Prognostic of NASA Turbofan Jet Engine

Fernando Chafim

05/01/2021

Contribuciones	Firma
Investigación previa	Fernando Chafim
Redacción de las respuestas	Fernando Chafim
Desarrollo código	Fernando Chafim

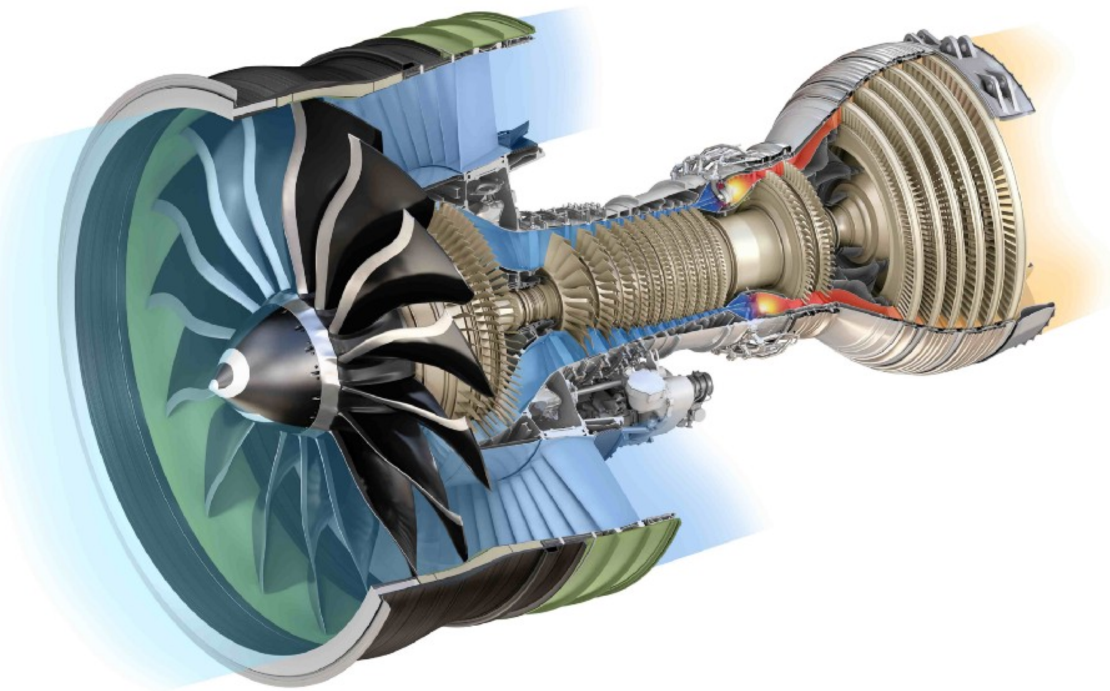
NASA TURBOFAN JET ENGINE

We define prognostics here exclusively as the estimation of remaining useful component life. The remaining useful life (RUL) estimates are in units of time (e.g., hours or cycles).

End-of-life can be subjectively determined as a function of operational thresholds that can be measured. These thresholds depend on user specifications to determine safe operational limits.

Prognostics is currently at the core of systems health management. Reliably estimating remaining life holds the promise for considerable cost savings (for example by avoiding unscheduled maintenance and by increasing equipment usage) and operational safety improvements. Remaining life estimates provide decision makers with information that allows them to change operational characteristics (such as load) which in turn may prolong the life of the component. It also allows planners to account for upcoming maintenance and set in motion a logistics process that supports a smooth transition from faulty equipment to fully functional.

```
knitr::include_graphics("img/turbofan1.jpeg")
```



2. Dataset

Los pronósticos y la gestión de la salud son un tema importante en la industria para predecir el estado de los activos para evitar tiempos de inactividad y fallas. Este conjunto de datos es la versión de Kaggle del muy conocido conjunto de datos públicos para el modelado de degradación de activos de la NASA. Incluye datos simulados Run-to-Failure de motores a reacción con turboventilador.

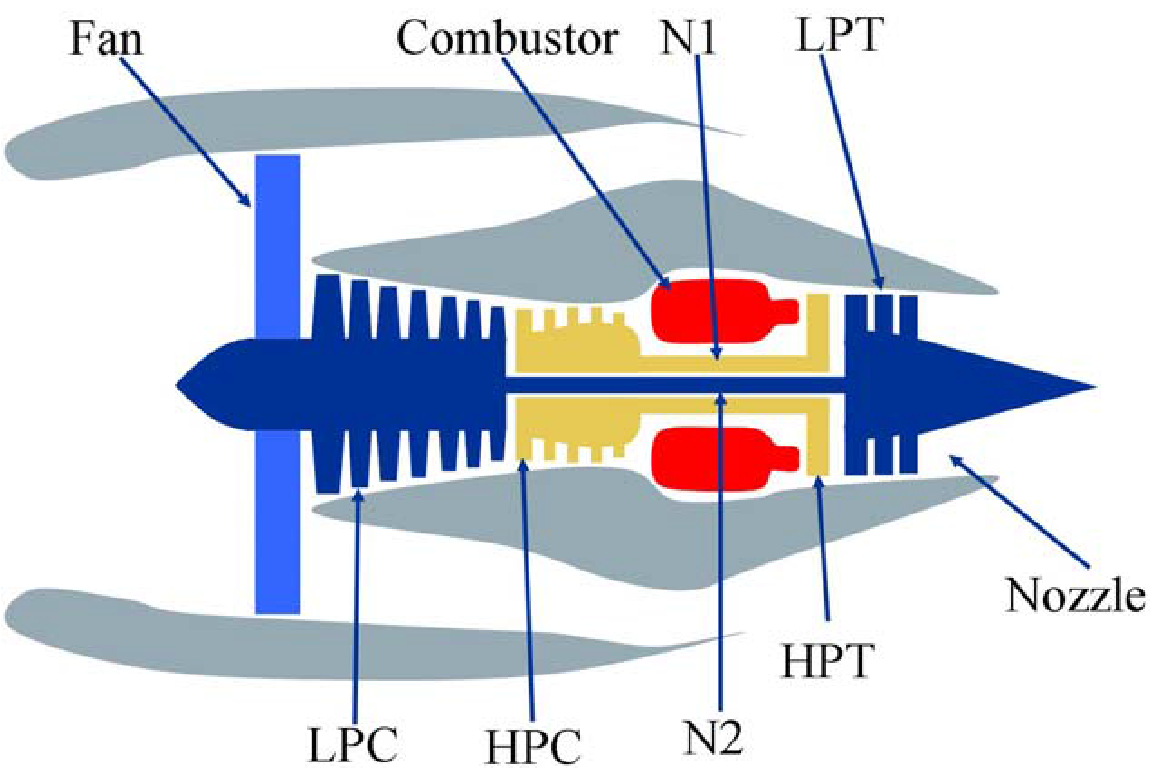
La simulación de la degradación del motor se llevó a cabo utilizando C-MAPSS. Se simularon cuatro conjuntos diferentes bajo diferentes combinaciones de condiciones operativas y modos de falla. Registra varios canales de sensores para caracterizar la evolución de la falla. El conjunto de datos fue proporcionado por el CoE de Pronósticos de NASA Ames.

Descripción

Symbol	Description	Units
Unit	-	-
Time	-	t
Altitude	Altitude	ft
Mach_N	Mach Number	M
SeaTemp	Sea Level temperature	°F
T2	Total temperature at **fan** inlet	°R
T2	Total temperature at **fan** inlet	°R
T24	Total temperature at **LPC** outlet	°R
T30	Total temperature at **HPC** outlet	°R
T50	Total temperature at **LPT** outlet	°R
P2	Pressure at **fan** inlet	psia
P15	Total pressure in **bypass-duct**	psia
P30	Total pressure at **HPC** outlet	psia

Symbol	Description	Units
Nf	Physical **fan** speed	rpm
Nc	Physical **core** speed	rpm
epr	Engine pressure ratio (P50/P2)	–
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	–
farB	Burner fuel-air ratio	–
htBleed	Bleed Enthalpy	–
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s
T48	(EGT) Total temperature at HPT outlet	°R
SmFan	Fan stall margin	–
SmLPC	LPC stall margin	–
SmHPC	HPC stall margin	–

```
knitr::include_graphics("img/turbofan2.png")
```



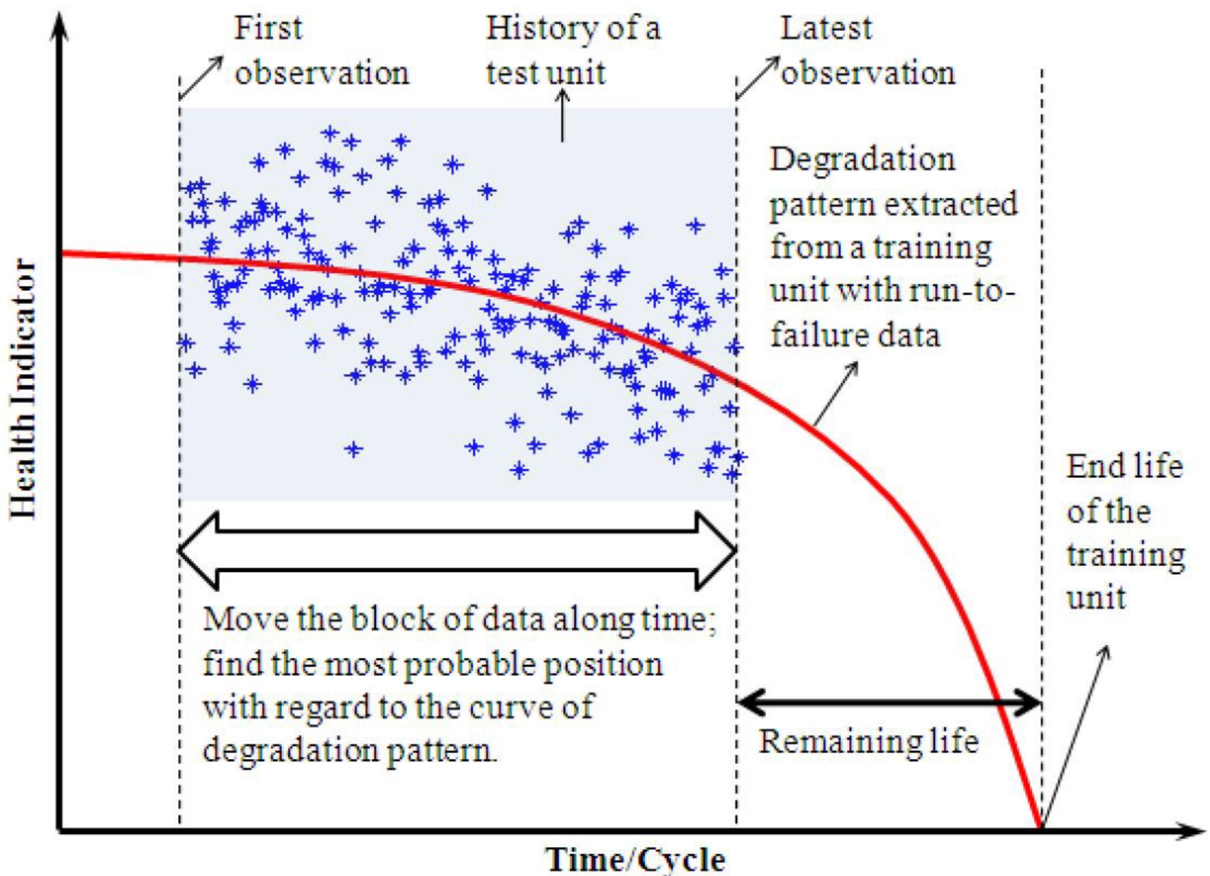
Importancia y objetivos de los análisis

Los pronósticos y la gestión de la salud son un tema importante en la industria para predecir el estado de los activos y evitar tiempos de inactividad y fallas. Este conjunto de datos es la versión de Kaggle del conjunto de datos públicos para el modelado de degradación de activos de la NASA. Incluye datos simulados Run-to-Failure de motores a reacción con turboventilador.

La simulación de la degradación del motor se llevó a cabo utilizando C-MAPSS. Se simularon cuatro conjuntos diferentes bajo diferentes combinaciones de condiciones operativas y modos de fallo. Registra varios canales de sensores para caracterizar la evolución de la falla. El conjunto de datos fue proporcionado por Prognostics CoE en NASA Ames.

En este conjunto de datos, el objetivo es predecir la **vida útil restante (RUL)** de cada motor. La vida útil restante (RUL) es el período de tiempo que es probable que funcione una máquina antes de que requiera reparación o reemplazo. Al tener en cuenta RUL, los ingenieros pueden programar el mantenimiento, optimizar la eficiencia operativa y evitar tiempos de inactividad no planificados. Por esta razón, estimar el RUL es una prioridad máxima en los programas de mantenimiento predictivo.

```
knitr::include_graphics("img/turbofan3.jpeg")
```



2. Integración y selección de los datos de interés a analizar.

Integración

La integración o fusión de los datos consiste en la combinación de datos procedentes de múltiples fuentes, con el fin de crear una estructura de datos coherente y única que contenga mayor cantidad de información.

```r
library(dplyr)
library(e1071)
library(zeallot)
library(stringr)
library(data.table)
library(DT)
library(DataExplorer)
library(ggcorrplot)
library(plotly)
library(anomalize)
library(zoo)
library(tibbletime)
library(h2o)
library(isofor)
library(boot)
library(mice)
library(ggpubr)
library(MASS)
library(TTR)
library(caret)
library(FNN)
library(dbscan)
library(outliers)
library(car)
```

```r
trainset <- fread("D:/UOC/Tipología y ciclo de vida de los datos/PRA2/CMaps/train_FD001.txt")
names(trainset) <- c('unit_number', 'time_in_cycles', 'altitude', 'mach_number', 'TRA', 'T2', 'T24', 'T3

testset <- fread("D:/UOC/Tipología y ciclo de vida de los datos/PRA2/CMaps/test_FD001.txt")
names(testset) <- c('unit_number', 'time_in_cycles', 'altitude', 'mach_number', 'TRA', 'T2', 'T24', 'T30

y_testset <- fread("D:/UOC/Tipología y ciclo de vida de los datos/PRA2/CMaps/RUL_FD001.txt") %>% pull()


read_logs <- function(path = "D:/UOC/Tipología y ciclo de vida de los datos/PRA2/CMaps"){

  files <- c("FD001.txt", "FD002.txt", "FD003.txt", "FD004.txt")
  sets <- c("train", "test")
  files <- apply(expand.grid(sets, files), 1, paste, collapse="_")
  f <- file.path(path, files)
  d <- lapply(f, fread)
  names(d) <- str_remove(files, ".txt")

  columns = c('unit_number','time_in_cycles','altitude','mach_number','TRA','T2','T24','T30','T50','P2'
           'Nc','epr','Ps30','phi','NRf','NRc','BPR','farB','htBleed','Nf_dmd','PCNfR_dmd','W31','W32')

  temp_func = function(df){
    names(df) <- columns
    return(df)
    }

  d <- lapply(d, temp_func)
```

```
  return(d)
}

#dfs <- read_logs()

#trainset <- bind_rows(
#  dfs$train_FD001#,
  #dfs$train_FD002,
  #dfs$train_FD003,
  #dfs$train_FD004
#  )

#testset <- bind_rows(
#  dfs$test_FD001#,
  #dfs$test_FD002,
  #dfs$test_FD003,
  #dfs$test_FD004
#  )

#trainset$set <- "train"
#testset$set <- "test"

df <- trainset
```

Selección

Features Selection 1

```
FindOutliers <- function(x) {
  if(class(x)%in% c("numeric", "integer")){
    lowerq = quantile(x, na.rm = TRUE)[2]
    upperq = quantile(x, na.rm = TRUE)[4]
    iqr = upperq - lowerq #Or use IQR(data)
    # we identify extreme outliers
    extreme.threshold.upper = (iqr * 3) + upperq
    extreme.threshold.lower = lowerq - (iqr * 3)
    result <- which(x > extreme.threshold.upper | x < extreme.threshold.lower)
    output <- length(result)
  } else {
    output <- 0
    }

  return(output)
}

DataProfiling <- function(df){

  column_names <- colnames(df)
  column_classes <- as.vector(sapply(df, function(x) class(x)))
  column_mean <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){mean(x, na.r
  column_sd <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){sd(x, na.rm = '
  column_median <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){median(x, n
  column_max <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){max(x, na.rm =
```

```r
  column_min <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){min(x, na.rm =
  column_nunique <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){length(un
  column_quantile0   <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){quant
  column_quantile25  <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){quant
  column_quantile50  <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){quant
  column_quantile75  <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){quant
  column_quantile100 <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){quant
  column_interquartile_range <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer"
  column_skewness <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){skewness
  column_kurtosis <- as.vector(sapply(df, function(x) if(class(x)%in% c("numeric", "integer")){kurtosis
  column_na <- as.vector(sapply(df, function(x) sum(is.na(x))))
  column_zero <- as.vector(sapply(df, function(x) sum(x==0)))
  column_outliers <- as.vector(sapply(df, FindOutliers))

  df_table <- tibble(
    names = column_names,
    classes = column_classes,
    min = column_min,
    max = column_max,
    n_unique = column_nunique,
    mean = column_mean,
    sd = column_sd,
    median = column_median,
    quantile_0 = column_quantile0,
    quantile_25 = column_quantile25,
    quantile_50 = column_quantile50,
    quantile_75 = column_quantile75,
    quantile_100 = column_quantile100,
    interquantile_range = column_interquartile_range,
    skewness = column_skewness,
    kurtosis = column_kurtosis,
    MissingValues = column_na,
    n_zero = column_zero,
    Outliers = column_outliers
    )

  return(df_table)

}

temp <- DataProfiling(df)
```

```r
unique_columns <- temp$names[temp$n_unique == 1 & !is.na(temp$n_unique)]
print(unique_columns)
```

```
## [1] "TRA"       "T2"        "P2"        "epr"       "farB"      "Nf_dmd"
## [7] "PCNfR_dmd"
```

```r
df <- df %>% dplyr::select(-one_of(unique_columns))
trainset <- trainset %>% dplyr::select(-one_of(unique_columns))
testset <- testset %>% dplyr::select(-one_of(unique_columns))
```

Preprocesado de los datos

Feature Engineering 1

Remaining Useful Life

```r
add_remaining_useful_life <- function(df){

  df <- df %>%
    group_by(unit_number) %>%
    mutate(max_time_in_cycles = max(time_in_cycles)) %>%
    ungroup() %>%
    mutate(
      RUL = max_time_in_cycles - time_in_cycles
      ) %>%
    dplyr::select(-max_time_in_cycles)

  return(df)
}
```

Cumulative terms

```r
add_cumulative_features <- function(df, columns){
  df <- df %>%
     as_tibble() %>%
     group_by(unit_number) %>%
 mutate(
     across(
       all_of(columns),
       list(
         cumsum = cumsum,
         cummin = cummin,
         cummax = cummax
         ),
       .names = "{.fn}_{.col}"
       )
     )
  return(df)
}
```

Lag terms

```r
add_lag <- function(df, columns, num = 1){

  for (i in 1:length(num)) {
    num_temp <- num[i]
    df <- df %>%
      as_tibble() %>%
      group_by(unit_number) %>%
      mutate(
        across(all_of(columns), ~lag(.x, n = num_temp, default = NA), .names = paste0("lag", num_temp,
```

```
        )
  }
  return(df)
}
```

Moving Average terms

```
add_moving_avg_features <- function(df, columns, num = 5){

  for (i in 1:length(num)) {
    num_temp <- num[i]
    df <- df %>%
      as_tibble() %>%
      group_by(unit_number) %>%
      mutate(
        across(
          all_of(columns),
          list(
         rollmean = ~rollapplyr(.x, num_temp, FUN = mean, fill=NA),
         rollsd = ~rollapplyr(.x, num_temp, FUN = sd, fill=NA),
         rolliqr = ~rollapplyr(.x, num_temp, FUN = IQR, fill=NA)
         ),
         .names = paste0("{.fn}", num_temp, "_{.col}")
          )
        )
  }
  return(df)
}
```

Interaction terms

```
add_interaction_terms <- function(df, columns){
  df_1 <- df %>% dplyr::select(all_of(columns))
  df_2 <- df %>% dplyr::select(!all_of(columns))
  df_interaction <- as.data.frame(model.matrix(~ .^2-1,df_1))
  df_output <- bind_cols(df_2, df_interaction)
  names(df_output) <- gsub(x = names(df_output), pattern = ":", replacement = "_")
  return(df_output)
}
```

Todos juntos

```
columns <- c("altitude", "mach_number","T24", "T30", "T50", "P15", "P30", "Nf", "Nc", "Ps30", "phi", "N

trainset2 <- trainset %>%
  add_remaining_useful_life() %>%
  add_interaction_terms(columns) %>%
  add_lag(columns=columns, num = 1:5) %>%
  add_moving_avg_features(columns = columns, num = c(3, 5)) %>%
  add_cumulative_features(columns=columns)
```

```r
y_trainset <- trainset2$RUL

testset2 <- testset %>%
  add_interaction_terms(columns) %>%
  add_lag(columns=columns, num = 1:5) %>%
  add_moving_avg_features(columns = columns, num = c(3, 5)) %>%
  add_cumulative_features(columns=columns)
```

Benchmark

```r
bechmark_performance <- function(x_train, y_train, x_test, y_test){

  df <- bind_cols(x_train, y= y_train)
  model_bench <- lm(y~., data=df)

  x_test <- x_test %>%
            group_by(unit_number) %>%
            filter(time_in_cycles == max(time_in_cycles))

  suppressWarnings({
    y_hat <- predict.lm(model_bench, x_test)
  })
  result <- postResample(y_hat, y_test)

  return(result)
}

bechmark_performance(x_train = trainset,
                     y_train = y_trainset,
                     x_test = testset,
                     y_test = y_testset)
```

```
##       RMSE   Rsquared        MAE
## 31.7404750  0.6525414 26.1234550
```

Análisis estadístico descriptivo

```r
temp <- DataProfiling(trainset2)

#datatable(temp, class = 'cell-border stripe', rownames = FALSE, filter = 'top',
#          options = list(
#            pageLength = 10,
#            autoWidth = TRUE,
#            scrollX = TRUE,
#            scrollY = TRUE
#            )
#          )
knitr::kable(temp, "pipe")
```

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	interquartile	skewness	kurtosis	Missing	zeros	outliers
unit_number	integer	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
time_in_cycles	integer	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
RUL	integer	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	100	0
altitude	numeric	-0.0087000	[illegible]	[illegible]	-0.0000089	[illegible]	[illegible]	-0.0087000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
mach_number	numeric	-0.0006000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0006000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	2070	0
T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0
T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-6.9158072	45.830611	0	0	406
P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	6
Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	805
Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1727151	0	0	0
phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	5
NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	689
BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1219704	0	0	0
htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0397424	0	0	0
W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
altitude_mach_number	numeric	-0.0000042	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0000042	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2459	282
altitude_T24	numeric	-5.5897500	[illegible]	[illegible]	-0.0056908	[illegible]	[illegible]	-5.5897500	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_T30	numeric	-13.7888100	[illegible]	[illegible]	-0.0141840	[illegible]	[illegible]	-13.7888100	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_T50	numeric	-12.2421860	[illegible]	[illegible]	-0.0123096	[illegible]	[illegible]	-12.2421860	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_P15	numeric	-0.1880070	[illegible]	[illegible]	-0.0001917	[illegible]	[illegible]	-0.1880070	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_P30	numeric	-4.8196260	[illegible]	[illegible]	-0.0049267	[illegible]	[illegible]	-4.8196260	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_Nf	numeric	-20.7762960	[illegible]	[illegible]	-0.0211828	[illegible]	[illegible]	-20.7762960	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_Nc	numeric	-78.8083410	[illegible]	[illegible]	-0.0806167	[illegible]	[illegible]	-78.8083410	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	Missing	n_Zeros	n_outliers
altitude_Ps30	numeric	-0.4114230	[illegible]	[illegible]	-0.0004149	[illegible]	[illegible]	-0.4114230	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_phi	numeric	-4.5464460	[illegible]	[illegible]	-0.0046274	[illegible]	[illegible]	-4.5464460	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_NRf	numeric	-20.7759480	[illegible]	[illegible]	-0.0211824	[illegible]	[illegible]	-20.7759480	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_NRc	numeric	-70.8311370	[illegible]	[illegible]	-0.0724228	[illegible]	[illegible]	-70.8311370	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_BPR	numeric	-0.0732209	[illegible]	[illegible]	-0.0000743	[illegible]	[illegible]	-0.0732209	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_htBleed	numeric	-3.4017000	[illegible]	[illegible]	-0.0034790	[illegible]	[illegible]	-3.4017000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_W31	numeric	-0.3404310	[illegible]	[illegible]	-0.0003466	[illegible]	[illegible]	-0.3404310	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
altitude_W32	numeric	-0.2039924	[illegible]	[illegible]	-0.0002100	[illegible]	[illegible]	-0.2039924	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	413	0
mach_number_T24	numeric	-0.3860820	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.3860820	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306335	0	2070	0
mach_number_T30	numeric	-0.9604440	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.9604440	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306194	0	2070	0
mach_number_T50	numeric	-0.8541240	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.8541240	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1305410	0	2070	0
mach_number_P15	numeric	-0.0129660	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0129660	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306454	0	2070	0
mach_number_P30	numeric	-0.3329220	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.3329220	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1305812	0	2070	0
mach_number_Nf	numeric	-1.4328840	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1.4328840	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306448	0	2070	0
mach_number_Nc	numeric	-5.4634800	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-5.4634800	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1307913	0	2070	0
mach_number_Ps30	numeric	-0.0287880	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0287880	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1307607	0	2070	0
mach_number_phi	numeric	-0.3134820	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.3134820	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306114	0	2070	0
mach_number_NRf	numeric	-1.4328840	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1.4328840	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306453	0	2070	0
mach_number_NRc	numeric	-4.9032600	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-4.9032600	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1308068	0	2070	0
mach_number_BPR	numeric	-0.0051035	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0051035	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306678	0	2070	0
mach_number_htBleed	numeric	-0.2376000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2376000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1306943	0	2070	0
mach_number_W31	numeric	-0.0234600	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0234600	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1303323	0	2070	0
mach_number_W32	numeric	-0.0141020	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0141020	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1302957	0	2070	0
T24_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0141206	0	0	0
T24_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1726714	0	0	0
T24_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0985438	0	0	0

names	classes	min … skewness, kurtosis	nMissing	nZeros	nOutliers
T24_P[illegible]	numeric	[illegible] - 0.1662027 0.6237733	0	0	0
T24_N[illegible]	numeric	[illegible] 0.1187944	0	0	0
T24_N[illegible]	numeric	[illegible]	1	0	806
T24_P[illegible]	numeric	[illegible] 0.1793184	0	0	0
T24_ph[illegible]	numeric	[illegible] - 0.1782 0.0335335	0	0	0
T24_NiR[illegible]	numeric	[illegible] 0.1196789	0	0	0
T24_NiR[illegible]	numeric	[illegible]	8	0	808
T24_BiR[illegible]	numeric	[illegible] 0.1314427	0	0	0
T24_ht[illegible]	numeric	[illegible] 0.0474604	0	0	0
T24_W[illegible]	numeric	[illegible] - 0.2956 0.9727637	0	0	0
T24_W[illegible]	numeric	[illegible] - 0.2928 0.3070770	0	0	0
T30_T[illegible]	numeric	[illegible] 0.1303795	0	0	0
T30_P[illegible]	numeric	[illegible]	3	0	0
T30_P[illegible]	numeric	[illegible]	9	0	0
T30_N[illegible]	numeric	[illegible]	2	0	0
T30_N[illegible]	numeric	[illegible]	3	0	124
T30_P[illegible]	numeric	[illegible] 0.1330004	0	0	0
T30_ph[illegible]	numeric	[illegible]	7	0	0
T30_NiR[illegible]	numeric	[illegible]	8	0	0
T30_NiR[illegible]	numeric	[illegible]	2	0	146
T30_BiR[illegible]	numeric	[illegible] 0.0878091	0	0	0
T30_ht[illegible]	numeric	[illegible] 0.0348660	0	0	0
T30_W[illegible]	numeric	[illegible] 0.0333409	0	0	0
T30_W[illegible]	numeric	[illegible] 0.0123499	0	0	0
T50_P[illegible]	numeric	[illegible] 0.1670909	0	0	0
T50_P[illegible]	numeric	[illegible] 0.1247147	0	0	0
T50_N[illegible]	numeric	[illegible] 0.1651892	0	0	0
T50_N[illegible]	numeric	[illegible]	9	0	16
T50_P[illegible]	numeric	[illegible] 0.1845018	0	0	0
T50_ph[illegible]	numeric	[illegible] 0.1189334	0	0	0
T50_NiR[illegible]	numeric	[illegible] 0.1650302	0	0	0

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	Missing	n_Zeros	Outliers
T50_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	18
T50_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1532927	0	0	0
T50_htBleed	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1415006	0	0	0
T50_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0262919	0	0	0
T50_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0558001	0	0	0
P15_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P15_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	5.3061050	32.24191[illegible]	[illegible]	0	410
P15_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	806
P15_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1754078	0	0	0
P15_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P15_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	5.2854979	32.08314[illegible]	[illegible]	0	410
P15_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	690
P15_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1236243	0	0	0
P15_htBleed	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0435012	0	0	0
P15_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P15_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P30_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P30_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	105
P30_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1206391	0	0	0
P30_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P30_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P30_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	51
P30_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0507883	0	0	0
P30_htBleed	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	1
P30_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
P30_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
Nf_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	816
Nf_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1737205	0	0	0
Nf_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
Nf_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	6

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	nMissing	n_Zeros	Outliers
Nf_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	704
Nf_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1234488	[illegible]	0	0	0
Nf_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.0425408	[illegible]	0	0	0
Nf_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
Nf_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
Nc_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	23
Nc_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	131
Nc_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	814
Nc_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	800
Nc_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	56
Nc_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	119
Nc_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.3580910	[illegible]	0	0	0
Nc_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.3584811	[illegible]	0	0	0
Ps30_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1262953	[illegible]	0	0	0
Ps30_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1736164	[illegible]	0	0	0
Ps30_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	29
Ps30_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1637307	[illegible]	0	0	0
Ps30_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1489774	[illegible]	0	0	0
Ps30_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.0483855	[illegible]	0	0	0
Ps30_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
phi_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
phi_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	78
phi_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.0502864	[illegible]	0	0	0
phi_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	1
phi_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
phi_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
NRf_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	705
NRf_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.1234894	[illegible]	0	0	0
NRf_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] 0.0425452	[illegible]	0	0	0
NRf_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
NRf_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
NRc_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	62
NRc_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	129

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	interquartile	skewness	kurtosis	Missing	n_Zeros	Outliers
NRc_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.3900639	[illegible]	0	0
NRc_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.3845375	[illegible]	0	0
BPR_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0970844	0	0	0
BPR_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0
BPR_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
htBleed_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0
htBleed_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0065185	[illegible]	0	0
W31_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
lag1_altitude	numeric	-0.0087000	0.0087000	158	-0.0000091	0.0021866	0.0000000	-0.0087000	-0.0015000	0.0000000	0.0015000	0.0087000	0.0030000	-0.0263183	[illegible]	100	NA	0
lag1_mach_number	numeric	-0.0006000	[illegible]	[illegible]	[illegible]	[illegible]	0.0000000	-0.0006000	-0.0002000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1307571	100	NA	0
lag1_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1201506	100	NA	0
lag1_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0045267	100	NA	0
lag1_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1978799	100	NA	0
lag1_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0000000	-6.8979842	[illegible]	100	NA	406
lag1_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100	NA	0
lag1_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
lag1_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	777
lag1_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2002886	100	NA	0
lag1_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100	NA	0
lag1_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
lag1_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	660
lag1_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1260110	100	NA	0
lag1_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0703201	100	NA	0
lag1_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100	NA	0
lag1_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100	NA	0
lag2_altitude	numeric	-0.0087000	0.0087000	158	-0.0000081	0.0021865	0.0000000	-0.0087000	-0.0015000	0.0000000	0.0015000	0.0087000	0.0030000	-0.0300	[illegible]	200	NA	0
lag2_mach_number	numeric	-0.0006000	[illegible]	[illegible]	[illegible]	[illegible]	0.0000000	-0.0006000	-0.0002000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.1319939	200	NA	0
lag2_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1325772	200	NA	0
lag2_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0106444	200	NA	0

names	classes	min … skewness (overprinted)	kurtosis	n	Missing Values	outliers
lag2_T50	numeric	[illegible]	-0.2081566	200	NA	0
lag2_P15	numeric	[illegible]	45.338[illegible] -6.8801151	[illegible]	NA	406
lag2_P30	numeric	[illegible]	-0.357[illegible]	200	NA	0
lag2_Nf	numeric	[illegible]	[illegible]	[illegible]	NA	0
lag2_Nc	numeric	[illegible]	[illegible]	[illegible]	NA	749
lag2_Ps30	numeric	[illegible]	-0.2236239	200	NA	0
lag2_phi	numeric	[illegible]	-0.401[illegible]	200	NA	0
lag2_NRf	numeric	[illegible]	[illegible]	[illegible]	NA	0
lag2_NRc	numeric	[illegible]	[illegible]	[illegible]	NA	628
lag2_BPR	numeric	[illegible]	-0.1393215	200	NA	0
lag2_htBleed	numeric	[illegible]	-0.0842956	200	NA	0
lag2_W31	numeric	[illegible]	-0.317[illegible]	200	NA	0
lag2_W32	numeric	[illegible]	-0.317[illegible]	200	NA	0
lag3_altitude	numeric	-0.0087000 0.0087000 [illegible] -0.0000073 [illegible] -0.0087000 0.0015000 [illegible]	-0.029[illegible]	300	NA	0
lag3_mach_number	numeric	-0.0006000 [illegible] -0.0006000 0.0002000 [illegible]	1.1340397	300	NA	0
lag3_T24	numeric	[illegible]	-0.1391064	300	NA	0
lag3_T30	numeric	[illegible]	-0.0127268	300	NA	0
lag3_T50	numeric	[illegible]	-0.2092889	300	NA	0
lag3_P15	numeric	[illegible]	45.091[illegible] -6.8621995	[illegible]	NA	406
lag3_P30	numeric	[illegible]	-0.341[illegible]	300	NA	0
lag3_Nf	numeric	[illegible]	[illegible]	[illegible]	NA	0
lag3_Nc	numeric	[illegible]	[illegible]	[illegible]	NA	727
lag3_Ps30	numeric	[illegible]	0.2494484	300	NA	0
lag3_phi	numeric	[illegible]	-0.380[illegible]	300	NA	0
lag3_NRf	numeric	[illegible]	[illegible]	[illegible]	NA	0
lag3_NRc	numeric	[illegible]	[illegible]	[illegible]	NA	602
lag3_BPR	numeric	[illegible]	-0.1415849	300	NA	0
lag3_htBleed	numeric	[illegible]	-0.0847392	300	NA	0
lag3_W31	numeric	[illegible]	-0.306[illegible]	300	NA	0
lag3_W32	numeric	[illegible]	-0.305[illegible]	300	NA	0

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	nMissing	nZeros	nOutliers
lag4_altitude	numeric	-0.0087000	0.0087000	159	-0.0000076	0.0021860	0.0000000	-0.0087000	-0.0015000	0.0000000	0.0015000	0.0087000	0.0030000	-0.0286019	6.8955486	400	NA	0
lag4_mach_number	numeric	-0.0006000	0.0006000	[illegible]	0.0000027	[illegible]	0.0000000	-0.0006000	-0.0002000	0.0000000	0.0003000	0.0006000	0.0005000	-0.0086669	1.1337862	400	NA	0
lag4_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1473674	400	NA	0
lag4_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0065031	400	NA	0
lag4_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2363099	400	NA	0
lag4_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0000000	-6.8442371	[illegible]	400	NA	406
lag4_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
lag4_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0067720	400	NA	0
lag4_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	703
lag4_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2623309	400	NA	0
lag4_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
lag4_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
lag4_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	575
lag4_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1428007	400	NA	0
lag4_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1017056	400	NA	0
lag4_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
lag4_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
lag5_altitude	numeric	-0.0087000	0.0087000	159	-0.0000068	[illegible]	0.0000000	-0.0087000	-0.0015000	0.0000000	0.0015000	0.0087000	0.0030000	[illegible]	[illegible]	500	NA	0
lag5_mach_number	numeric	-0.0006000	0.0006000	[illegible]	0.0000027	[illegible]	0.0000000	-0.0006000	-0.0002000	0.0000000	0.0003000	0.0006000	0.0005000	[illegible]	1.1348525	500	NA	0
lag5_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1430997	500	NA	0
lag5_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0186507	500	NA	0
lag5_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2555627	500	NA	0
lag5_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0000000	-6.8262275	[illegible]	500	NA	406
lag5_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	500	NA	0
lag5_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0240081	500	NA	0
lag5_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	500	NA	676
lag5_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2672464	500	NA	0
lag5_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	500	NA	0

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	n	Missing_Values	Outliers
lag5_NRm	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0101762	[illegible]	500	NA	0
lag5_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	548
lag5_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1434169	[illegible]	500	NA	0
lag5_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0992936	[illegible]	500	NA	0
lag5_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	500	NA	0
lag5_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	500	NA	0
rollmean3_altitude	numeric	-0.0063000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0063000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	1
rollsd3_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean3_mach_number	numeric	-0.0005333	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0005333	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.3356897	[illegible]	200	NA	0
rollsd3_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.7656307	[illegible]	200	NA	0
rolliqr3_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	0
rollmean3_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.1897775	[illegible]	200	NA	0
rollsd3_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean3_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0140140	[illegible]	200	NA	0
rollsd3_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean3_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.2474249	[illegible]	200	NA	0
rollsd3_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rolliqr3_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rollmean3_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.7133815	24.917	2109	NA	3284
rollsd3_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2119	NA	1100
rolliqr3_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2119	NA	1100
rollmean3_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	0
rollsd3_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rolliqr3_P30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rollmean3_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1
rollsd3_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
rolliqr3_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
rollmean3_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	873
rollsd3_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean3_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.2594036	[illegible]	200	NA	0
rollsd3_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0

