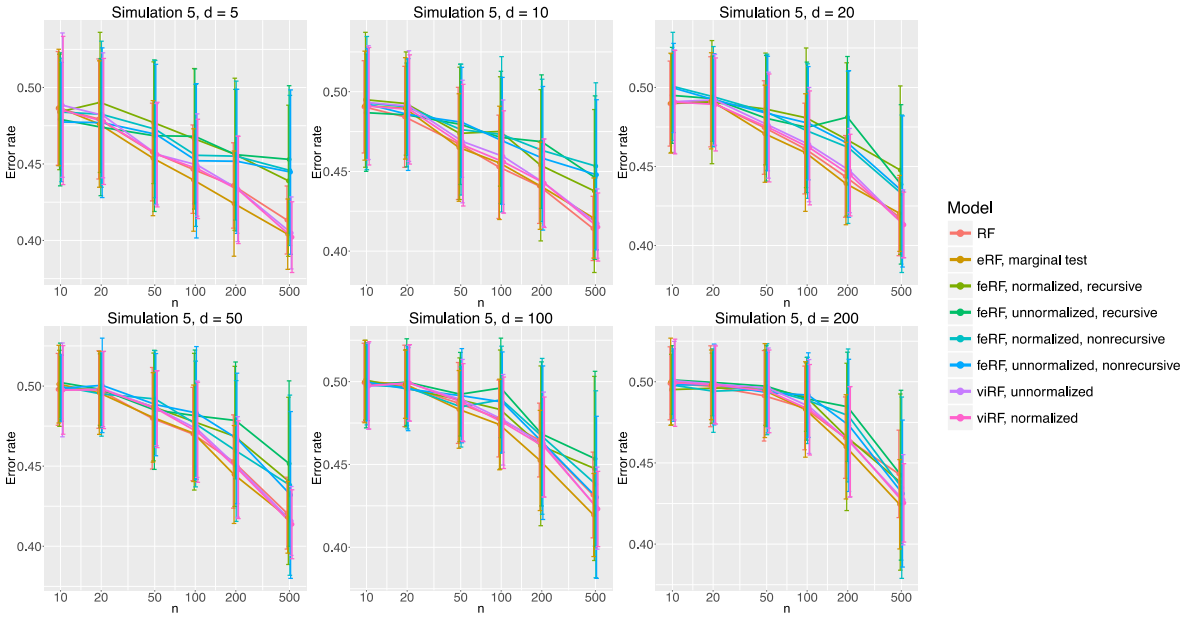

Figure S10. Error rate in classification simulation Model 5.

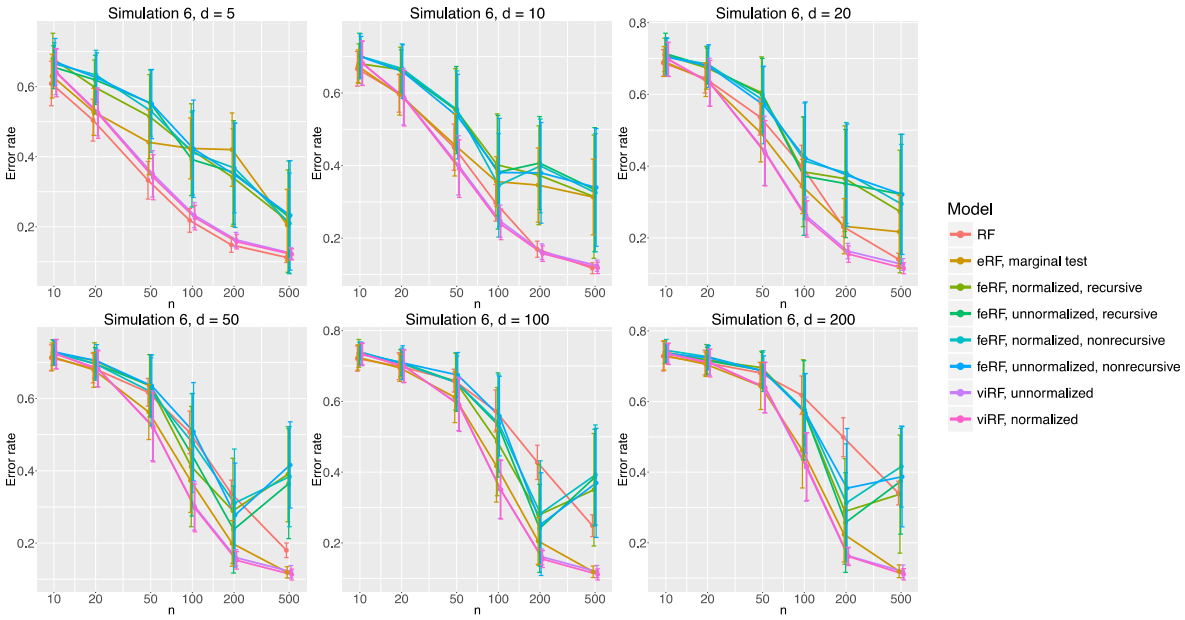

Figure S11. Error rate in classification simulation Model 6.

