

Estimating Variance of Simple Defined Variable Effect directly

Felix Kapulla

```
knitr::opts_chunk$set(fig.width=14, fig.height=8)
```

```
library(Matrix)
library(tidyverse)
library(ggplot2)
library(ggpubr)
library(ranger)
library(MixMatrix)
library(mvtnorm)
library(stringr)
library(parallel)

source('C:/Users/feix_/iCloudDrive/Studium Master/CQM - Thesis Internship/Thesis-VariableEffects/Baselin

# cores <- detectCores()
# clust <- makeCluster(cores-1)
# parallel::clusterEvalQ(clust,
#                          expr = {source('C:/Users/feix_/iCloudDrive/Studium Master/CQM - Thesis Interns
```

Simulation

```
n <- c(400, 4000) ; num.trees <- 2000 ; repeats <- 300; cor <- c(0, 0.8); k <- c(0.2, 1)
formulas <- c("2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5")
scenarios <- data.frame(expand.grid(n, num.trees, formulas, repeats, cor, k))
colnames(scenarios) = c("N", "N_Trees", "Formula", "Repeats", "Correlation", "k")
scenarios[, "Formula"] <- as.character(scenarios[, "Formula"]) ### Formula became Factor
scenarios <- split(scenarios, seq(nrow(scenarios)))

system.time(result <- lapply(X = scenarios, FUN = sim_multi))
```

```
##      user      system elapsed
## 42339.00   1733.19  12221.58
```

```
#Run Simulation
# system.time(result <- parLapply(cl = clust,
#                                X = scenarios,
#                                fun = sim_multi))

#stopCluster(clust)
```

```
print_results(result)
```

```
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      1.765698 4.160138 -0.2800863 2.035495 -1.846411
## Mean(s) of simulated LM Variable Effect(s):
##      2.002335 4.000064 -0.5011545 2.199185 -2.002412
## Mean(s) of True Variable Effect(s):
##      2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
##      0.8494947 1.580444 0.2570895 0.9696275 0.8538932 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.8823533 1.799978 0.2939875 1.017027 0.8927729
## Median of simulated Variable Effects (RF):
##      1.629731 4.07471 -0.2568132 1.897095 -1.737883 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##      0.8812027 1.839226 0.2882644 1.023365 0.9056035 .
## Number of Smaller Nulls:
##      32 5 90 28 31
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      2.15934 4.562708 -0.3314276 2.386773 -2.143212
## Mean(s) of simulated LM Variable Effect(s):
##      1.999956 3.999899 -0.4997669 2.20002 -2.000552
## Mean(s) of True Variable Effect(s):
##      2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
##      0.7150549 1.021212 0.1871322 0.7632865 0.6561337 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.9343821 1.160741 0.5914964 1.028114 0.9444246
## Median of simulated Variable Effects (RF):
##      2.089034 4.538783 -0.3204083 2.294994 -2.13397 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##      1.030416 1.310006 0.5031846 1.079875 1.051001 .
## Number of Smaller Nulls:
##      79 66 83 71 82
##
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
##      Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      1.526242 3.613951 -0.2198187 1.7492 -0.9984387
## Mean(s) of simulated LM Variable Effect(s):
##      1.99842 4.002544 -0.5000666 2.20251 -2.005675
## Mean(s) of True Variable Effect(s):
##      2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
##      0.8900724 1.204663 0.4067772 0.9078672 0.4919783 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.9003605 1.331032 0.4378309 0.9775285 0.5155451
## Median of simulated Variable Effects (RF):
```

```

## 1.431997 3.652251 -0.2159904 1.774963 -0.9343169 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.9160294 1.309709 0.4516191 0.9922027 0.5242374 .
## Number of Smaller Nulls:
## 4 0 45 6 28
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## 1.942769 3.922745 -0.373521 2.121884 -1.65685
## Mean(s) of simulated LM Variable Effect(s):
## 1.998732 3.998882 -0.4998101 2.198625 -1.998448
## Mean(s) of True Variable Effect(s):
## 2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
## 0.7143828 0.8874226 0.3612267 0.7503952 0.5498966 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.8178707 0.9564779 0.5928801 0.9061791 0.6676561
## Median of simulated Variable Effects (RF):
## 1.92614 3.937387 -0.3803599 2.111175 -1.638986 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.9235312 1.053381 0.6562049 1.009068 0.7537339 .
## Number of Smaller Nulls:
## 74 59 78 60 73
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
## Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## 1.620681 3.730387 -0.2630816 1.851318 -1.633554
## Mean(s) of simulated LM Variable Effect(s):
## 1.999565 3.999589 -0.4963104 2.1981 -2.001831
## Mean(s) of True Variable Effect(s):
## 2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
## 0.302139 0.3742583 0.1247736 0.3055296 0.2649387 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.3085468 0.397935 0.1375476 0.3321858 0.3190966
## Median of simulated Variable Effects (RF):
## 1.627814 3.730098 -0.2613386 1.834667 -1.628406 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.321045 0.4119004 0.1471011 0.347162 0.3364447 .
## Number of Smaller Nulls:
## 17 6 72 7 15
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
## Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## 1.871921 3.945221 -0.3184287 2.103979 -1.874091
## Mean(s) of simulated LM Variable Effect(s):
## 1.999416 4.000789 -0.5012892 2.201464 -1.999033
## Mean(s) of True Variable Effect(s):
## 2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2064484 0.2480334 0.09067848 0.2259108 0.203166 .