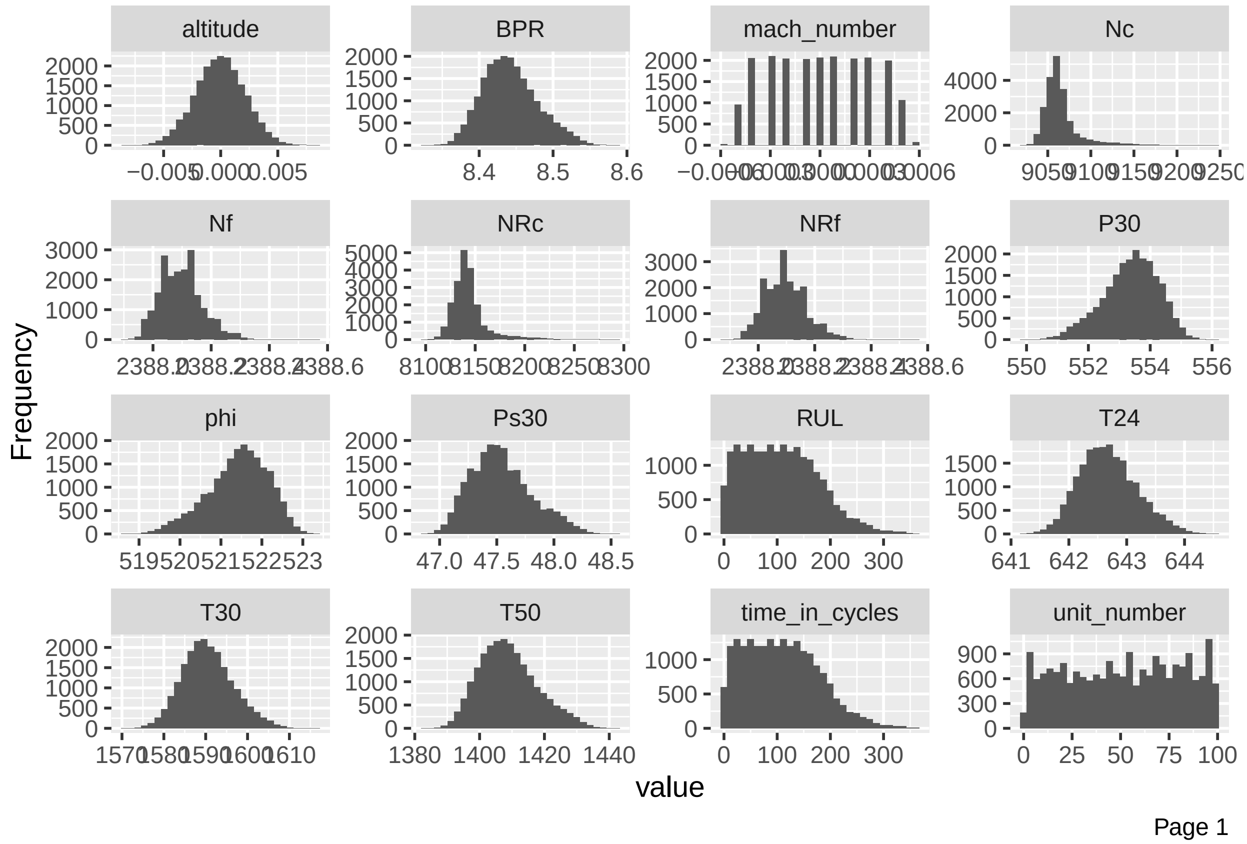
names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	nMissing	nZero	nOutliers
rollmean3_phio	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	0
rollsd3_phio	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	5
rolliqr3_phio	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	6
rollmean3_NRef	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1
rollsd3_NRef	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
rolliqr3_NRef	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	9
rollmean3_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	694
rollsd3_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rolliqr3_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rollmean3_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.1902186	200	NA	0
rollsd3_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rolliqr3_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rollmean3_hBed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0995197	200	NA	0
rollsd3_hBed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	6
rolliqr3_hBed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	4
rollmean3_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	0
rollsd3_W30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr3_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean3_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	200	NA	0
rollsd3_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	6
rolliqr3_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rollmean5_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr5_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	5
rollmean5_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.2374906	400	NA	0
rollsd5_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rolliqr5_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.3081020	400	NA	0
rollmean5_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.2287542	400	NA	0
rollsd5_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rolliqr5_T24	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rollmean5_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0437974	400	NA	0
rollsd5_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1
rolliqr5_T30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1
rollmean5_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.2991102	400	NA	0
rollsd5_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr5_T50	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rollmean5_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-4.0135028	[illegible]	NA	4766
rollsd5_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1682
rolliqr5_P15	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	257

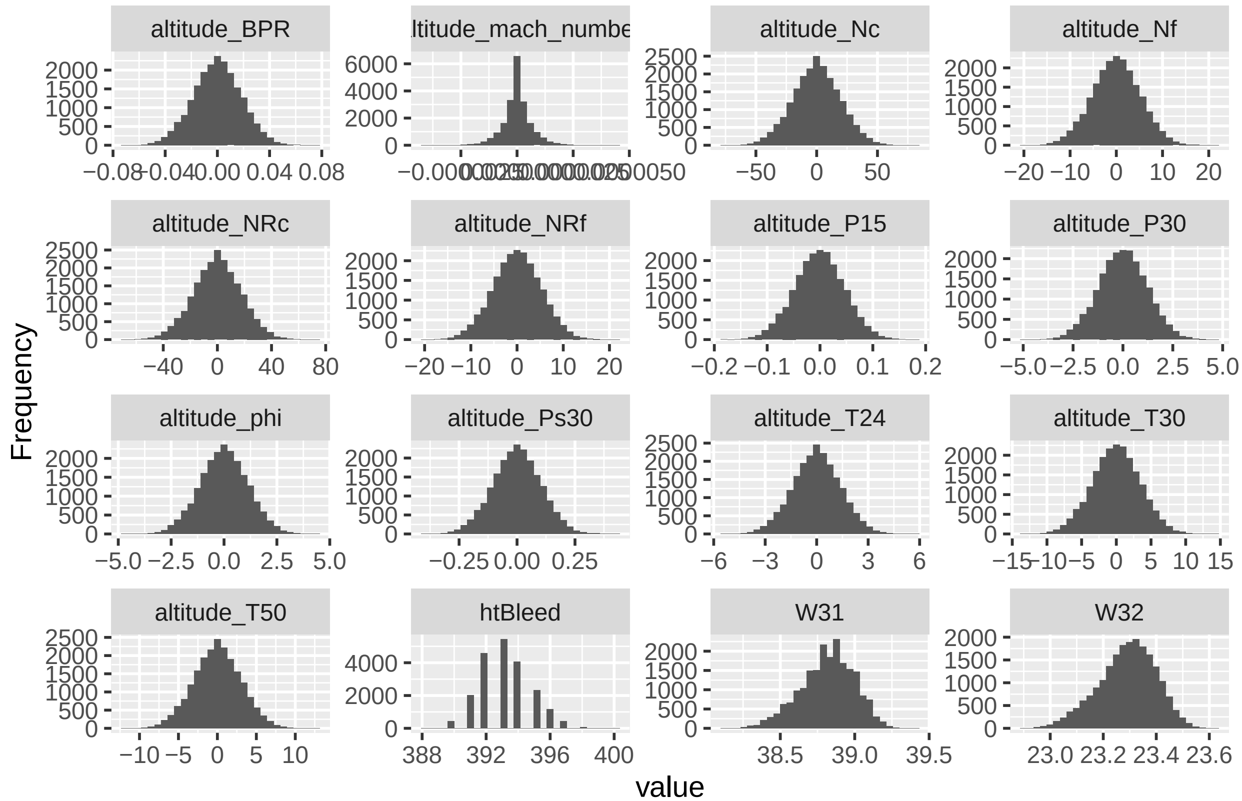
names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	Missing	n_Zeros	Outliers
rollmean5_B30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_B30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rolliqr5_B30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rollmean5_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollsd5_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
rolliqr5_Nf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	25
rollmean5_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	880
rollsd5_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rolliqr5_Nc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	5
rollmean5_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rolliqr5_Ps30	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean5_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	3
rolliqr5_phi	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	4
rollmean5_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollsd5_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
rolliqr5_NRf	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	16
rollmean5_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	667
rollsd5_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	11
rolliqr5_NRc	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	16
rollmean5_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	2
rolliqr5_BPR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	4
rollmean5_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	1
rolliqr5_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	0
rollmean5_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rolliqr5_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	10
rollmean5_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	400	NA	0
rollsd5_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	5
rolliqr5_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	NA	7
cumsum_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8	59
cummin_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8	114
cummax_altitude	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	11	110
cumsum_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	39	172
cummin_mach_number	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	38	169

names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	iqr	skewness	kurtosis	Missing	n_zeros	Outliers
cummax	numeric	0.0006000	[illegible]	[illegible]	0.0004780	[illegible]	0.0006000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-3.6813017	28.0070540	[illegible]	45	7255
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2186950	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	3
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1538808	0	0	0
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2169694	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0656799	0	0	2
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	1
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2151697	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.5077545	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0876014	0	0	0
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2190194	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1.8668090	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-50.7493925	[illegible]	[illegible]	[illegible]	8
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2198910	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2190353	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.9805937	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	6
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2180763	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	962
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2157606	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1.0079017	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1914289	0	0	0
cumsum_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2199107	0	0	0
cummin_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cummax_[illegible]	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0

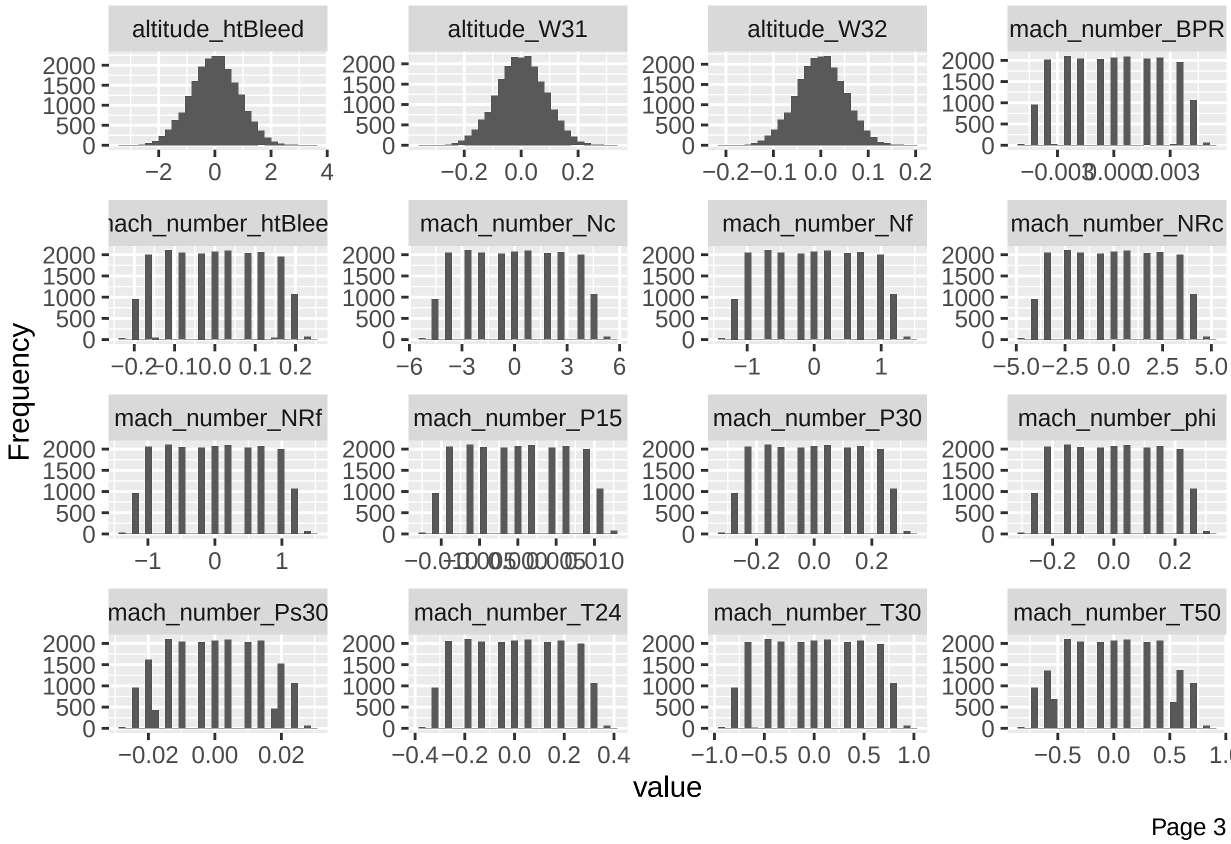
names	classes	min	max	n_unique	mean	sd	median	quantile_0	quantile_25	quantile_50	quantile_75	quantile_100	interquartile	skewness	kurtosis	Missing	nZeros	Outliers
cumsum_NR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2190301	0	0	0
cummin_NR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.9052250	0	0	0
cummax_NR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	4
cumsum_NRd	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2185192	0	0	0
cummin_NRd	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cummax_NRd	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	721
cumsum_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2170711	0	0	0
cummin_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1382855	0	0	3
cummax_BRR	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.1640209	0	0	0
cumsum_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2169563	0	0	0
cummin_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	2
cummax_htBleed	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0
cumsum_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2219665	0	0	0
cummin_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	1
cummax_W31	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	3
cumsum_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.2211579	0	0	0
cummin_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0
cummax_W32	numeric	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0

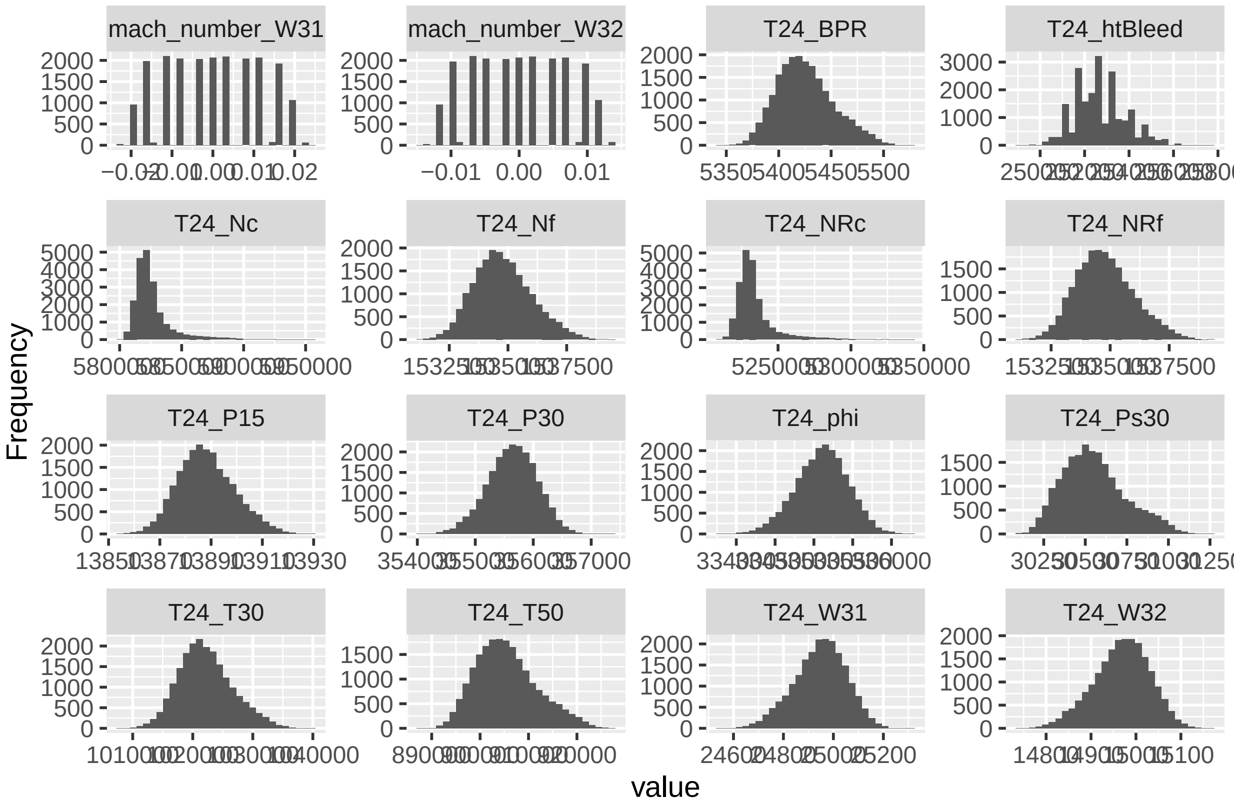
```r
plot_histogram(trainset2)
```


altitude
BPR
mach_number
Nc
Nf
NRc
NRf
P30
phi
Ps30
RUL
T24
T30
T50
time_in_cycles
unit_number
Frequency
value
Page 1

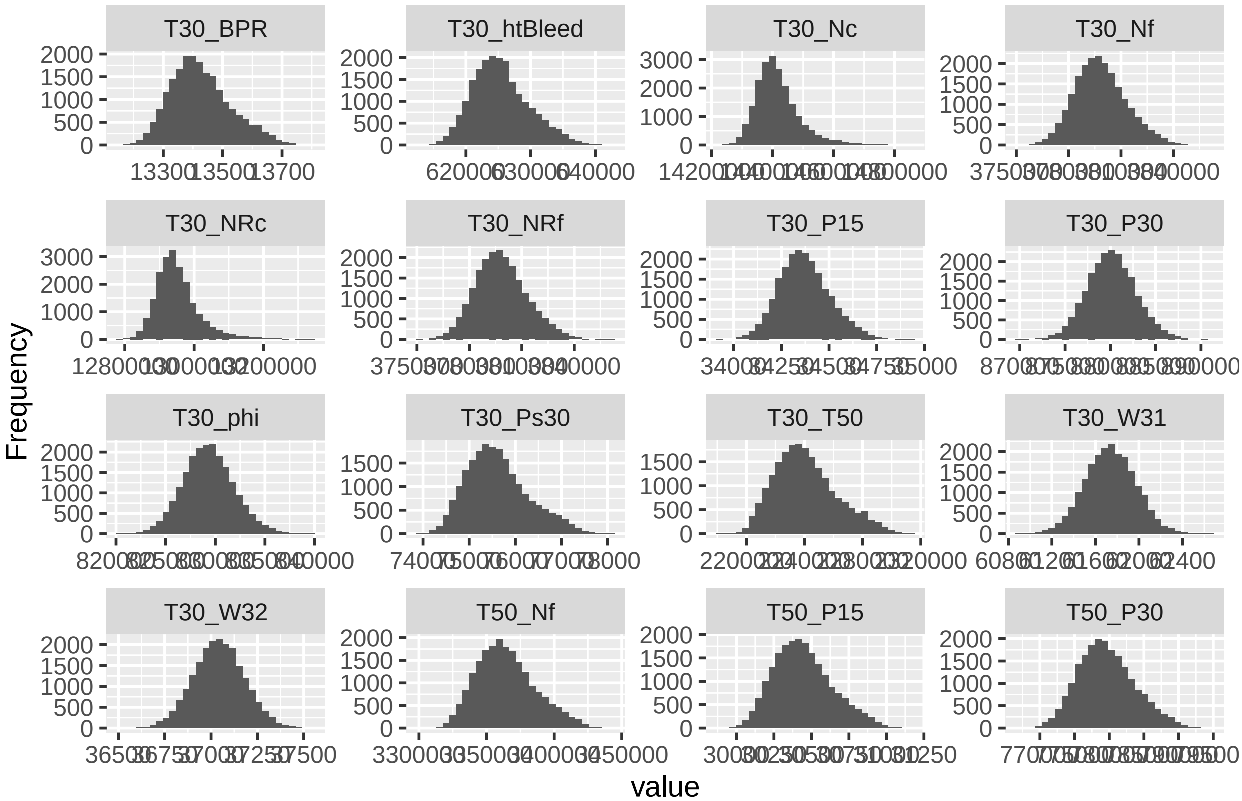

altitude_BPR
ltitude_mach_numbe
altitude_Nc
altitude_Nf
altitude_NRc
altitude_NRf
altitude_P15
altitude_P30
altitude_phi
altitude_Ps30
altitude_T24
altitude_T30
altitude_T50
htBleed
W31
W32
Frequency
value

Page 2

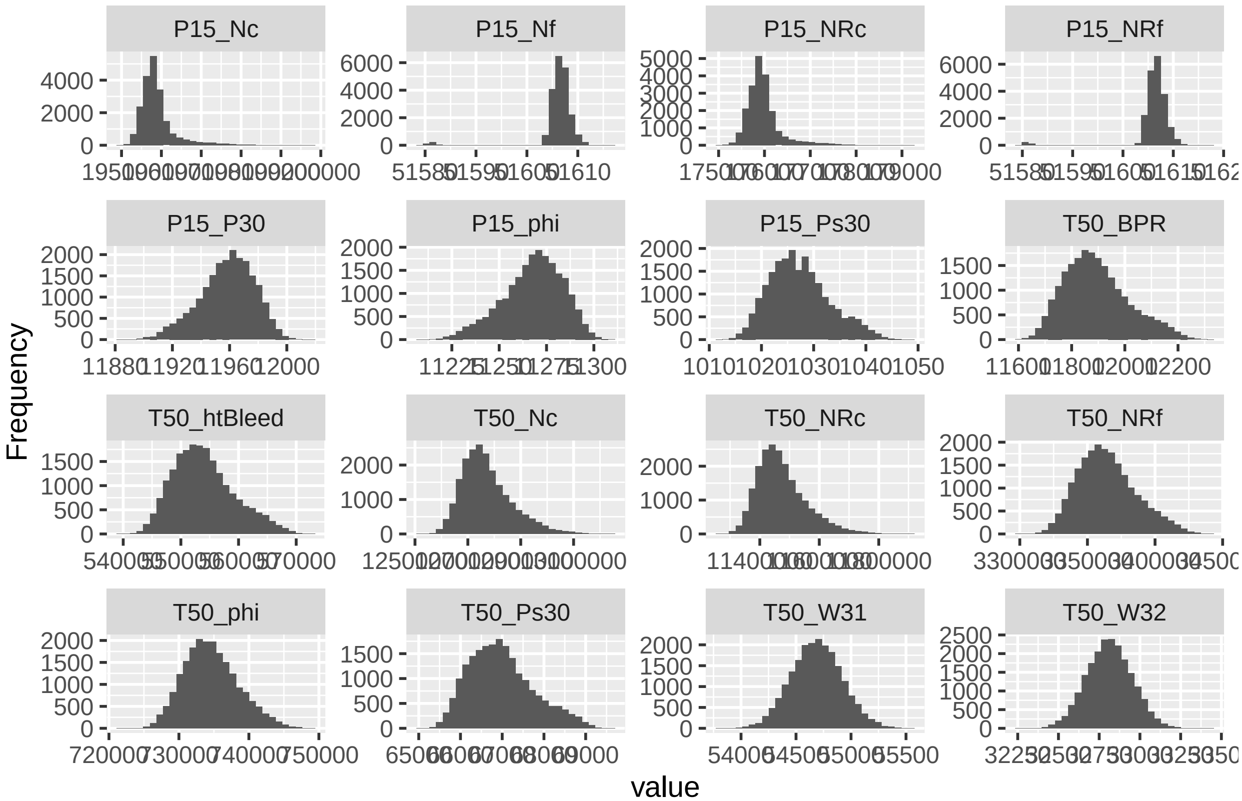

altitude_htBleed
altitude_W31
altitude_W32
mach_number_BPR
ıach_number_htBlee
mach_number_Nc
mach_number_Nf
mach_number_NRc
mach_number_NRf
mach_number_P15
mach_number_P30
mach_number_phi
mach_number_Ps30
mach_number_T24
mach_number_T30
mach_number_T50
Frequency
value
Page 3


mach_number_W31
mach_number_W32
T24_BPR
T24_htBleed
T24_Nc
T24_Nf
T24_NRc
T24_NRf
T24_P15
T24_P30
T24_phi
T24_Ps30
T24_T30
T24_T50
T24_W31
T24_W32
Frequency
value

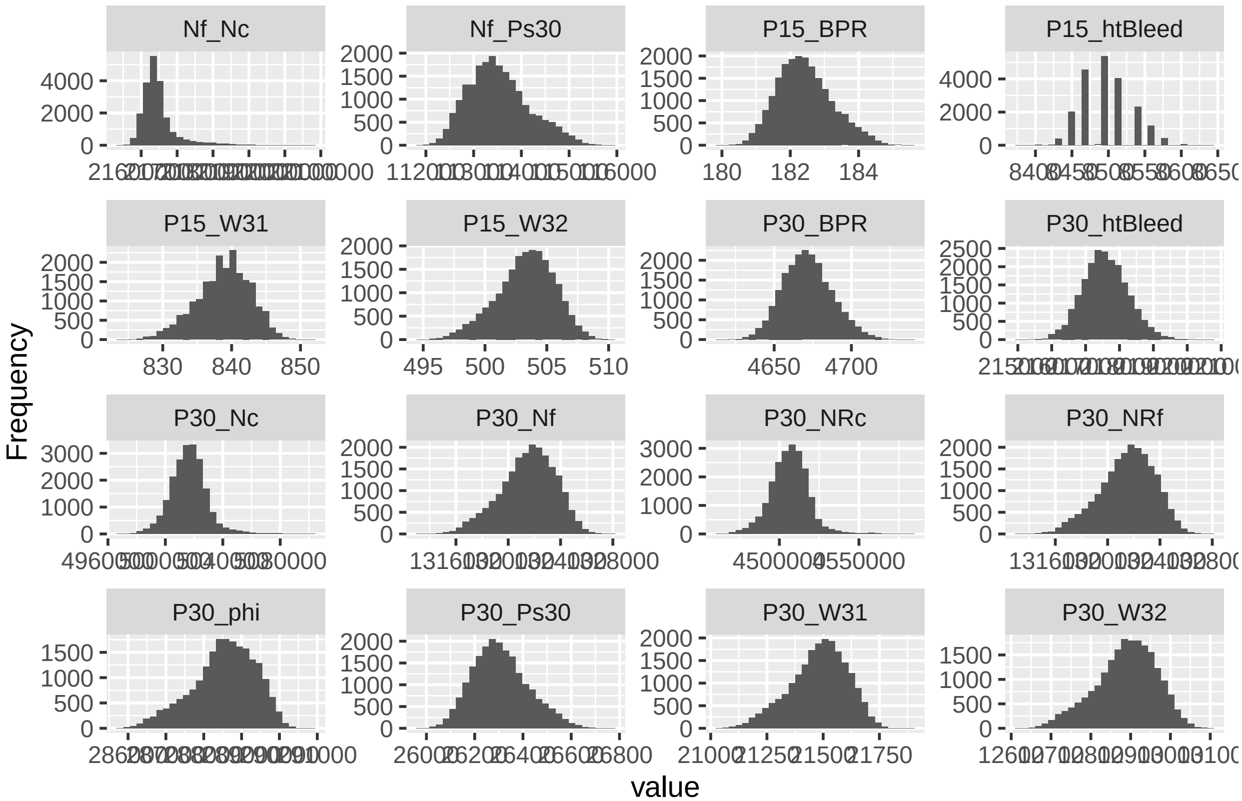
Page 4


T30_BPR
T30_htBleed
T30_Nc
T30_Nf
T30_NRc
T30_NRf
T30_P15
T30_P30
T30_phi
T30_Ps30
T30_T50
T30_W31
T30_W32
T50_Nf
T50_P15
T50_P30
Frequency
value

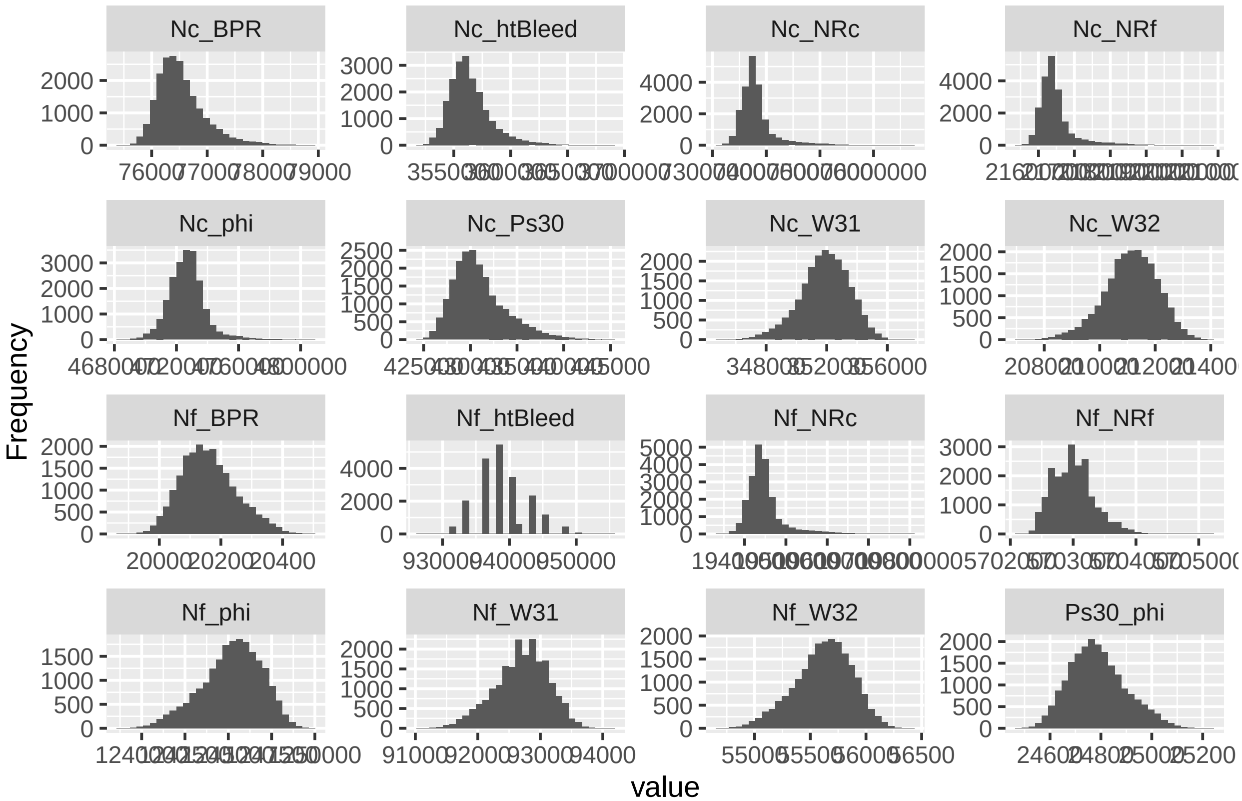
Page 5


P15_Nc
P15_Nf
P15_NRc
P15_NRf
P15_P30
P15_phi
P15_Ps30
T50_BPR
T50_htBleed
T50_Nc
T50_NRc
T50_NRf
T50_phi
T50_Ps30
T50_W31
T50_W32
Frequency
value

Page 6

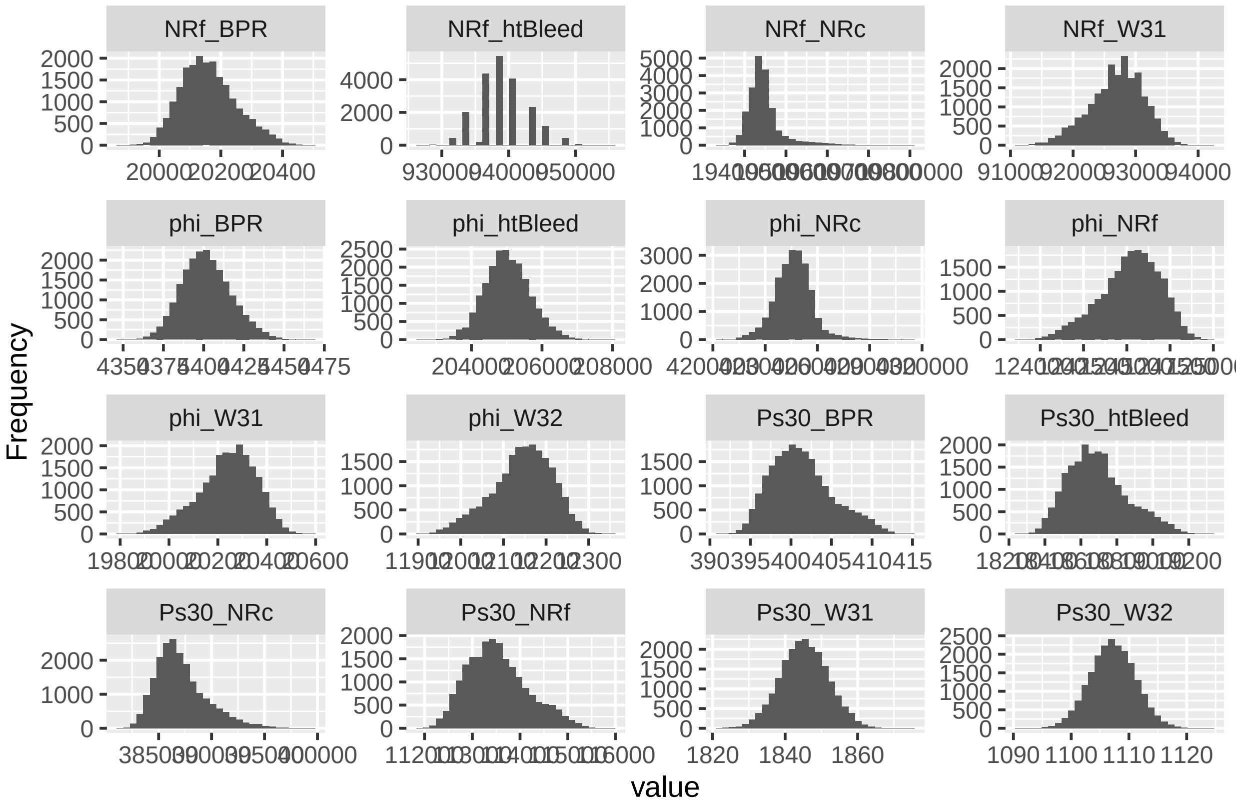

Nf_Nc
Nf_Ps30
P15_BPR
P15_htBleed
P15_W31
P15_W32
P30_BPR
P30_htBleed
P30_Nc
P30_Nf
P30_NRc
P30_NRf
P30_phi
P30_Ps30
P30_W31
P30_W32
Frequency
value

Page 7

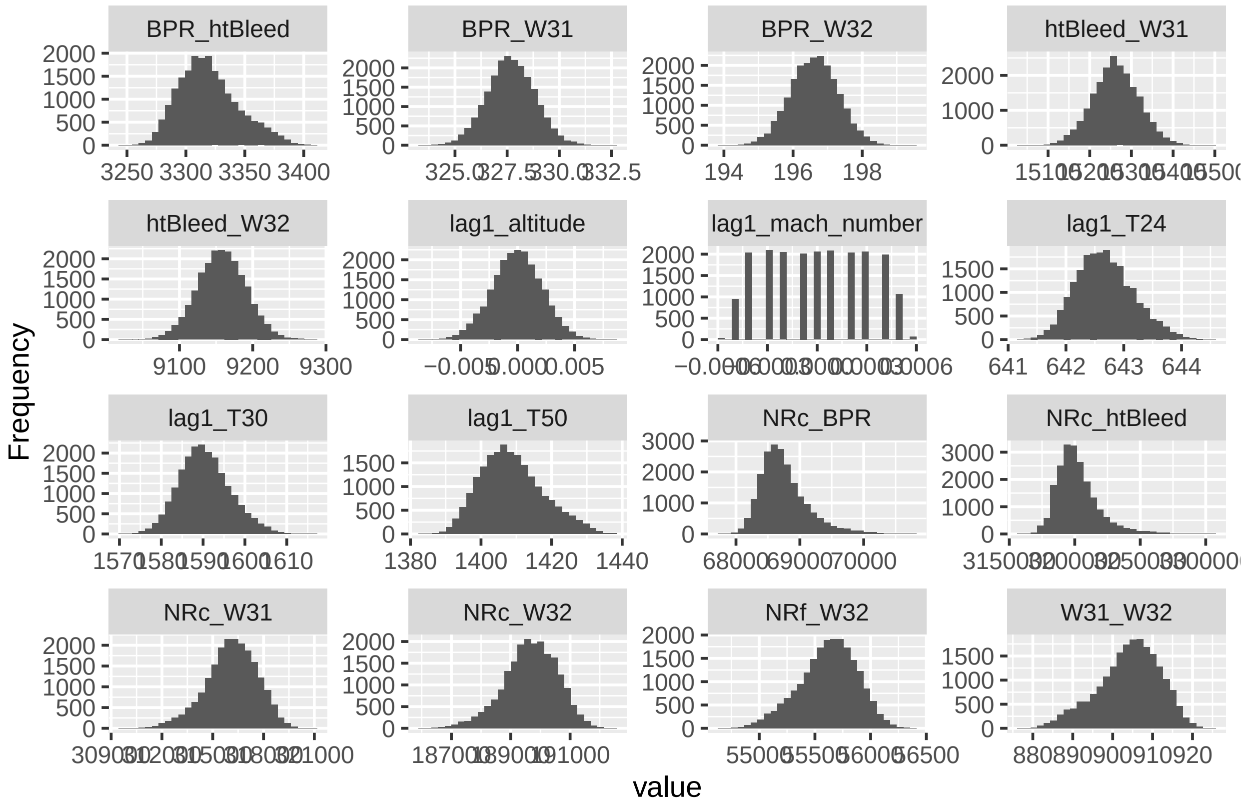


Nc_BPR
Nc_htBleed
Nc_NRc
Nc_NRf
Nc_phi
Nc_Ps30
Nc_W31
Nc_W32
Nf_BPR
Nf_htBleed
Nf_NRc
Nf_NRf
Nf_phi
Nf_W31
Nf_W32
Ps30_phi
Frequency
value

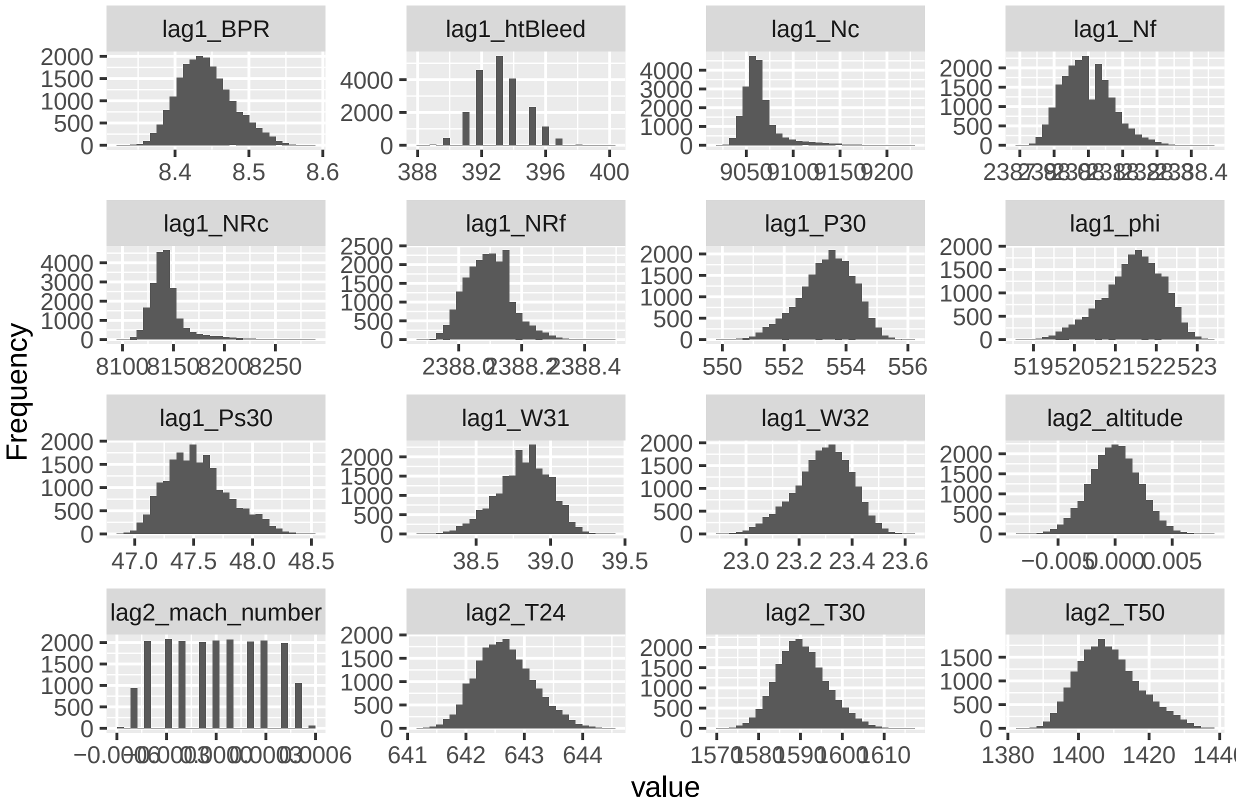
Page 8


NRf_BPR
NRf_htBleed
NRf_NRc
NRf_W31
phi_BPR
phi_htBleed
phi_NRc
phi_NRf
phi_W31
phi_W32
Ps30_BPR
Ps30_htBleed
Ps30_NRc
Ps30_NRf
Ps30_W31
Ps30_W32
Frequency
value

