05_figure2_individuals

randy

2023-10-24

```r
source("R/00_functions.R")
# load("randy_margin_model_random_intercept_slope.Rdata")
load("R/train_test.rda")
```

setting

```r
train_data <- train %>%
  mutate(id = as.character(id),
         group = "Training")
test_data <- test %>%
  mutate(id = as.character(id),
         group = "Testing")

test_baseline <-
  test_data %>%
  group_by(id) %>%
  slice(1L)

id_test <- unique(test_data$id)

knots = c(10, 12, 15)
pred_time = c(3, 6, 9, 12)
bks <- brokenstick::brokenstick(ht ~ time | id,
                                data = train_data,
                                knots = knots)
dataset_baseline <- train_data %>%
      group_by(id) %>%
      slice(1L) %>%
      dplyr::select(-time) %>%
      dplyr::select(baseline = ht, everything())
print("2.baseline works")
```

```
#> [1] "2.baseline works"
```

```r
print("this is the dataset_baseline--------------------")
```

```
#> [1] "this is the dataset_baseline--------------------"
```

```
bks_pred_knots <- predict(bks,
                          x = pred_time,
                          shape = "long",
                          include_data = FALSE) %>%
  dplyr::select(id, time, `.pred`)
print("this is bks_pred_knots for training --------------------")
```

```
#> [1] "this is bks_pred_knots for training --------------------"
```

```
train_pred <-
  bks_pred_knots %>%
  # pivot_wider(names_from = time,
  #             values_from = .pred) %>%
  select_if(not_all_na) %>%
  # dplyr::select(target = as.character(match_time),
  #               everything()) %>%
  full_join(dataset_baseline, by = c("id")) %>%
  mutate(timef = paste0("time", time))

## checkpoint ------------------------------------------------------------------
bks_pred_knots <- predict(bks,
                          x = pred_time,
                          shape = "long",
                          # group = train_data$id,
                          include_data = FALSE) %>%
  dplyr::select(id, time, `.pred`)
print("this is bks_pred_knots for training --------------------")
```

```
#> [1] "this is bks_pred_knots for training --------------------"
```

```
train_pred <- bks_pred_knots %>%
  select_if(not_all_na)

train_new <- full_join(train_pred, dataset_baseline)
## add one factor time_var variable

test_new <- test_baseline %>%
  mutate(time = list(pred_time)) %>%
  unnest() %>%
  rename(baseline = ht)
```

```
#> Warning: ‘cols‘ is now required when using ‘unnest()‘.
#> i Please use ‘cols = c(time)‘.
```

```
## 1.2 linear fitting
lm_bks_train <- lm("`.pred` ~ as.factor(time) * sex * genotype * baseline",
                   data = train_new)
predicted_train <- predict(lm_bks_train)
predicted_test<- predict(lm_bks_train, newdata = test_new)
```

```r
lb_train <- train_new %>%
  ungroup() %>%
  mutate(lm_bks_target = predicted_train) %>%
  dplyr::select(lm_bks_target) %>%
  cbind(train_pred) %>%
  as.data.frame() %>%
  dplyr::select(contains(c("id", "time", "lm_bks_target"))) %>%
  as.matrix() %>%
  as.data.frame() %>%
  mutate(lm_bks_target = as.numeric(lm_bks_target))


test_new[, "lm_bks_target"] = as.numeric(predicted_test)
lb_test <- test_new %>%
  dplyr::select(contains(c("id", "time", "lm_bks")))
```

testing

```r
id = 159662
# id =  154526
test_eld_n10 <- pred_matching(
                 lb_data = lb_train,
                 lb_test = lb_test,
                 train_data = train_data,
                 test_data = test_data,
                 match_methods = "euclidean",
                 match_num = 10,
                 gamlss_formula = "ht ~ cs(time, df = 3)",
                 gamsigma_formula = "~ cs(time, df = 1)",
                 match_plot = TRUE,
                 predict_plot = TRUE,
                 sbj = id)
```

```
#>
#>  using euclidean distance
```

```
#> Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
#> i Please use ‘linewidth‘ instead.
#> This warning is displayed once every 8 hours.
#> Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
#> generated.