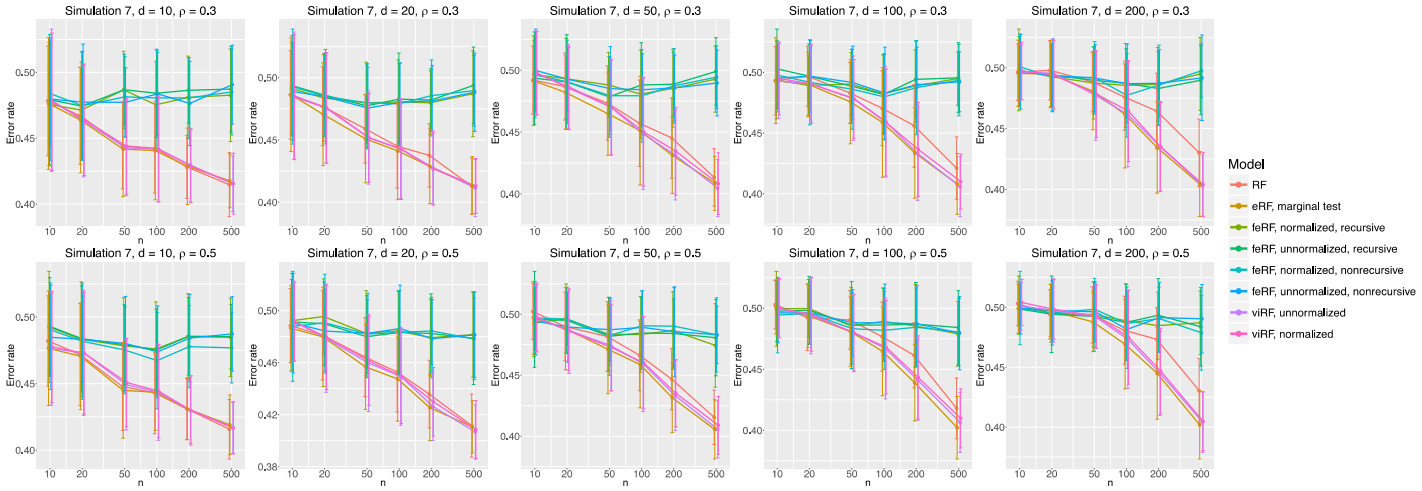

Figure S12. Error rate in classification simulation Model 7.

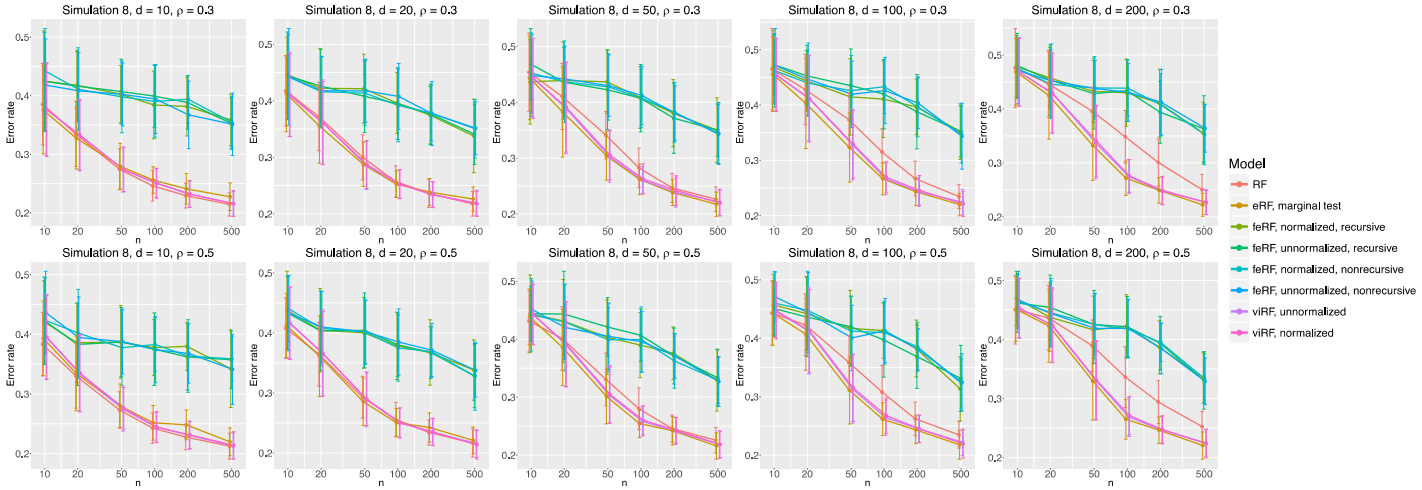

Figure S13. Error rate in classification simulation Model 8.

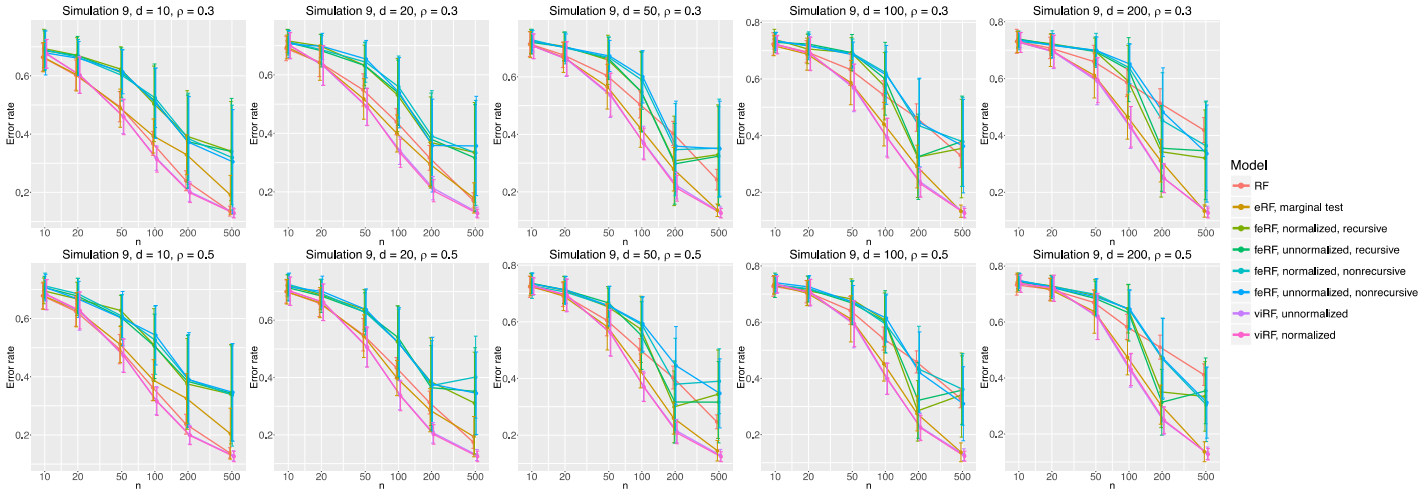

Figure S14. Error rate in classification simulation Model 9.