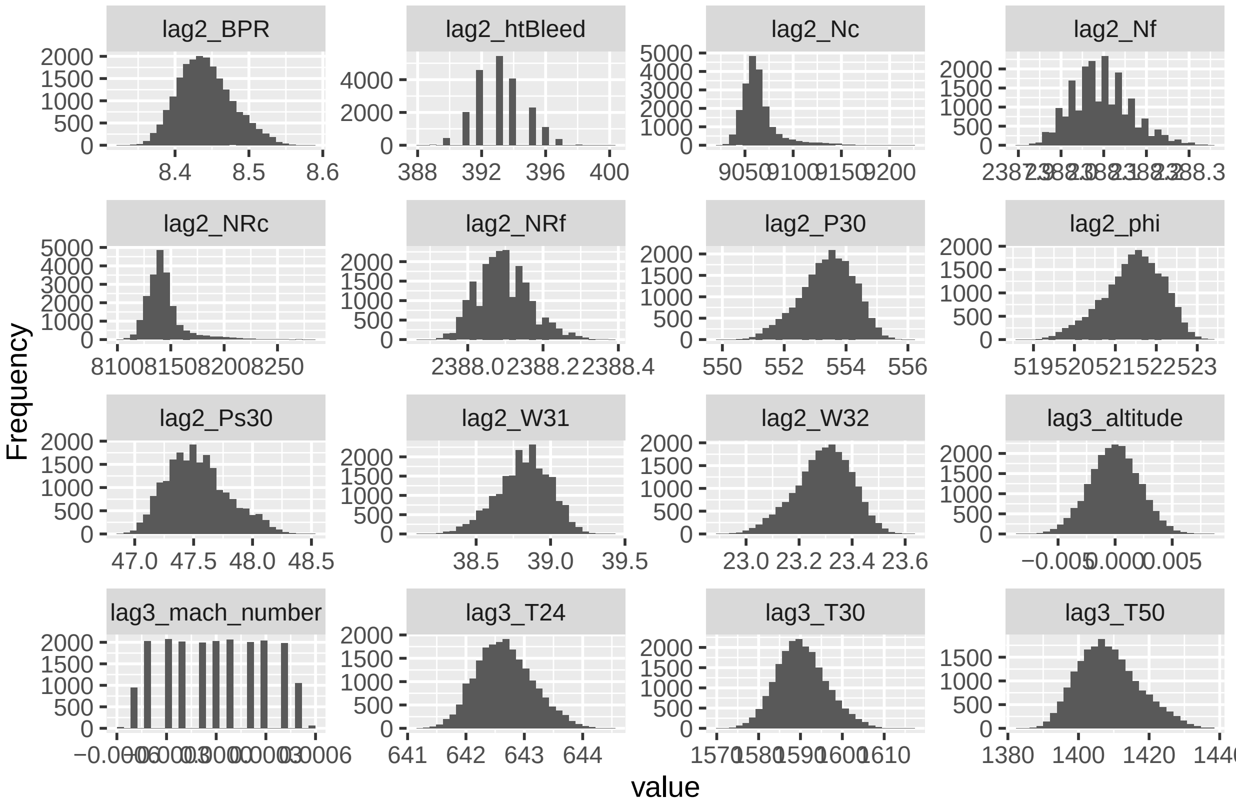
Page 9


BPR_htBleed
BPR_W31
BPR_W32
htBleed_W31
htBleed_W32
lag1_altitude
lag1_mach_number
lag1_T24
lag1_T30
lag1_T50
NRc_BPR
NRc_htBleed
NRc_W31
NRc_W32
NRf_W32
W31_W32
Frequency
value

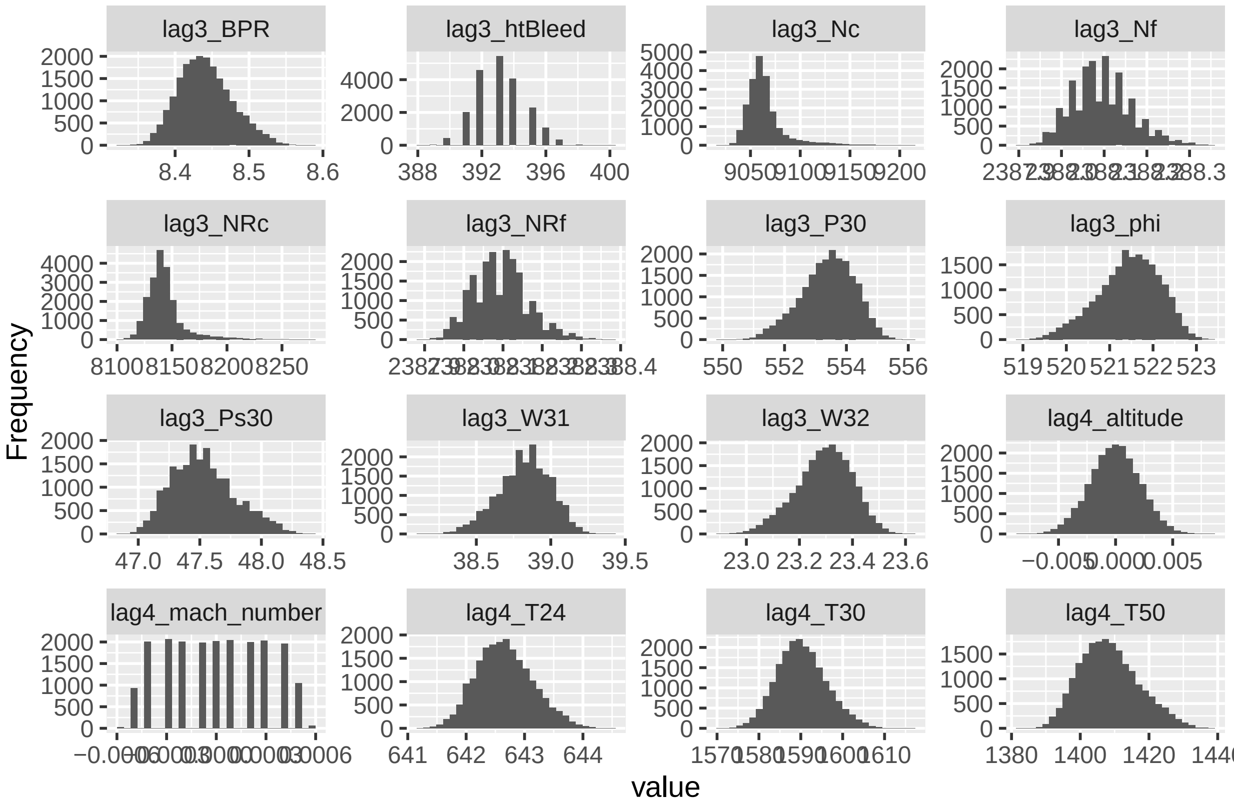
Page 10


lag1_BPR
lag1_htBleed
lag1_Nc
lag1_Nf
lag1_NRc
lag1_NRf
lag1_P30
lag1_phi
lag1_Ps30
lag1_W31
lag1_W32
lag2_altitude
lag2_mach_number
lag2_T24
lag2_T30
lag2_T50
Frequency
value

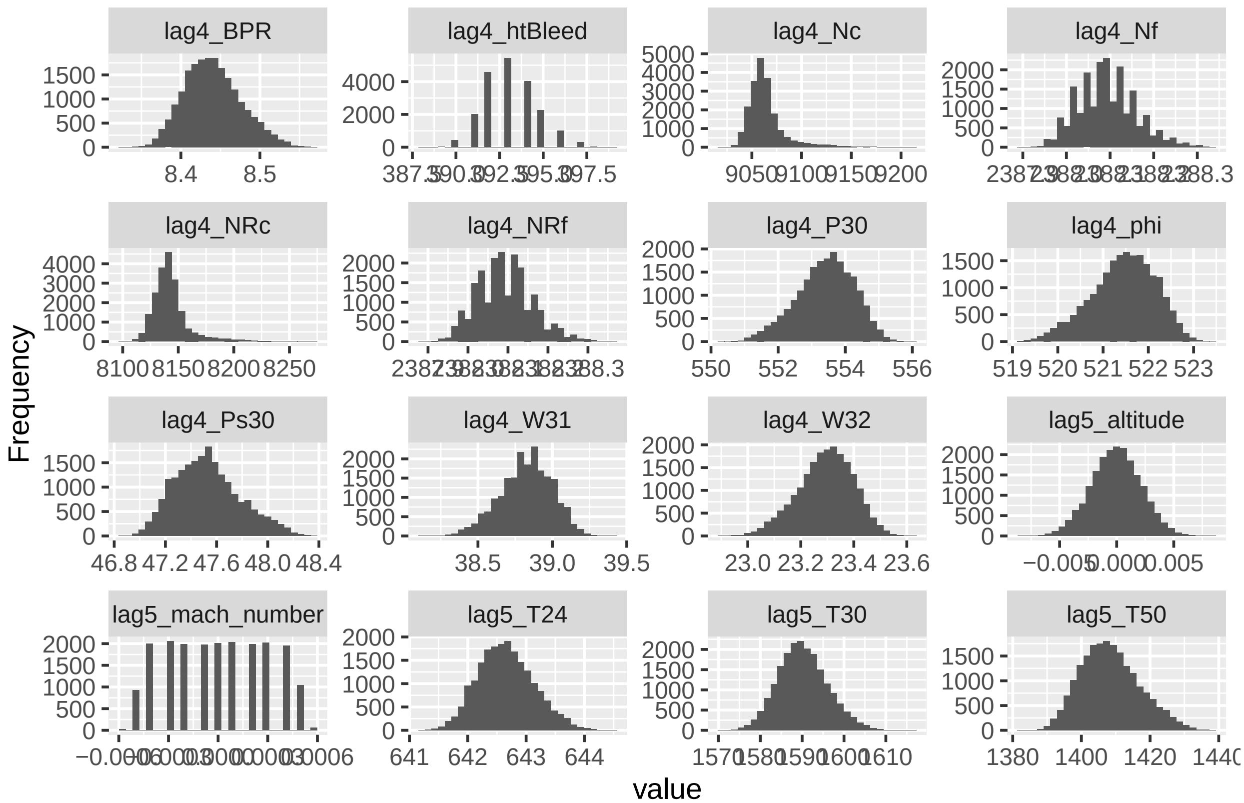
Page 11


lag2_BPR
lag2_htBleed
lag2_Nc
lag2_Nf
lag2_NRc
lag2_NRf
lag2_P30
lag2_phi
lag2_Ps30
lag2_W31
lag2_W32
lag3_altitude
lag3_mach_number
lag3_T24
lag3_T30
lag3_T50
Frequency
value

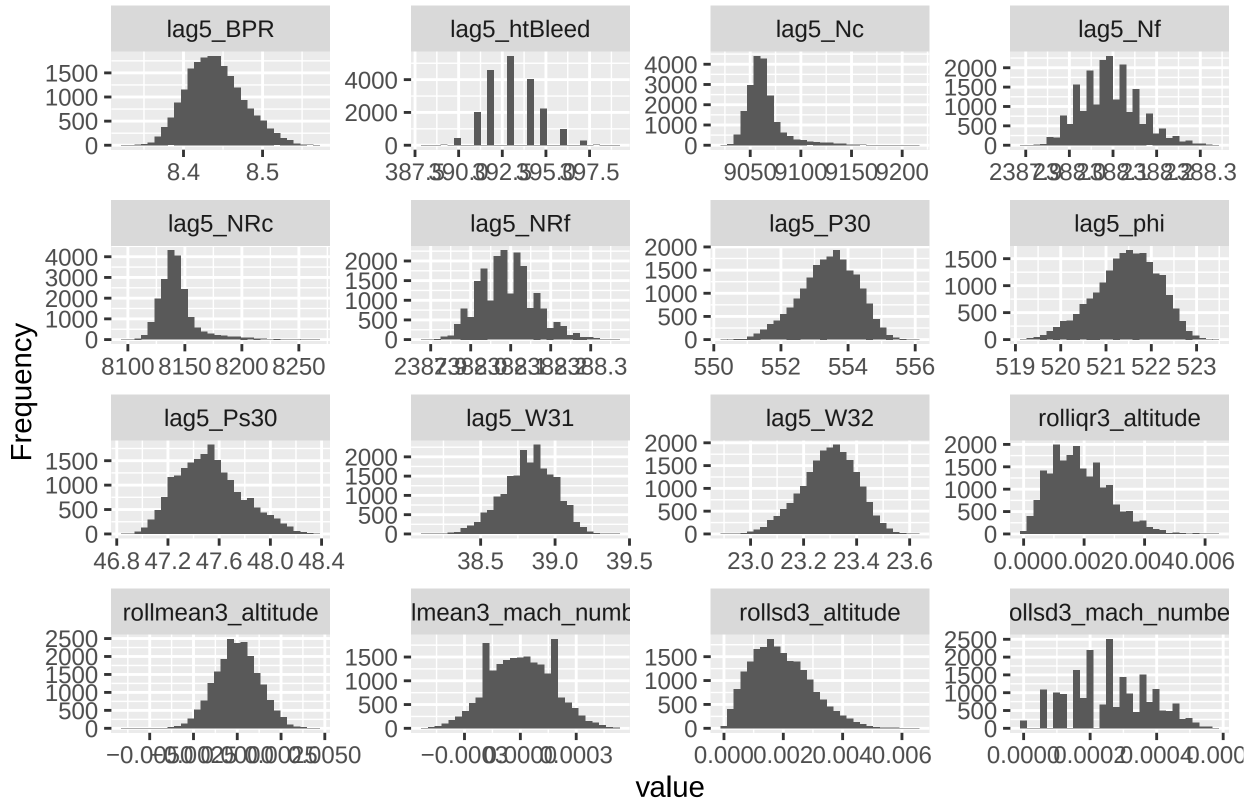
Page 12

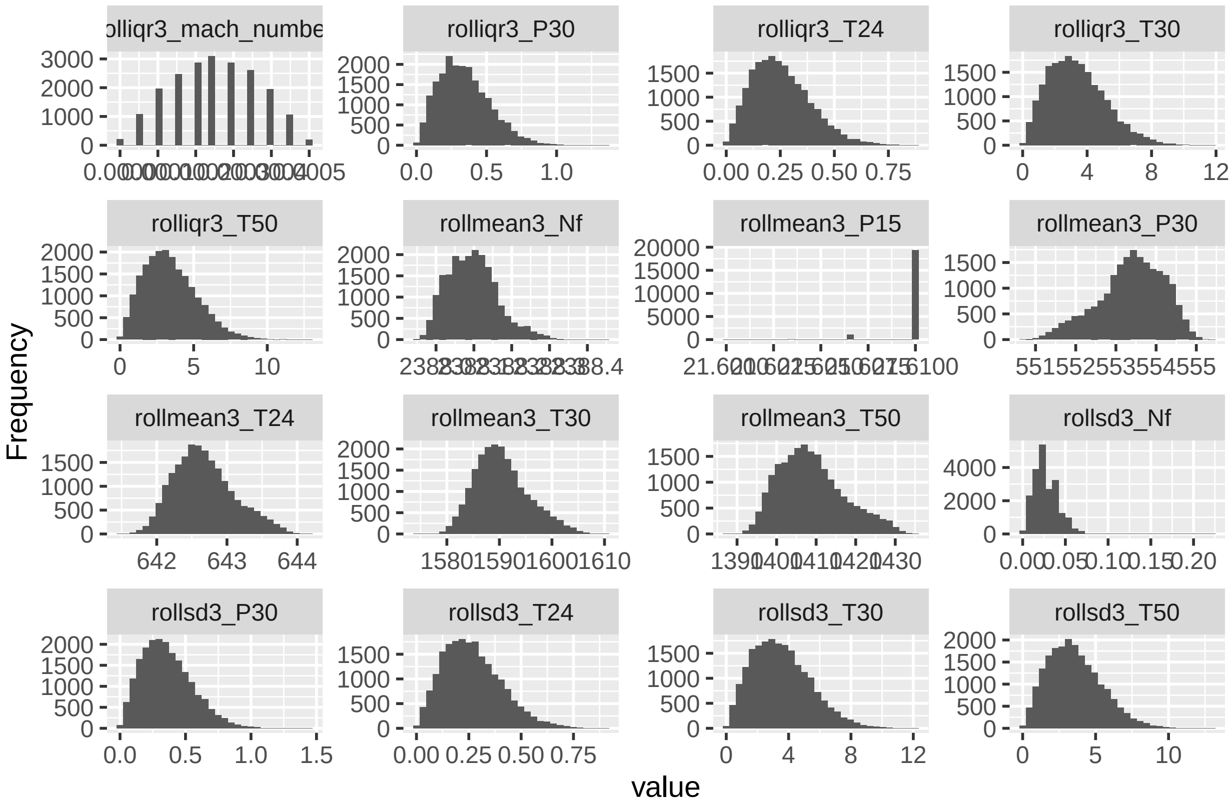

lag3_BPR
lag3_htBleed
lag3_Nc
lag3_Nf
lag3_NRc
lag3_NRf
lag3_P30
lag3_phi
lag3_Ps30
lag3_W31
lag3_W32
lag4_altitude
lag4_mach_number
lag4_T24
lag4_T30
lag4_T50
Frequency
value

Page 13

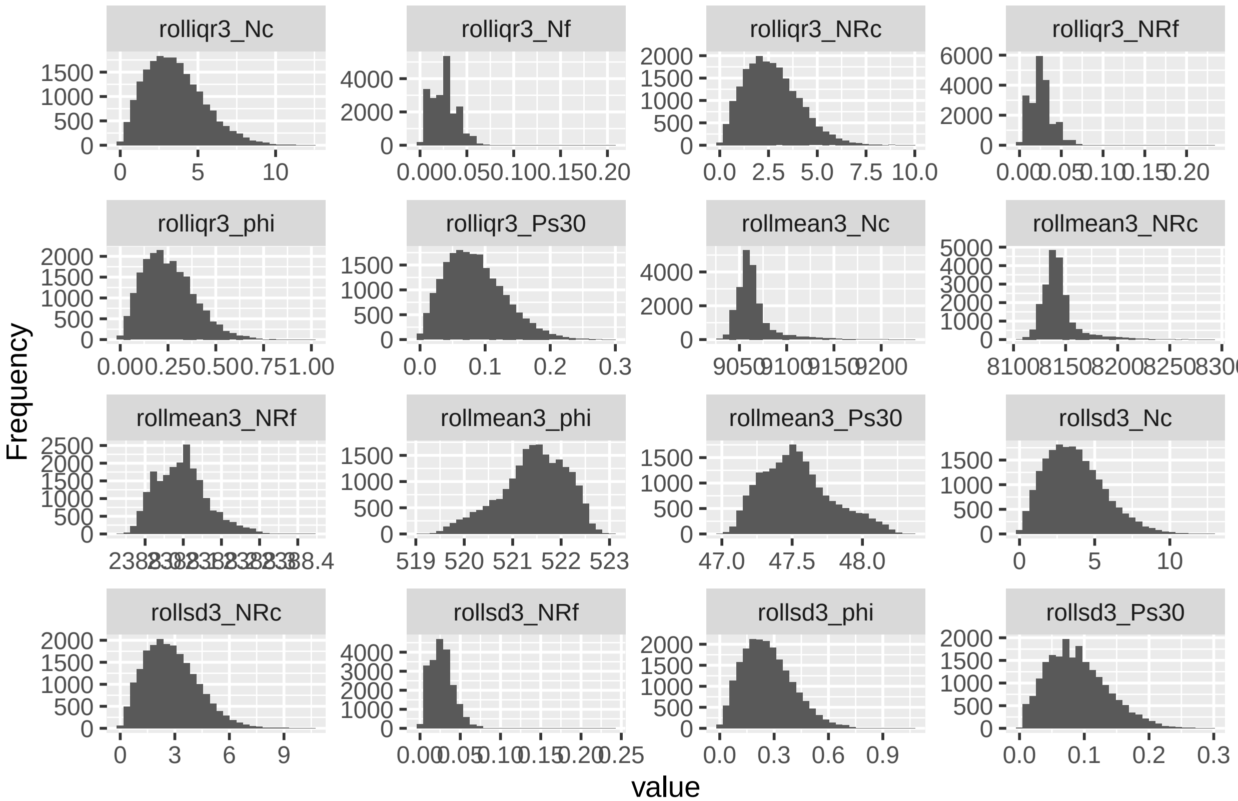

lag4_BPR
lag4_htBleed
lag4_Nc
lag4_Nf
lag4_NRc
lag4_NRf
lag4_P30
lag4_phi
lag4_Ps30
lag4_W31
lag4_W32
lag5_altitude
lag5_mach_number
lag5_T24
lag5_T30
lag5_T50
Frequency
value

Page 14

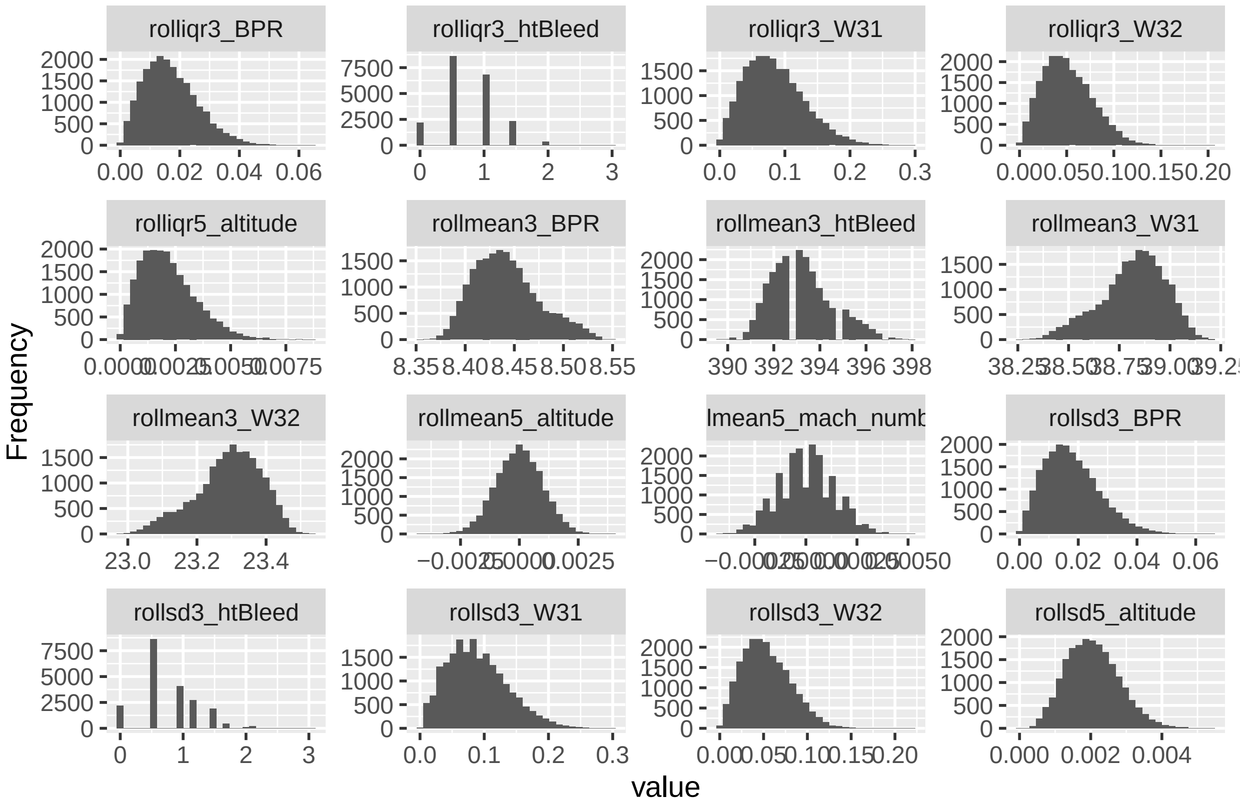

lag5_BPR
lag5_htBleed
lag5_Nc
lag5_Nf
lag5_NRc
lag5_NRf
lag5_P30
lag5_phi
lag5_Ps30
lag5_W31
lag5_W32
rolliqr3_altitude
rollmean3_altitude
lmean3_mach_numb
rollsd3_altitude
ollsd3_mach_numbe
Frequency
value


rolliqr3_P30
rolliqr3_T24
rolliqr3_T30
rolliqr3_T50
rollmean3_Nf
rollmean3_P15
rollmean3_P30
rollmean3_T24
rollmean3_T30
rollmean3_T50
rollsd3_Nf
rollsd3_P30
rollsd3_T24
rollsd3_T30
rollsd3_T50
Frequency
value

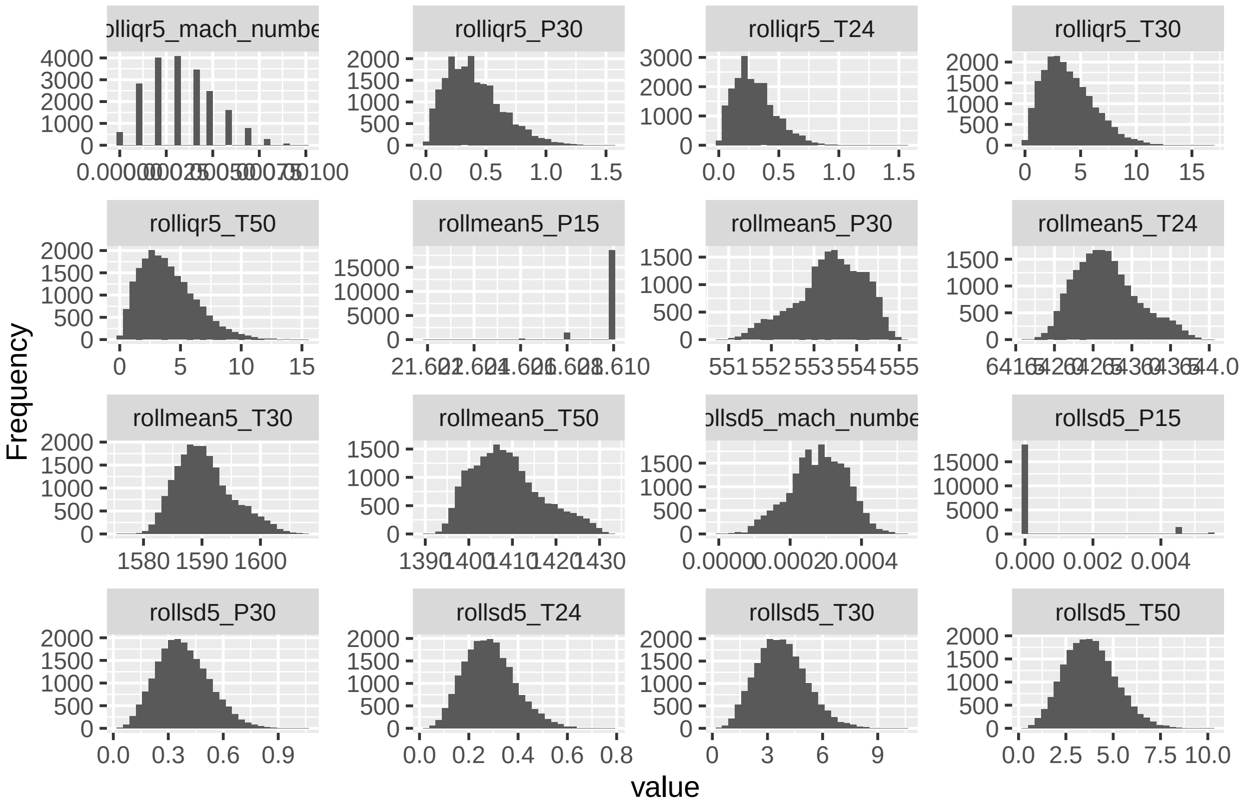
Page 16


rolliqr3_Nc
rolliqr3_Nf
rolliqr3_NRc
rolliqr3_NRf
rolliqr3_phi
rolliqr3_Ps30
rollmean3_Nc
rollmean3_NRc
rollmean3_NRf
rollmean3_phi
rollmean3_Ps30
rollsd3_Nc
rollsd3_NRc
rollsd3_NRf
rollsd3_phi
rollsd3_Ps30
Frequency
value

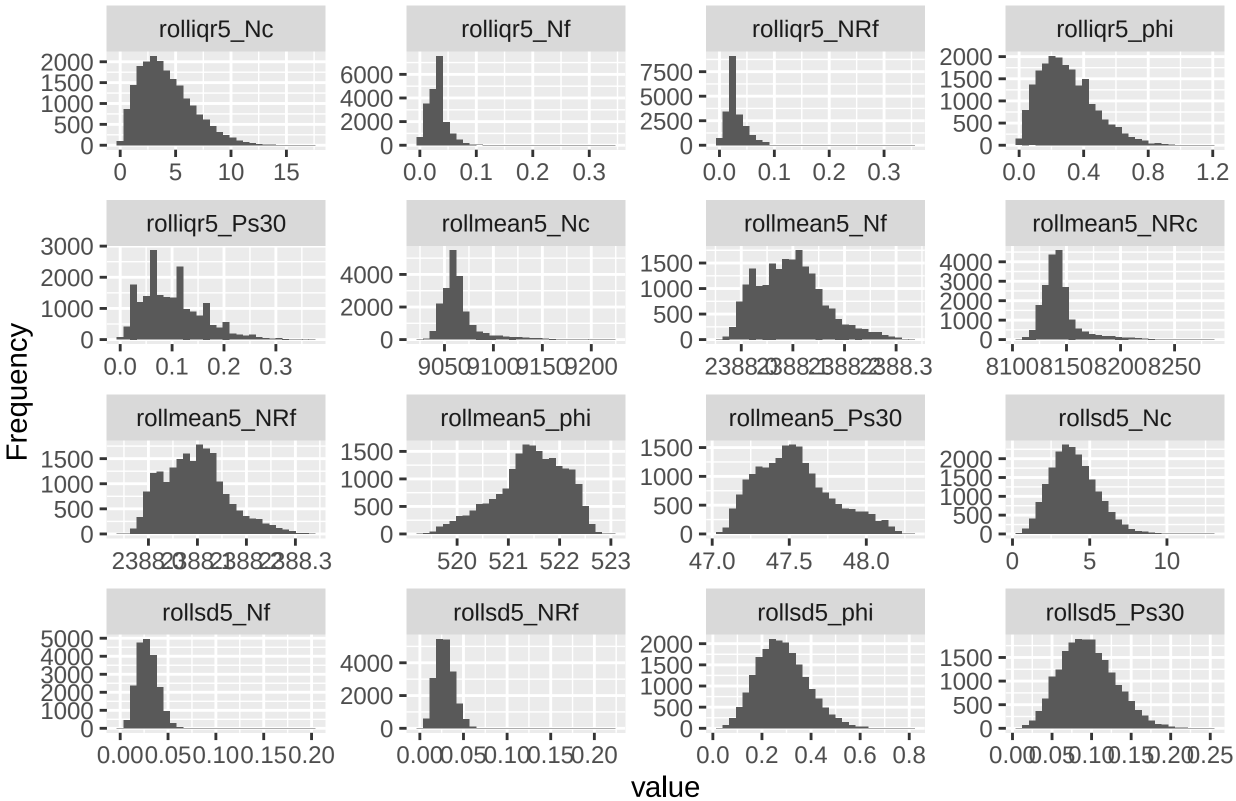
Page 17


rolliqr3_BPR
rolliqr3_htBleed
rolliqr3_W31
rolliqr3_W32
rolliqr5_altitude
rollmean3_BPR
rollmean3_htBleed
rollmean3_W31
rollmean3_W32
rollmean5_altitude
lmean5_mach_numb
rollsd3_BPR
rollsd3_htBleed
rollsd3_W31
rollsd3_W32
rollsd5_altitude
Frequency
value

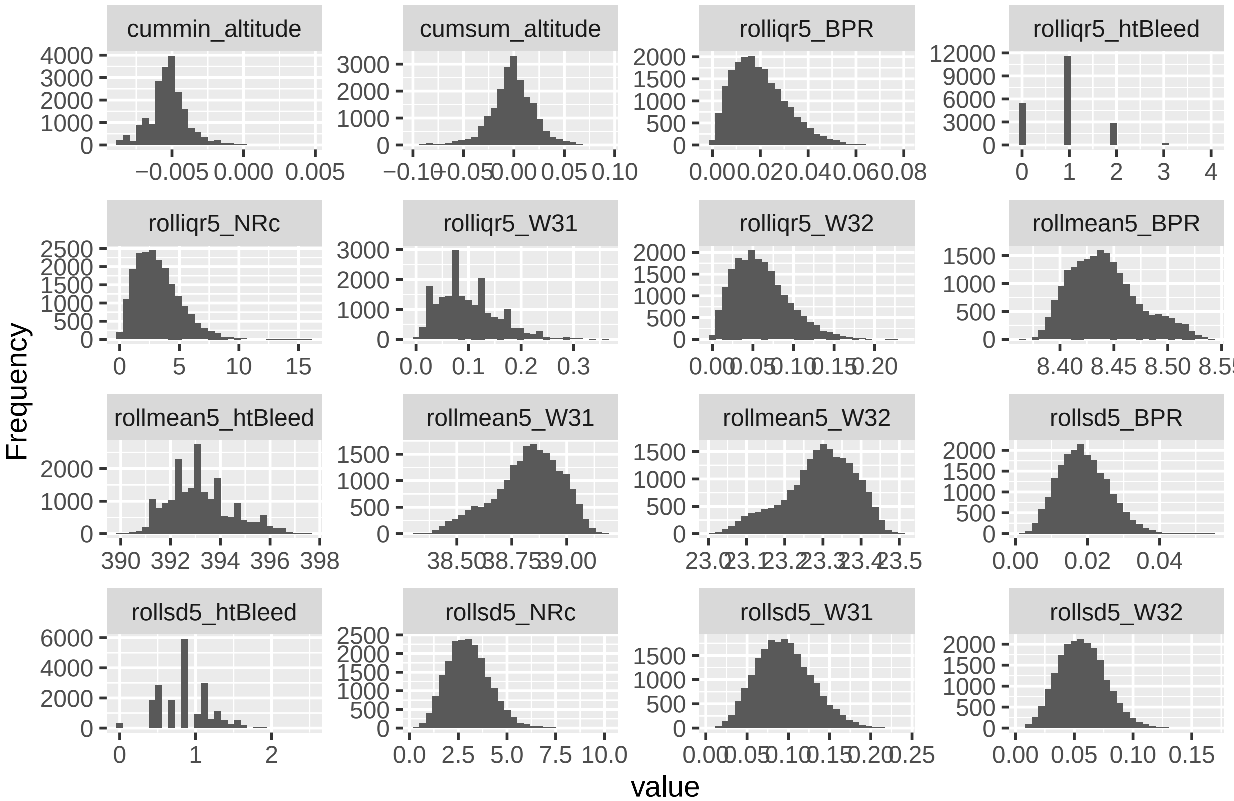
Page 18


rolliqr5_mach_numbe
rolliqr5_P30
rolliqr5_T24
rolliqr5_T30
rolliqr5_T50
rollmean5_P15
rollmean5_P30
rollmean5_T24
rollmean5_T30
rollmean5_T50
ollsd5_mach_numbe
rollsd5_P15
rollsd5_P30
rollsd5_T24
rollsd5_T30
rollsd5_T50
Frequency
value

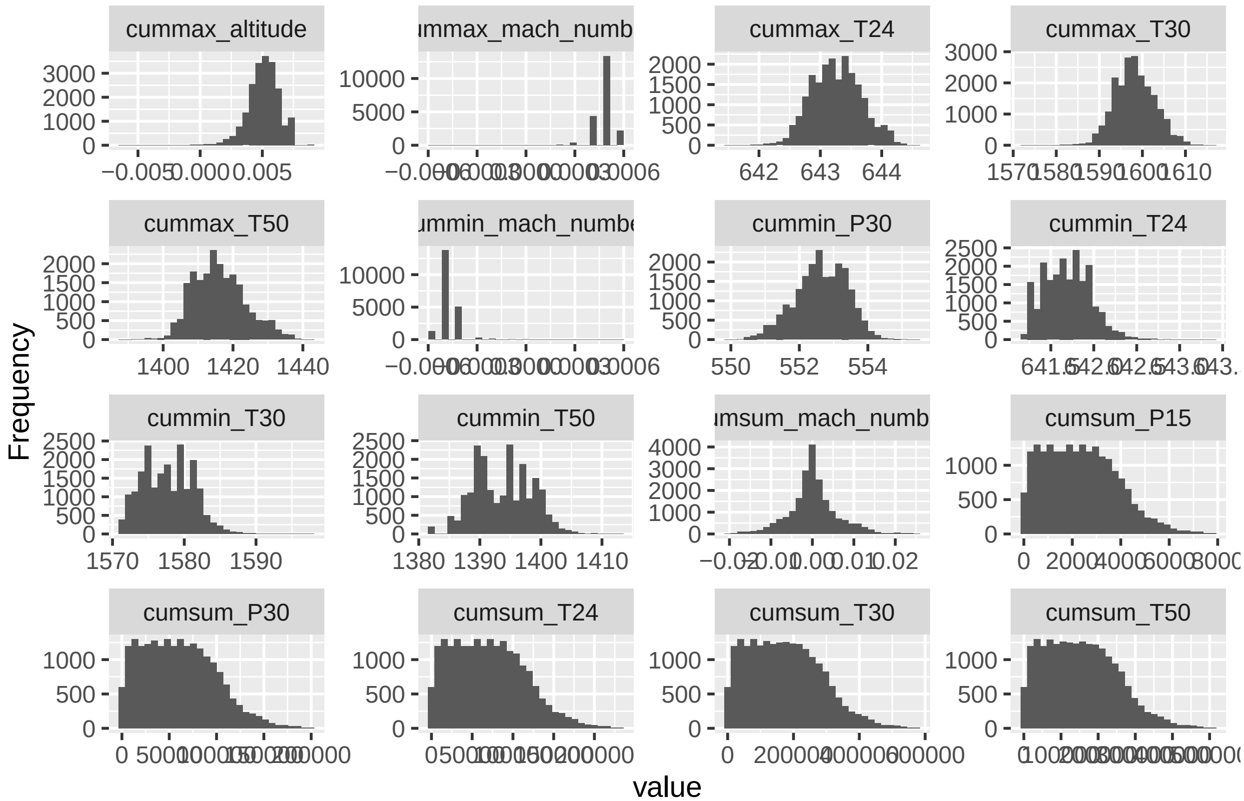
Page 19


rolliqr5_Nc
rolliqr5_Nf
rolliqr5_NRf
rolliqr5_phi
rolliqr5_Ps30
rollmean5_Nc
rollmean5_Nf
rollmean5_NRc
rollmean5_NRf
rollmean5_phi
rollmean5_Ps30
rollsd5_Nc
rollsd5_Nf
rollsd5_NRf
rollsd5_phi
rollsd5_Ps30
Frequency
value