```

```

## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2431238 0.2851551 0.1725158 0.2435186 0.248131
## Median of simulated Variable Effects (RF):
## 1.874888 3.943481 -0.3132151 2.115864 -1.868591 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.2581177 0.3214974 0.1946313 0.271109 0.2715056 .
## Number of Smaller Nulls:
## 75 66 87 70 72
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## 1.416837 3.200286 -0.06073655 1.54687 -0.6981361
## Mean(s) of simulated LM Variable Effect(s):
## 2.011312 4.000078 -0.5038853 2.198191 -2.004073
## Mean(s) of True Variable Effect(s):
## 2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2784019 0.3176799 0.1771393 0.2649613 0.185488 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2749744 0.32482 0.1770215 0.283014 0.1911751
## Median of simulated Variable Effects (RF):
## 1.387367 3.201169 -0.05212429 1.55451 -0.6903183 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.2816415 0.3226017 0.1895593 0.2928259 0.2018363 .
## Number of Smaller Nulls:
## 10 12 29 8 19
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula = 2*x.1+4*x.2-0.5*x.3+2.2*x.4-2*x.5 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## 1.660347 3.648601 -0.2882246 1.840207 -1.303673
## Mean(s) of simulated LM Variable Effect(s):
## 2.000822 3.997865 -0.497928 2.200873 -2.001006
## Mean(s) of True Variable Effect(s):
## 2 4 -0.5 2.2 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2138037 0.2083314 0.1320871 0.2056709 0.1641011 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2048572 0.2367852 0.1742315 0.2246341 0.1958679
## Median of simulated Variable Effects (RF):
## 1.653723 3.648114 -0.2956053 1.838443 -1.29399 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.226464 0.2665232 0.1993862 0.2462304 0.2075676 .
## Number of Smaller Nulls:
## 64 67 76 54 66