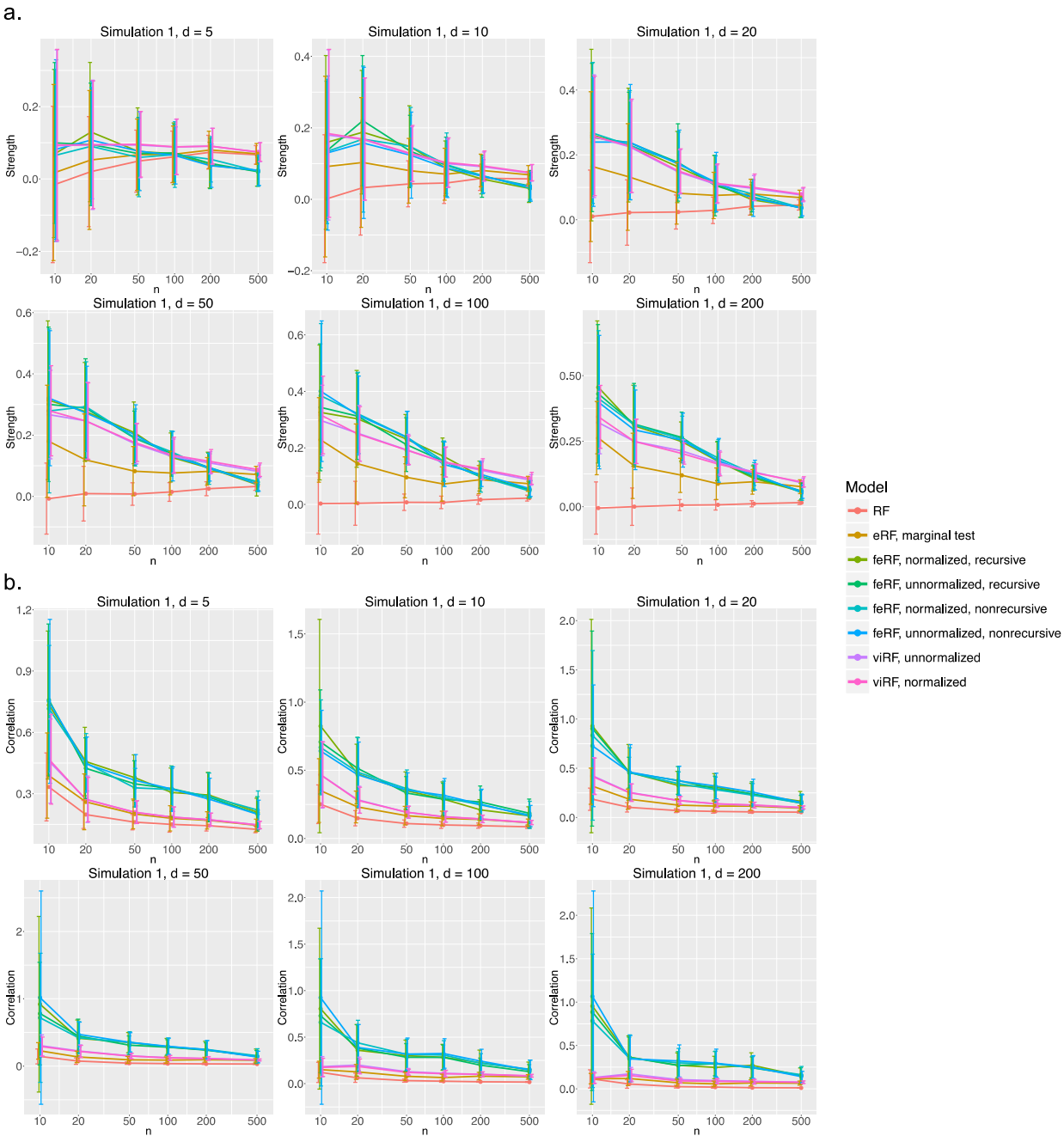

Figure S15. Strength and correlation estimations in classification simulation Model 1.