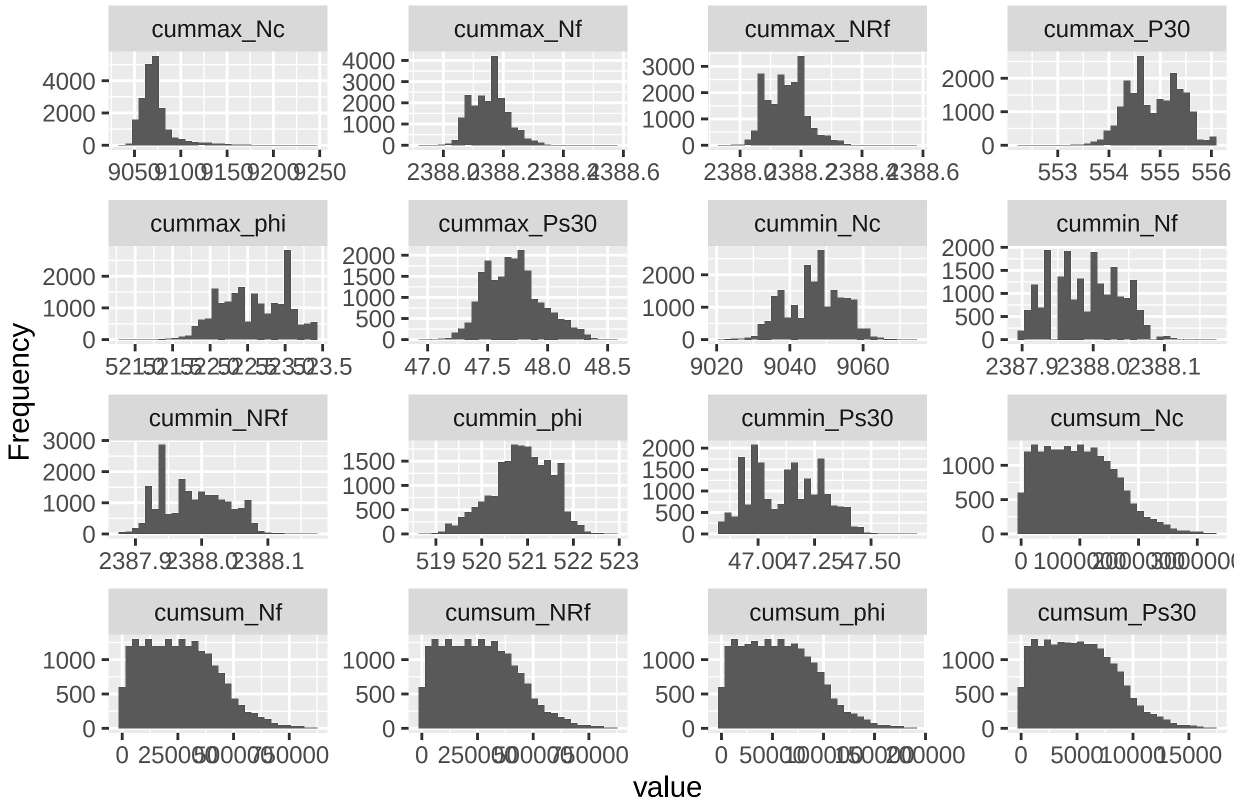
Page 20


cummin_altitude
cumsum_altitude
rolliqr5_BPR
rolliqr5_htBleed
rolliqr5_NRc
rolliqr5_W31
rolliqr5_W32
rollmean5_BPR
rollmean5_htBleed
rollmean5_W31
rollmean5_W32
rollsd5_BPR
rollsd5_htBleed
rollsd5_NRc
rollsd5_W31
rollsd5_W32
Frequency
value

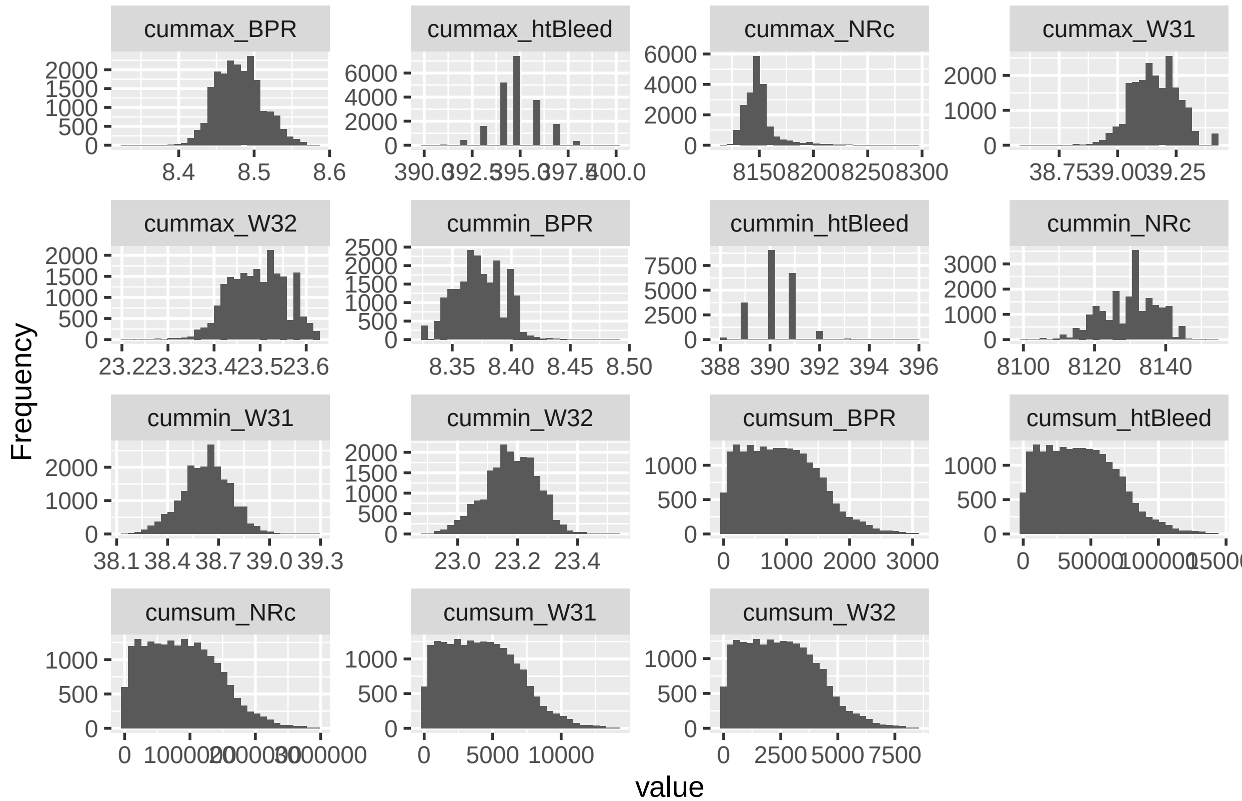
Page 21


cummax_altitude
ummax_mach_numb
cummax_T24
cummax_T30
cummax_T50
ummin_mach_numbe
cummin_P30
cummin_T24
cummin_T30
cummin_T50
imsum_mach_numb
cumsum_P15
cumsum_P30
cumsum_T24
cumsum_T30
cumsum_T50
Frequency
value

Page 22


cummax_Nc
cummax_Nf
cummax_NRf
cummax_P30
cummax_phi
cummax_Ps30
cummin_Nc
cummin_Nf
cummin_NRf
cummin_phi
cummin_Ps30
cumsum_Nc
cumsum_Nf
cumsum_NRf
cumsum_phi
cumsum_Ps30
Frequency
value

Page 23



Page 24

3. Limpieza de los datos

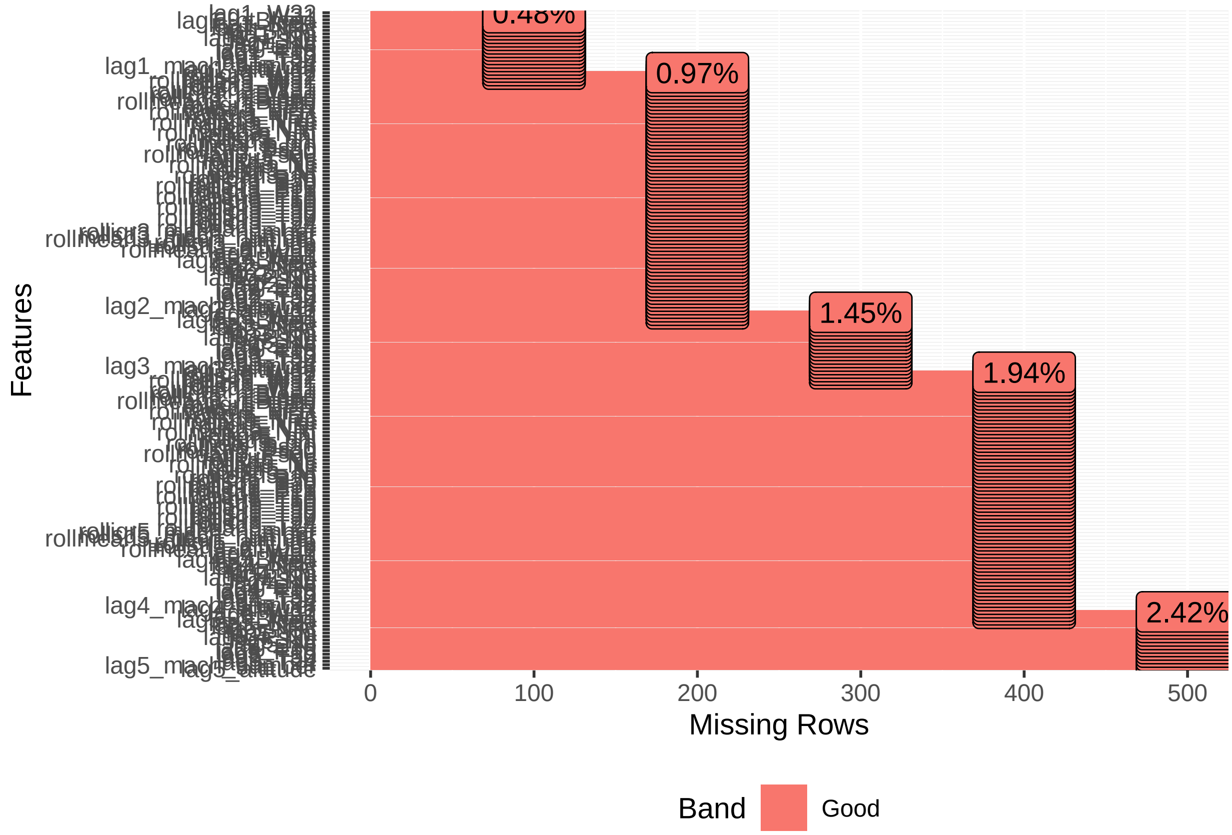
3.2. Identificación y tratamiento de valores extremos.

Nuestro dataset no tiene Missing values, pero hay valores extremos. Para

3.1. Missing Values o Ceros

Como tenemos muchas variables, filtramos las variables sin na y generamos un gráfica para conocer las columnas con NA y sus patrones.

```
columns_without_na <- colnames(trainset2[, sapply(trainset2, Negate(anyNA)), drop = FALSE])
df_columns_with_na <- trainset2[ , -which(names(trainset2) %in% columns_without_na)]
plot_missing(df_columns_with_na)
```



Como se esperaba, las columnas con lag y roll contienen NAs. Pero esto tiene que ser hasta un time_in_cycles máximo de 5. Vamos a confirmar esto:

```
df_columns_with_na$time_in_cycles <- trainset2$time_in_cycles

borra <- df_columns_with_na %>% filter_all(any_vars(is.na(.)))

table(borra$time_in_cycles)
```

```
##
##   1   2   3   4   5
## 100 100 100 100 100
```

Merece la pena hacer algún tipo de imputación de valor?

Para contestar a esta pregunta, vamos a ver si hay alguna máquina en nuestros datos que tenga fallado en este período:

```
trainset2 %>% filter(RUL == 0) %>% pull(time_in_cycles) %>% sort() %>% unique()
```

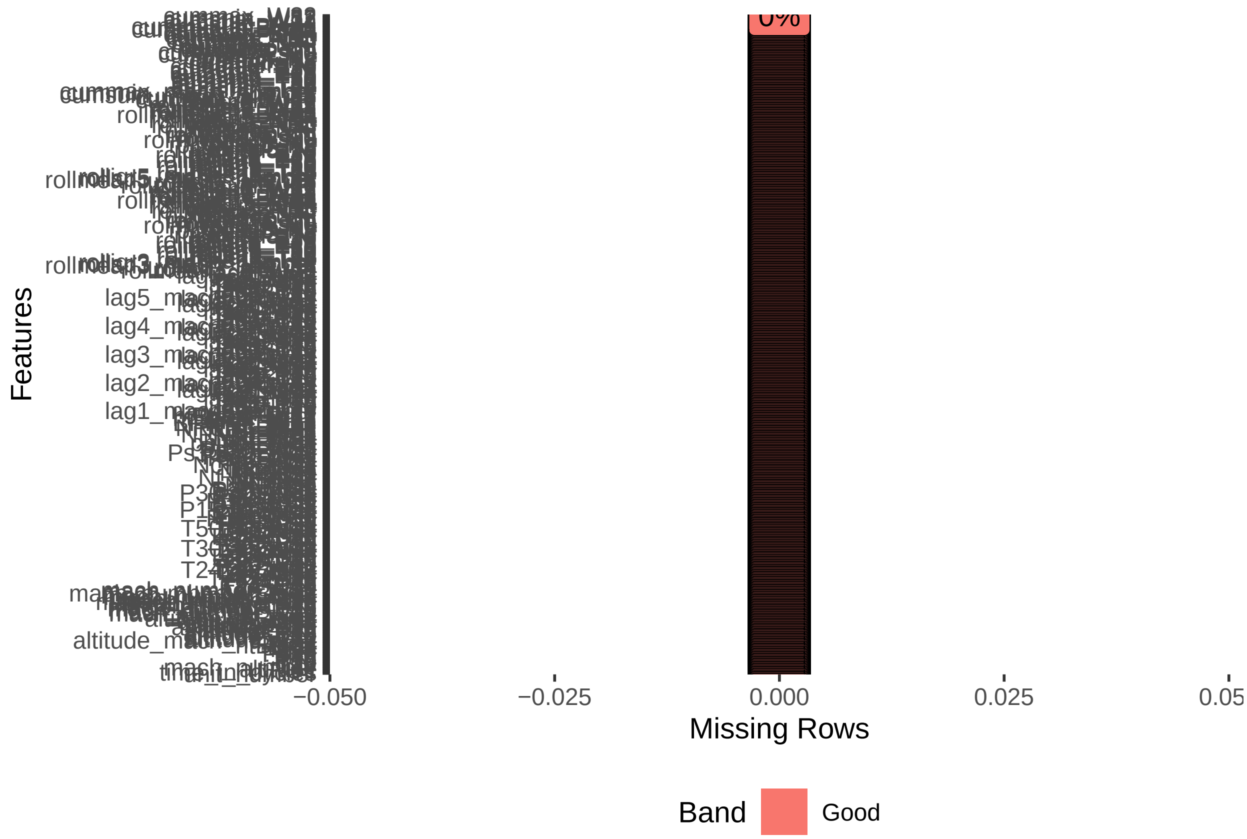
```
##  [1] 128 135 137 147 150 153 154 155 156 158 163 165 166 168 170 172 174 178 179
## [20] 180 181 185 188 189 191 192 193 194 195 196 198 199 200 201 202 207 208 209
## [39] 210 213 214 215 216 217 222 229 230 231 234 240 256 257 258 259 267 269 275
## [58] 276 278 283 287 293 313 336 341 362
```

```r
testset2 %>%
  group_by(unit_number) %>%
  filter(time_in_cycles == max(time_in_cycles)) %>%
  pull(time_in_cycles) %>%
  sort() %>%
  unique()
```

```
##  [1]  31  34  37  39  46  48  49  50  54  55  56  68  71  72  73  74  76  78  83
## [20]  88  89  94  97  98 101 105 106 110 112 113 121 123 125 126 130 131 133 134
## [39] 135 136 137 140 143 144 145 146 147 148 150 152 155 156 158 159 160 162 164
## [58] 165 166 168 171 172 176 177 184 186 187 189 192 195 196 198 203 205 213 217
## [77] 232 234 244 303
```

Como se puede ver no existe ninguna necesidad de usar las observaciones de 1 a 5. Por lo tanto, vamos a eliminar estas observaciones con NA sin perder ninguna información importante para nuestro proyecto.

```r
trainset2 <- trainset2 %>% na.omit()
testset2 <- testset2 %>% na.omit()

plot_missing(trainset2)
```



3.2. Identificación y tratamiento de valores extremos

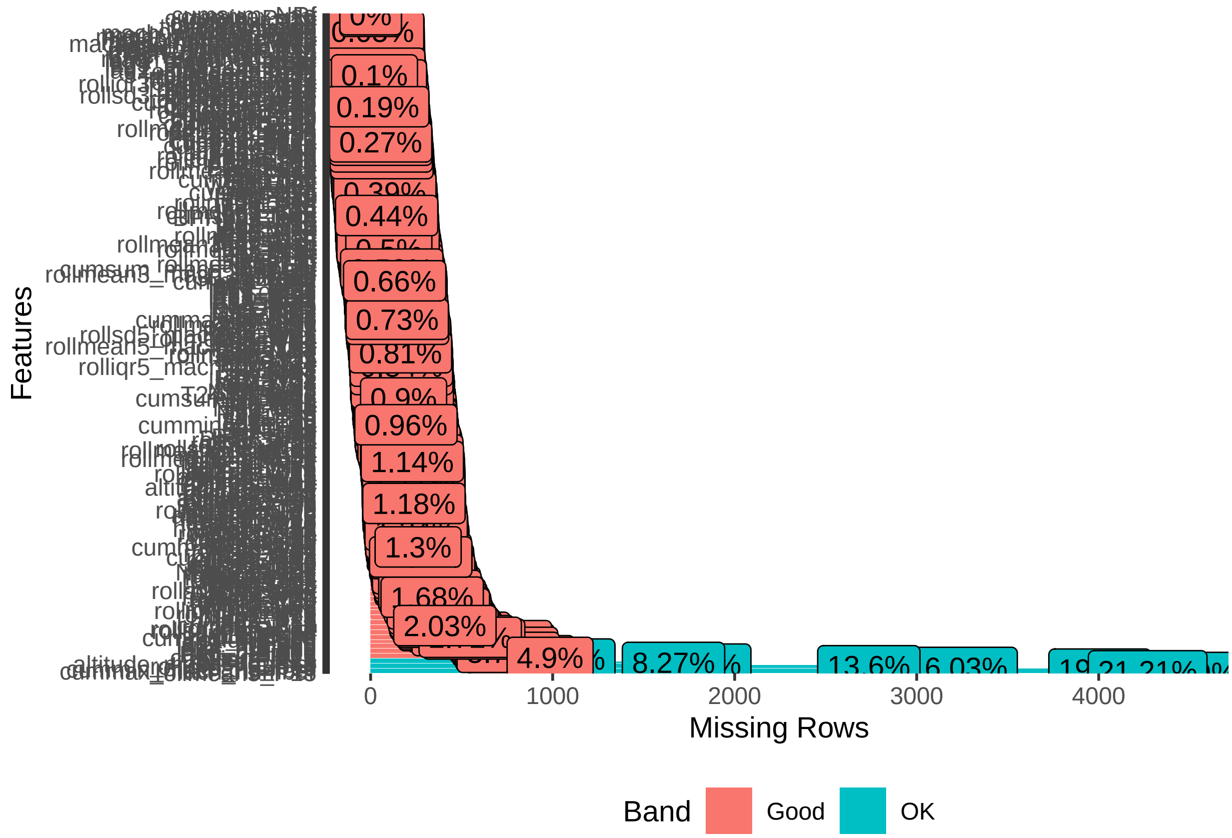
El tratamiento de valores extremos es importante porque estos pueden sesgar / cambiar drásticamente las estimaciones y predicciones de ajuste. Estos datos se encuentran muy alejados de la distribución normal de una variable o población.

Nuestros datos son basados en simulación de sensores. Esto significa a priori que no hay sospechas de que las observaciones que se desvían tanto del resto fueron generadas mediante un mecanismo distinto.

Estos datos son legítimos y en princípio no hace falta quitar de nuestra muestra. Pero podemos generar un algorítimo para quitar outliers basada en un criterio univariable y otros multivariable.

Nc numeric NRc numeric P15 numeric Nf numeric NRf numeric T30 numeric

```r
remove_outliers <- function(x, na.rm = TRUE, ...) {
  qnt <- quantile(x, probs=c(.25, .75), na.rm = na.rm, ...)
  H <- 1.5 * IQR(x, na.rm = na.rm)
  y <- x
  y[x < (qnt[1] - H)] <- NA
  y[x > (qnt[2] + H)] <- NA
  y
}

temp <- trainset2 %>%
  group_by(time_in_cycles) %>%
  mutate_at(vars(-group_cols()), remove_outliers)

plot_missing(temp)
```



```r
columns = c("altitude", "mach_number", "T24", "T30", "T50", "P15", "P30", "Nf", "Nc", "Ps30", "phi", "NR


OutlierAnomalyDetection <- function(df, columns = "all", method = "all"){
  if(any(class(df) == "grouped_df")){
    df <- df %>% ungroup()
  }
  if(any(columns == "all")){
    df_temp <- df
  } else {
    df_temp <- df %>% dplyr::select(one_of(columns))
  }

  if(method %in% c("all", "cook")){
    mod <- lm(RUL ~., data = df)
    cooksd <- cooks.distance(mod)
    #influential <- as.numeric(names(cooksd)[(cooksd > (4/sample_size))])
    #df <- df[-influential, ]
    df$cook_score <- as.vector(cooksd)
  }
  if(method %in% c("all", "mahalanobis")){
    score <- as.vector(mahalanobis(df_temp, colMeans(df_temp), cov(df_temp)))
    df$mahalanobis_score <- score
  }
  if(method %in% c("all", "PCA")){
    # Cálculo de PCA
    pca <- prcomp(
         x = df_temp,
         center = TRUE,
         scale. = TRUE
       )
    comp <- seq_along(pca$sdev)
    recon <- as.matrix(pca$x[, comp]) %*% t(as.matrix(pca$rotation[, comp]))
    # Se revierte la trasformación centrado y escalado.
    recon <- scale(recon , center = FALSE, scale = 1/pca$scale)
    recon <- scale(recon , center = -1*pca$center, scale = FALSE)
    # Cálculo del error cuadrático medio de recostrucción
    reconstruction_error <- apply(X = (recon - df_temp)^2, MARGIN = 1, FUN = mean)
    # Se añade el error de reconstrucción al dataframe original.
    df$reconstruction_error_pca <- reconstruction_error
  }
  if(method %in% c("all", "knn")){
    #
    df_scaled <- scale(df_temp)
    df_knn <- get.knn(data = df_scaled, k = 5)
    # Averagedistancetonearestneighbors
    df_score <- rowMeans(df_knn$nn.dist)
    df$knn_score <- df_score

  }
  if(method %in% c("all", "dbscan")){
    #
    df$dbscan_score <- lof(scale(df_temp), k = 5)
```

```r
}
#if(method %in% c("all", "oneclasssvm")){
#  replications = 5
#  # function to obtain R-Squared from the data
#  oneclasssvm <- function(data, indices) {
#    x <- data[indices,] # allows boot to select sample
#    y <- rep("oneclass", nrow(x))
#    model <- svm(x,type='one-classification') #train an one-classification model
#    # the output is Boolean
#    pred <- predict(model, x) #create predictions
#    return(pred)
#    }
#  # bootstrapping
#  results <- boot(
#    data=df_temp,
#    statistic=oneclasssvm,
#    R=replications,
#    parallel = "multicore",
#    ncpus = 5
#    )
#  temp <- as.data.frame(results$t)
#  df$oneclasssvm_score <- as.vector(colSums(temp))/replications
#}
if(method %in% c("all", "isolationforest")){
  #
  df_tree <- iForest(X=df_temp, nt = 100)
  # Scores near 1 indicate anomalies (small path length)
  df_score <- predict(df_tree, newdata = df_temp)
  df$isolationforest_score <- df_score

}
if(method %in% c("all", "autoencoder")){
  h2o.init(ip = "localhost",
       # Si emplea un único core los resultados son reproducibles.
       # Si se dispone de muchos datos, mejor emplear varios cores.
       nthreads = 1
       )
  # Se eliminan los datos del cluster por si ya había sido iniciado.
  h2o.removeAll()
  h2o.no_progress()
  df_h2o <- as.h2o(
              x = df_temp,
              destination_frame = "df_h2o"
          )
  # División de las observaciones en conjunto de entrenamiento y validación.
  df_h2o_split <- h2o.splitFrame(data = df_h2o, ratios = 0.8, seed = 123)
  df_h2o_train <- df_h2o_split[[1]]
  df_h2o_validacion <- df_h2o_split[[2]]
  # Se define la variables empleadas por el autoencoder
  predictores  <- setdiff(h2o.colnames(df_h2o), "y")
  # Entrenamiento del modelo autoencoder con 2 neuronas en la capa oculta.
  autoencoder <- h2o.deeplearning(
    x = predictores,
```

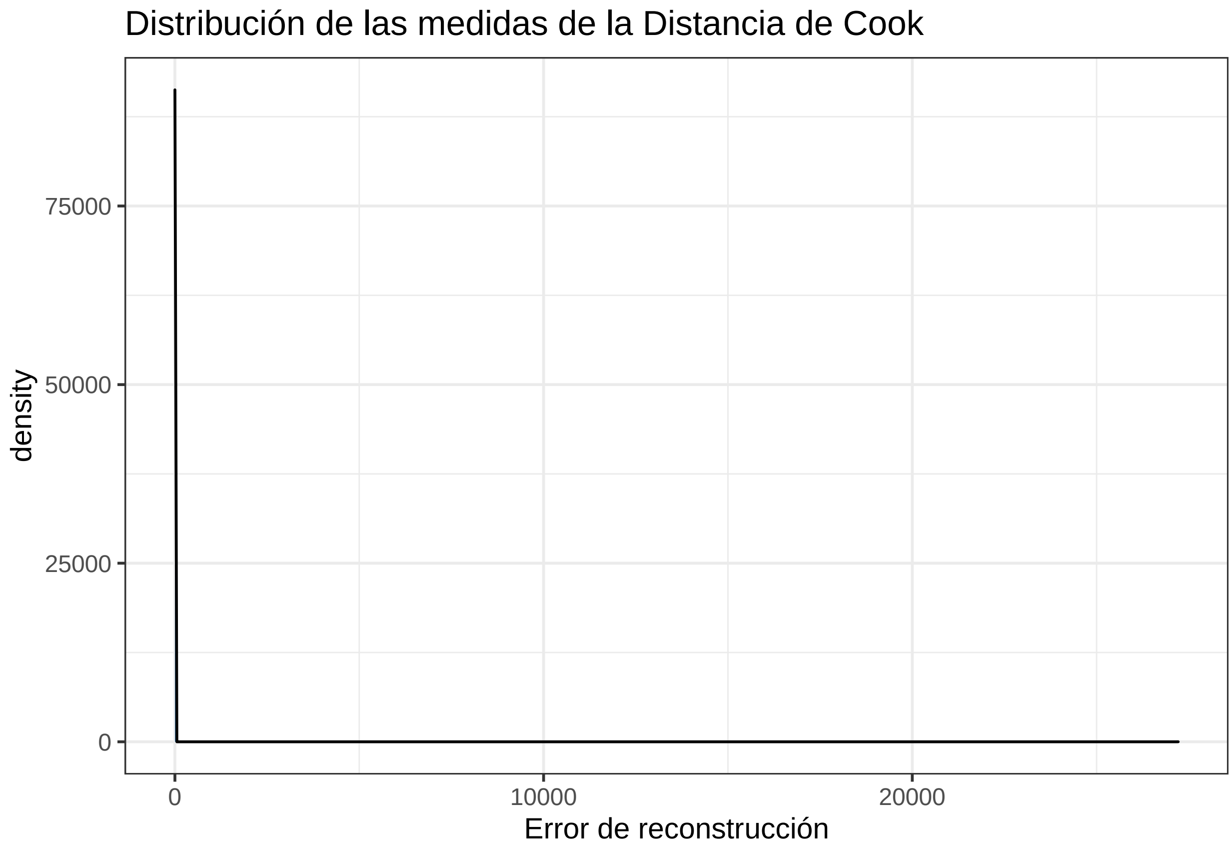
```r
      training_frame   = df_h2o_train,
      validation_frame = df_h2o_validacion,
      activation     = "Tanh",
      autoencoder    = TRUE,
      hidden         = c(2),
      epochs         = 50,
      ignore_const_cols    = FALSE,
      score_each_iteration = TRUE,
      # Seed solo válido cuando se emplea un único core
      seed = 999
      )

    reconstruction_error <- h2o.anomaly(
                              object = autoencoder,
                              data   = df_h2o[, predictores],
                              per_feature = FALSE
                            )
    reconstruction_error <- as.data.frame(reconstruction_error)
    reconstruction_error <- reconstruction_error[, 1]
    # Se añade el error de reconstrucción al dataframe original.
    df$anomaly_score_autoencoder <- reconstruction_error

  }

  return(df)
}

temp <- OutlierAnomalyDetection(df = trainset2, columns, method = "all")
```

```
##  Connection successful!
##
## R is connected to the H2O cluster:
##     H2O cluster uptime:         47 minutes 11 seconds
##     H2O cluster timezone:       Europe/Paris
##     H2O data parsing timezone:  UTC
##     H2O cluster version:        3.32.0.1
##     H2O cluster version age:    2 months and 28 days
##     H2O cluster name:           H2O_started_from_R_ferna_wkw207
##     H2O cluster total nodes:    1
##     H2O cluster total memory:   7.07 GB
##     H2O cluster total cores:    8
##     H2O cluster allowed cores:  1
##     H2O cluster healthy:        TRUE
##     H2O Connection ip:          localhost
##     H2O Connection port:        54321
##     H2O Connection proxy:       NA
##     H2O Internal Security:      FALSE
##     H2O API Extensions:         Amazon S3, Algos, AutoML, Core V3, TargetEncoder, Core V4
##     R Version:                  R version 4.0.3 (2020-10-10)
```

```r
df_temp <- temp %>% ungroup() %>% dplyr::select(cook_score, mahalanobis_score, reconstruction_error_pca
```

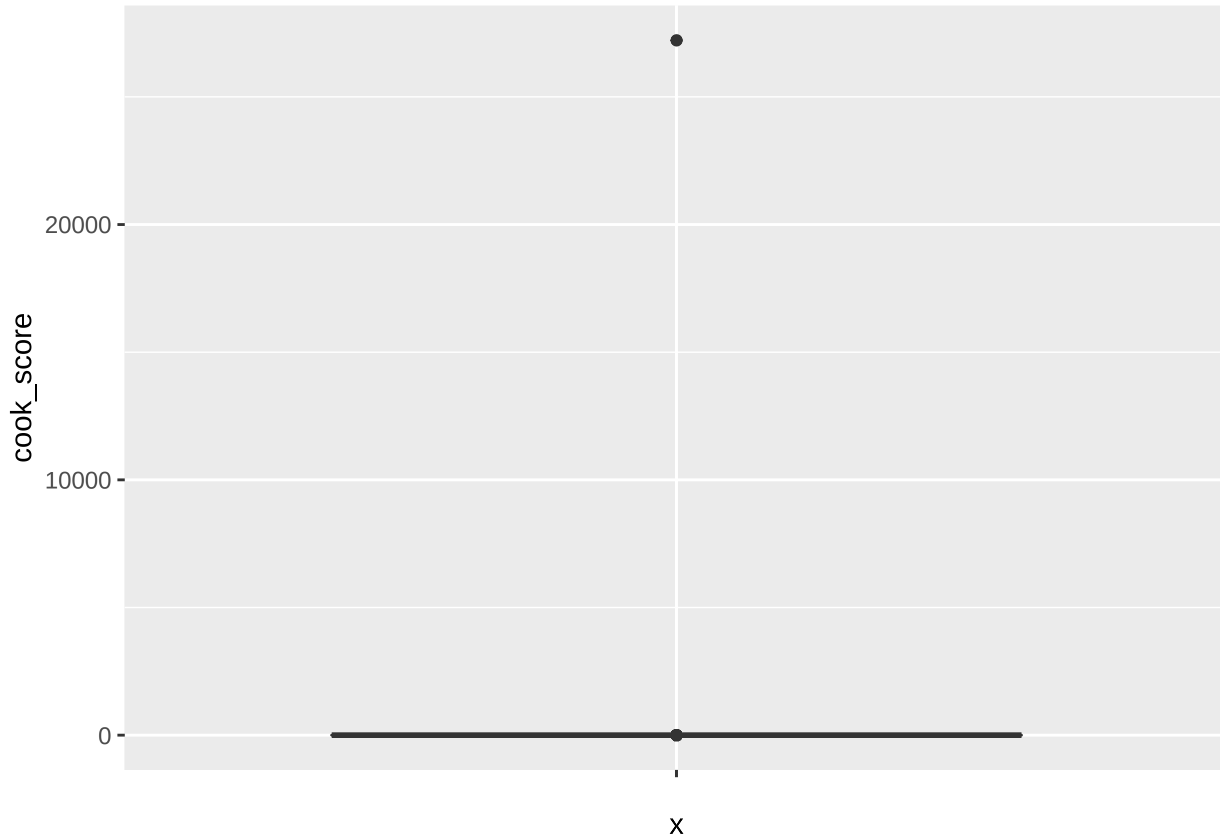
```r
# Compute a correlation matrix
corr <- round(cor(df_temp), 1)
# Compute a matrix of correlation p-values
p.mat <- cor_pmat(df_temp)

# Visualize the lower triangle of the correlation matrix
# Barring the no significant coefficient
corr.plot <- ggcorrplot(
  corr, hc.order = TRUE, type = "lower", lab = TRUE, outline.col = "white",
  p.mat = p.mat
  )

corr.plot
```


```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = cook_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Distribución de las medidas de la Distancia de Cook",
       x = "Error de reconstrucción") +
  theme_bw()
```



Según la distribución de la distancia de Cook, podemos ver claramente que existen outliers en los valores más grandes.

```
ggplot(data = df_temp, aes(x = "", y = cook_score)) +
  geom_boxplot() #+
```




```
#  coord_cartesian(ylim = c(0, 150)) # I set the y axis scale so the plot looks better.
```

Podemos aplicar la prueba de Grubbs. La prueba de Grubbs detecta un valor atípico a la vez (valor más alto o más bajo), por lo que las hipótesis nula y alternativa son las siguientes:

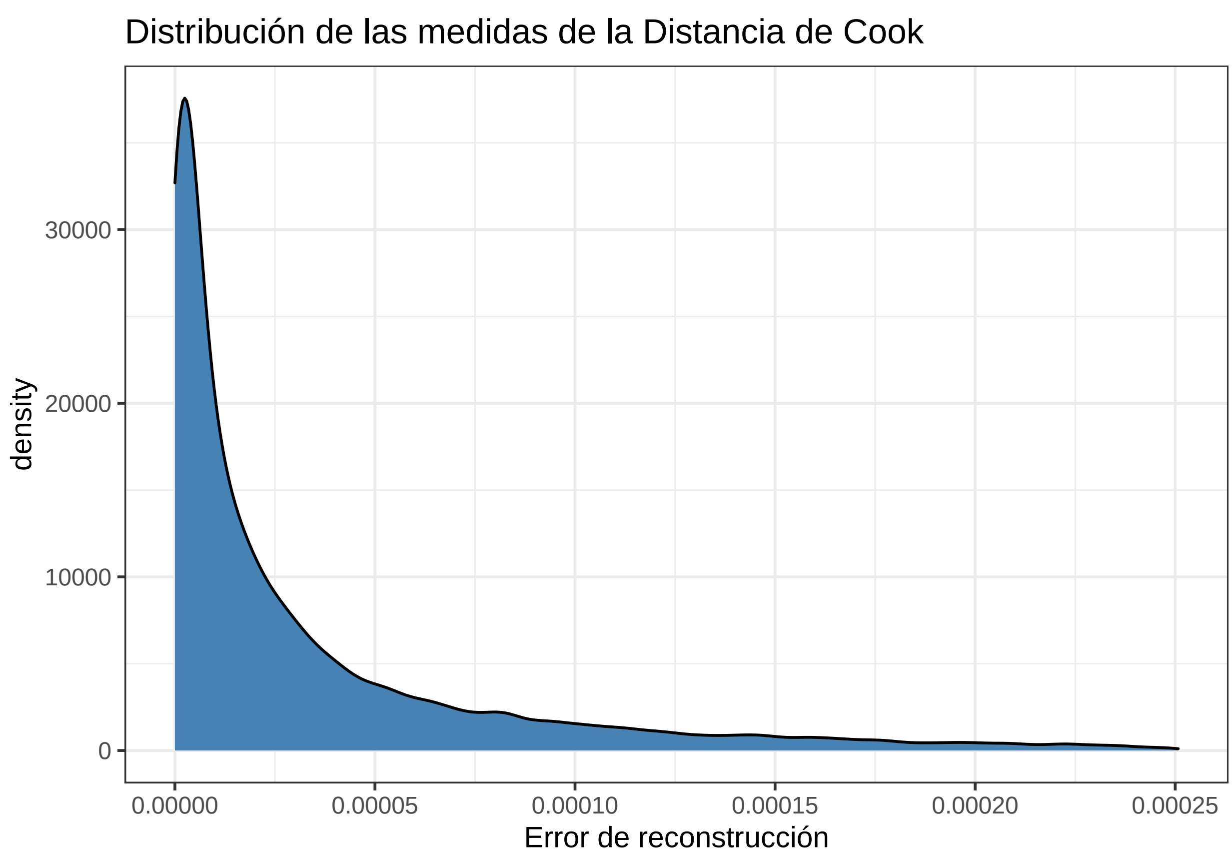
- H_0: El valor más alto **no** es un outlier
- H_1: El valor más alto es un outlier

El valor de p $< 0{,}05$ significa que al nivel de significancia del 5%, rechazamos la hipótesis de que el valor más no es un valor atípico.

Abajo generamos un loop dónde evaluamos en cada turno el valor más grande, si es outliers, se lo eliminamos y aplicamos otra vez la prueba hasta que no exista ningún outlier según Grubbs.