```

```

effect_plots <- plot_effects(result)
se_plot <- plot_se(result)

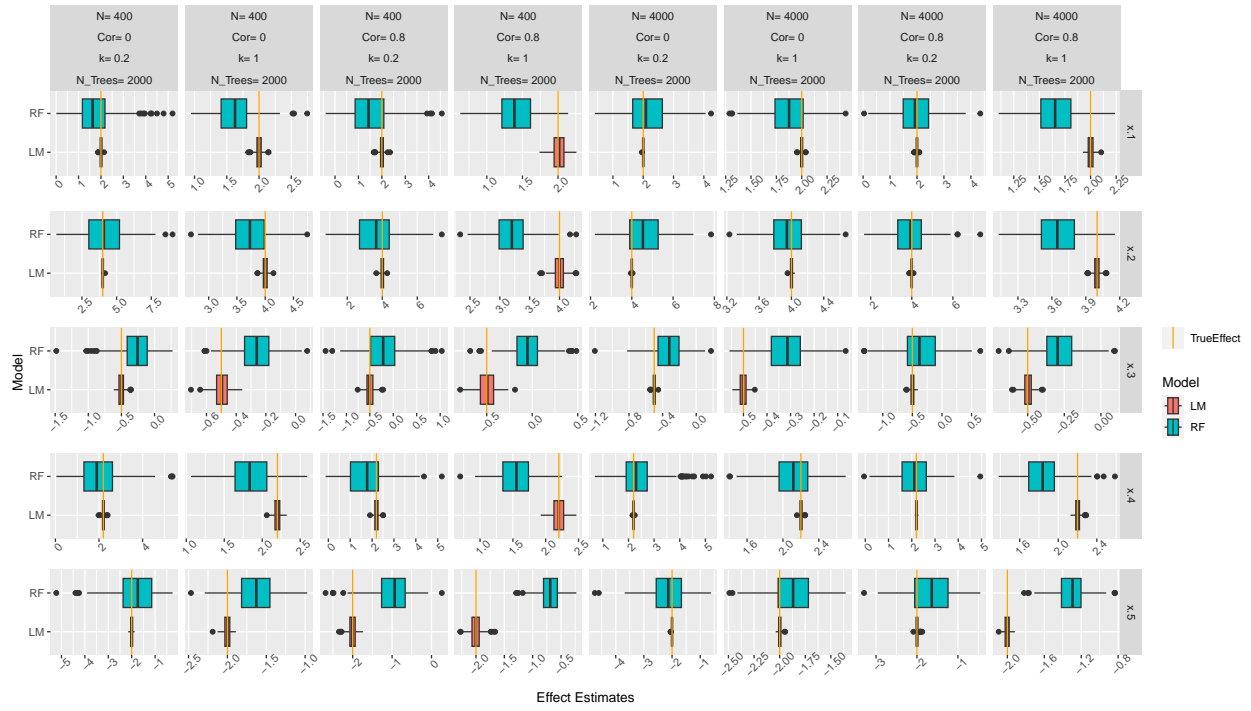
```

```

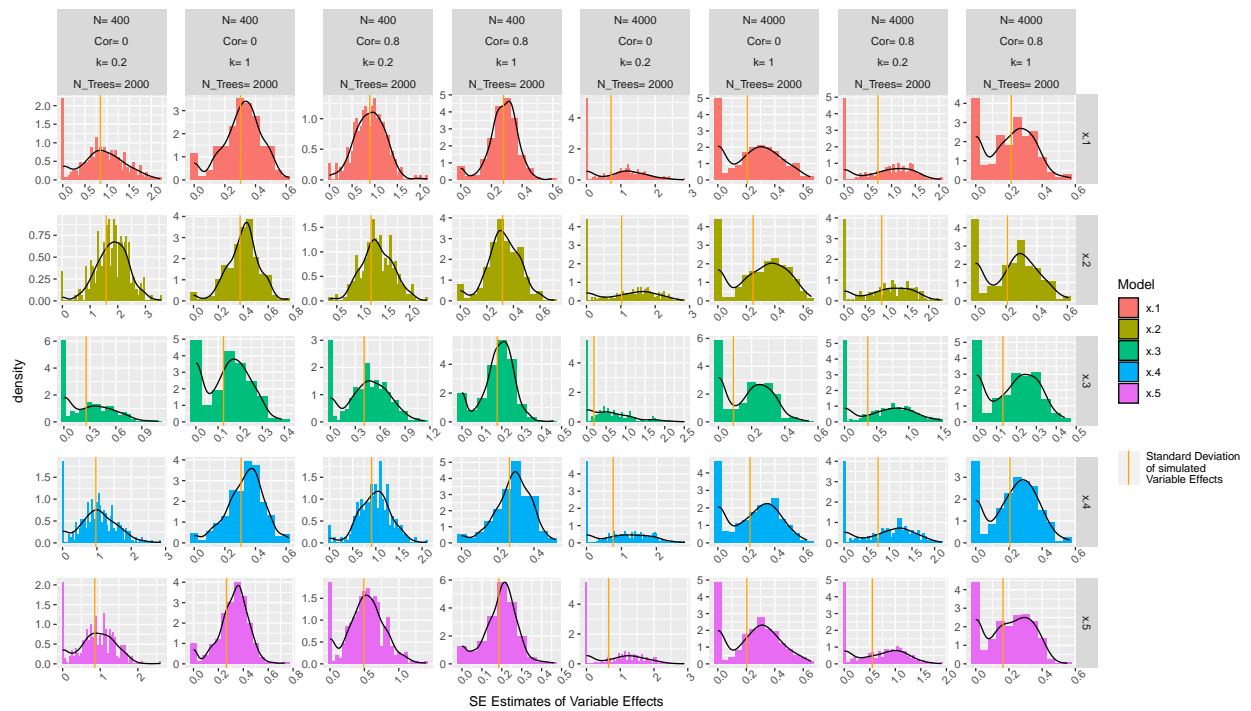
## 'summarise()' has grouped output by 'N', 'cor', 'k', 'num.trees'. You can
## override using the '.groups' argument.