```

```
#>
#>  plotting matching paired individual trajectories
#> GAMLSS-RS iteration 1: Global Deviance = 1860.635
#> GAMLSS-RS iteration 2: Global Deviance = 1859.725
#> GAMLSS-RS iteration 3: Global Deviance = 1859.769
#> GAMLSS-RS iteration 4: Global Deviance = 1859.769
#>
#>  gamlss model fitting is done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for observed time points are done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for predicted time points are done
```

```
test_mhl_n10 <- pred_matching(
                   lb_data = lb_train,
                   lb_test = lb_test,
                   train_data = train_data,
                   test_data = test_data,
                   match_methods = "mahalanobis",
                   match_num = 10,
                   gamlss_formula = "ht ~ cs(time, df = 3)",
                   gamsigma_formula = "~ cs(time, df = 1)",
                 match_plot = TRUE,
                 predict_plot = TRUE,
                   sbj = id)
```

```
#>
#>  using mahalanobis distance with matching number
#>
#>  plotting matching paired individual trajectories
#> GAMLSS-RS iteration 1: Global Deviance = 1860.635
#> GAMLSS-RS iteration 2: Global Deviance = 1859.725
#> GAMLSS-RS iteration 3: Global Deviance = 1859.769
#> GAMLSS-RS iteration 4: Global Deviance = 1859.769
#>
#>  gamlss model fitting is done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for observed time points are done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for predicted time points are done
```

```
## mahalanobis p value matching
test_mhl_p90 <- try(pred_matching(
                 lb_data = lb_train,
                 lb_test = lb_test,
                 train_data = train_data,
                 test_data = test_data,
                 match_methods = "mahalanobis",
                 match_alpha = 0.95,
                 gamlss_formula = "ht ~ cs(time, df = 3)",
                 gamsigma_formula = "~ cs(time, df = 1)",
                 match_plot = TRUE,
                 predict_plot = TRUE,
                 sbj = id))
```

```
#>
#>  using mahalanobis distance with F test p value
#>
#>  plotting matching paired individual trajectories
#> GAMLSS-RS iteration 1: Global Deviance = 38822.62
#> GAMLSS-RS iteration 2: Global Deviance = 38868.58
#> GAMLSS-RS iteration 3: Global Deviance = 38875.06
#> GAMLSS-RS iteration 4: Global Deviance = 38876.06
#> GAMLSS-RS iteration 5: Global Deviance = 38876.24
#> GAMLSS-RS iteration 6: Global Deviance = 38876.28
#> GAMLSS-RS iteration 7: Global Deviance = 38876.28
#> GAMLSS-RS iteration 8: Global Deviance = 38876.29
#> GAMLSS-RS iteration 9: Global Deviance = 38876.29
#>
#>  gamlss model fitting is done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for observed time points are done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for predicted time points are done
```

```r
## mahalanobis p value matching
test_mhl_p80 <- try(pred_matching(
                 lb_data = lb_train,
                 lb_test = lb_test,
                 train_data = train_data,
                 test_data = test_data,
                 match_methods = "mahalanobis",
                 match_alpha = 0.8,
                 gamlss_formula = "ht ~ cs(time, df = 3)",