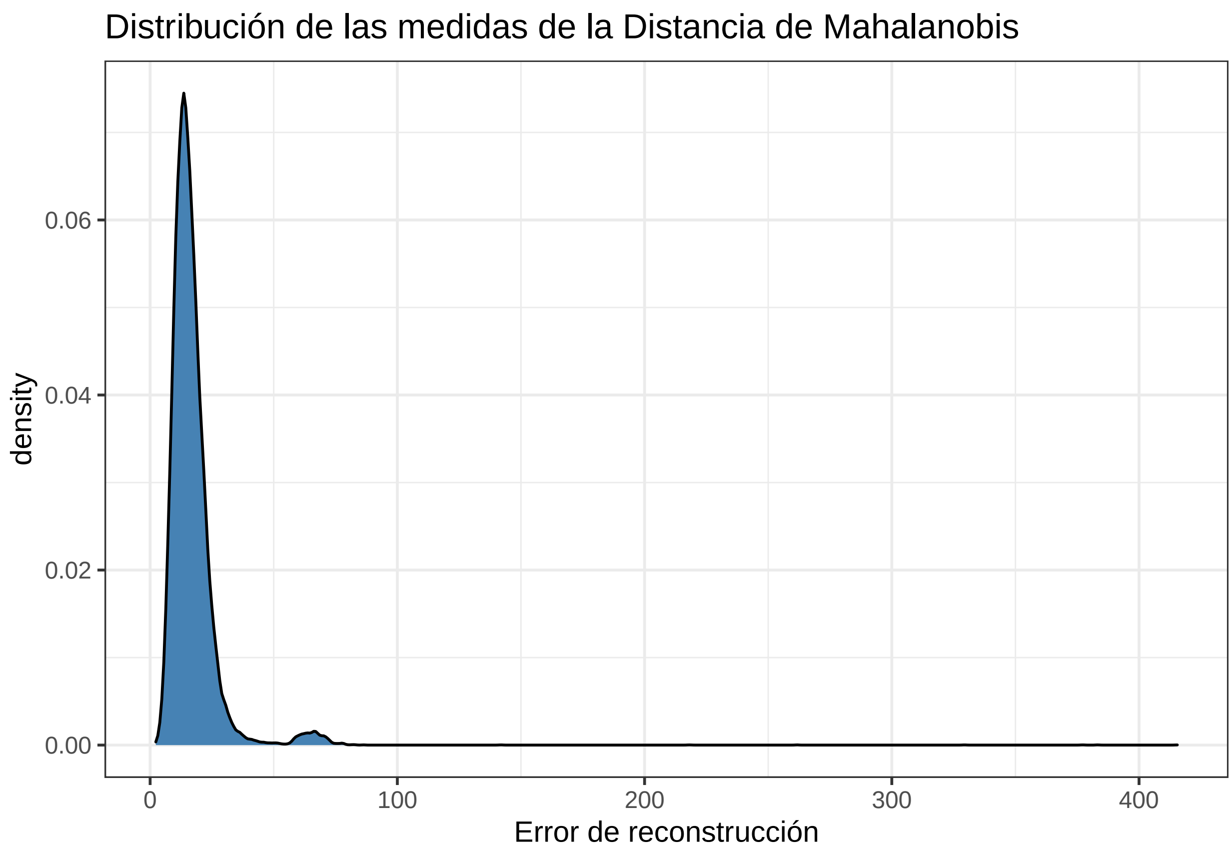
```
actual_p_value <- 0
p_value <- 0.05
outliers <- c()
df_no_cook_outlier <- df_temp

while (actual_p_value < p_value) {

  grubbs_test <- grubbs.test(df_no_cook_outlier$cook_score)
  actual_p_value <- grubbs_test$p.value
  value <- gsub(pattern = "highest value ", "", x = grubbs_test$alternative)
  value <- as.numeric(gsub(pattern = " is an outlier", "", x = value))
```

```r
  outliers <- c(outliers, value)
  if(actual_p_value < p_value){
    df_no_cook_outlier <- df_no_cook_outlier %>% filter(cook_score < max(cook_score))
    #print(max(df_no_cook_outlier$cook_score))
    #print(paste0("The value ", value, " is an outlier and it's been removed."))
  } else {
    break
  }

}


tibble(df_no_cook_outlier) %>%
  ggplot(aes(x = cook_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Distribución de las medidas de la Distancia de Cook",
       x = "Error de reconstrucción") +
  theme_bw()
```

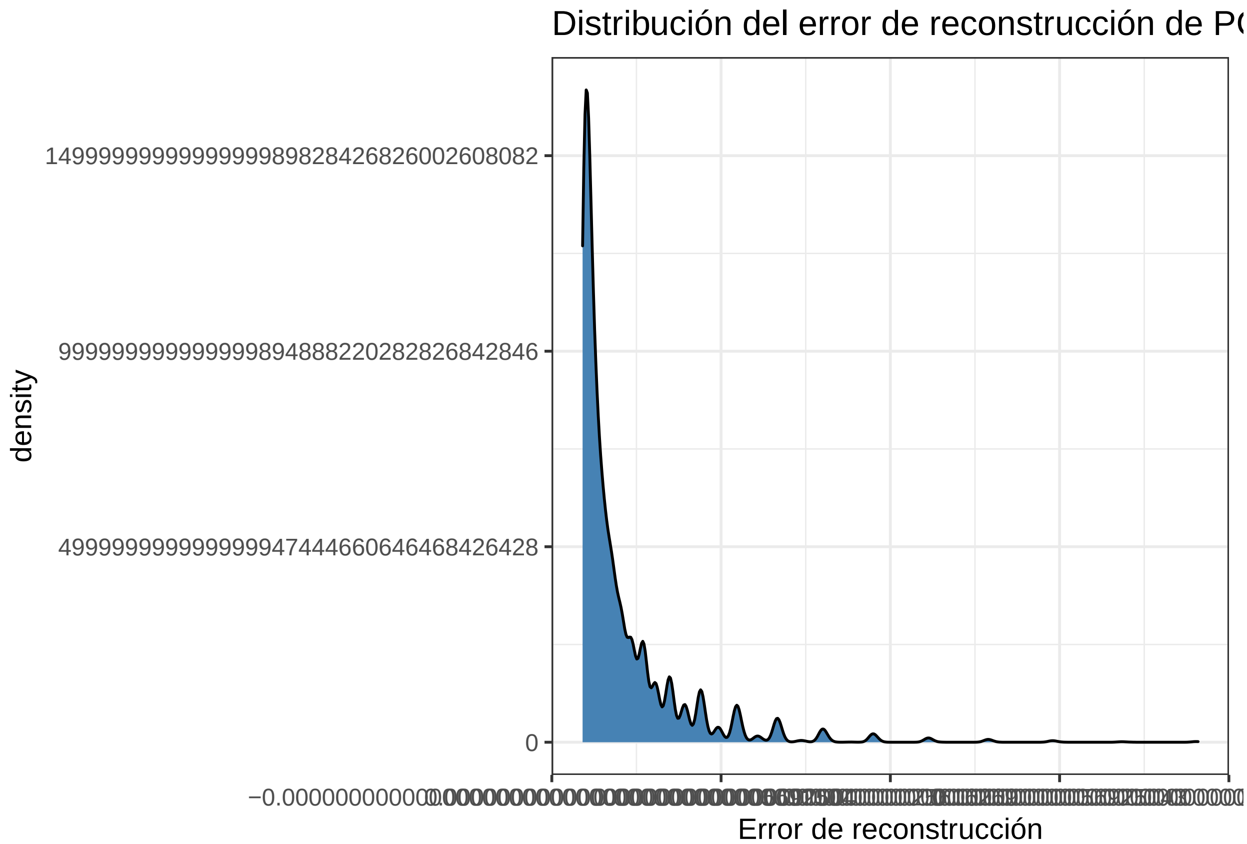



```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = mahalanobis_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Distribución de las medidas de la Distancia de Mahalanobis",
       x = "Error de reconstrucción") +
```

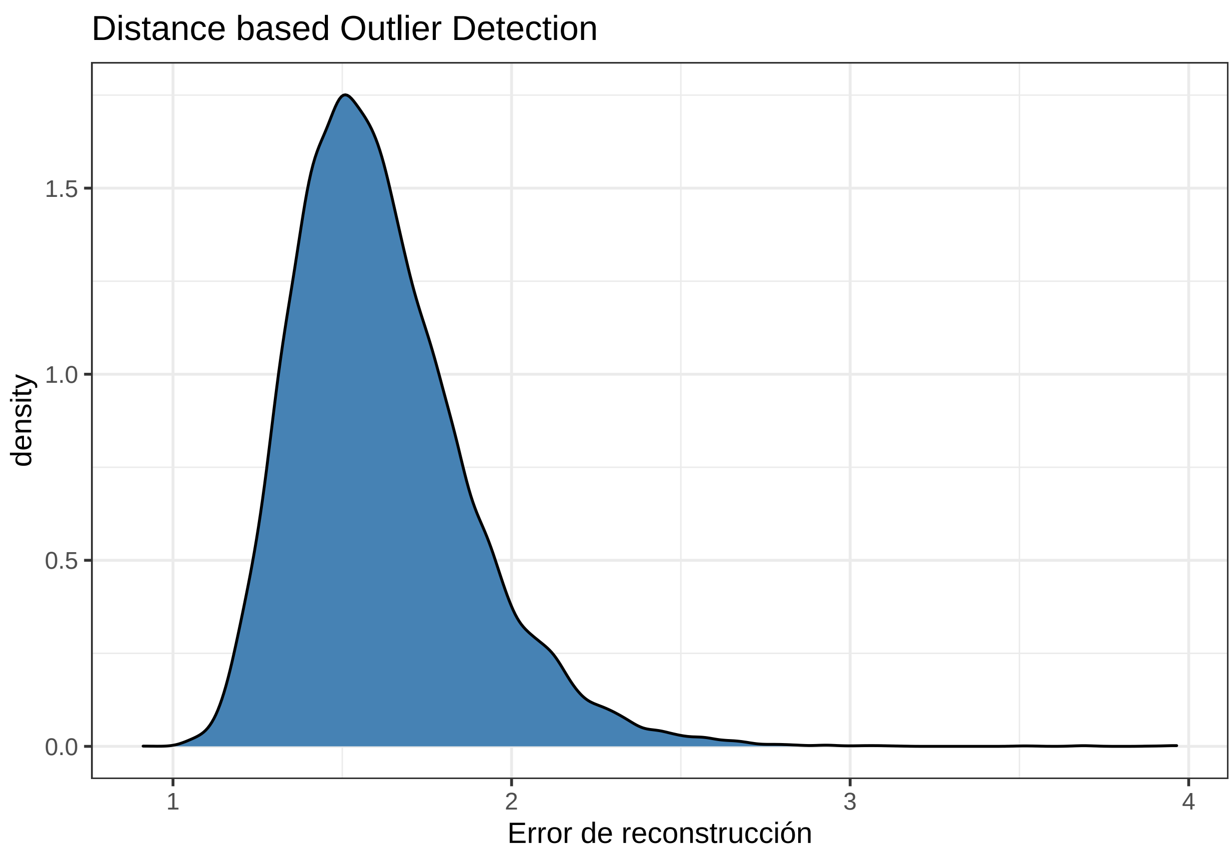
```
  theme_bw()
```



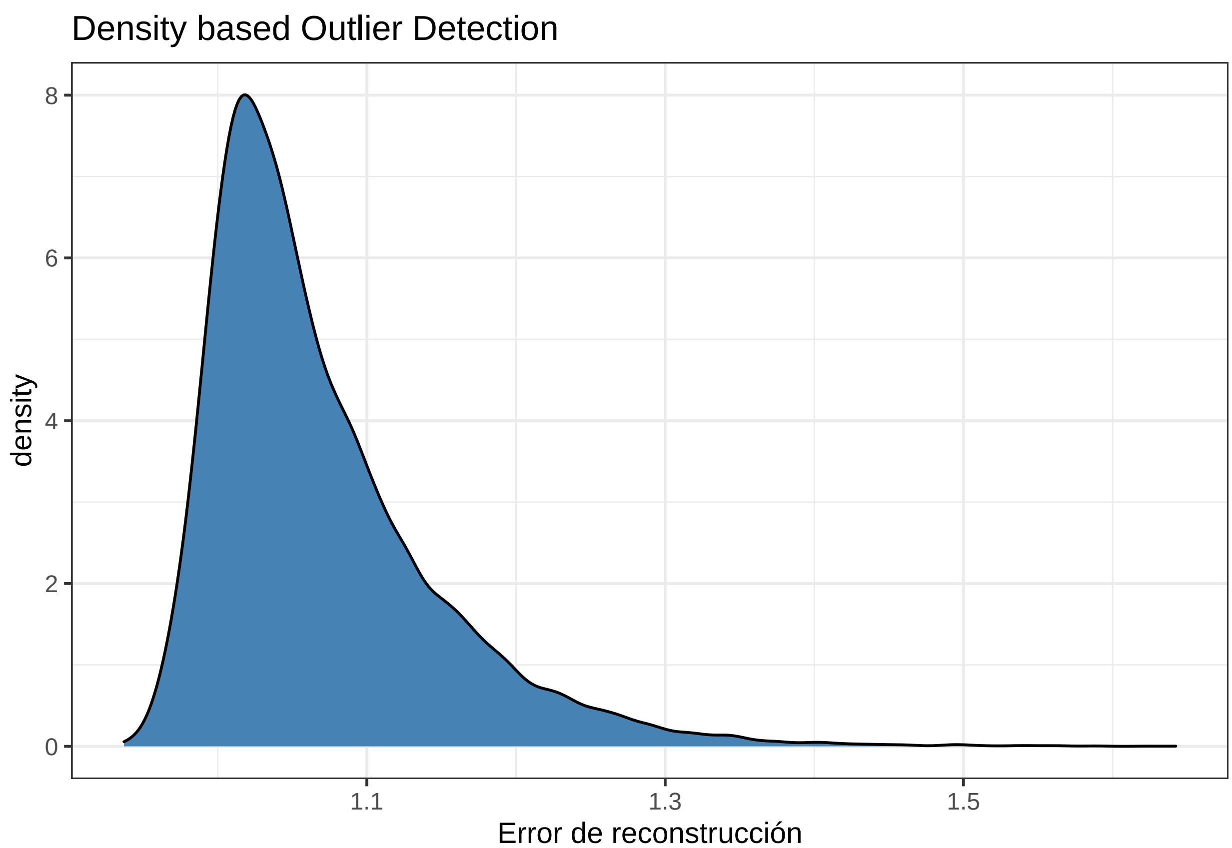

```
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = reconstruction_error_pca)) +
  geom_density(fill = "steelblue") +
  labs(title = "Distribución del error de reconstrucción de PCA",
       x = "Error de reconstrucción") +
  theme_bw()
```



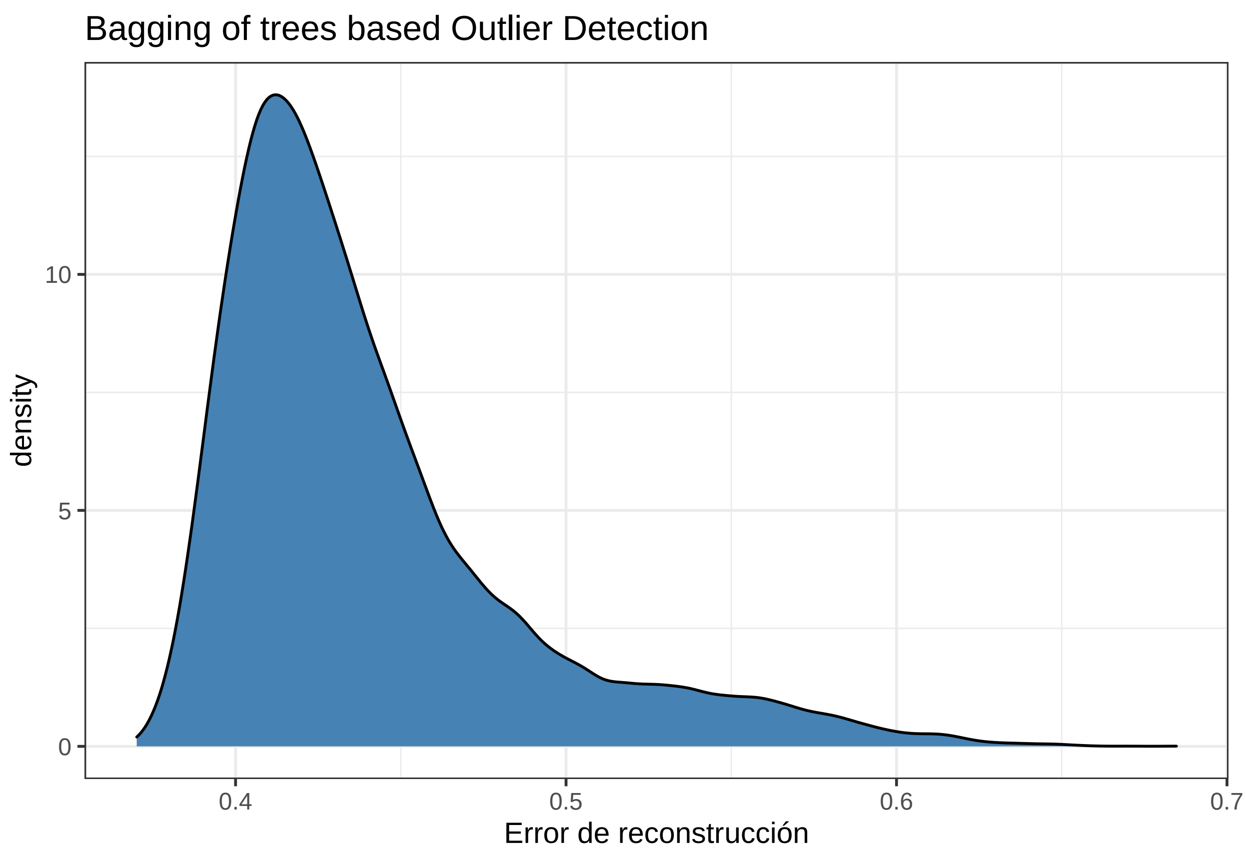

```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = knn_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Distance based Outlier Detection",
       x = "Error de reconstrucción") +
  theme_bw()
```



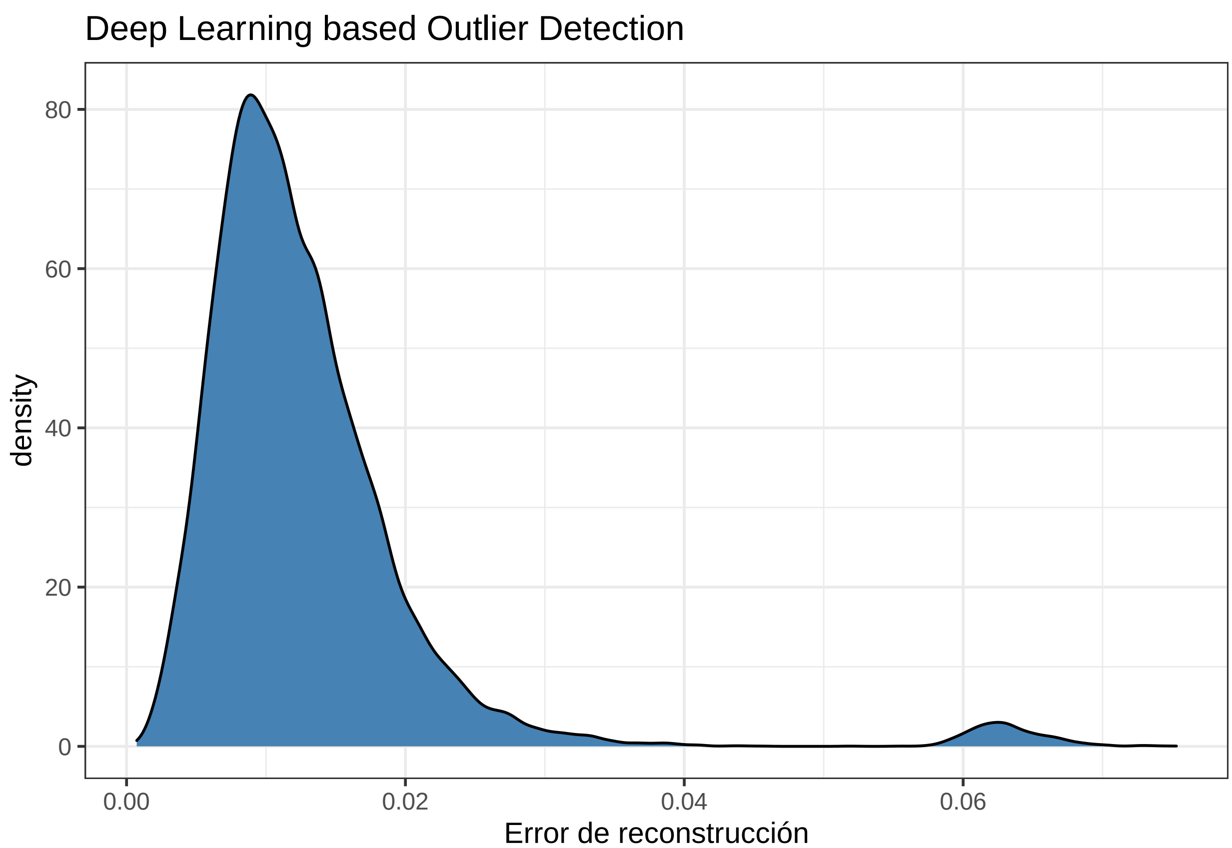

```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = dbscan_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Density based Outlier Detection",
       x = "Error de reconstrucción") +
  theme_bw()
```




```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = isolationforest_score)) +
  geom_density(fill = "steelblue") +
  labs(title = "Bagging of trees based Outlier Detection",
       x = "Error de reconstrucción") +
  theme_bw()
```




```r
# Distribución del error de reconstrucción.
tibble(df_temp) %>%
  ggplot(aes(x = anomaly_score_autoencoder)) +
  geom_density(fill = "steelblue") +
  labs(title = "Deep Learning based Outlier Detection",
       x = "Error de reconstrucción") +
  theme_bw()
```



4. Análisis de los datos

4.1. Selección de los grupos de datos que se quieren analizar/comparar (planificación de los análisis a aplicar).

Nuestro análisis se centrará en las máquinas cuando estas fallaron (RUL = 0). Además vamos a examinar dos grupos distintos, uno son las máquinas que fallaron con menos de 200 vuelos y el otro grupo son las máquinas con 200 o más vuelos.

```
df <- trainset2 %>%
  ungroup() %>%
  filter(RUL == 0) %>%
  mutate(
    label = if_else(time_in_cycles < 200, "bad", "good")
  ) %>%
  dplyr::select(-RUL, -time_in_cycles, -unit_number)



table(df$label)
```

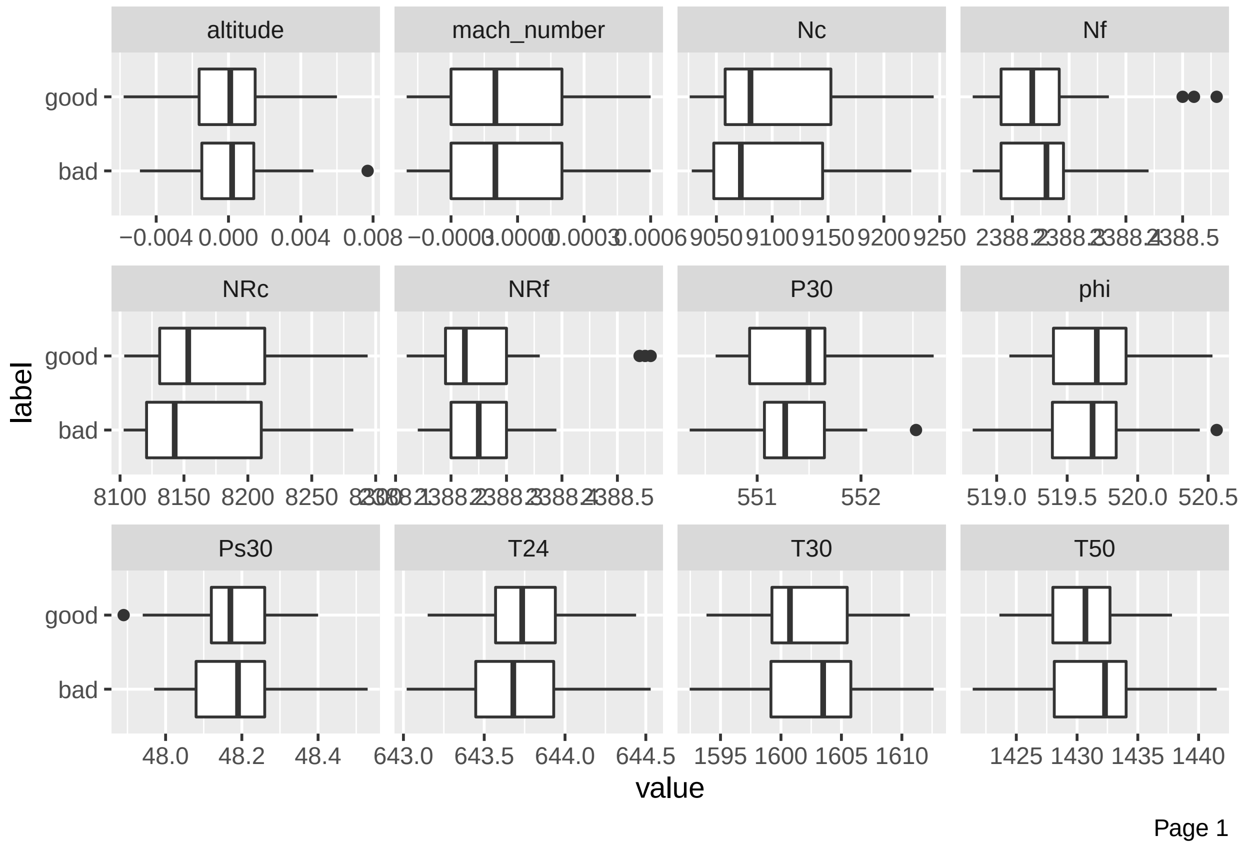
```
##
##  bad good
##   52   48
```

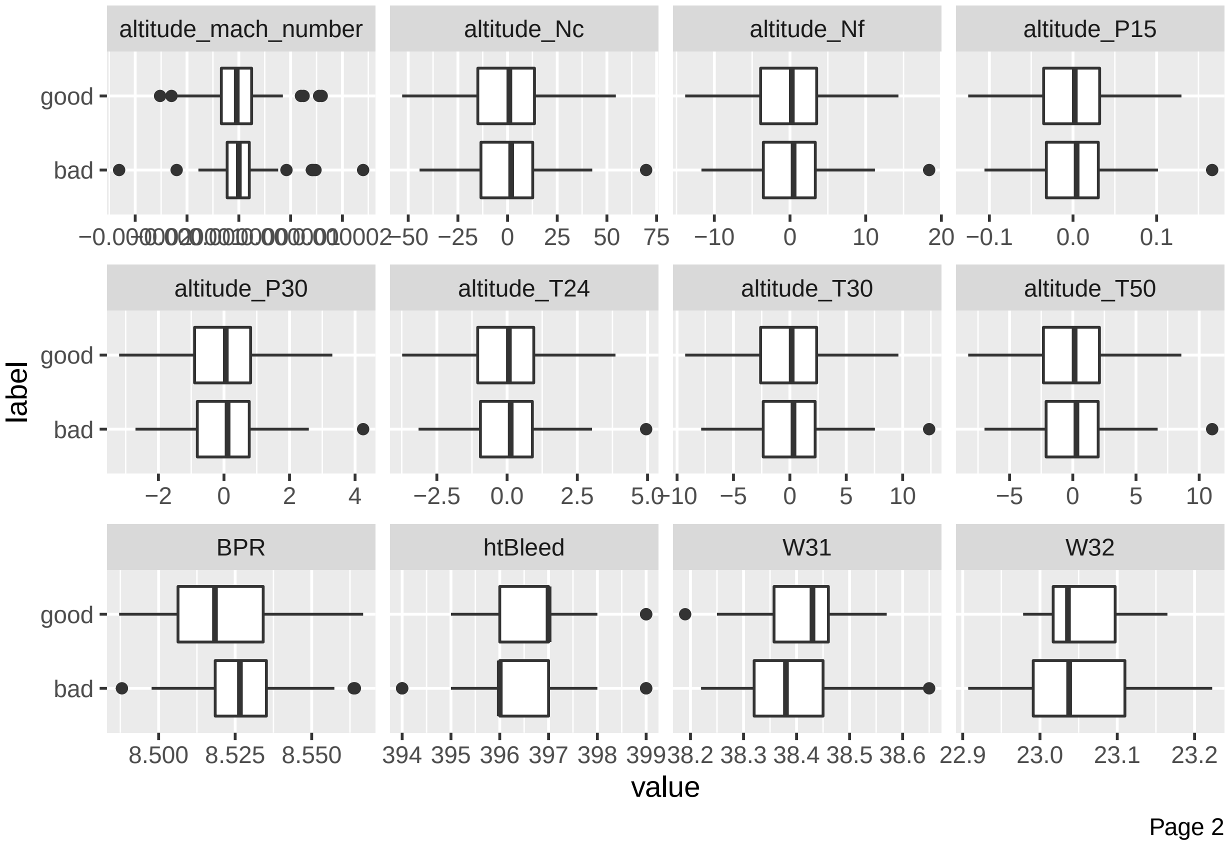
```r
temp <- DataProfiling(df)

unique_columns <- temp$names[temp$n_unique == 1 & !is.na(temp$n_unique)]
print(unique_columns)
```

```
##  [1] "P15"          "lag1_P15"     "lag2_P15"     "lag3_P15"     "lag4_P15"
##  [6] "lag5_P15"     "rollsd3_P15"  "rolliqr3_P15" "rollsd5_P15"  "rolliqr5_P15"
## [11] "cummax_P15"
```

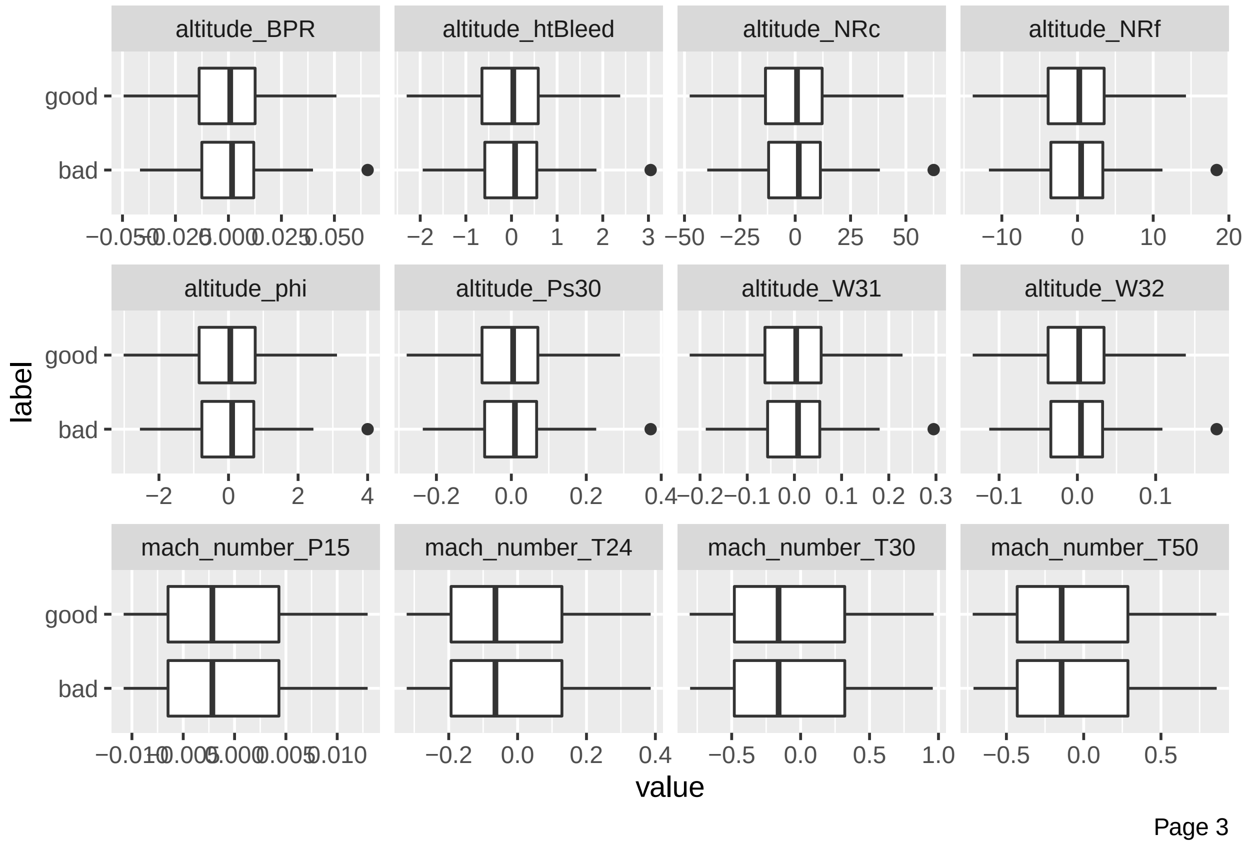
```r
df <- df %>% dplyr::select(-one_of(unique_columns))
```

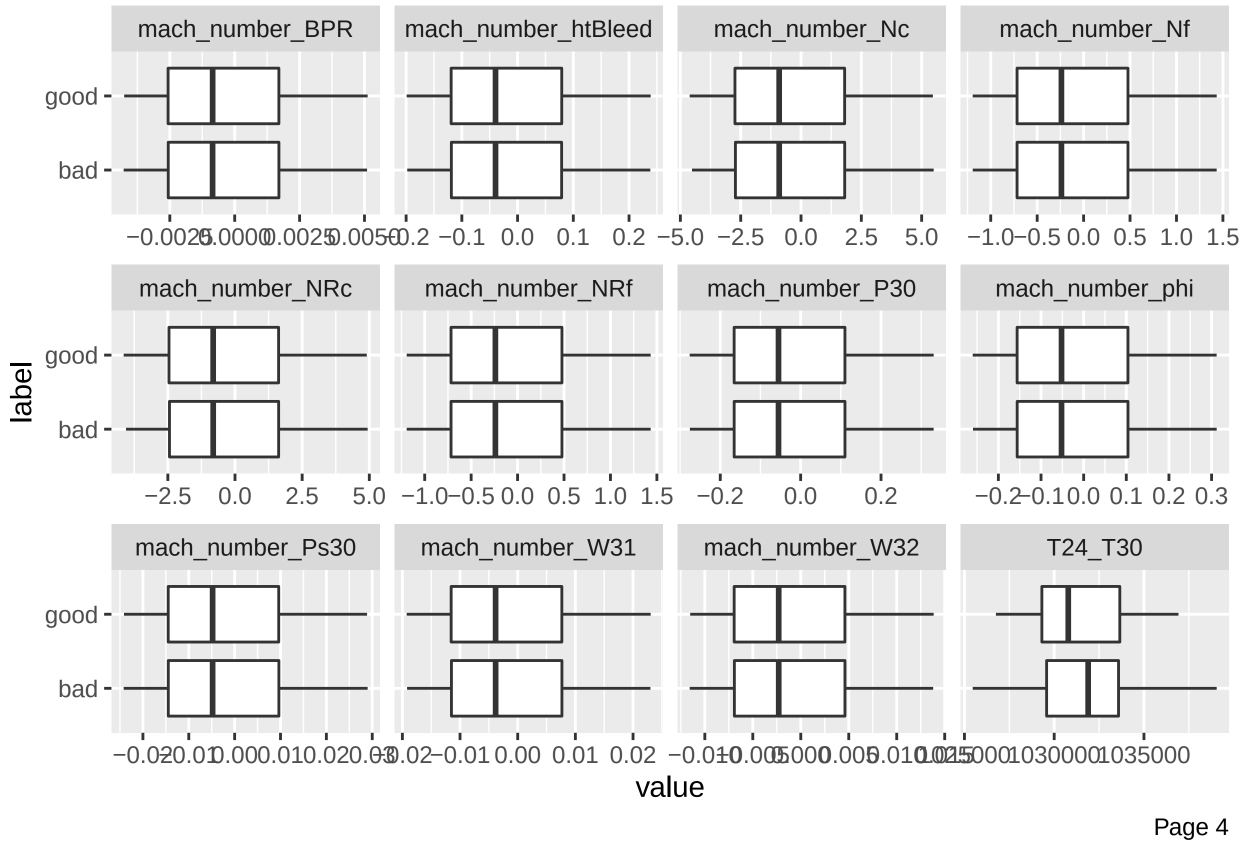
```r
plot_boxplot(df, by = "label")
```

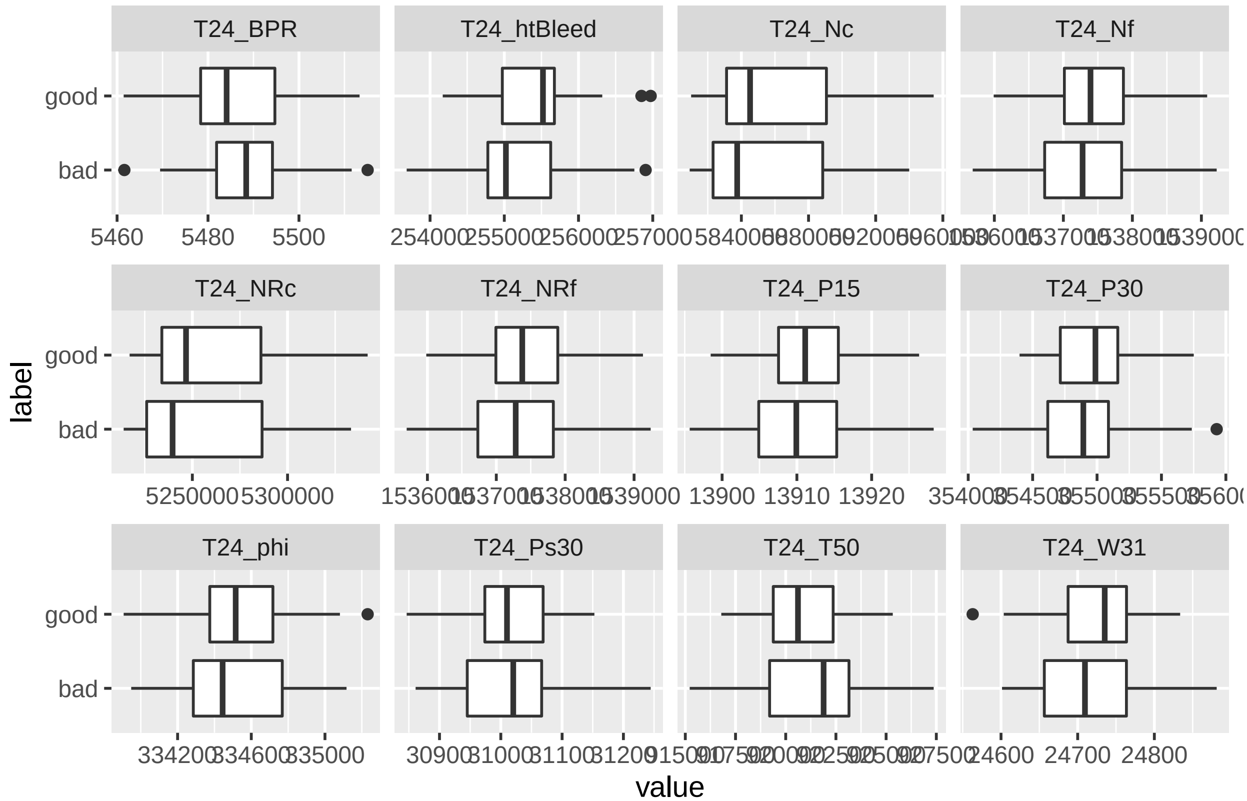




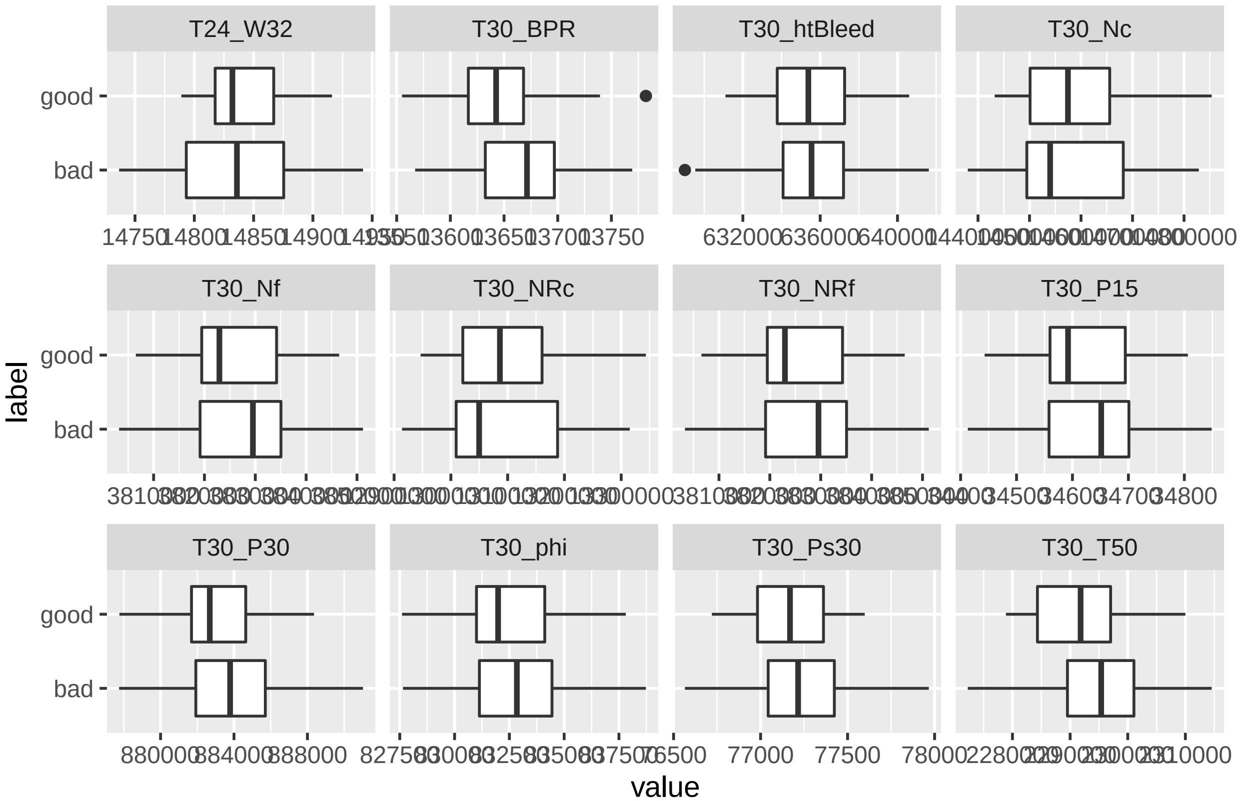
altitude_mach_number
altitude_Nc
altitude_Nf
altitude_P15
altitude_P30
altitude_T24
altitude_T30
altitude_T50
BPR
htBleed
W31
W32
good
bad
label
value
Page 2


altitude_BPR
altitude_htBleed
altitude_NRc
altitude_NRf
altitude_phi
altitude_Ps30
altitude_W31
altitude_W32
mach_number_P15
mach_number_T24
mach_number_T30
mach_number_T50
good
bad
label
value
Page 3