```

effect_plots



se_plot



```
##### Simulation Setup
n <- c(400, 4000) ; num.trees <- 2000 ; repeats <- 300; cor <- c(0, 0.8); k <- c(0.2, 1)
formulas <- "-0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)"
scenarios <- data.frame(expand.grid(n, num.trees, formulas, repeats, cor, k))
colnames(scenarios) = c("N", "N_Trees", "Formula", "Repeats", "Correlation", "k")
scenarios[, "Formula"] <- as.character(scenarios[, "Formula"]) ### Formula became Factor
scenarios <- split(scenarios, seq(nrow(scenarios)))

system.time(result <- lapply(X = scenarios, FUN = sim_multi))
```

```
##      user      system elapsed
## 44567.89   1234.30  11338.95
```

```
print_results(result)
```

```
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4) ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.1562217 3.102419 0.002651791 0.934037
## Mean(s) of simulated LM Variable Effect(s):
##   -1.450824 3.00752 0.009146743 1.634032
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668
## Standard Error of simulated Variable Effects (RF):
##   0.3750851 1.244659 0.2890007 0.7719354 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.4225392 1.407134 0.3233565 0.7786444
## Median of simulated Variable Effects (RF):
##   -0.1044294 3.133029 0.006812927 0.7989819 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.4209205 1.363937 0.3136983 0.7518801 .
## Number of Smaller Nulls:
##   18 0 29 6
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4) ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.1415862 3.019131 0.02506283 1.091027
## Mean(s) of simulated LM Variable Effect(s):
##   -1.493408 3.000081 -0.00555536 1.647097
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668
## Standard Error of simulated Variable Effects (RF):
##   0.4527778 0.9837263 0.3156663 0.8441504 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.5123811 1.063988 0.4228515 0.851087
## Median of simulated Variable Effects (RF):
##   -0.1211503 2.923169 0.01440579 1.024585 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.5372544 1.089251 0.4554145 0.8971101 .
## Number of Smaller Nulls:
##   50 26 61 30
```

```

##
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
##      Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      -0.04000734 2.974037 -0.001816456 0.879645
## Mean(s) of simulated LM Variable Effect(s):
##      -1.453838 3.001001 -0.005774327 1.627799
## Mean(s) of True Variable Effect(s):
##      -0.02 3 0 1.00668
## Standard Error of simulated Variable Effects (RF):
##      0.469775 1.211724 0.5072089 0.8054308 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.5009119 1.217201 0.5326407 0.8552199
## Median of simulated Variable Effects (RF):
##      -0.03695104 2.90757 -0.04016296 0.8140563 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##      0.5032084 1.189682 0.5227502 0.8461156 .
## Number of Smaller Nulls:
##      14 0 3 2
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
##      Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      -0.05996332 3.072403 -0.02960176 0.9663213
## Mean(s) of simulated LM Variable Effect(s):
##      -1.497457 2.996022 0.001632659 1.649611
## Mean(s) of True Variable Effect(s):
##      -0.02 3 0 1.00668
## Standard Error of simulated Variable Effects (RF):
##      0.4864397 0.9224415 0.4975599 0.8537466 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.5226245 0.9851745 0.5529222 0.8271817
## Median of simulated Variable Effects (RF):
##      -0.01682697 3.093829 -0.03081717 0.8790254 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##      0.5572689 1.000186 0.5843954 0.8874955 .
## Number of Smaller Nulls:
##      66 15 51 39
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
##      Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##      -0.4567763 3.014113 0.007479166 1.143642
## Mean(s) of simulated LM Variable Effect(s):
##      -1.465661 3.00417 -0.003846221 1.617922
## Mean(s) of True Variable Effect(s):
##      -0.5 3 0 1.175201
## Standard Error of simulated Variable Effects (RF):
##      0.2066638 0.2847246 0.1697098 0.2613133 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##      0.2284497 0.3205926 0.1803074 0.3005662
## Median of simulated Variable Effects (RF):
##      -0.4468365 3.004734 0.006058431 1.152851 .
## Median of Standard Errors Estimates of Variable Effects (RF):