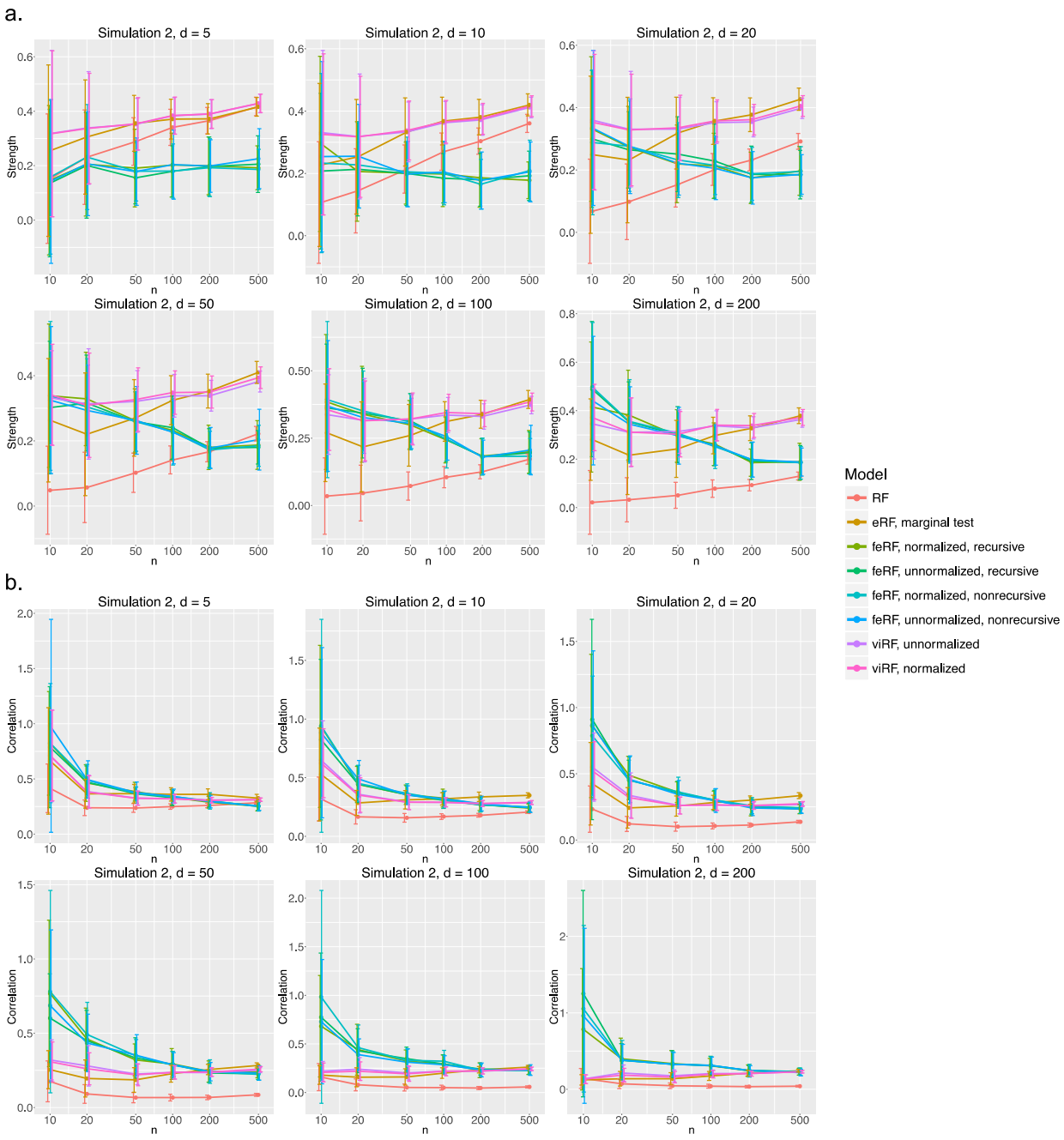

Figure S16. Strength and correlation estimations in classification simulation Model 2.

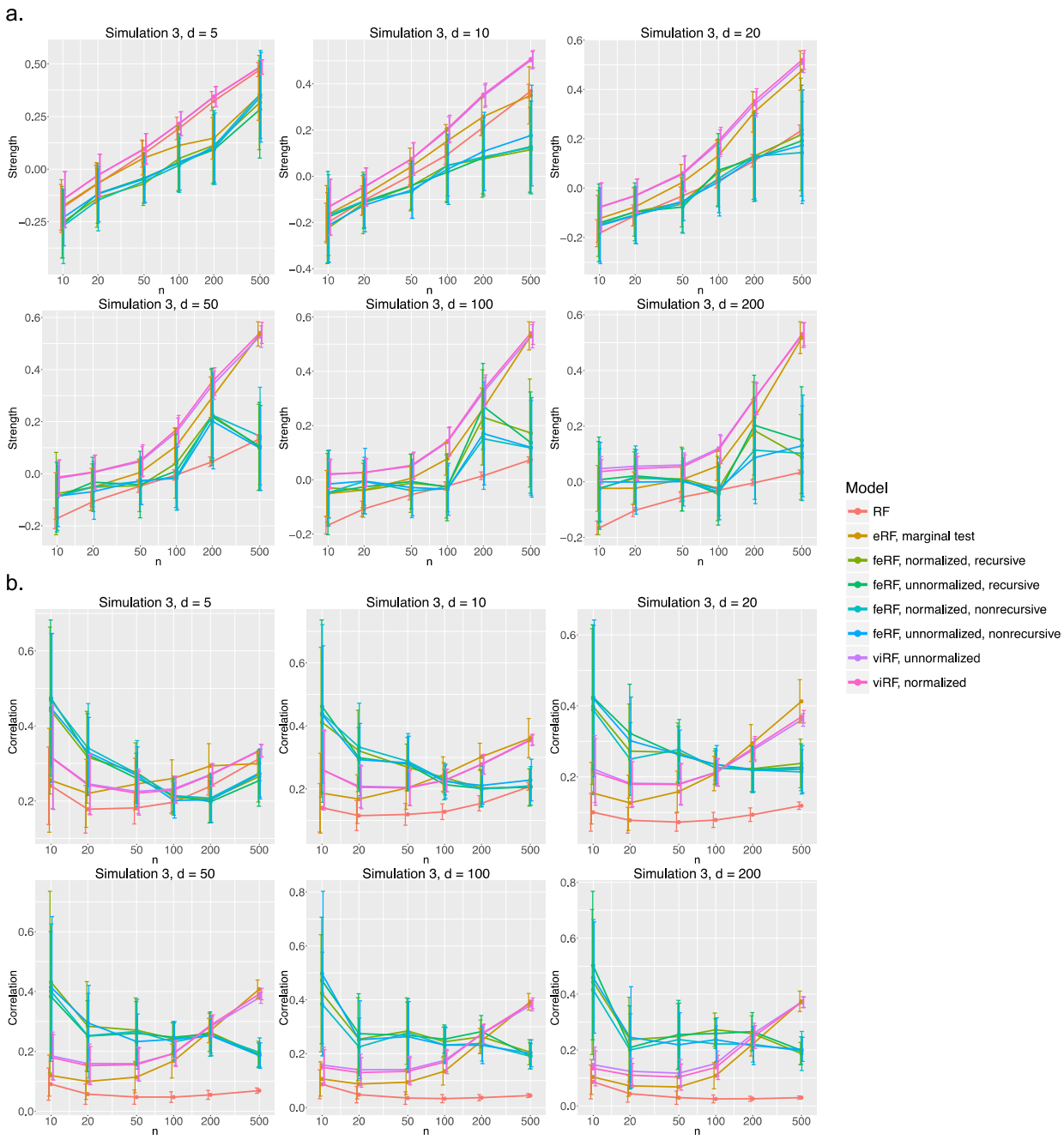

Figure S17. Strength and correlation estimations in classification simulation Model 3.