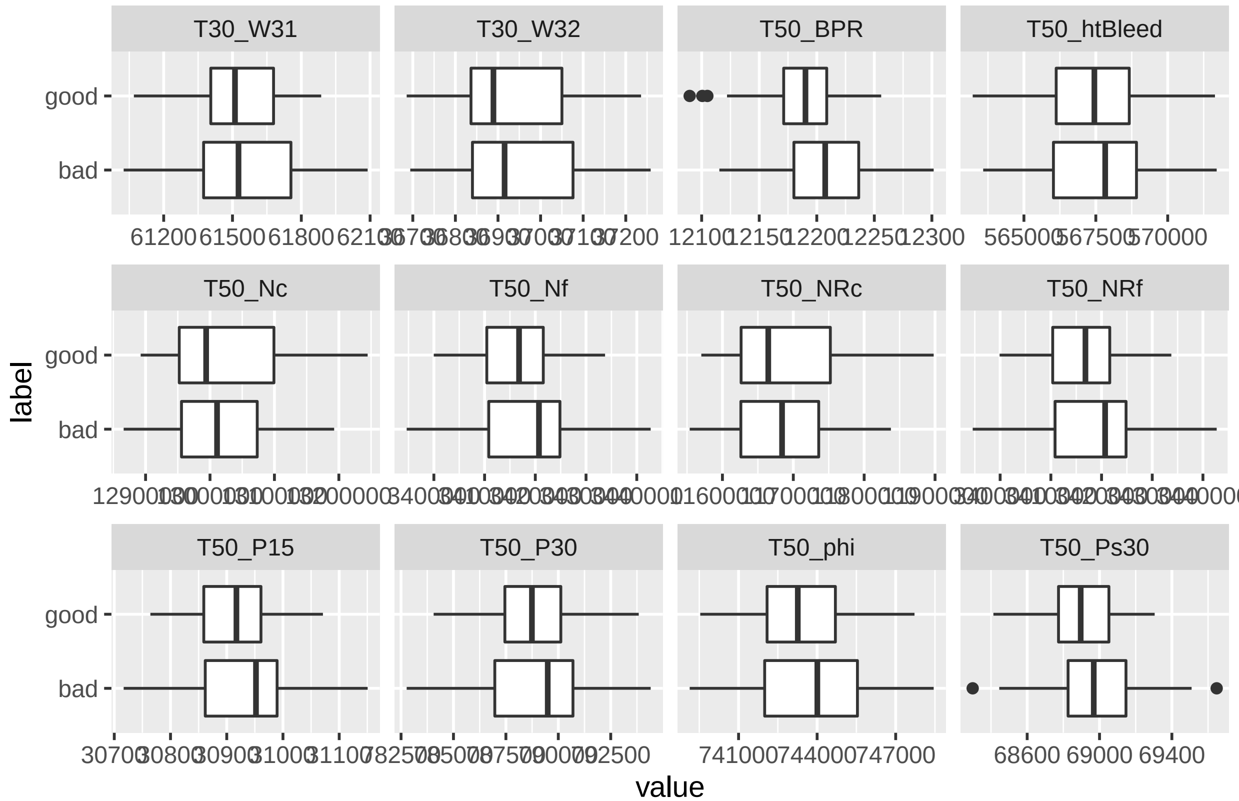

mach_number_BPR
mach_number_htBleed
mach_number_Nc
mach_number_Nf
mach_number_NRc
mach_number_NRf
mach_number_P30
mach_number_phi
mach_number_Ps30
mach_number_W31
mach_number_W32
T24_T30
good
bad
label
value
Page 4


T24_BPR
T24_htBleed
T24_Nc
T24_Nf
T24_NRc
T24_NRf
T24_P15
T24_P30
T24_phi
T24_Ps30
T24_T50
T24_W31
good
bad
label
value
5460
5480
5500
5250000
5300000
13900
13910
13920
24600
24700
24800

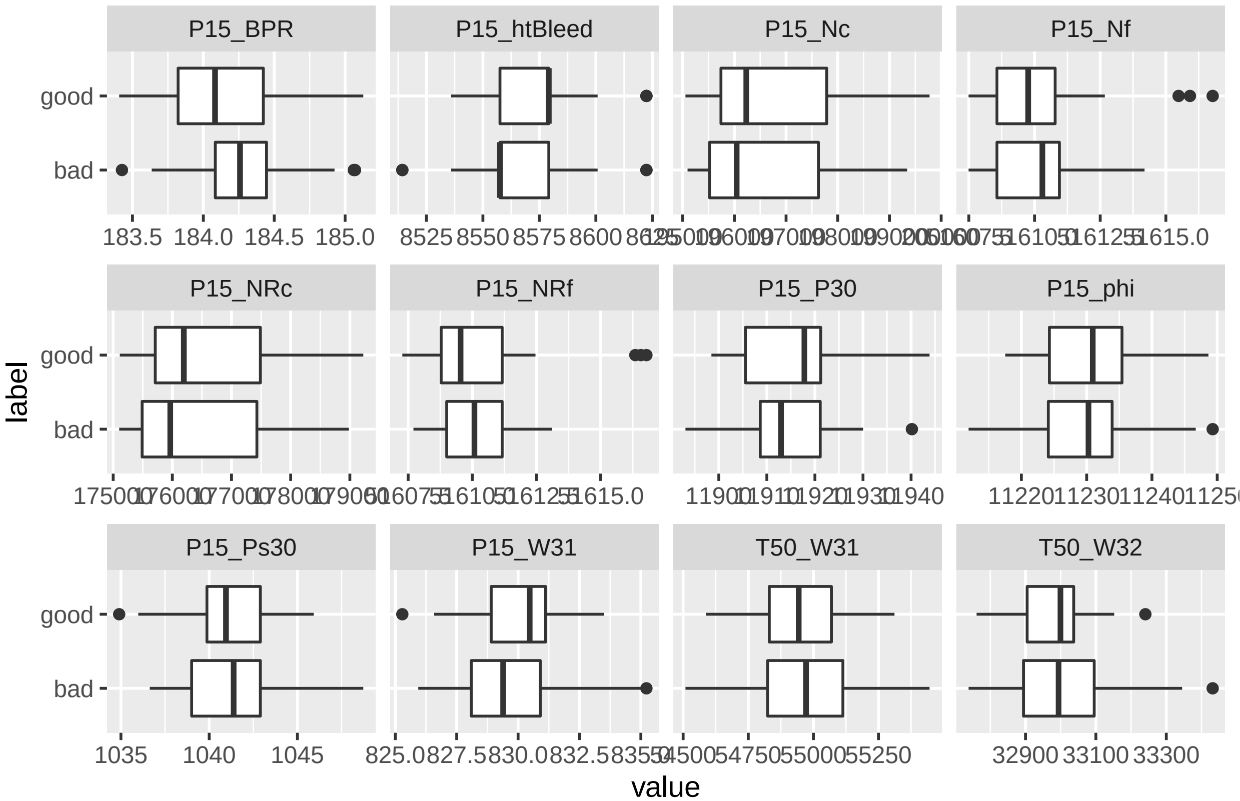
Page 5


T24_W32
T30_BPR
T30_htBleed
T30_Nc
T30_Nf
T30_NRc
T30_NRf
T30_P15
T30_P30
T30_phi
T30_Ps30
T30_T50
good
bad
label
value

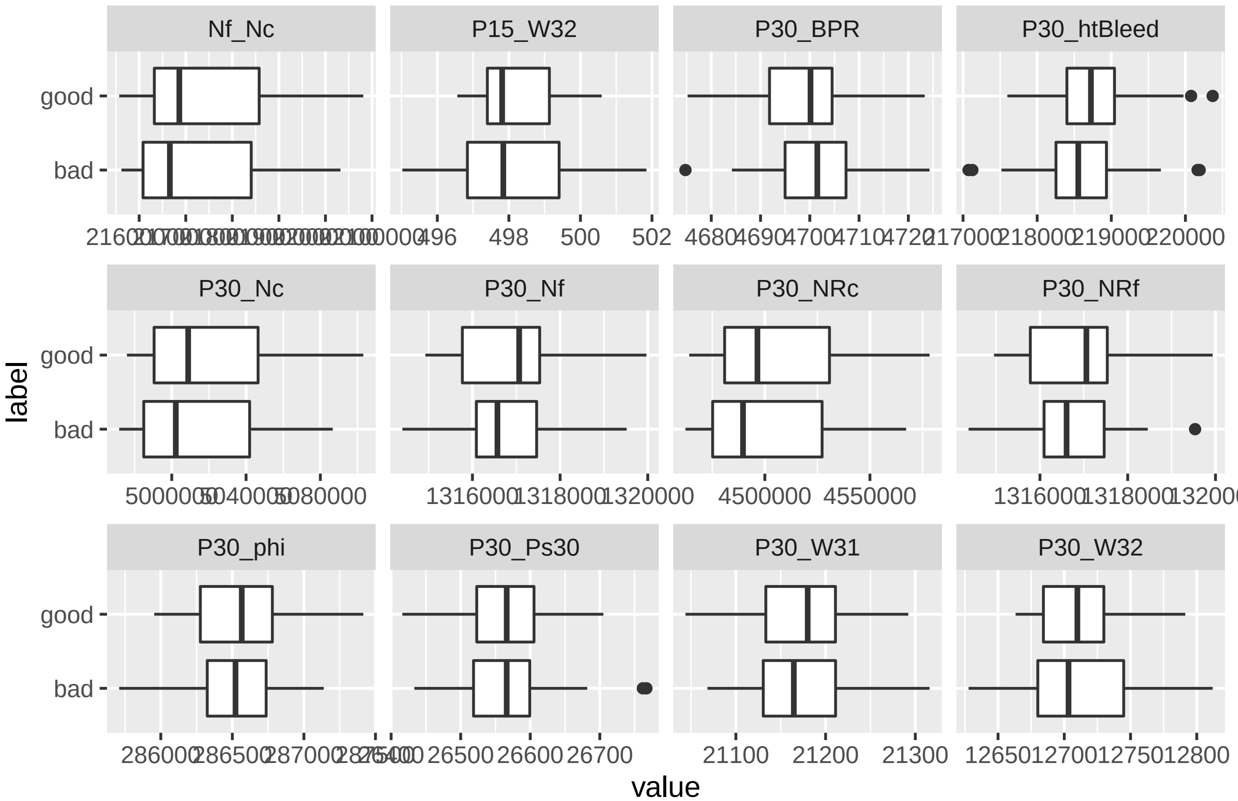
Page 6


T30_W31
T30_W32
T50_BPR
T50_htBleed
T50_Nc
T50_Nf
T50_NRc
T50_NRf
T50_P15
T50_P30
T50_phi
T50_Ps30
good
bad
label
value
61200
61500
61800
12100
12150
12200
12250
12300
565000
567500
570000
741000
744000
747000
68600
69000
69400

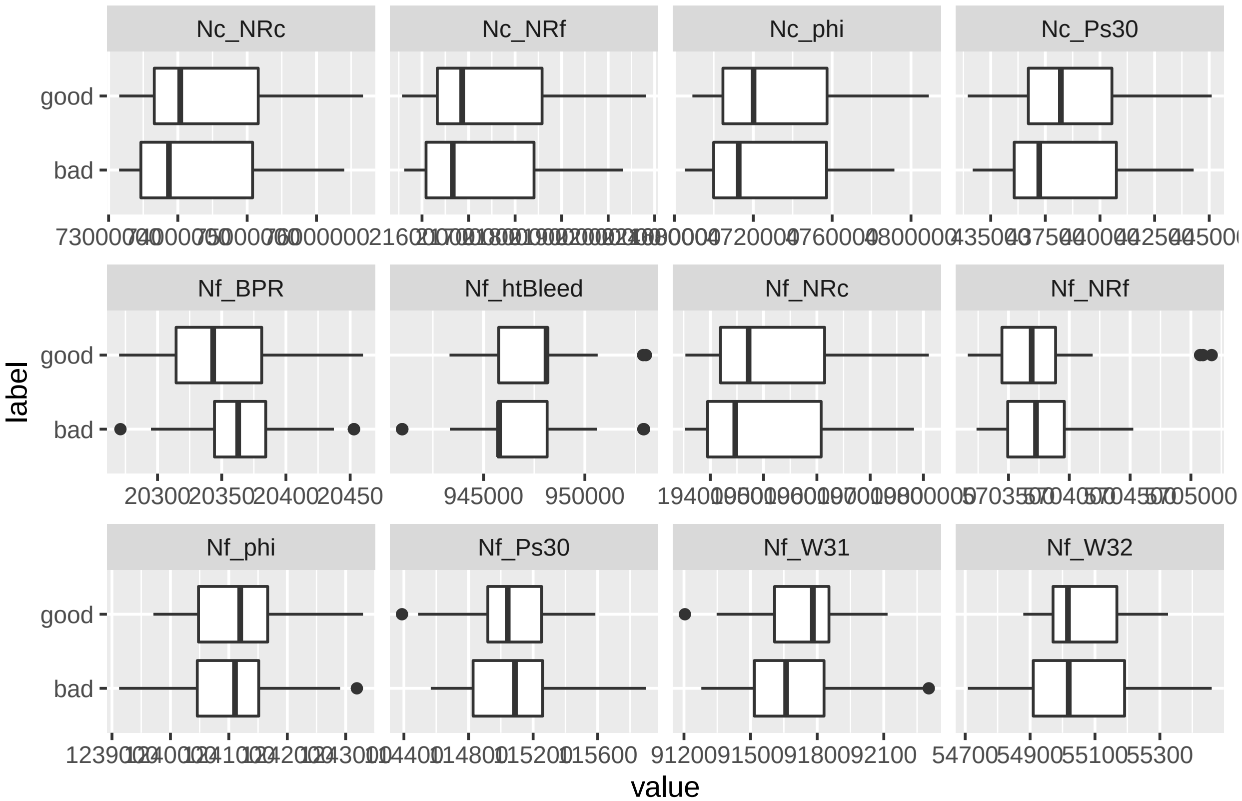
Page 7


P15_BPR
P15_htBleed
P15_Nc
P15_Nf
P15_NRc
P15_NRf
P15_P30
P15_phi
P15_Ps30
P15_W31
T50_W31
T50_W32
good
bad
label
value

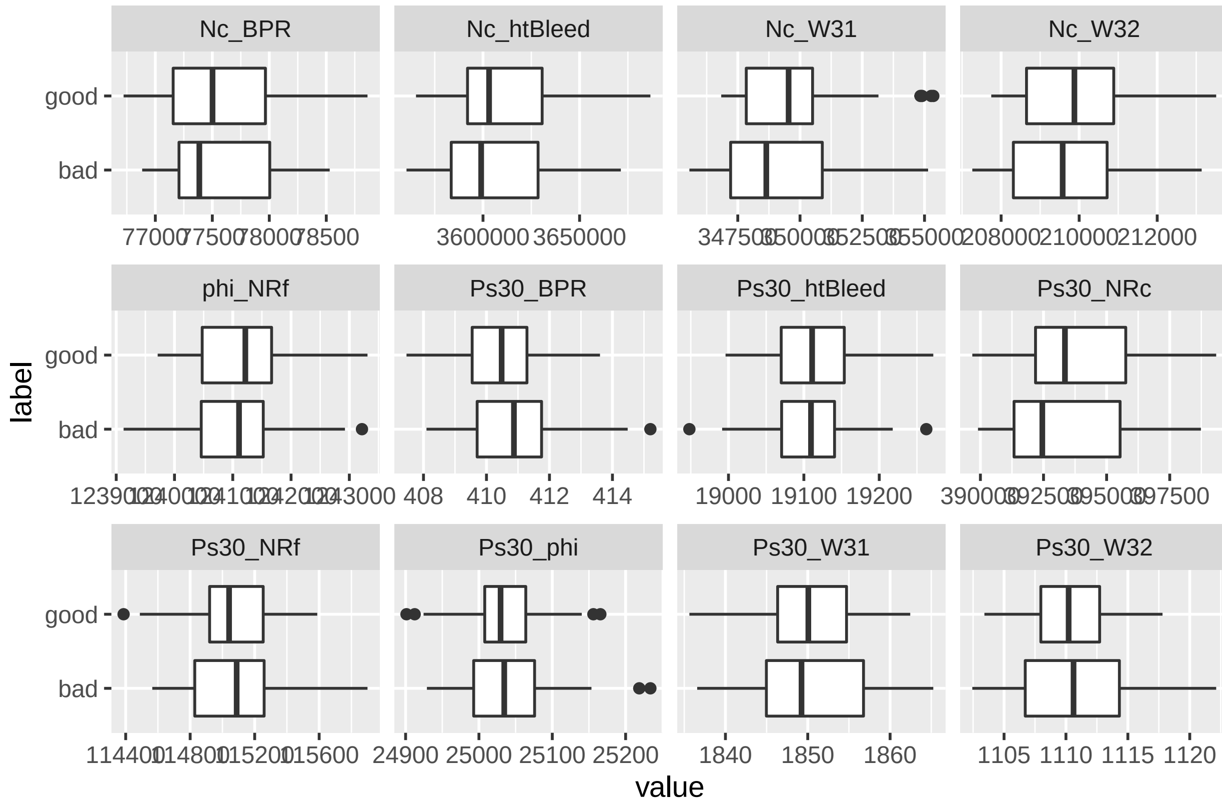
Page 8


Nf_Nc
P15_W32
P30_BPR
P30_htBleed
P30_Nc
P30_Nf
P30_NRc
P30_NRf
P30_phi
P30_Ps30
P30_W31
P30_W32
good
bad
label
value

Page 9

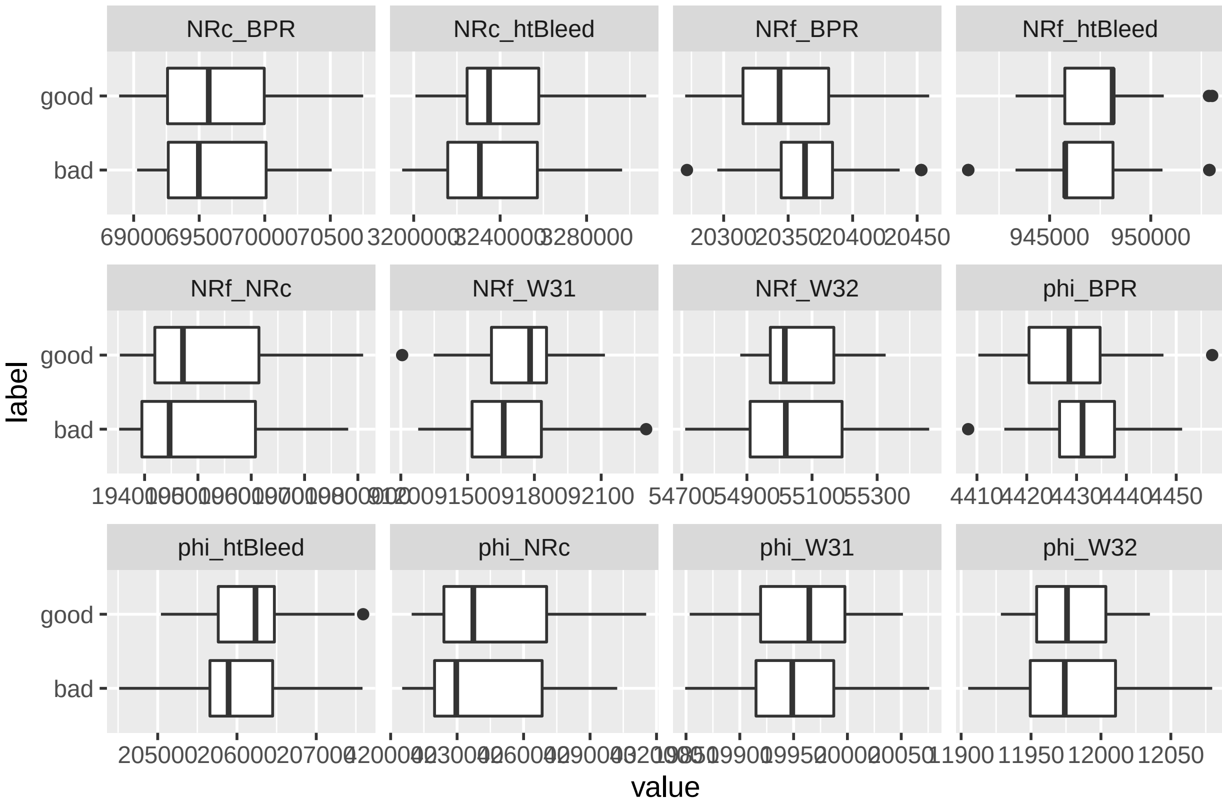

Nc_NRc
Nc_NRf
Nc_phi
Nc_Ps30
Nf_BPR
Nf_htBleed
Nf_NRc
Nf_NRf
Nf_phi
Nf_Ps30
Nf_W31
Nf_W32
good
bad
label
value
20300 20350 20400 20450
945000 950000
54700 54900 55100 55300

Page 10

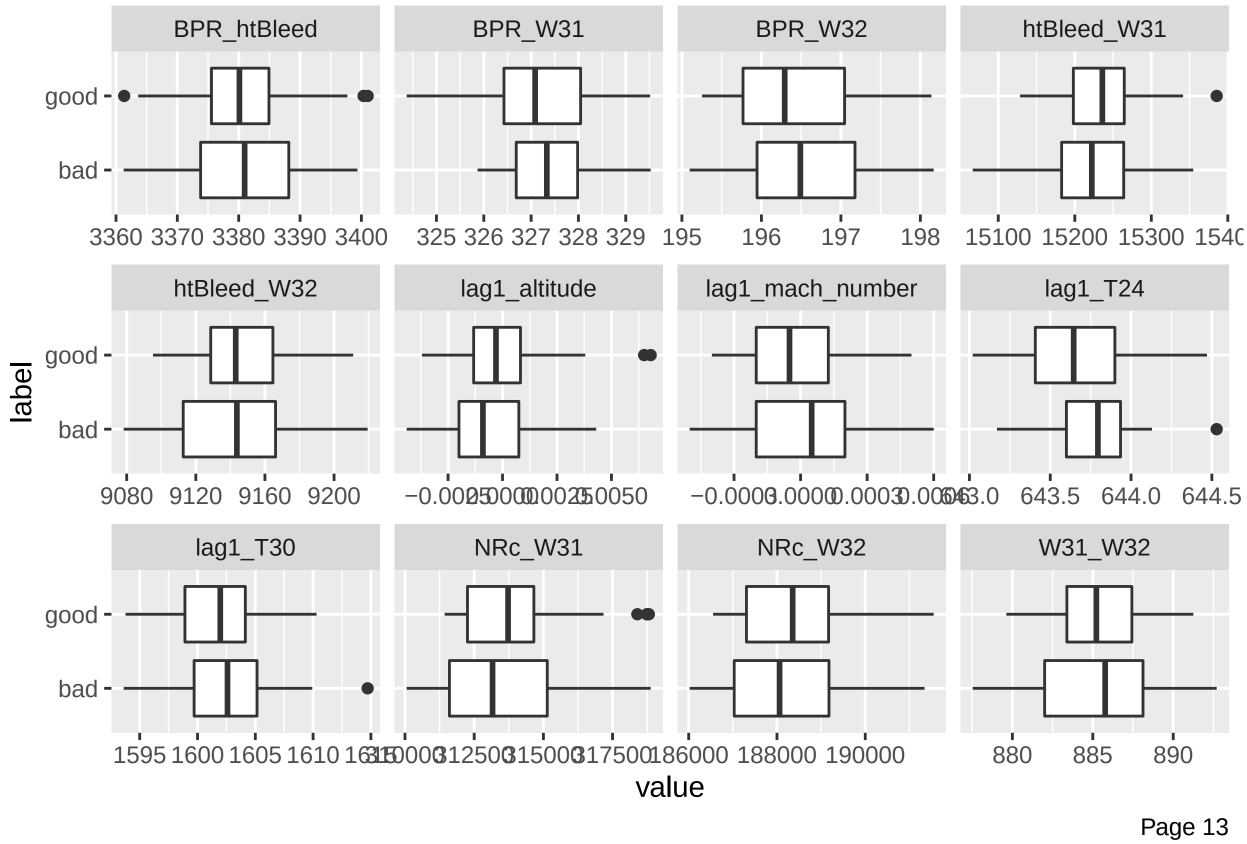


Nc_BPR
Nc_htBleed
Nc_W31
Nc_W32
phi_NRf
Ps30_BPR
Ps30_htBleed
Ps30_NRc
Ps30_NRf
Ps30_phi
Ps30_W31
Ps30_W32
good
bad
label
value

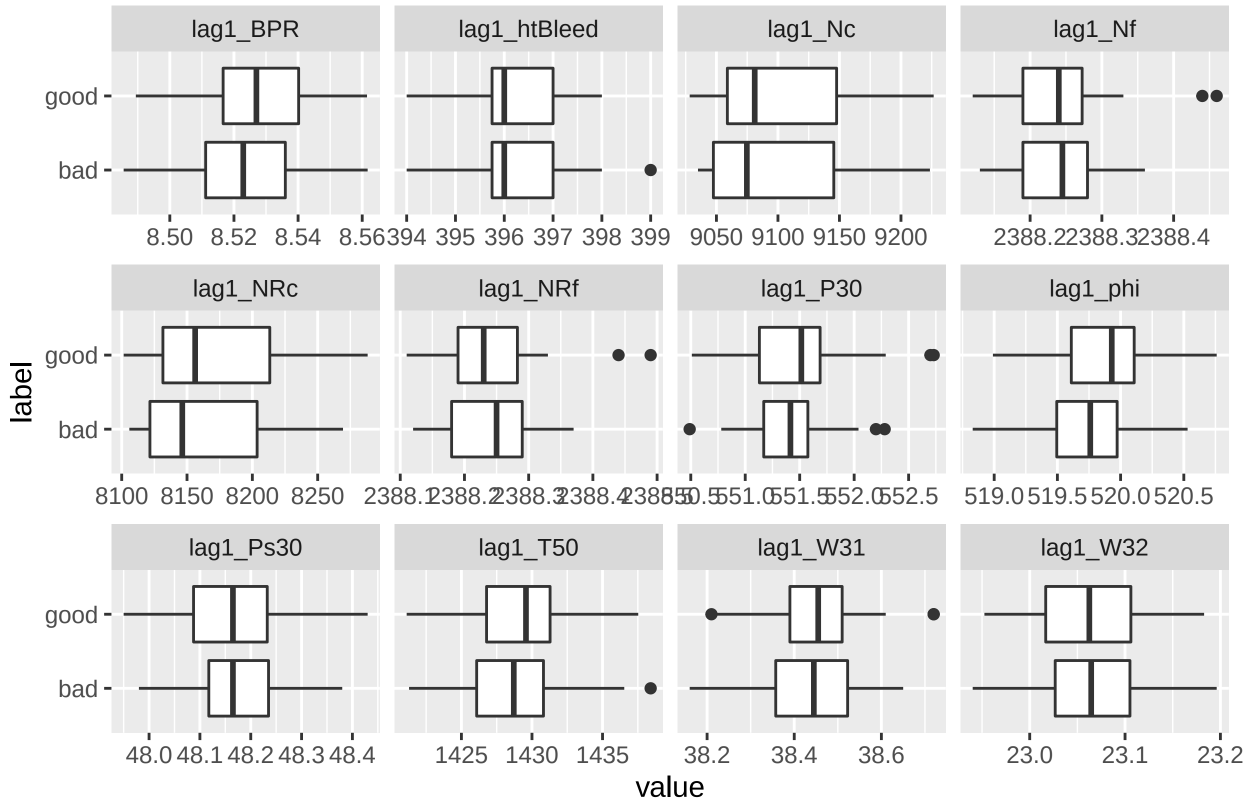
Page 11

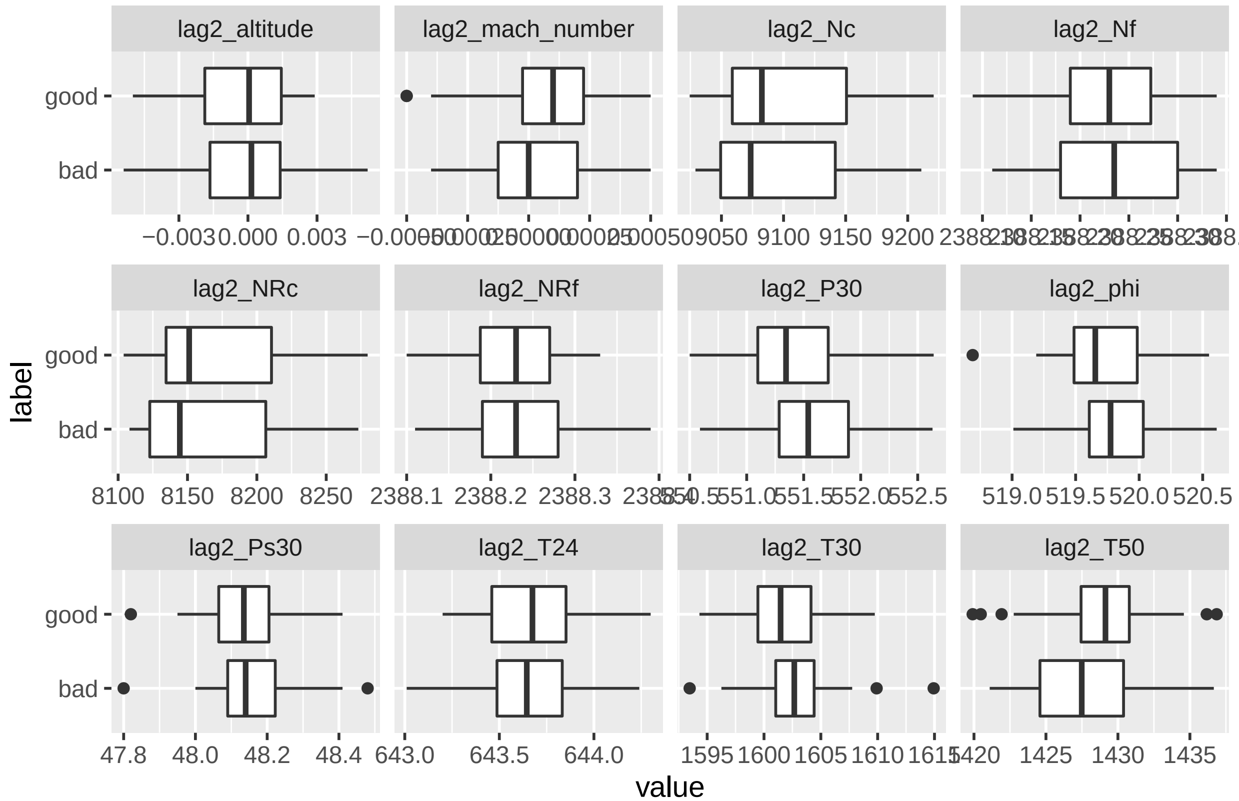

NRc_BPR
NRc_htBleed
NRf_BPR
NRf_htBleed
NRf_NRc
NRf_W31
NRf_W32
phi_BPR
phi_htBleed
phi_NRc
phi_W31
phi_W32
good
bad
label
value
3200000
3240000
3280000
20300
20350
20400
20450
945000
950000
54700
54900
55100
55300
4410
4420
4430
4440
4450
205000
206000
11900
11950
12000
12050

Page 12

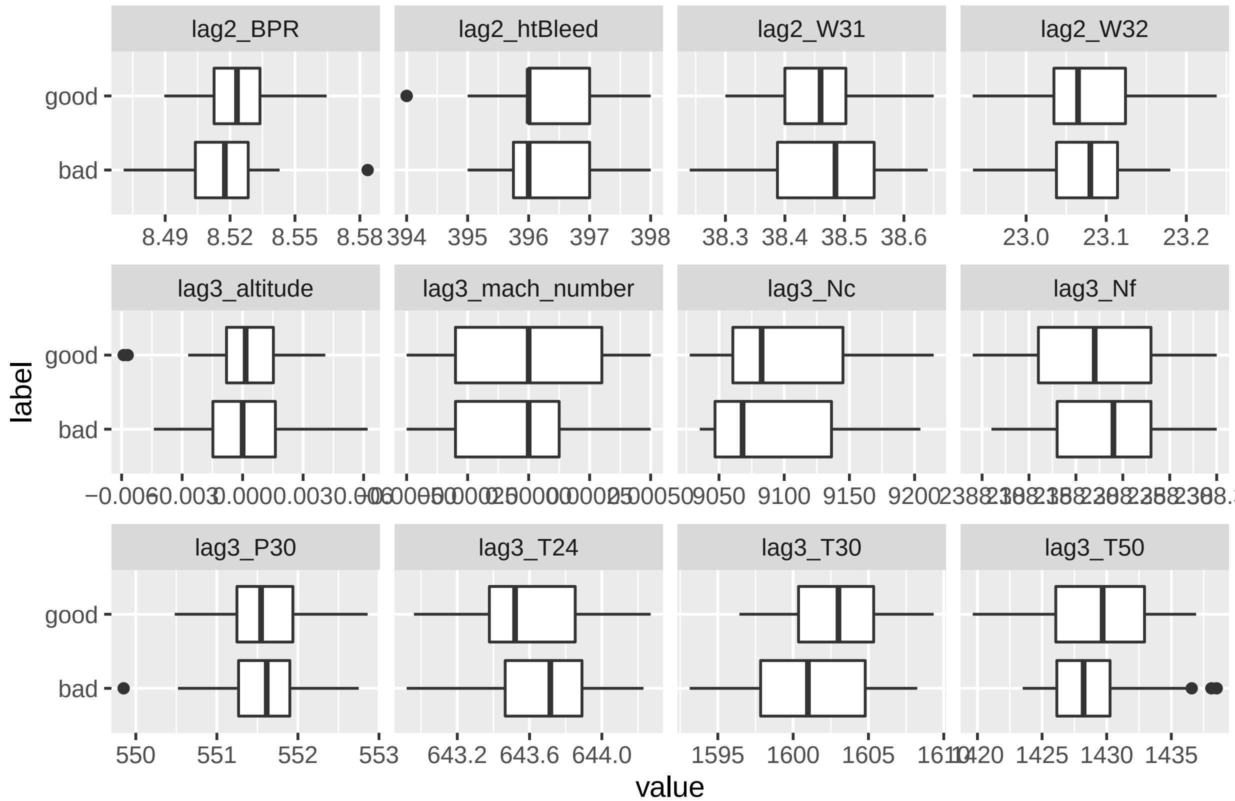


BPR_htBleed
BPR_W31
BPR_W32
htBleed_W31
htBleed_W32
lag1_altitude
lag1_mach_number
lag1_T24
lag1_T30
NRc_W31
NRc_W32
W31_W32
good
bad
label
value
3360 3370 3380 3390 3400
325 326 327 328 329
195 196 197 198
15100 15200 15300
9080 9120 9160 9200
643.5 644.0 644.5
1595 1600 1605 1610
186000 188000 190000
880 885 890
Page 13

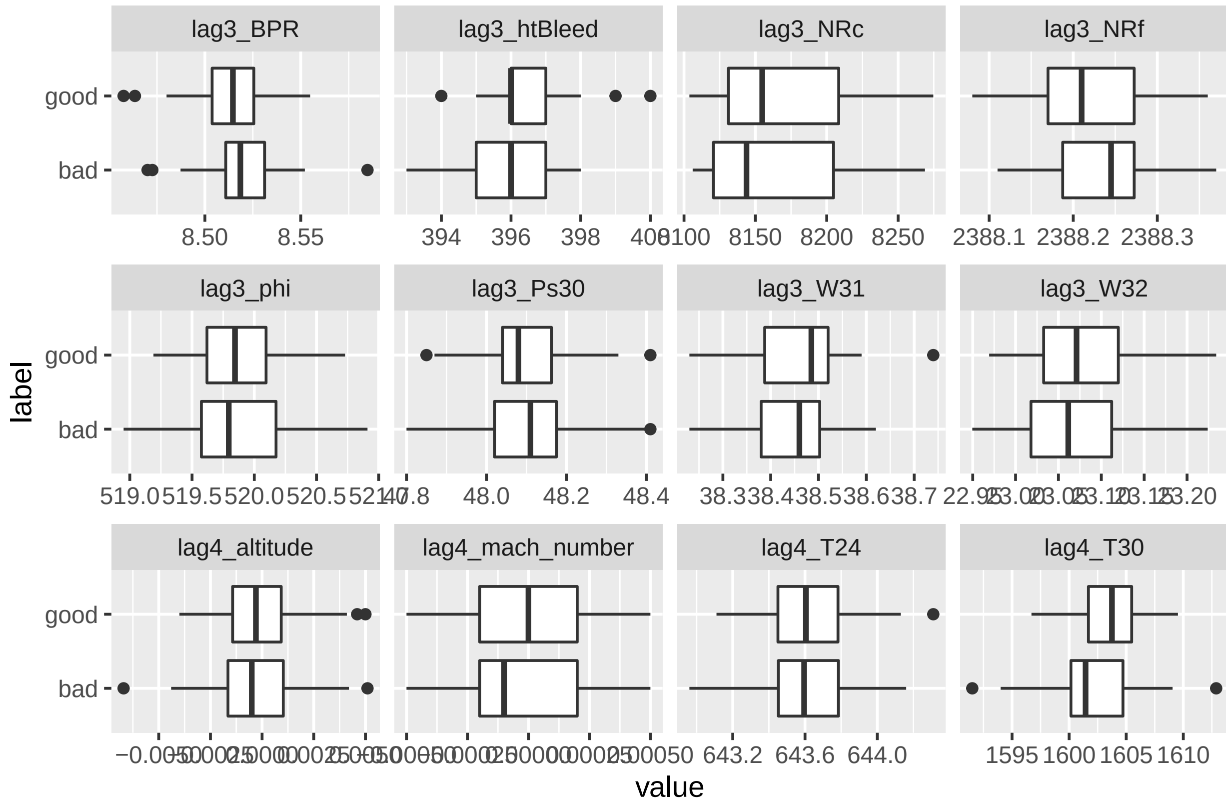

lag1_BPR
lag1_htBleed
lag1_Nc
lag1_Nf
lag1_NRc
lag1_NRf
lag1_P30
lag1_phi
lag1_Ps30
lag1_T50
lag1_W31
lag1_W32
good
bad
label
value


lag2_altitude
lag2_mach_number
lag2_Nc
lag2_Nf
lag2_NRc
lag2_NRf
lag2_P30
lag2_phi
lag2_Ps30
lag2_T24
lag2_T30
lag2_T50
good
bad
label
value