                 gamsigma_formula = "~ cs(time, df = 1)",
                 match_plot = TRUE,
                 predict_plot = TRUE,
                 sbj = id))
```

```
#>
#>  using mahalanobis distance with F test p value
#>
#>  plotting matching paired individual trajectories
#> GAMLSS-RS iteration 1: Global Deviance = 72869.76
#> GAMLSS-RS iteration 2: Global Deviance = 72903.07
#> GAMLSS-RS iteration 3: Global Deviance = 72906.36
#> GAMLSS-RS iteration 4: Global Deviance = 72906.79
#> GAMLSS-RS iteration 5: Global Deviance = 72906.86
#> GAMLSS-RS iteration 6: Global Deviance = 72906.87
#> GAMLSS-RS iteration 7: Global Deviance = 72906.88
#> GAMLSS-RS iteration 8: Global Deviance = 72906.88
#> GAMLSS-RS iteration 9: Global Deviance = 72906.88
```

```
#>
#>  gamlss model fitting is done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for observed time points are done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for predicted time points are done
```

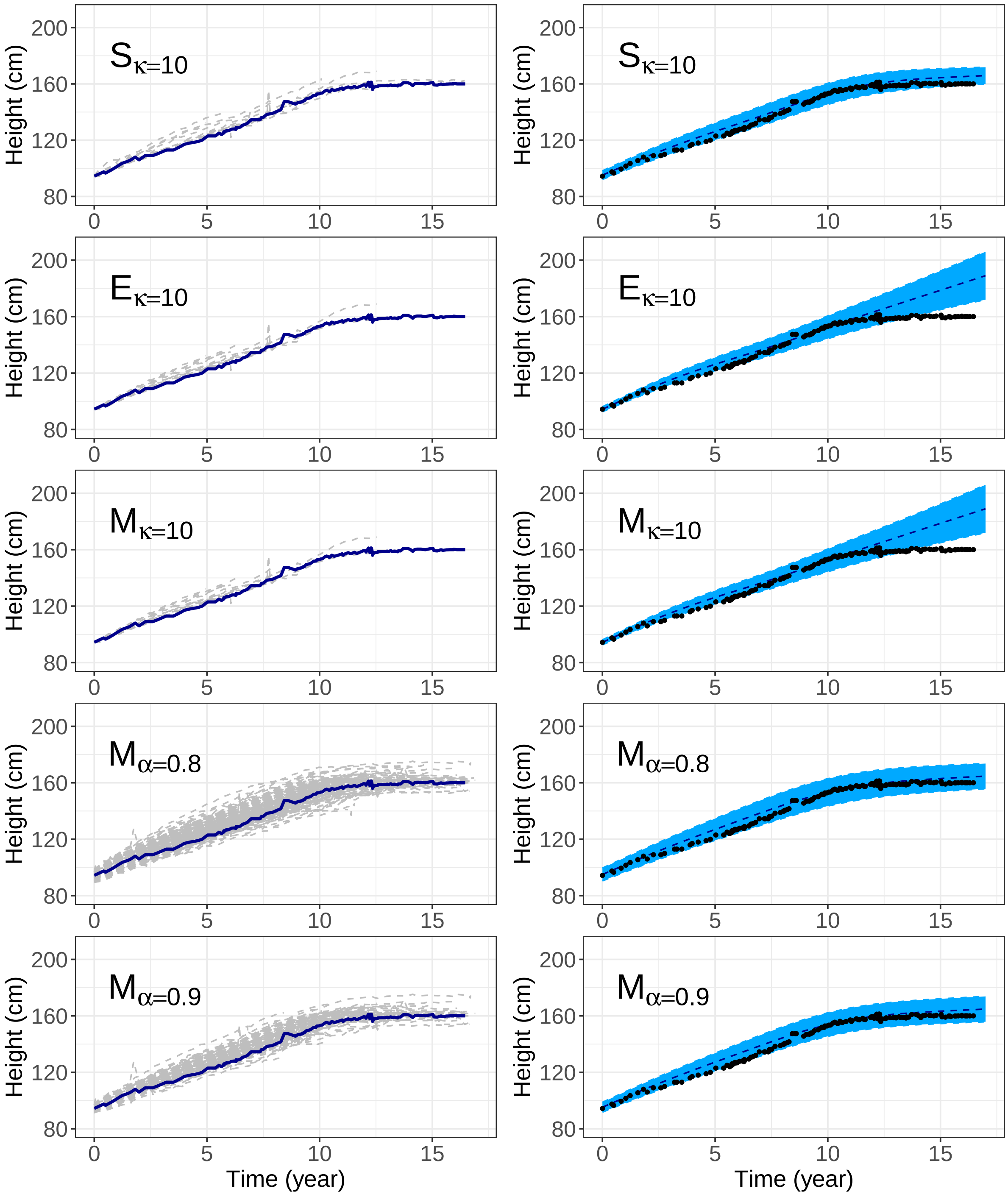
```r
test_sgl10_n10 <- pred_matching(
                     lb_data = lb_train,
                     lb_test = lb_test,
                     train_data = train_data,
                     test_data = test_data,
                     match_methods = "single",
                     match_num = 10,
                     match_time = 12,
                     match_alpha = NULL,
                     gamlss_formula = "ht ~ cs(time, df = 3)",
                     gamsigma_formula = "~ cs(time, df = 1)",
                 match_plot = TRUE,
                 predict_plot = TRUE,
                     sbj = id)
```

```
#>
#>  using single critical time point matching
#>
#>  plotting matching paired individual trajectories
#> GAMLSS-RS iteration 1: Global Deviance = 2548.712
#> GAMLSS-RS iteration 2: Global Deviance = 2553
#> GAMLSS-RS iteration 3: Global Deviance = 2554.012
#> GAMLSS-RS iteration 4: Global Deviance = 2554.432
#> GAMLSS-RS iteration 5: Global Deviance = 2554.623
#> GAMLSS-RS iteration 6: Global Deviance = 2554.727
#> GAMLSS-RS iteration 7: Global Deviance = 2554.758
#> GAMLSS-RS iteration 8: Global Deviance = 2554.791
#> GAMLSS-RS iteration 9: Global Deviance = 2554.811
#> GAMLSS-RS iteration 10: Global Deviance = 2554.823
#> GAMLSS-RS iteration 11: Global Deviance = 2554.831
#> GAMLSS-RS iteration 12: Global Deviance = 2554.836
#> GAMLSS-RS iteration 13: Global Deviance = 2554.838
#> GAMLSS-RS iteration 14: Global Deviance = 2554.84
#> GAMLSS-RS iteration 15: Global Deviance = 2554.841
#> GAMLSS-RS iteration 16: Global Deviance = 2554.842
#>
#>  gamlss model fitting is done
#> new prediction
#> new prediction
#>
#>  gamlss model prediction for observed time points are done
#> new prediction
#> new prediction
```

```
#>
#>  gamlss model prediction for predicted time points are done
```

```r
sgl10_n10_m <- test_sgl10_n10$matching_trajectory +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$S_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))


eld_n10_m <- test_eld_n10$matching_trajectory +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$E_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))