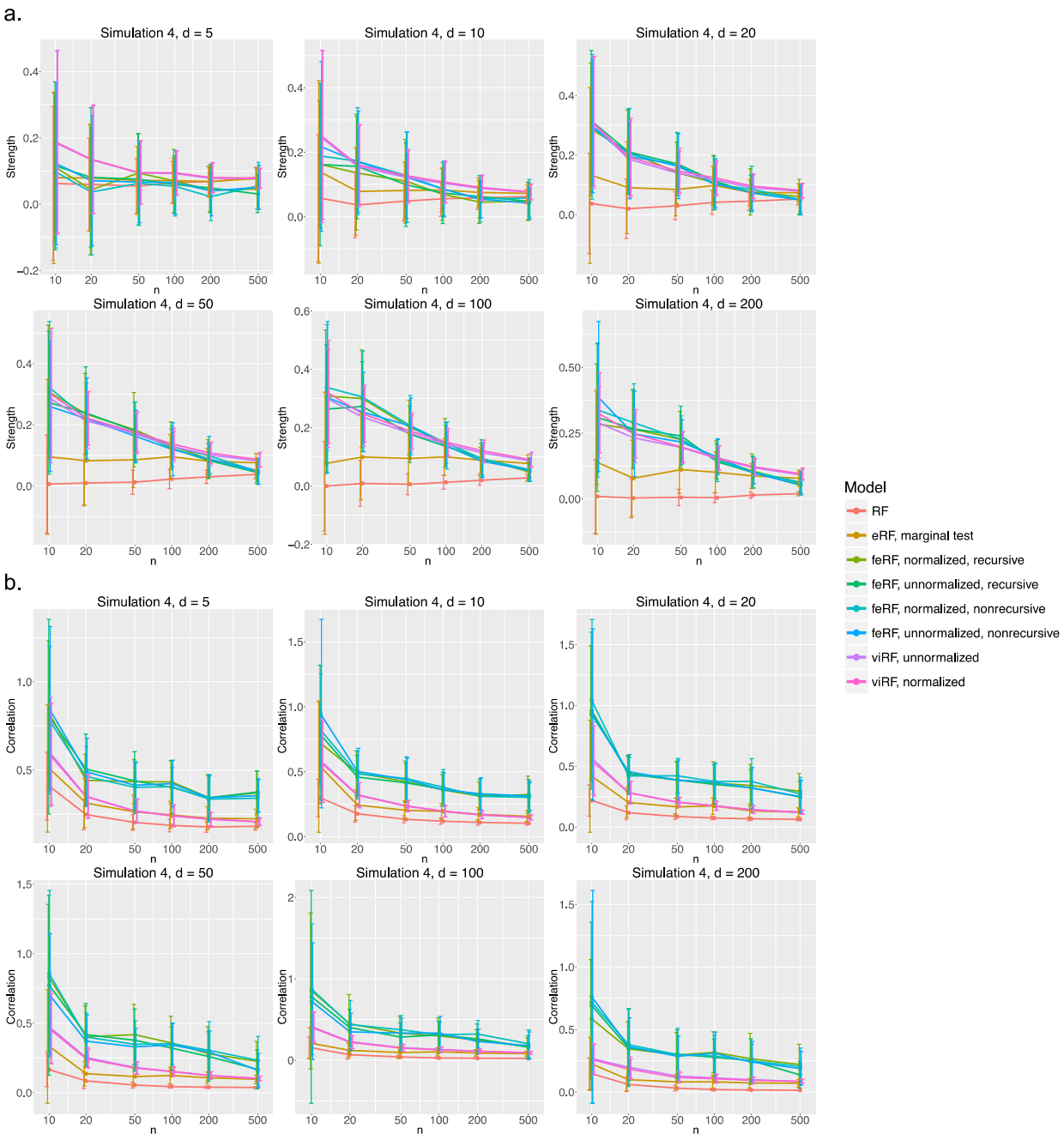

Figure S18. Strength and correlation estimations in classification simulation Model 4.