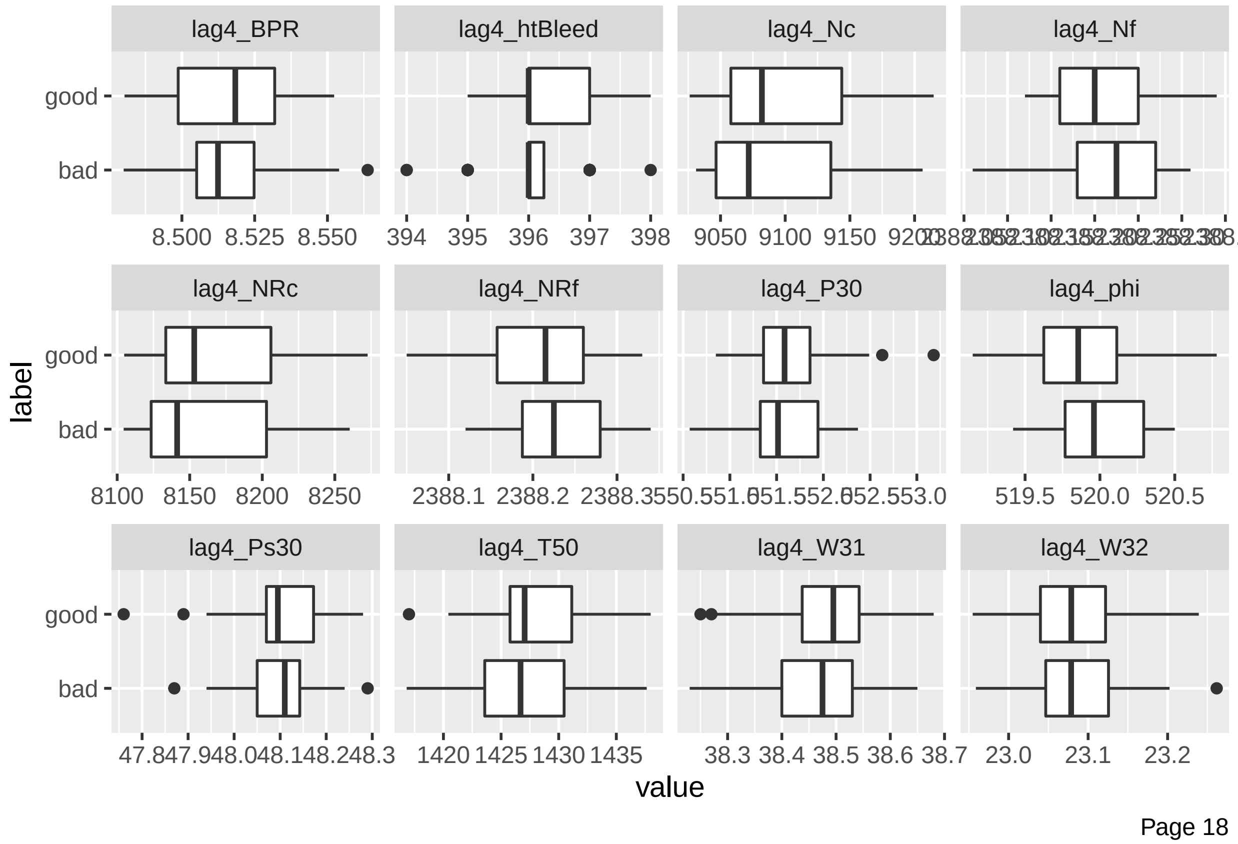
Page 15

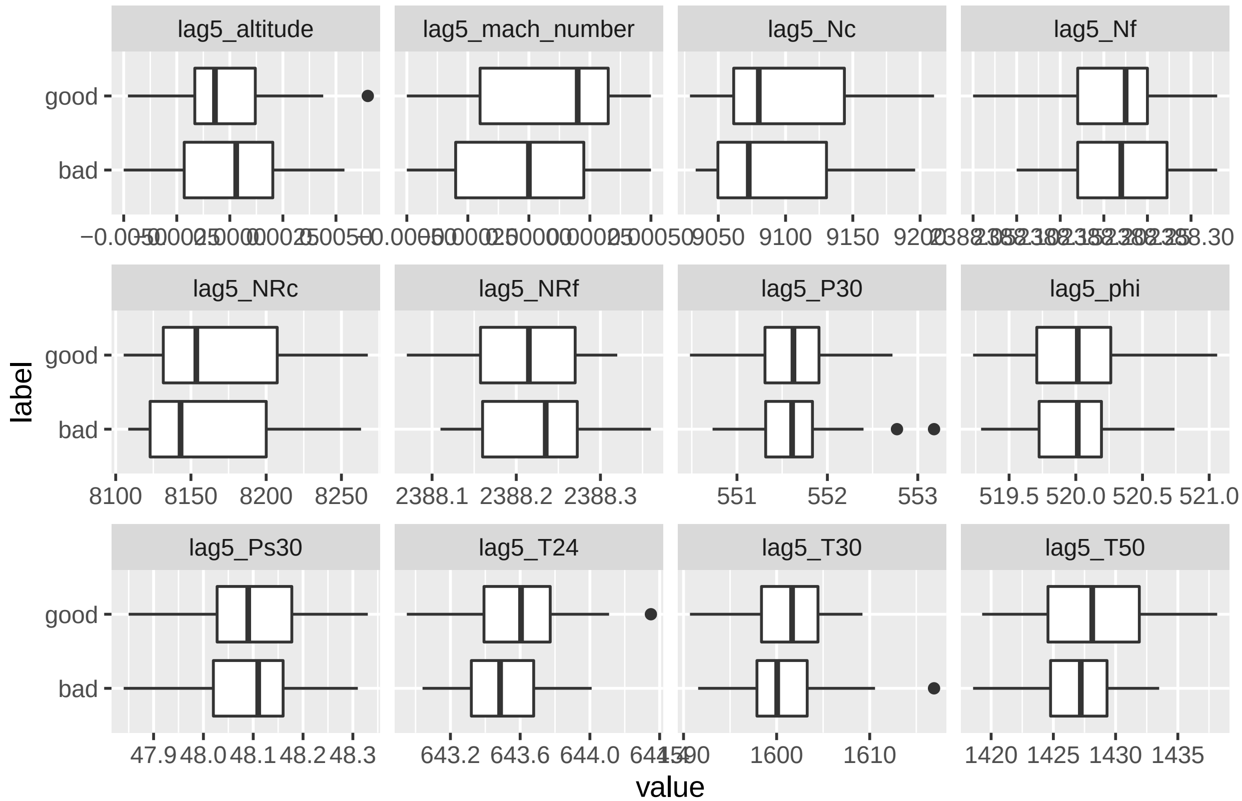

lag2_BPR
lag2_htBleed
lag2_W31
lag2_W32
lag3_altitude
lag3_mach_number
lag3_Nc
lag3_Nf
lag3_P30
lag3_T24
lag3_T30
lag3_T50
good
bad
label
value
8.49 8.52 8.55 8.58
394 395 396 397 398
38.3 38.4 38.5 38.6
23.0 23.1 23.2
9050 9100 9150 9200
550 551 552 553
643.2 643.6 644.0
1595 1600 1605 1610
1420 1425 1430 1435

Page 16

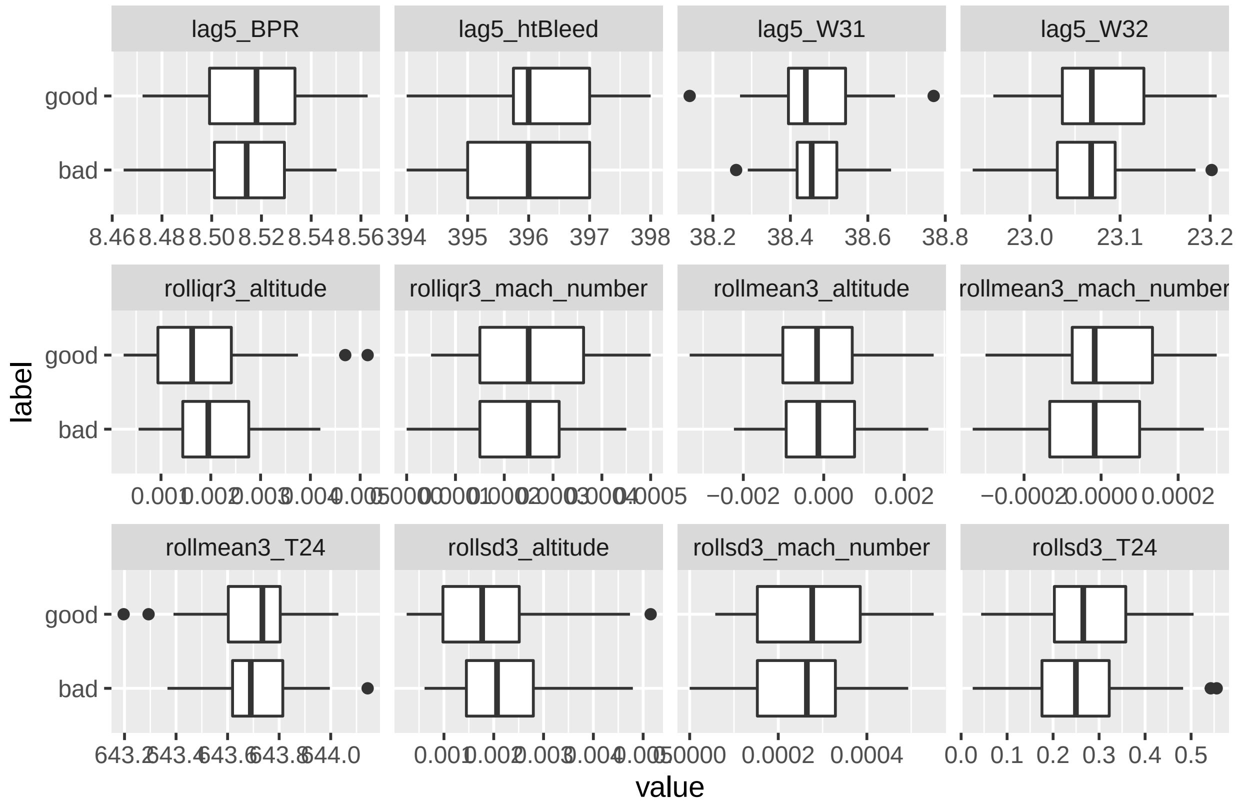

lag3_BPR
lag3_htBleed
lag3_NRc
lag3_NRf
lag3_phi
lag3_Ps30
lag3_W31
lag3_W32
lag4_altitude
lag4_mach_number
lag4_T24
lag4_T30
good
bad
label
value

Page 17

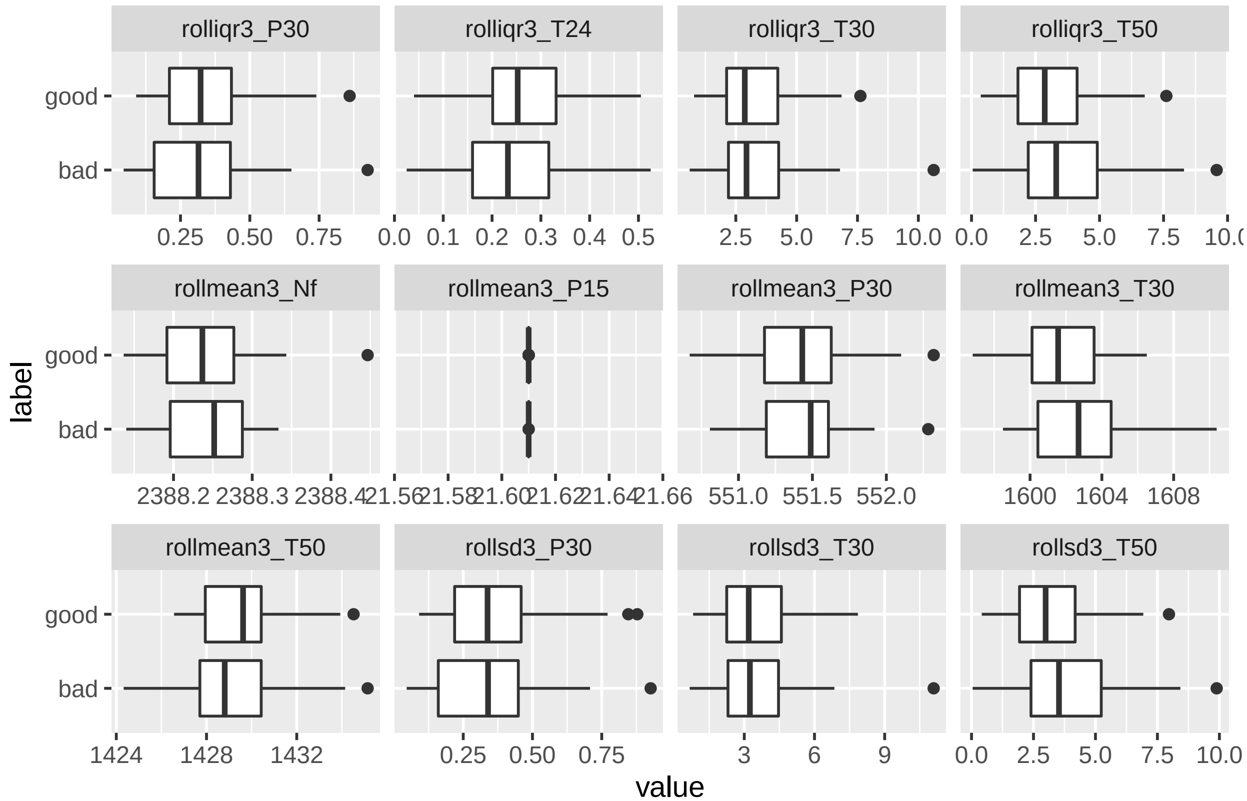

lag4_BPR
lag4_htBleed
lag4_Nc
lag4_Nf
lag4_NRc
lag4_NRf
lag4_P30
lag4_phi
lag4_Ps30
lag4_T50
lag4_W31
lag4_W32
good
bad
label
value
8.500 8.525 8.550
394 395 396 397 398
9050 9100 9150 9200
8100 8150 8200 8250
2388.1 2388.2 2388.3
519.5 520.0 520.5
1420 1425 1430 1435
38.3 38.4 38.5 38.6 38.7
23.0 23.1 23.2
Page 18


lag5_altitude
lag5_mach_number
lag5_Nc
lag5_Nf
lag5_NRc
lag5_NRf
lag5_P30
lag5_phi
lag5_Ps30
lag5_T24
lag5_T30
lag5_T50
good
bad
label
value
9050 9100 9150
8100 8150 8200 8250
2388.1 2388.2 2388.3
551 552 553
519.5 520.0 520.5 521.0
47.9 48.0 48.1 48.2 48.3
643.2 643.6 644.0
1600 1610
1420 1425 1430 1435

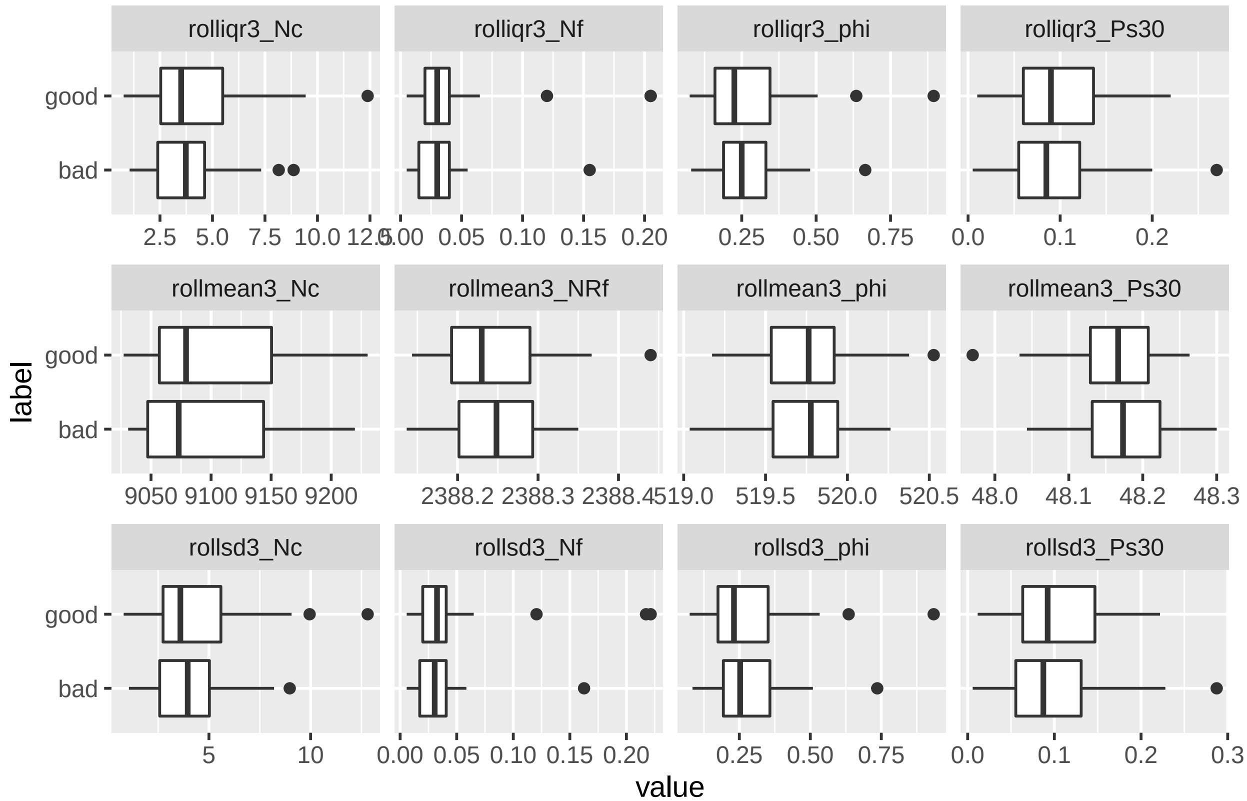
Page 19


lag5_BPR
lag5_htBleed
lag5_W31
lag5_W32
rolliqr3_altitude
rolliqr3_mach_number
rollmean3_altitude
rollmean3_mach_number
rollmean3_T24
rollsd3_altitude
rollsd3_mach_number
rollsd3_T24
good
bad
label
value

Page 20

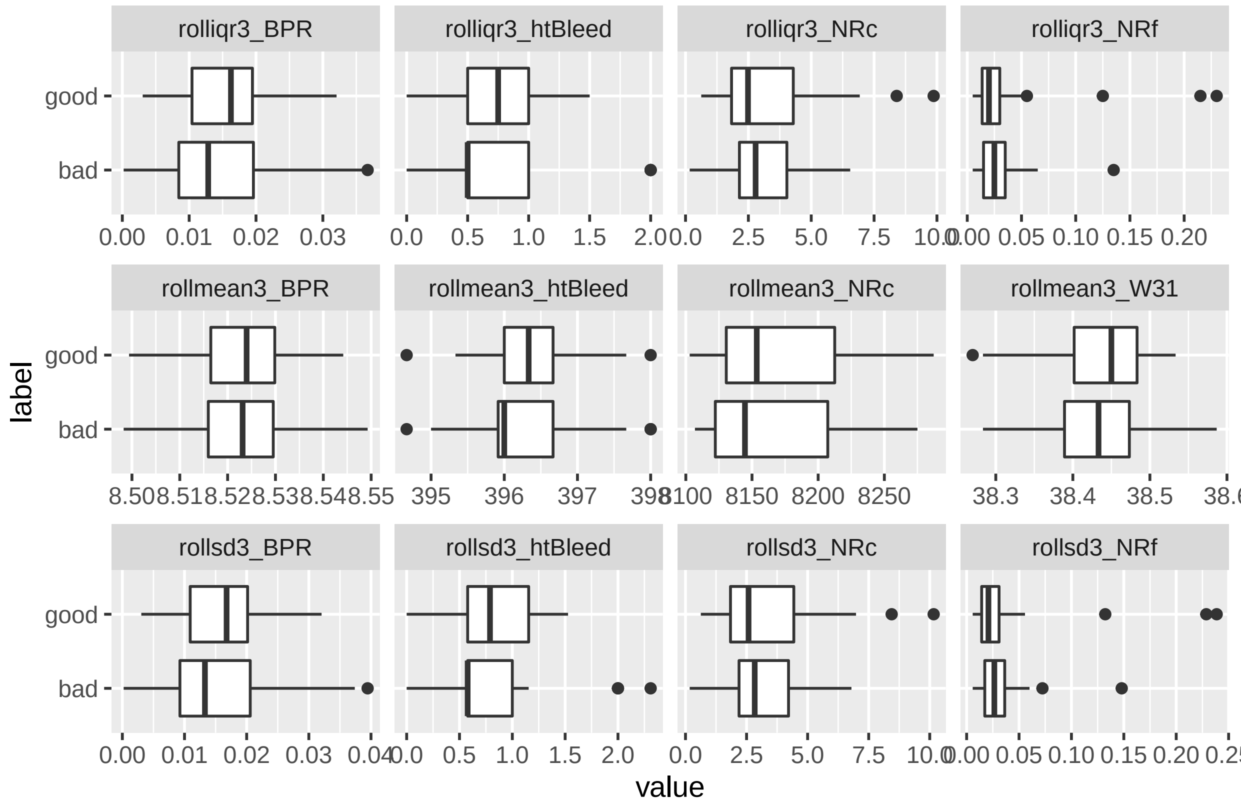

rolliqr3_P30
rolliqr3_T24
rolliqr3_T30
rolliqr3_T50
rollmean3_Nf
rollmean3_P15
rollmean3_P30
rollmean3_T30
rollmean3_T50
rollsd3_P30
rollsd3_T30
rollsd3_T50
good
bad
label
value
0.25 0.50 0.75
0.0 0.1 0.2 0.3 0.4 0.5
2.5 5.0 7.5 10.0
2388.2 2388.3 2388.4
21.56 21.58 21.60 21.62 21.64 21.66
551.0 551.5 552.0
1600 1604 1608
1424 1428 1432
3 6 9
0.0 2.5 5.0 7.5 10.0

Page 21

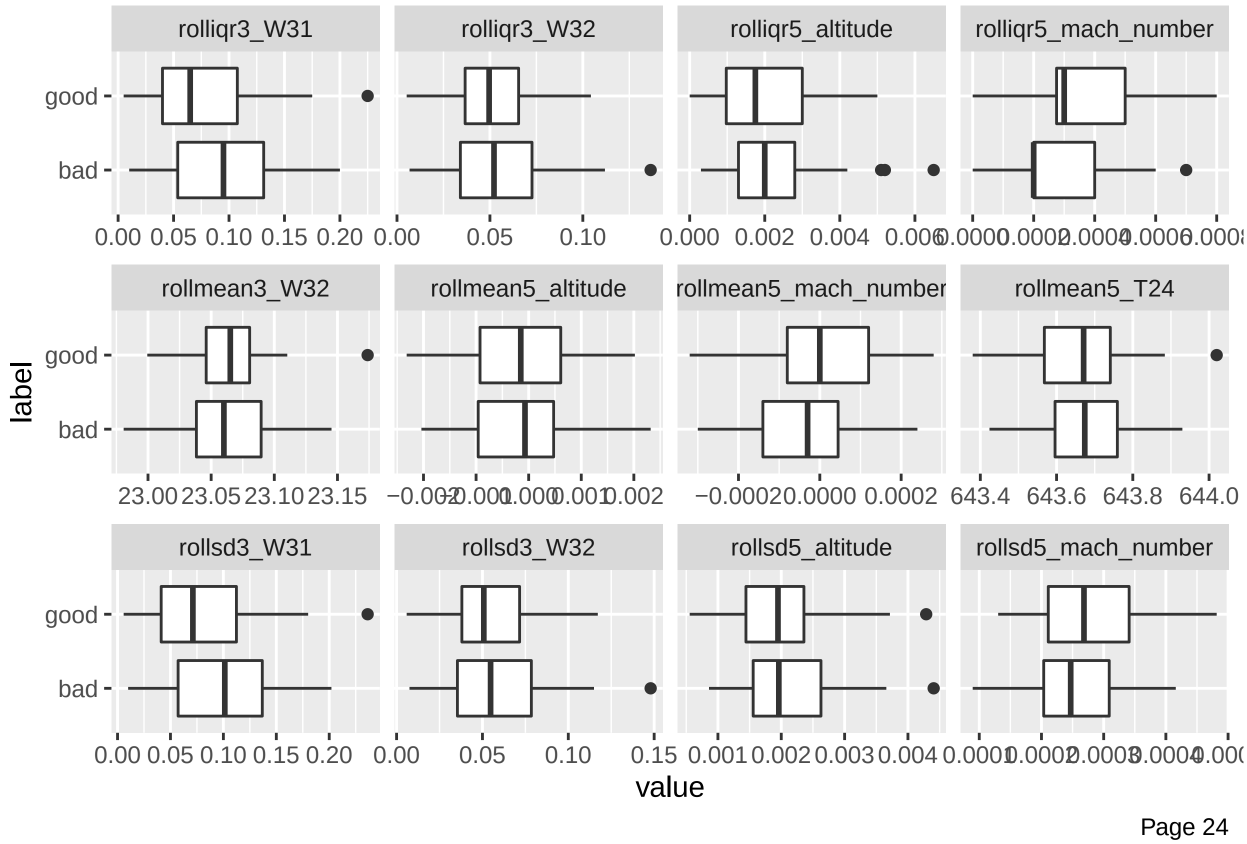


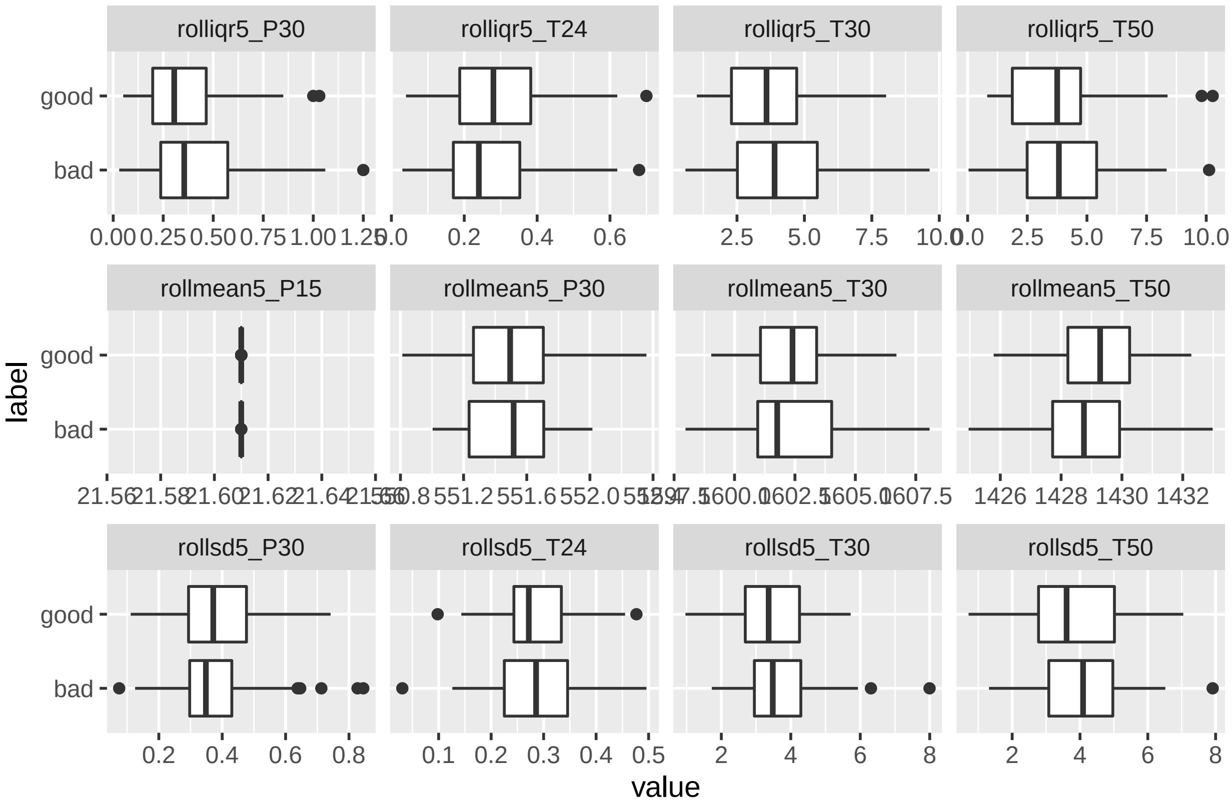
rolliqr3_Nc
rolliqr3_Nf
rolliqr3_phi
rolliqr3_Ps30
rollmean3_Nc
rollmean3_NRf
rollmean3_phi
rollmean3_Ps30
rollsd3_Nc
rollsd3_Nf
rollsd3_phi
rollsd3_Ps30
good
bad
label
value

Page 22

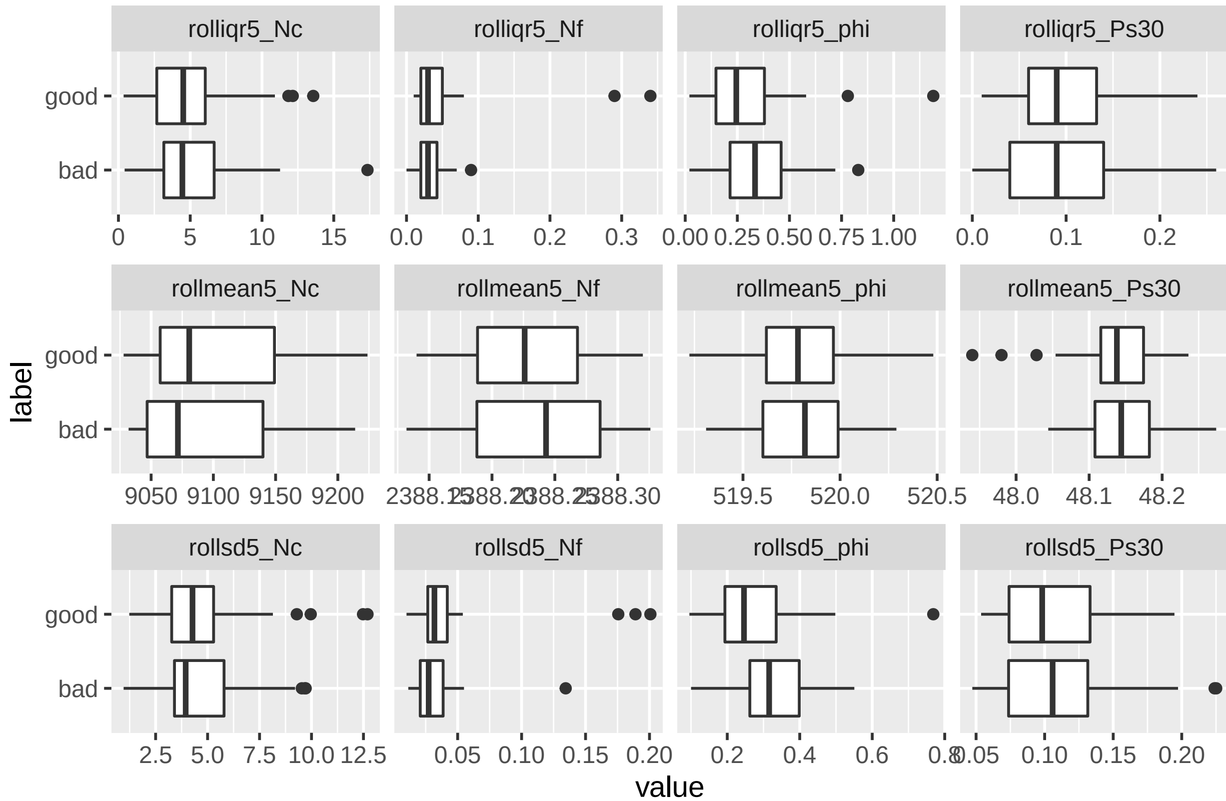

rolliqr3_BPR
rolliqr3_htBleed
rolliqr3_NRc
rolliqr3_NRf
rollmean3_BPR
rollmean3_htBleed
rollmean3_NRc
rollmean3_W31
rollsd3_BPR
rollsd3_htBleed
rollsd3_NRc
rollsd3_NRf
good
bad
label
value

Page 23


rolliqr3_W31
rolliqr3_W32
rolliqr5_altitude
rolliqr5_mach_number
rollmean3_W32
rollmean5_altitude
rollmean5_mach_number
rollmean5_T24
rollsd3_W31
rollsd3_W32
rollsd5_altitude
rollsd5_mach_number
good
bad
label
value
Page 24

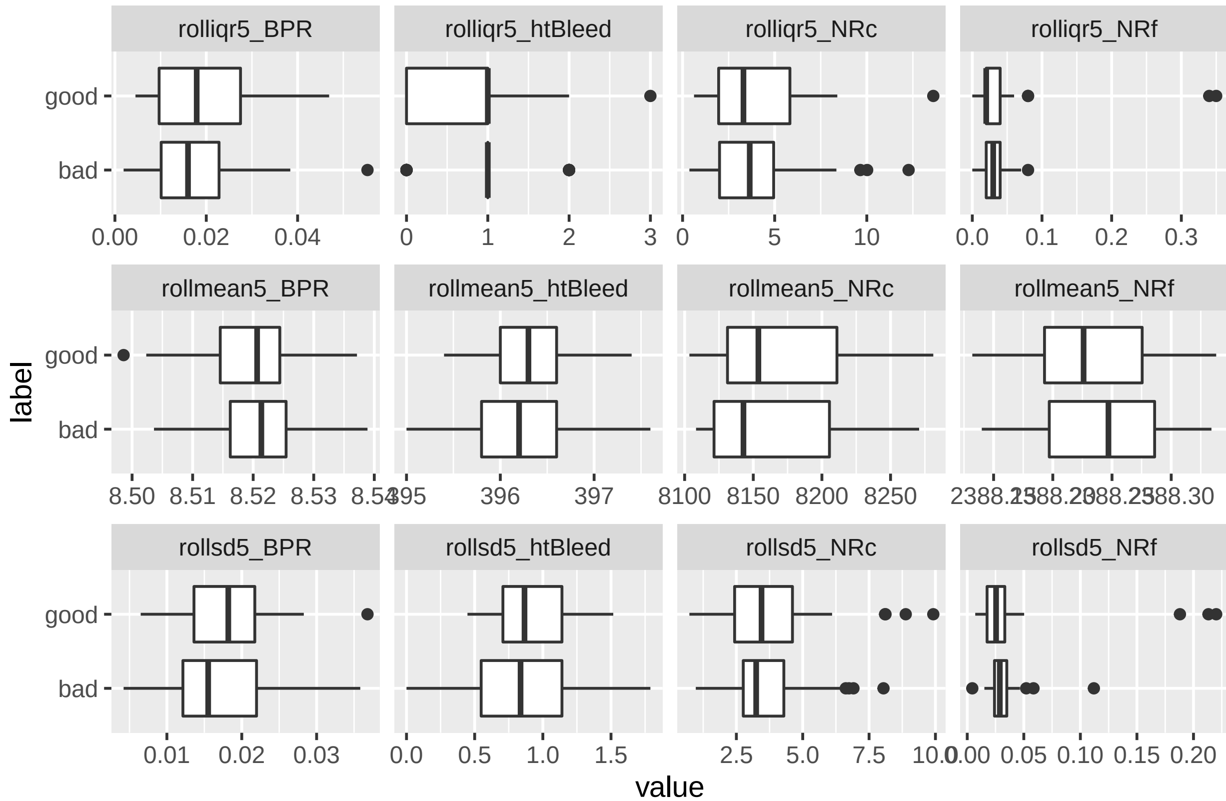

rolliqr5_P30
rolliqr5_T24
rolliqr5_T30
rolliqr5_T50
rollmean5_P15
rollmean5_P30
rollmean5_T30
rollmean5_T50
rollsd5_P30
rollsd5_T24
rollsd5_T30
rollsd5_T50
good
bad
label
value

Page 25

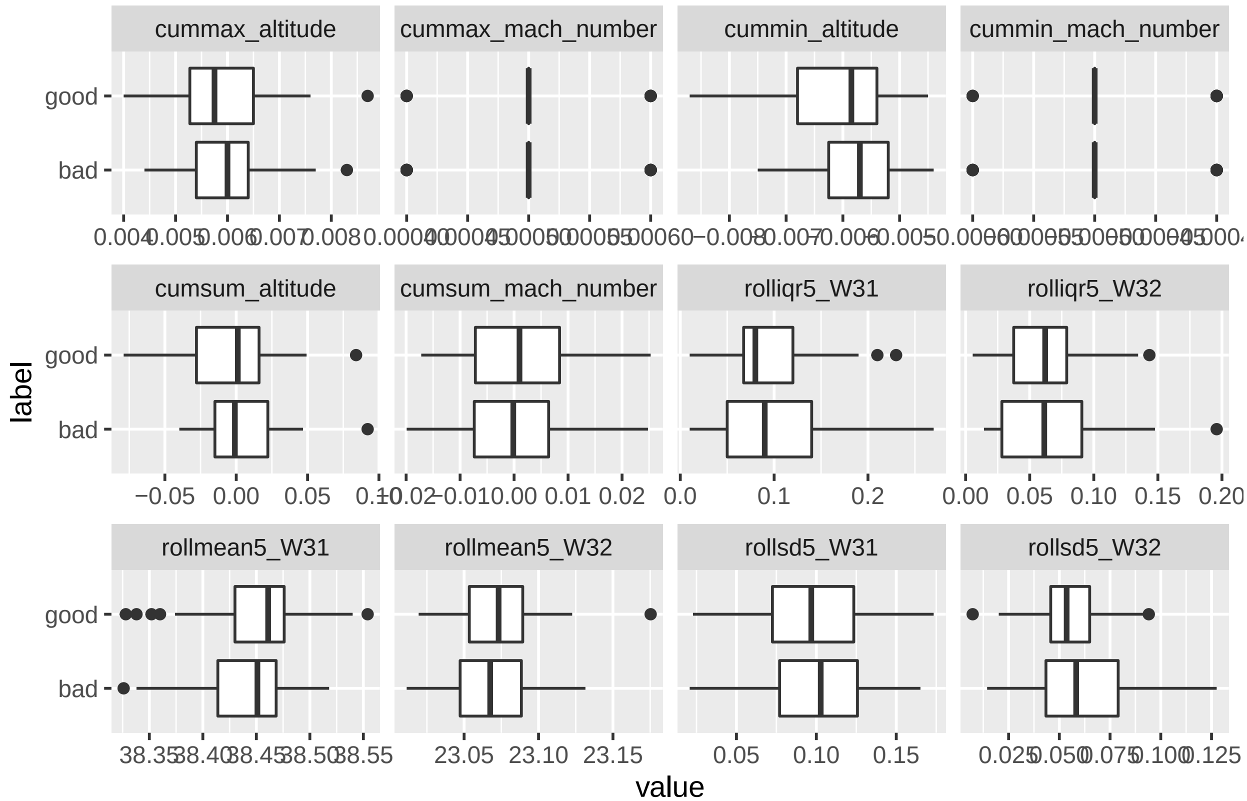


rolliqr5_Nc
rolliqr5_Nf
rolliqr5_phi
rolliqr5_Ps30
rollmean5_Nc
rollmean5_Nf
rollmean5_phi
rollmean5_Ps30
rollsd5_Nc
rollsd5_Nf
rollsd5_phi
rollsd5_Ps30
good
bad
label
value