mhl_n10_m <- test_mhl_n10$matching_trajectory +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))



mhl_p90_m <- test_mhl_p90$matching_trajectory +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\alpha = 0.9}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))

mhl_p80_m <- test_mhl_p80$matching_trajectory +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\alpha = 0.8}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))

sgl10_n10_t <- test_sgl10_n10$predictive_centiles +
  ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$S_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))


eld_n10_t <- test_eld_n10$predictive_centiles +
    ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$E_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))


mhl_n10_t <- test_mhl_n10$predictive_centiles +
      ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\kappa = 10}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))


mhl_p90_t <- test_mhl_p90$predictive_centiles +
      ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\alpha = 0.90}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))

mhl_p80_t <- test_mhl_p80$predictive_centiles +
      ylim(c(80, 210)) +
  labs(title = parse(text = latex2exp::TeX("$M_{\\alpha = 0.80}$")),
       y = "Height (cm)",
       x = "Time (year)") +
  theme(axis.text.x = element_text(size = 15),
        axis.text.y = element_text(size = 15),
        title = element_text(size = 20),
        axis.title.x = element_text(size = 16),
        axis.title.y = element_text(size = 16)) +
  theme(plot.margin=unit(c(-1.7, 0.2, 0, 0.2), 'cm'),
        plot.title = element_text(hjust = 0.1, vjust = -10))
```

```
library(ggpubr)

figure1 <- ggarrange(sgl10_n10_m, sgl10_n10_t,
                    # + theme(axis.text.x=element_blank()),
                    eld_n10_m, eld_n10_t,
                    # + theme(axis.text.x=element_blank()),
                    mhl_n10_m, mhl_n10_t,
                    # + theme(axis.text.x=element_blank()),
                    mhl_p80_m, mhl_p80_t,
                    mhl_p90_m, mhl_p90_t,
                    # + theme(axis.text.x=element_blank()),
                    ncol = 2, nrow = 5) +
  theme(plot.margin = margin(2, 0.5, 0.5, 0.5, "cm"))


figure1
```



```r
ggsave(filename = paste0("figure/S05_trajectory_final_plots_id_", id, "_", Sys.Date(), "_2alpha.png"))
```

‘‘‘r

```
sessionInfo()
```

```
#> R version 4.2.2 (2022-10-31)
#> Platform: aarch64-apple-darwin20 (64-bit)
#> Running under: macOS 14.0
#> 
#> Matrix products: default
#> BLAS:   /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRblas.0.dylib
#> LAPACK: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRlapack.dylib
#> 
#> locale:
#> [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
#> 
#> attached base packages:
#> [1] parallel  splines   stats     graphics  grDevices utils     datasets 
#> [8] methods   base     
#> 
#> other attached packages:
#>  [1] ggpubr_0.6.0        JMbayes_0.8-85      rstan_2.21.8       
#>  [4] StanHeaders_2.26.27 doParallel_1.0.17   iterators_1.0.14   
#>  [7] foreach_1.5.2       survival_3.5-5      gamlss_5.1-7       
#> [10] gamlss.dist_6.0-5   MASS_7.3-60         gamlss.data_6.0-2  
#> [13] nlme_3.1-162        flextable_0.9.2     gtsummary_1.7.1    