```

```

## 0.2254082 0.3099767 0.1779625 0.293341 .
## Number of Smaller Nulls:
## 4 1 6 0
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.4493646 3.003746 0.002189407 1.216095
## Mean(s) of simulated LM Variable Effect(s):
## -1.494477 3.000812 -0.001290465 1.653668
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201
## Standard Error of simulated Variable Effects (RF):
## 0.1780044 0.2195641 0.1673658 0.2080786 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2033913 0.2255008 0.1806238 0.2277421
## Median of simulated Variable Effects (RF):
## -0.4370713 2.989858 0.0001581352 1.222878 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.22064 0.2289078 0.188577 0.234487 .
## Number of Smaller Nulls:
## 31 17 36 16
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.1882025 2.770228 0.1390613 1.04218
## Mean(s) of simulated LM Variable Effect(s):
## -1.466087 3.012388 0.006381881 1.63057
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201
## Standard Error of simulated Variable Effects (RF):
## 0.1958283 0.2573603 0.1967099 0.2545071 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2056799 0.294524 0.2219109 0.2794088
## Median of simulated Variable Effects (RF):
## -0.1878599 2.789704 0.1393925 1.025781 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.1999657 0.2852174 0.2135883 0.2723528 .
## Number of Smaller Nulls:
## 3 0 0 0
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4)$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.369312 2.957836 0.02974153 1.126295
## Mean(s) of simulated LM Variable Effect(s):
## -1.49282 3.004281 -0.001801194 1.652452
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201
## Standard Error of simulated Variable Effects (RF):
## 0.1834441 0.2193697 0.1859169 0.2120078 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.1870038 0.2230574 0.1872613 0.2186765

```

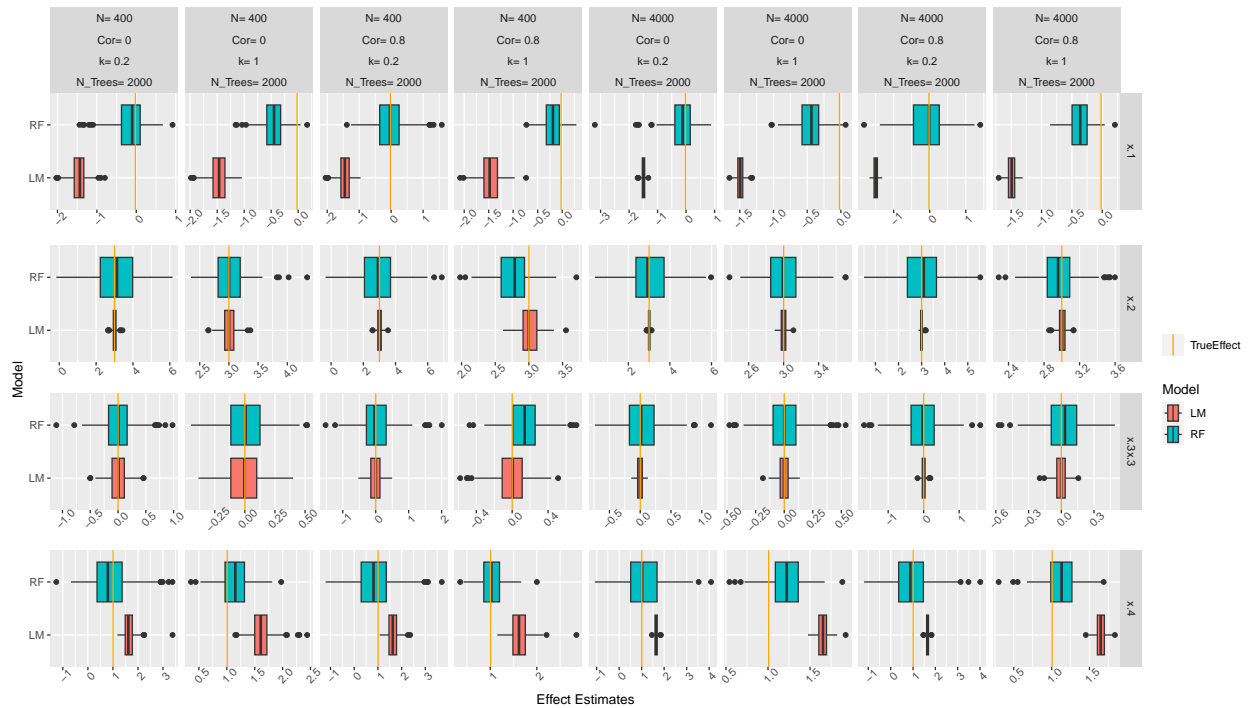


```
## Median of simulated Variable Effects (RF):
##   -0.3628256 2.95802 0.03390052 1.13166 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.2033234 0.2310488 0.1987329 0.2333356 .
## Number of Smaller Nulls:
##   39 22 41 21
```