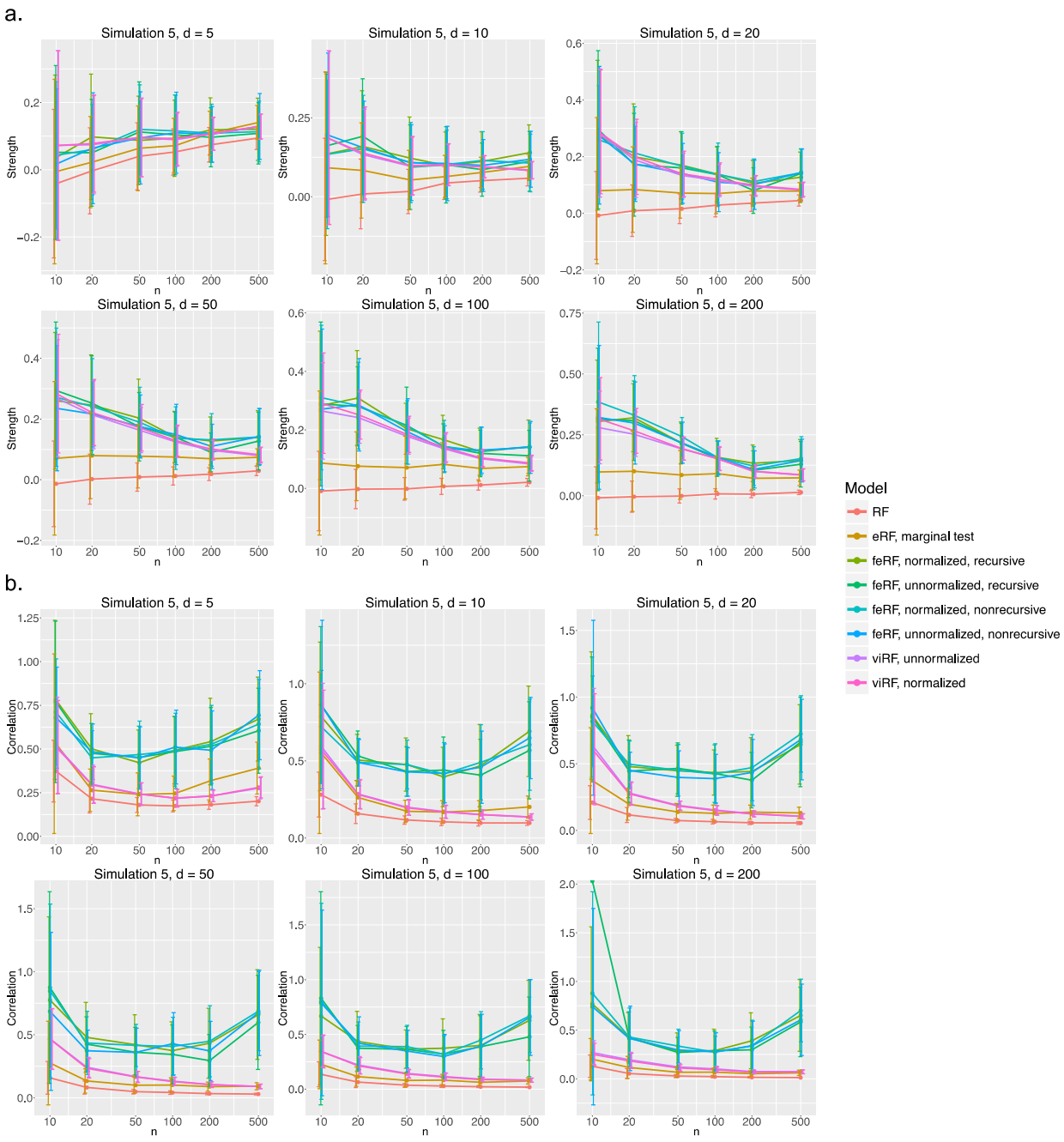

Figure S19. Strength and correlation estimations in classification simulation Model 5.

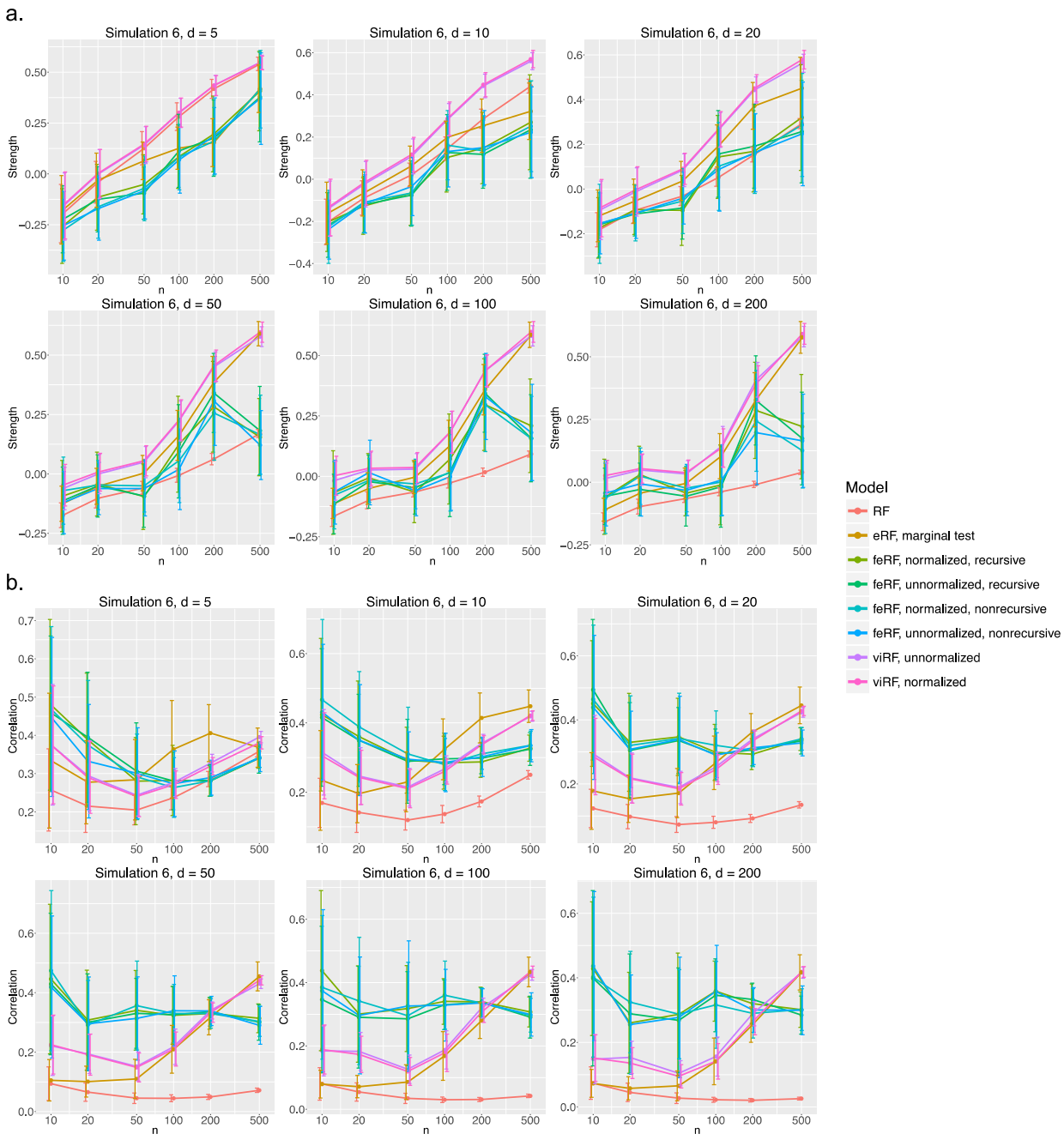

Figure S20. Strength and correlation estimations in classification simulation Model 6.