#> [16] lubridate_1.9.2     forcats_1.0.0       stringr_1.5.0      
#> [19] dplyr_1.1.2         purrr_1.0.1         readr_2.1.4        
#> [22] tidyr_1.3.0         tibble_3.2.1        ggplot2_3.4.3      
#> [25] tidyverse_2.0.0     here_1.0.1         
#> 
#> loaded via a namespace (and not attached):
#>   [1] uuid_1.1-0              backports_1.4.1         Hmisc_5.1-0            
#>   [4] systemfonts_1.0.4       inline_0.3.19           digest_0.6.33          
#>   [7] htmltools_0.5.5         fansi_1.0.4             magrittr_2.0.3         
#>  [10] checkmate_2.2.0         cluster_2.1.4           tzdb_0.4.0             
#>  [13] RcppParallel_5.1.7      matrixStats_1.0.0       officer_0.6.2          
#>  [16] askpass_1.1             timechange_0.2.0        gfonts_0.2.0           
#>  [19] prettyunits_1.1.1       colorspace_2.1-0        textshaping_0.3.6      
#>  [22] xfun_0.39               callr_3.7.3             crayon_1.5.2           
#>  [25] jsonlite_1.8.7          lme4_1.1-34             glue_1.6.2             
#>  [28] gtable_0.3.3            brokenstick_2.5.0       car_3.1-2              
#>  [31] pkgbuild_1.4.2          abind_1.4-5             scales_1.2.1           
#>  [34] fontquiver_0.2.1        DBI_1.1.3               rstatix_0.7.2          
#>  [37] Rcpp_1.0.11             xtable_1.8-4            htmlTable_2.4.1        
#>  [40] foreign_0.8-84          latex2exp_0.9.6         freshr_1.0.2           
#>  [43] Formula_1.2-5           stats4_4.2.2            fontLiberation_0.1.0   
#>  [46] htmlwidgets_1.6.2       ellipsis_0.3.2          farver_2.1.1           
#>  [49] pkgconfig_2.0.3         loo_2.6.0               keep_1.0               
#>  [52] nnet_7.3-19             utf8_1.2.3              crul_1.4.0             
#>  [55] labeling_0.4.2          tidyselect_1.2.0        rlang_1.1.1            
#>  [58] later_1.3.1             munsell_0.5.0           tools_4.2.2            
#>  [61] cli_3.6.1               generics_0.1.3          jagsUI_1.5.2           
#>  [64] broom_1.0.5             evaluate_0.21           fastmap_1.1.1          
#>  [67] yaml_2.3.7              ragg_1.2.5              processx_3.8.2         
#>  [70] knitr_1.43              zip_2.3.0               mime_0.12              
#>  [73] xml2_1.3.5              compiler_4.2.2          rstudioapi_0.15.0      
#>  [76] curl_5.0.1              ggsignif_0.6.4          gt_0.9.0
#>  [79] broom.helpers_1.13.0    stringi_1.7.12          highr_0.10
#>  [82] ps_1.7.5                gdtools_0.3.3           lattice_0.21-8
#>  [85] Matrix_1.5-3            fontBitstreamVera_0.1.1 nloptr_2.0.3
#>  [88] vctrs_0.6.3             pillar_1.9.0            lifecycle_1.0.3
#>  [91] cowplot_1.1.1           data.table_1.14.8       httpuv_1.6.11
#>  [94] R6_2.5.1                promises_1.2.0.1        gridExtra_2.3
#>  [97] matrixsampling_2.0.0    rjags_4-14              codetools_0.2-19
#> [100] boot_1.3-28.1           openssl_2.1.0           rprojroot_2.0.3
#> [103] withr_2.5.0             httpcode_0.3.0          hms_1.1.3
#> [106] grid_4.2.2              rpart_4.1.19            coda_0.19-4
#> [109] minqa_1.2.5             rmarkdown_2.23          carData_3.0-5
#> [112] shiny_1.7.4.1           base64enc_0.1-3
```