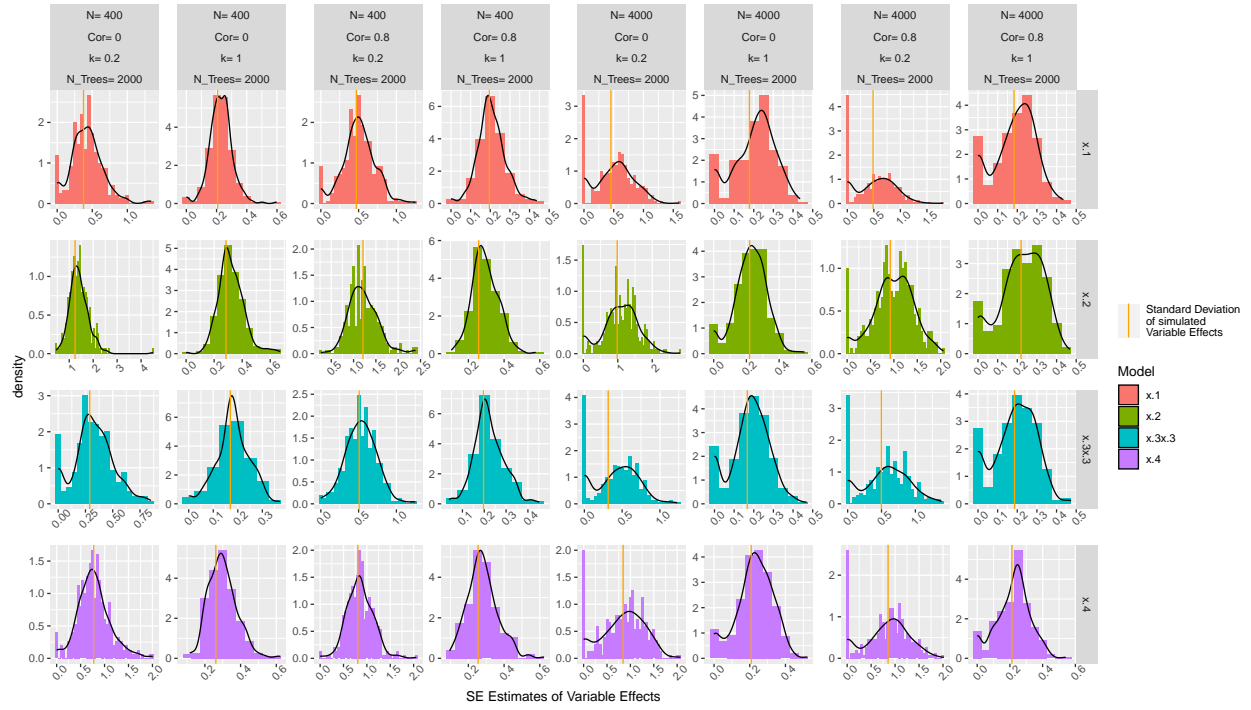
```
effect_plots <- plot_effects(result)
se_plot <- plot_se(result)
```

```
## 'summarise()' has grouped output by 'N', 'cor', 'k', 'num.trees'. You can
## override using the '.groups' argument.
```

```
effect_plots
```



```
se_plot
```



```
##### Simulation Setup
n <- c(400, 4000) ; num.trees <- 2000 ; repeats <- 300 ; cor <- c(0, 0.8) ; k <- c(0.2, 1)
formulas <- "-0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4"
scenarios <- data.frame(expand.grid(n, num.trees, formulas, repeats, cor, k))
colnames(scenarios) = c("N", "N_Trees", "Formula", "Repeats", "Correlation", "k")
scenarios[, "Formula"] <- as.character(scenarios[, "Formula"]) ### Formula became Factor
scenarios <- split(scenarios, seq(nrow(scenarios)))

system.time(result <- lapply(X = scenarios, FUN = sim_multi))
```

```
##      user      system elapsed
## 45124.75   1384.37  12369.25
```

```
print_results(result)
```

```
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.3930512 2.940552 -0.004642126 0.7753538 -1.143109
## Mean(s) of simulated LM Variable Effect(s):
##   -1.483667 2.992709 -0.004173174 1.633668 -1.991801
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668 -2
## Standard Error of simulated Variable Effects (RF):
##   0.5050929 1.193422 0.3184247 0.7068629 0.8086891 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.5870907 1.366507 0.3367847 0.7075613 3.808673
## Median of simulated Variable Effects (RF):
##   -0.359115 2.947306 -0.004124017 0.6685353 -1.031712 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.5839813 1.320796 0.3470301 0.7027218 3.771208 .
## Number of Smaller Nulls:
##   11 0 39 11 0
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.1884349 3.115669 0.009758058 0.9875865 -1.841083
## Mean(s) of simulated LM Variable Effect(s):
##   -1.493246 3.000621 -0.0009537506 1.644995 -2.007155
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668 -2
## Standard Error of simulated Variable Effects (RF):
##   0.586581 0.9135359 0.3311035 0.7097422 1.478373 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.6716635 1.008918 0.3749408 0.7682952 3.039295
## Median of simulated Variable Effects (RF):
##   -0.1508823 3.121792 0.01140838 0.9250205 -1.762566 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.7282611 1.05823 0.3824123 0.8337543 3.056489 .
## Number of Smaller Nulls:
##   45 30 68 40 8
```