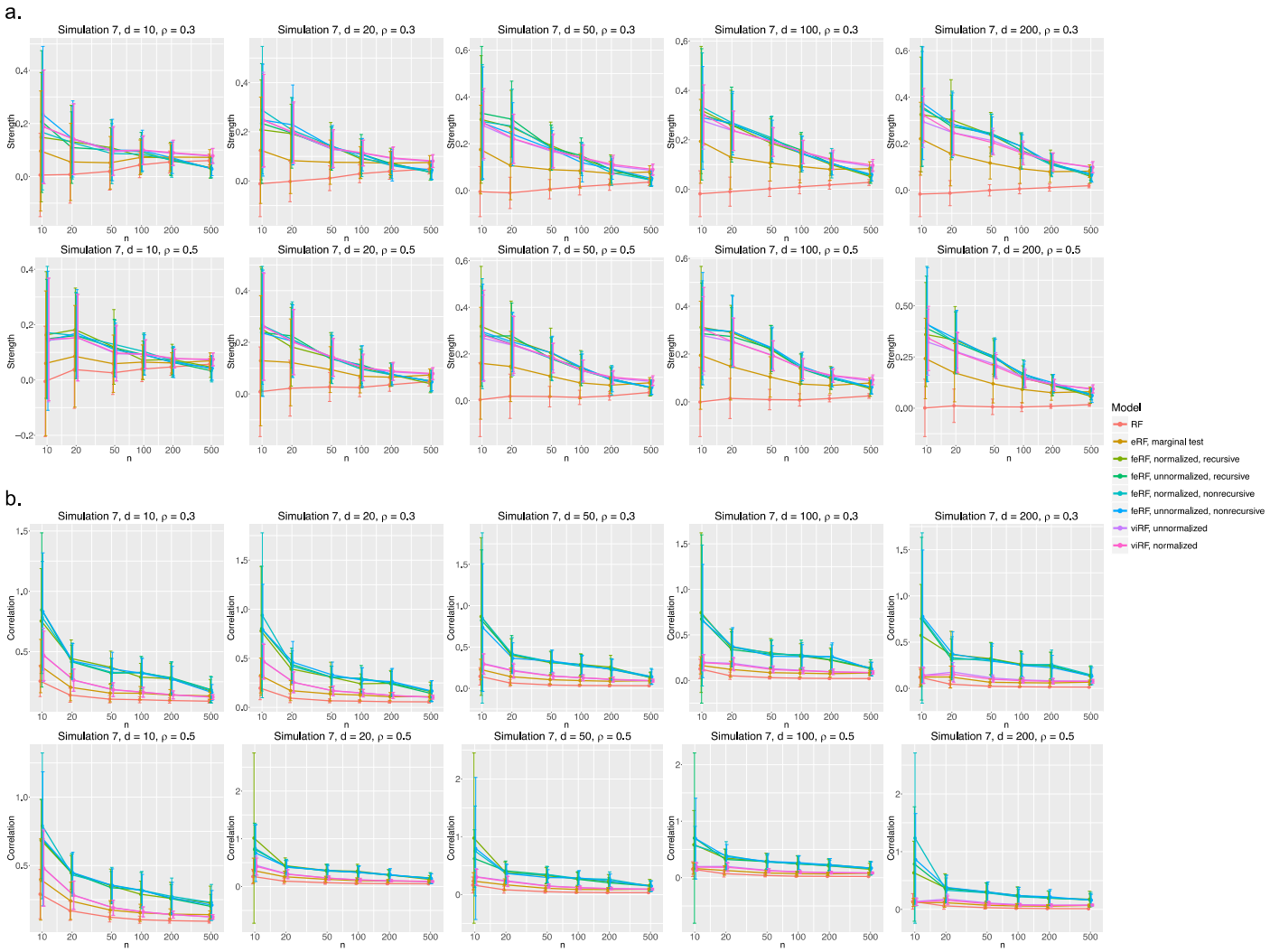

Figure S21. Strength and correlation estimations in classification simulation Model 7.