```

##
## Setting: N = 400 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.1150211 2.986642 0.07169844 0.8996221 -0.9548162
## Mean(s) of simulated LM Variable Effect(s):
##   -1.448904 2.983978 0.0176315 1.638859 -2.311563
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668 -2
## Standard Error of simulated Variable Effects (RF):
##   0.5256045 1.150078 0.5234379 0.7279262 0.948319 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.5519138 1.261285 0.565142 0.815152 3.37505
## Median of simulated Variable Effects (RF):
##   -0.07618163 3.016077 0.03879535 0.778919 -0.8424775 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.5390097 1.233902 0.5568727 0.8084366 3.263698 .
## Number of Smaller Nulls:
##   4 0 5 2 0
##
## Setting: N = 4000 ; k = 0.2 N_Trees = 2000 ; Correlation = 0.8 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.1148555 3.029524 0.0206135 0.9939978 -1.760414
## Mean(s) of simulated LM Variable Effect(s):
##   -1.493485 3.003046 0.0007864512 1.639235 -2.285786
## Mean(s) of True Variable Effect(s):
##   -0.02 3 0 1.00668 -2
## Standard Error of simulated Variable Effects (RF):
##   0.5908879 0.9843089 0.4760935 0.8042422 1.802603 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.6196806 0.9918503 0.5411358 0.8535615 2.976614
## Median of simulated Variable Effects (RF):
##   -0.1046045 3.029926 0.02388686 0.9995556 -1.581787 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.6789564 1.023498 0.5851544 0.8711268 3.010002 .
## Number of Smaller Nulls:
##   48 19 59 33 11
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
##      Formula = -0.5*x.1^3+3*x.2+0.5*sqrt(abs(x.3))*x.3^2+exp(x.4)+2*x.1*x.4 ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
##   -0.6487884 2.981257 -0.01455935 1.063978 -1.044002
## Mean(s) of simulated LM Variable Effect(s):
##   -1.456299 3.001222 0.01120132 1.643666 -1.984955
## Mean(s) of True Variable Effect(s):
##   -0.5 3 0 1.175201 -2
## Standard Error of simulated Variable Effects (RF):
##   0.2912973 0.302632 0.1754866 0.3125536 0.191601 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
##   0.2996861 0.330565 0.1984615 0.3373983 0.2840965
## Median of simulated Variable Effects (RF):
##   -0.6358929 2.971965 0.001129179 1.053435 -1.039769 .
## Median of Standard Errors Estimates of Variable Effects (RF):