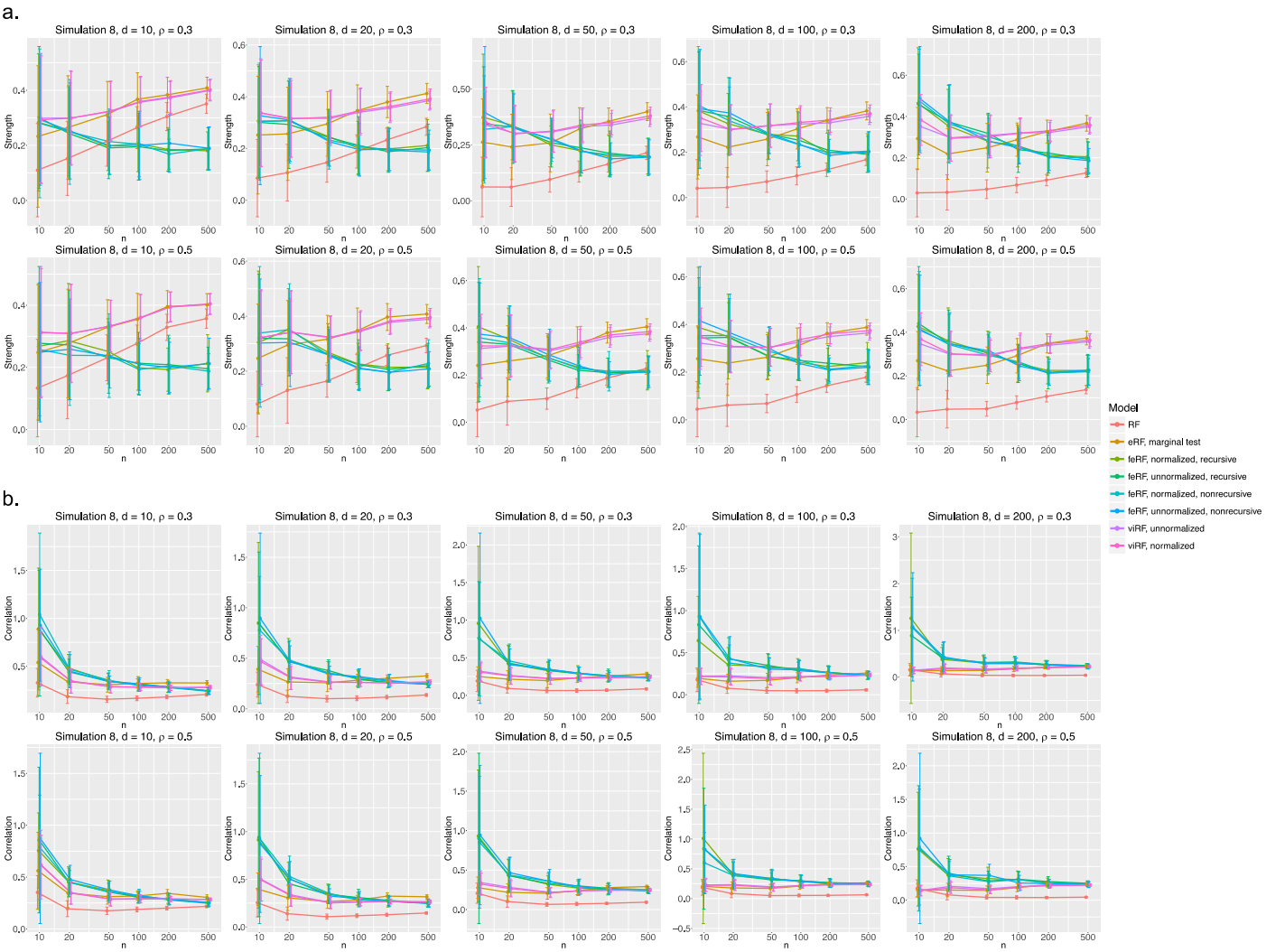

Figure S22. Strength and correlation estimations in classification simulation Model 8.

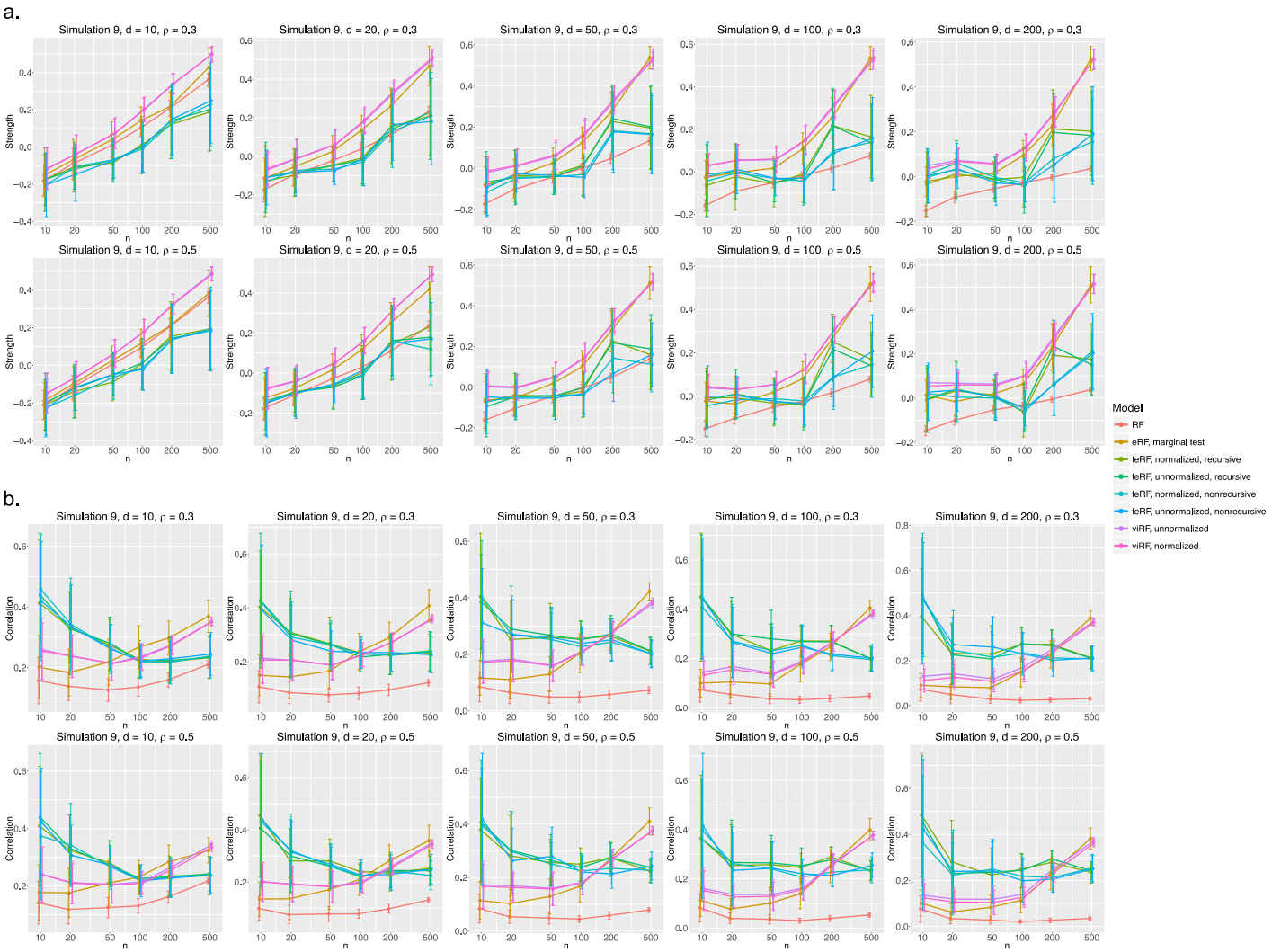

Figure S23. Strength and correlation estimations in classification simulation Model 9.