```

```

## 0.3006386 0.3223058 0.2040467 0.3315643 0.2791841 .
## Number of Smaller Nulls:
## 1 0 6 0 0
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4) + 2 \cdot x_1 \cdot x_4$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.5439176 2.996005 -0.002707696 1.159552 -1.738637
## Mean(s) of simulated LM Variable Effect(s):
## -1.499335 3.001102 0.006435655 1.644912 -1.997311
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2020786 0.2110858 0.1425932 0.2216575 0.1596363 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.2332624 0.2168693 0.1557404 0.2421136 0.2106214
## Median of simulated Variable Effects (RF):
## -0.5373407 3.00315 -0.01160946 1.157458 -1.740282 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.2451908 0.2197421 0.1627759 0.2513942 0.2189123 .
## Number of Smaller Nulls:
## 30 19 50 25 10
##
## Setting: N = 400 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4) + 2 \cdot x_1 \cdot x_4$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.3411433 2.831839 0.1230824 1.13267 -0.6124932
## Mean(s) of simulated LM Variable Effect(s):
## -1.460177 2.993721 0.003020572 1.636002 -2.26547
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2685001 0.289264 0.2166005 0.3105779 0.1476834 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.264097 0.3077683 0.2394876 0.3459258 0.2357229
## Median of simulated Variable Effects (RF):
## -0.3211209 2.834094 0.124438 1.133558 -0.6196198 .
## Median of Standard Errors Estimates of Variable Effects (RF):
## 0.2530752 0.2964167 0.2305697 0.3352539 0.226106 .
## Number of Smaller Nulls:
## 0 0 1 0 0
##
## Setting: N = 4000 ; k = 1 N_Trees = 2000 ; Correlation = 0.8 ;
## Formula =  $-0.5 \cdot x_1^3 + 3 \cdot x_2 + 0.5 \cdot \sqrt{\text{abs}(x_3)} \cdot x_3^2 + \exp(x_4) + 2 \cdot x_1 \cdot x_4$  ; N_Trees = 2000
## Mean(s) of simulated RF Variable Effect(s):
## -0.5070049 3.009949 0.02285546 1.137109 -1.236571
## Mean(s) of simulated LM Variable Effect(s):
## -1.488514 3.00024 0.001945576 1.640555 -2.290594
## Mean(s) of True Variable Effect(s):
## -0.5 3 0 1.175201 -2
## Standard Error of simulated Variable Effects (RF):
## 0.2264456 0.2299142 0.2059167 0.2357768 0.1587578 .
## Mean of Standard Errors Estimates of Variable Effects (RF):
## 0.217131 0.2387044 0.2029618 0.2420133 0.1882779

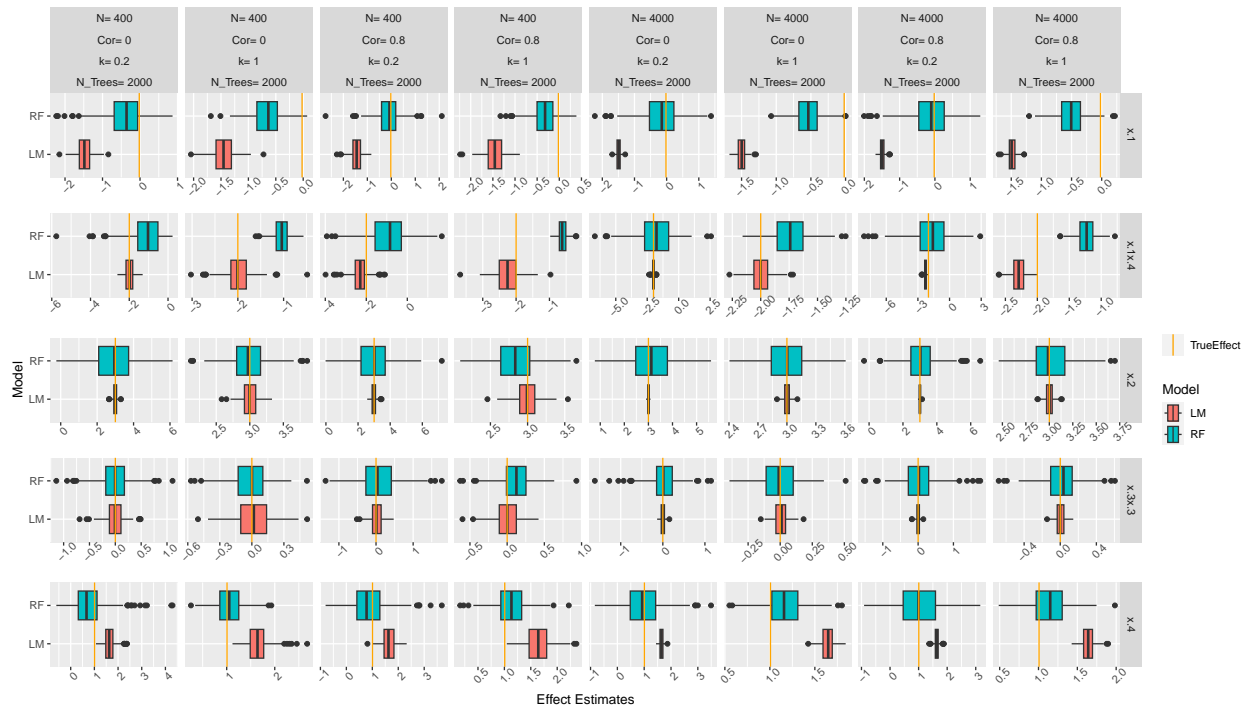
```

```
## Median of simulated Variable Effects (RF):
##   -0.5056281 2.986084 0.03573558 1.15093 -1.227985 .
## Median of Standard Errors Estimates of Variable Effects (RF):
##   0.2272654 0.245898 0.2123682 0.2474276 0.1884245 .
## Number of Smaller Nulls:
##   30 15 28 17 6
```

```
effect_plots <- plot_effects(result)
se_plot <- plot_se(result)
```

```
## 'summarise()' has grouped output by 'N', 'cor', 'k', 'num.trees'. You can
## override using the '.groups' argument.
```

effect_plots



se_plot