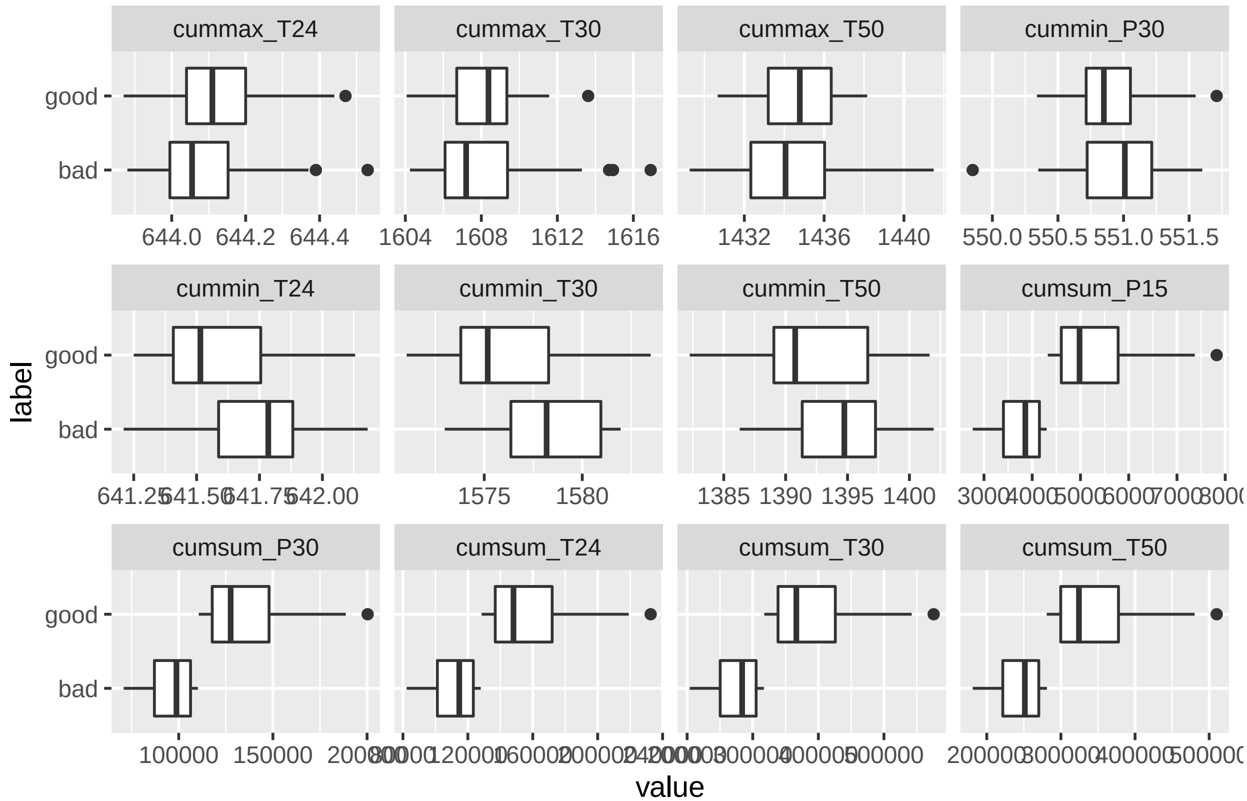
Page 26

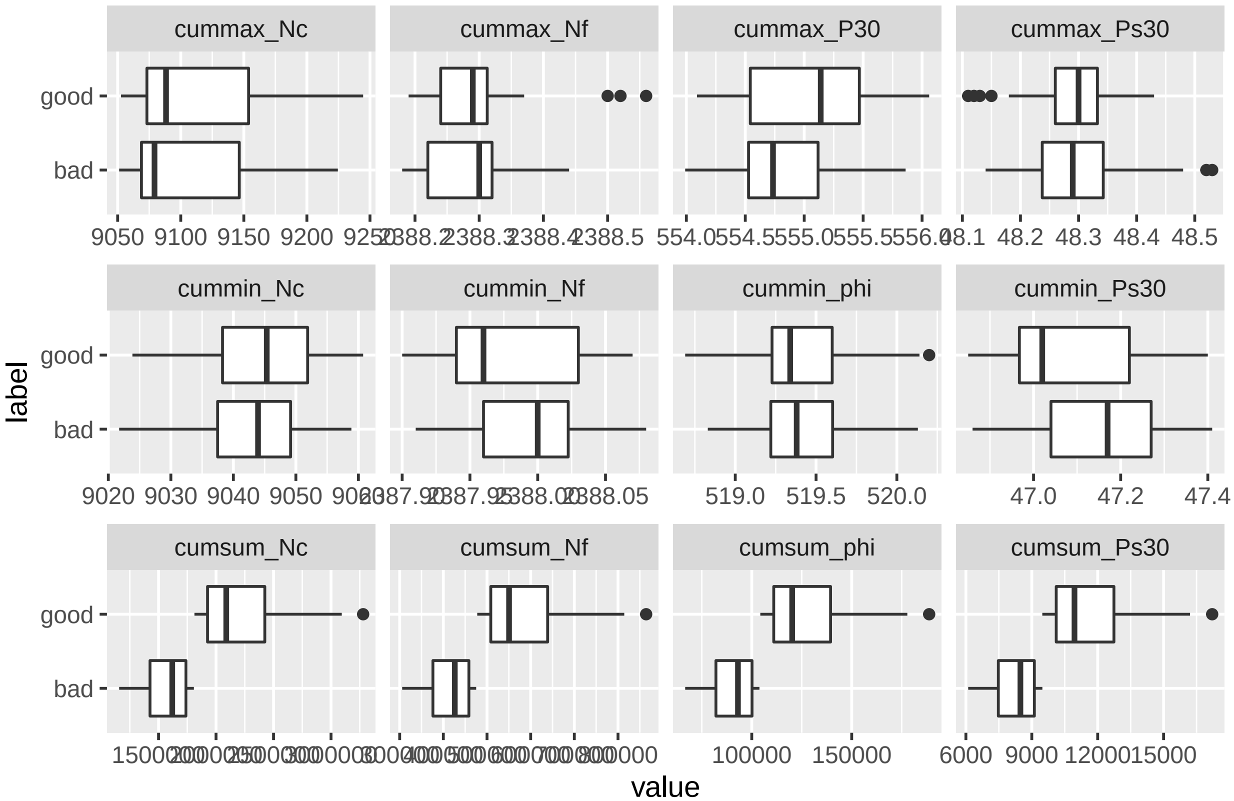

rolliqr5_BPR
rolliqr5_htBleed
rolliqr5_NRc
rolliqr5_NRf
rollmean5_BPR
rollmean5_htBleed
rollmean5_NRc
rollmean5_NRf
rollsd5_BPR
rollsd5_htBleed
rollsd5_NRc
rollsd5_NRf
good
bad
label
value

Page 27

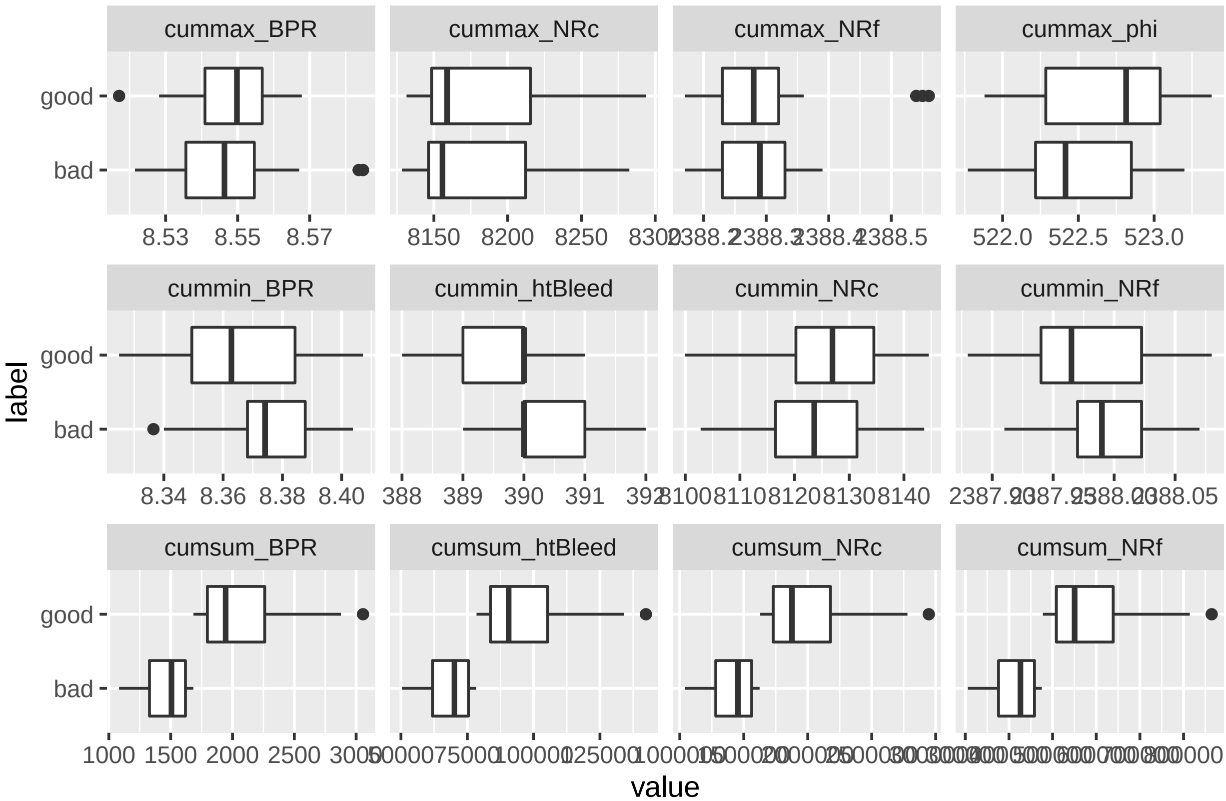

cummax_altitude
cummax_mach_number
cummin_altitude
cummin_mach_number
cumsum_altitude
cumsum_mach_number
rolliqr5_W31
rolliqr5_W32
rollmean5_W31
rollmean5_W32
rollsd5_W31
rollsd5_W32
good
bad
label
value

Page 28

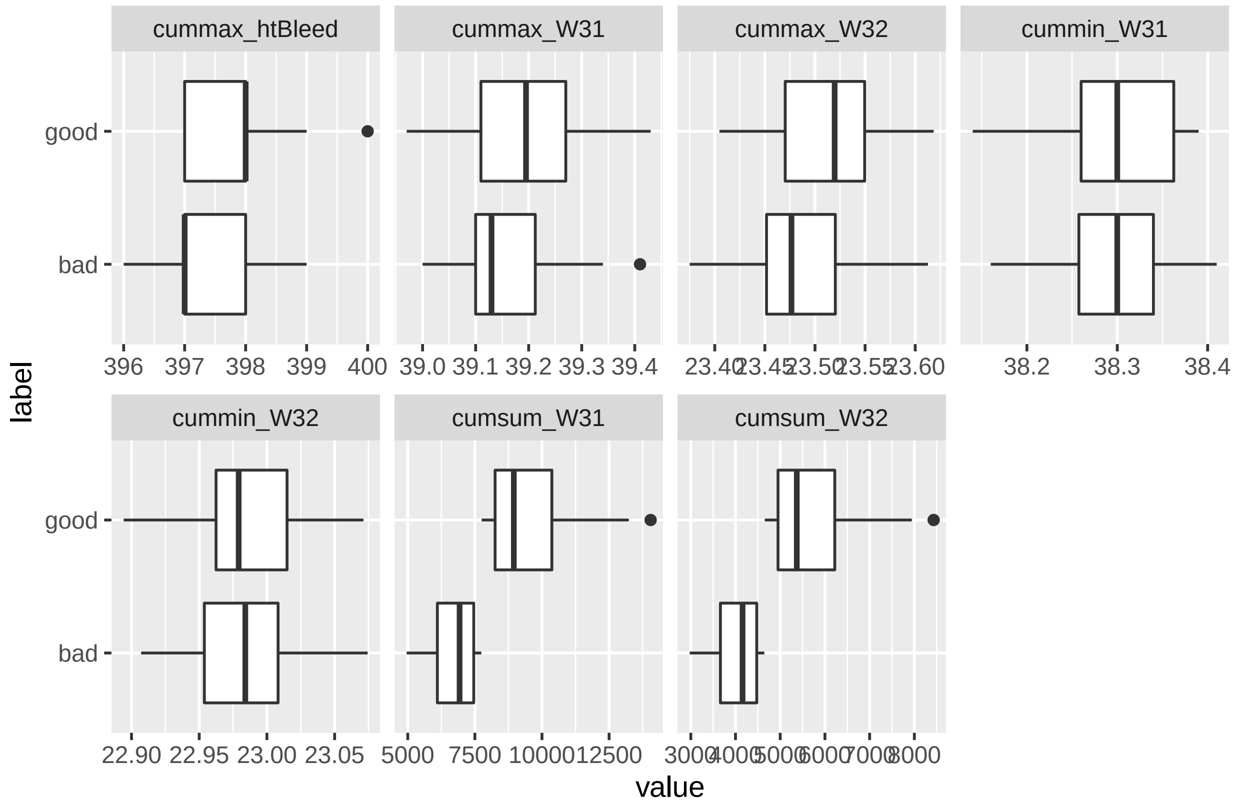

cummax_T24
cummax_T30
cummax_T50
cummin_P30
cummin_T24
cummin_T30
cummin_T50
cumsum_P15
cumsum_P30
cumsum_T24
cumsum_T30
cumsum_T50
good
bad
label
value


cummax_Nc
cummax_Nf
cummax_P30
cummax_Ps30
cummin_Nc
cummin_Nf
cummin_phi
cummin_Ps30
cumsum_Nc
cumsum_Nf
cumsum_phi
cumsum_Ps30
good
bad
label
value
9050
9100
9150
9200
2388.5
48.2
48.3
48.4
48.5
9020
9030
9040
9050
2388.05
519.0
519.5
520.0
47.0
47.2
47.4
100000
150000
6000
9000
15000

Page 30


cummax_BPR
cummax_NRc
cummax_NRf
cummax_phi
cummin_BPR
cummin_htBleed
cummin_NRc
cummin_NRf
cumsum_BPR
cumsum_htBleed
cumsum_NRc
cumsum_NRf
good
bad
label
value
8.53
8.55
8.57
8150
8200
8250
522.0
522.5
523.0
8.34
8.36
8.38
8.40
388
389
390
391
1000
1500
2000
2500

Page 31



Page 32

4.2 Comprobación de la normalidad y homogeneidad de la varianza.

Primero, vamos a examinar si los datos siguen una distribución normal,

```
normalityChecker <- function(df){
  df$label = NULL
  result <- data.frame(variables = names(df), p_value = NA, result = NA)

  for (i in 1:nrow(result)) {
    variable <- df[,i]
    prueba <- shapiro.test(pull(variable))
    result$p_value[i] <- prueba$p.value
    result$result[i] <- if_else(prueba$p.value >= 0.05, "normally distributed", "not normally distribut

  }

  return(result)
}

pruebas <- normalityChecker(df)

table(pruebas$result)
```

```
##
##     normally distributed not normally distributed
```

```
## 193 187
```

Ahora vamos a examinar la homocedasticidad.

Para las variables que siguen una distribución normal usaremos el **test de Levene** y el **test de Fligner-Killeen** para las columnas que no cumplen con la condición de normalidad.

```r
homoscedasticityChecker <- function(df_in, normalitychecker_table){
  group = as.factor(df_in$label)
  df_in$label = NULL
  names(normalitychecker_table) <- c("variables", "normal_p_value", "normal_result")
  normalitychecker_table$homoscedasticity_p_value = NA
  normalitychecker_table$homoscedasticity_result = NA

  for (i in 1:nrow(normalitychecker_table)) {
    y <- pull(df_in[,i])
    temp_df <- data.frame(y = y, group = group)

    if(normalitychecker_table[i,3] == "normally distributed"){
      result_test <- leveneTest(y = y, group = group, data = temp_df)
      p_value = result_test$`Pr(>F)`[1]
    } else {
      result_test <- fligner.test(x = y, g = group, data = temp_df)
      p_value = result_test$p.value
    }
    normalitychecker_table$homoscedasticity_p_value[i] = p_value
    normalitychecker_table$homoscedasticity_result[i] = if_else(p_value < 0.05, "statistically differen
  }
  return(normalitychecker_table)
}

pruebas2 <- homoscedasticityChecker(df = df, normalitychecker_table = pruebas)

table(pruebas2$homoscedasticity_result)
```

```
##
## not statistically different variances     statistically different variances
##                                   343                                    37
```

```r
table(pruebas2$normal_result, pruebas2$homoscedasticity_result)
```

```
##
##                          not statistically different variances
##   normally distributed                                     179
##   not normally distributed                                 164
##
##                          statistically different variances
##   normally distributed                                 14
##   not normally distributed                             23
```

4.3 Aplicación de pruebas estadísticas para comparar los grupos de datos.

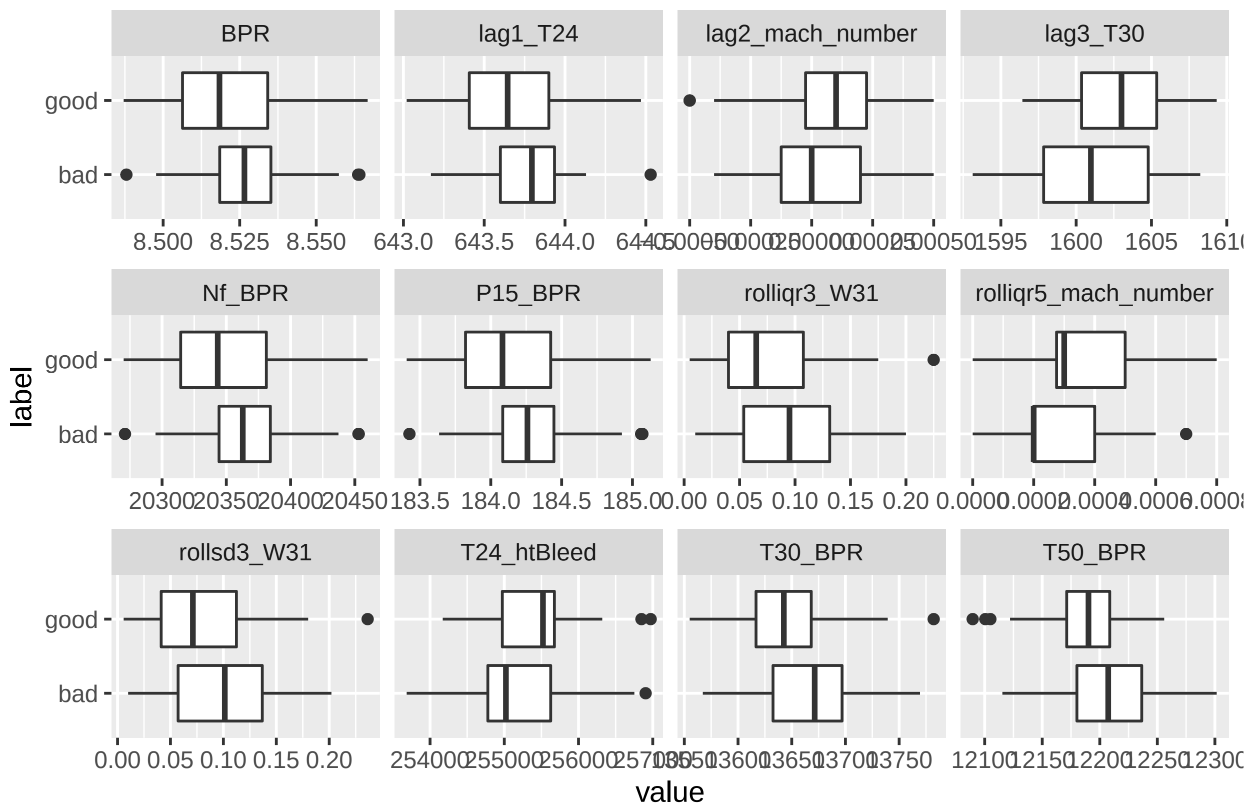
Queremos saber si el promedio de nuestra variables para el grupo de máquinas buenas (m_A) es significativamente diferente al de las máquinas malas (m_B).

En este caso, tenemos dos grupos de muestras no relacionados (es decir, independientes o no apareados). Por lo tanto, es posible utilizar una prueba t independiente para evaluar si las medias son diferentes.

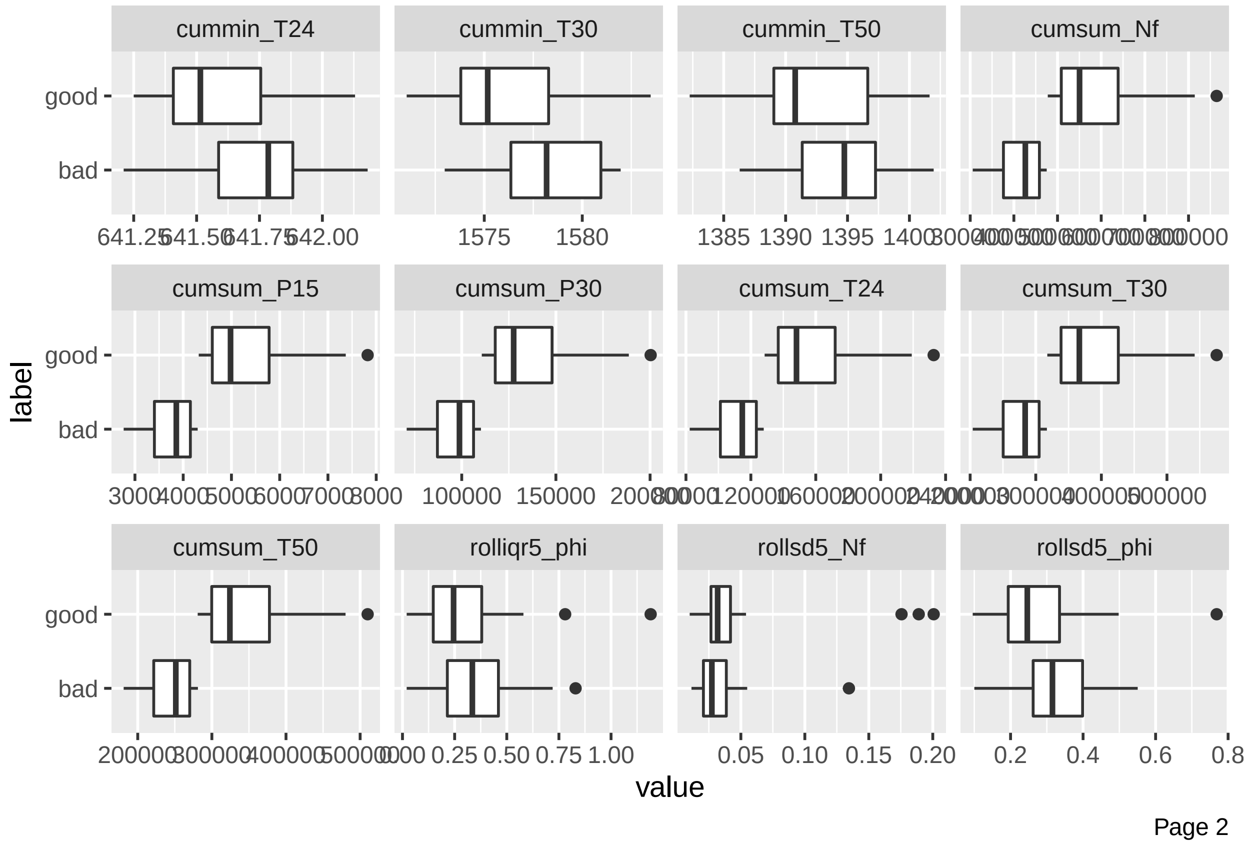
```
significantDifferenceChecker <- function(df_in, normalitychecker_table){

  group = as.factor(df_in$label)
  df_in$label = NULL
  normalitychecker_table$mean_difference_p_value = NA
  normalitychecker_table$mean_difference_result = NA

  for (i in 1:nrow(normalitychecker_table)) {
    y <- pull(df_in[,i])
    temp_df <- data.frame(y = y, group = group)

    if(normalitychecker_table[i,3] == "normally distributed" & normalitychecker_table[i,5] == "statisti
      result_test <- t.test(x = temp_df$y[group == "good"], y = temp_df$y[group == "bad"])
    } else {
      result_test <- wilcox.test(x = temp_df$y[group == "good"], y = temp_df$y[group == "bad"])
    }

    normalitychecker_table$mean_difference_p_value[i] = result_test$p.value
    normalitychecker_table$mean_difference_result[i] = if_else(result_test$p.value < 0.05, "statistical
  }

  return(normalitychecker_table)
}

pruebas3 <- significantDifferenceChecker(df_in = df, normalitychecker_table = pruebas2)

table(pruebas3$mean_difference_result)
```

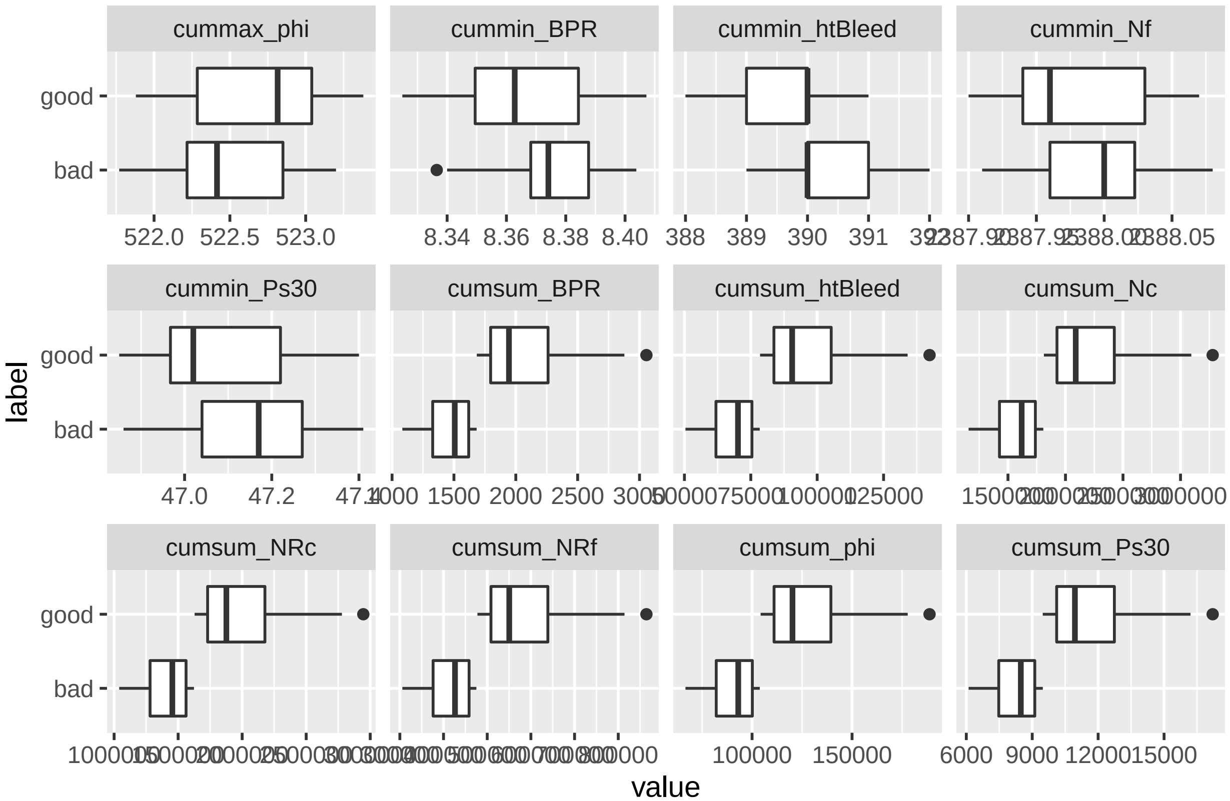
```
##
## not statistically different means     statistically different means
##                               339                                41
```

```
columns_significant <- pruebas3 %>% filter(mean_difference_result == "statistically different means") %

plot_boxplot(df[,c(columns_significant,"label")], by = "label")
```

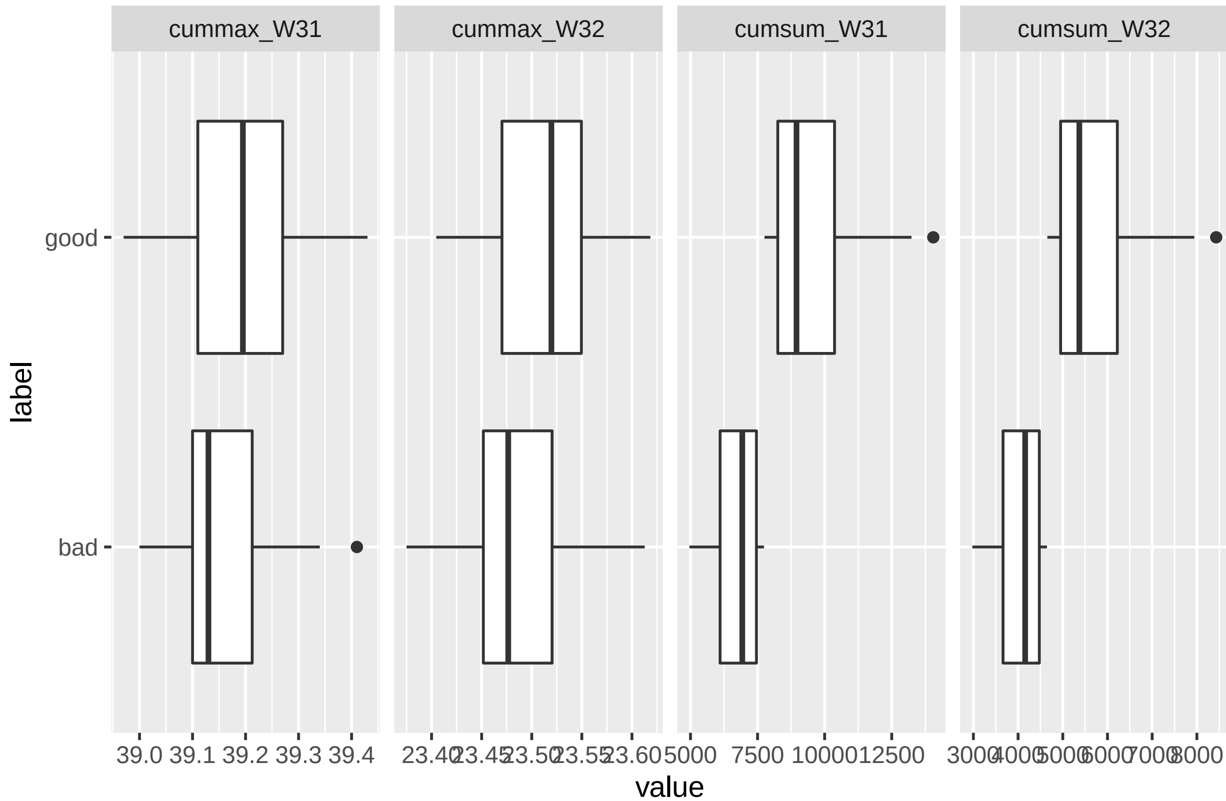

BPR
lag1_T24
lag2_mach_number
lag3_T30
Nf_BPR
P15_BPR
rolliqr3_W31
rolliqr5_mach_number
rollsd3_W31
T24_htBleed
T30_BPR
T50_BPR
good
bad
label
value

Page 1


cummin_T24
cummin_T30
cummin_T50
cumsum_Nf
cumsum_P15
cumsum_P30
cumsum_T24
cumsum_T30
cumsum_T50
rolliqr5_phi
rollsd5_Nf
rollsd5_phi
good
bad
label
value
Page 2

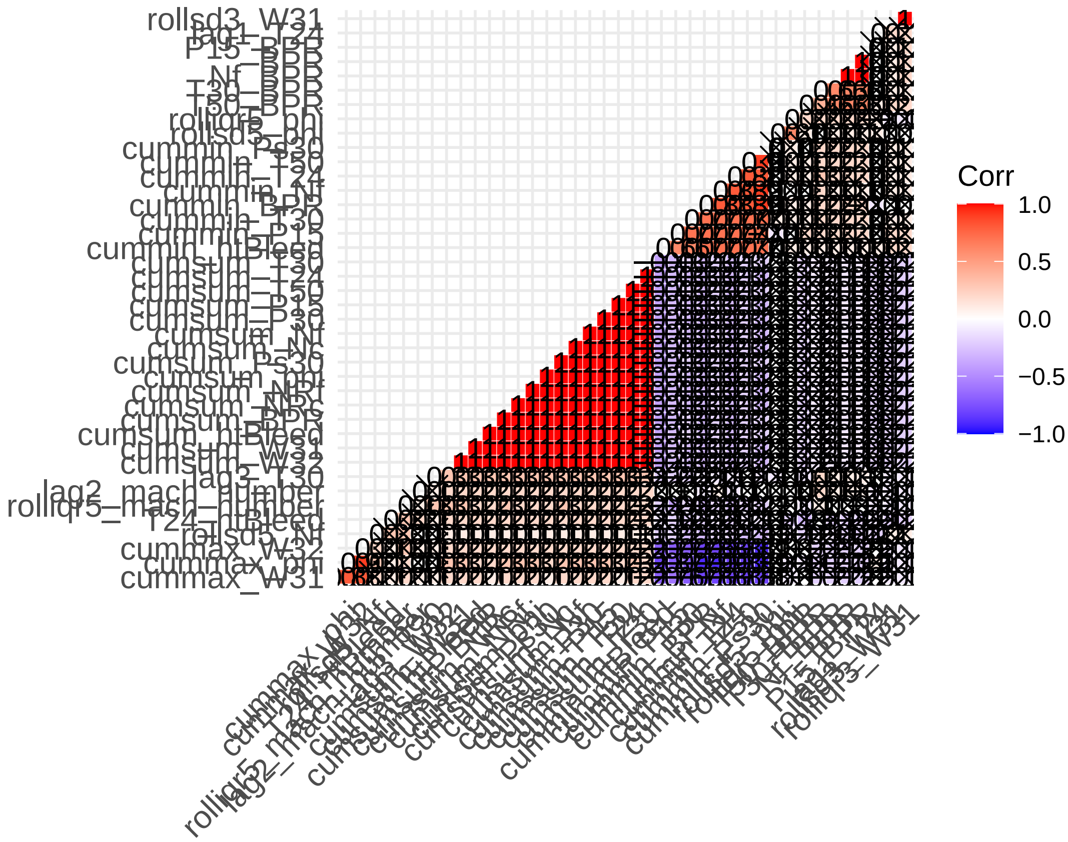

cummax_phi
cummin_BPR
cummin_htBleed
cummin_Nf
cummin_Ps30
cumsum_BPR
cumsum_htBleed
cumsum_Nc
cumsum_NRc
cumsum_NRf
cumsum_phi
cumsum_Ps30
good
bad
label
value
522.0
522.5
523.0
8.34
8.36
8.38
8.40
388
389
390
391
47.0
47.2
1500
2000
2500
100000
125000
100000
150000
6000
9000
12000
15000

Page 3



Ahora vamos a estudiar el problema de multicolinealidad o de cuando dos variables son altamente correlacionadas, al punto de ser una combinación lineal. Además, examinaremos si la correlación es significativamente diferente de cero. Usaremos la prueba de Spearman porque gran parte de las variables no son normalmente distribuídas.

```
df_temp <- df[,columns_significant]


# Compute a correlation matrix
corr <- round(cor(df_temp), 1)
# Compute a matrix of correlation p-values
p.mat <- cor_pmat(df_temp, method="spearman")

# Visualize the lower triangle of the correlation matrix
# Barring the no significant coefficient
corr.plot <- ggcorrplot(
  corr, hc.order = TRUE, type = "lower", lab = TRUE, outline.col = "white",
  p.mat = p.mat
  )

corr.plot
```



El resultado indica que hay muchas correlaciones que no son estadísticamente diferente de cero y, por otro lado, algunas variables son combinación lineales. Este última condición sería un problema en un modelo de regresión lineal.

```r
df_temp <- df[,c(columns_significant[!grepl('cumsum', columns_significant)], "label")]
df_temp$label <- as.factor(df_temp$label)

model <- glm(label ~., data = df_temp, family = binomial) %>%
  stepAIC(trace = FALSE)

# Summarize the final selected model
summary(model)
```

```
##
## Call:
## glm(formula = label ~ T30_BPR + lag2_mach_number + lag3_T30 +
##     rollsd3_W31 + rolliqr3_W31 + rollsd5_Nf + rollsd5_phi + cummin_T30 +
##     cummax_phi + cummin_htBleed, family = binomial, data = df_temp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9656  -0.7588  -0.1976   0.6786   2.3426
##
## Coefficients:
##                       Estimate  Std. Error z value Pr(>|z|)
## (Intercept)      2785.571308 1038.209952   2.683  0.00730 **
```

```
## T30_BPR            -0.017296    0.006311  -2.741  0.00613 **
## lag2_mach_number 3426.358515 1208.903800   2.834  0.00459 **
## lag3_T30            0.114932    0.082391   1.395  0.16303
## rollsd3_W31      -126.389551   65.188692  -1.939  0.05252 .
## rolliqr3_W31      125.561192   67.570453   1.858  0.06314 .
## rollsd5_Nf         20.661018   12.321427   1.677  0.09357 .
## rollsd5_phi        -5.280365    2.381735  -2.217  0.02662 *
## cummin_T30         -0.315528    0.146278  -2.157  0.03100 *
## cummax_phi         -3.105803    1.304948  -2.380  0.01731 *
## cummin_htBleed     -1.567459    0.629747  -2.489  0.01281 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 138.469  on 99  degrees of freedom
## Residual deviance:  92.621  on 89  degrees of freedom
## AIC: 114.62
##
## Number of Fisher Scoring iterations: 5
```

Hemos usado una selección de variable paso a paso basada en el criterio de información de akaike. Un modelo con un AIC bajo se caracteriza por una baja complejidad (minimiza p) y un buen ajuste. Las variables más importantes para explicar si una máquina tiene buen o malo rendimiento son:

- T30_BPR

- lag2_mach_number
- lag3_T30

- rollsd3_W31

- rolliqr3_W31

- rollsd5_Nf

- rollsd5_phi

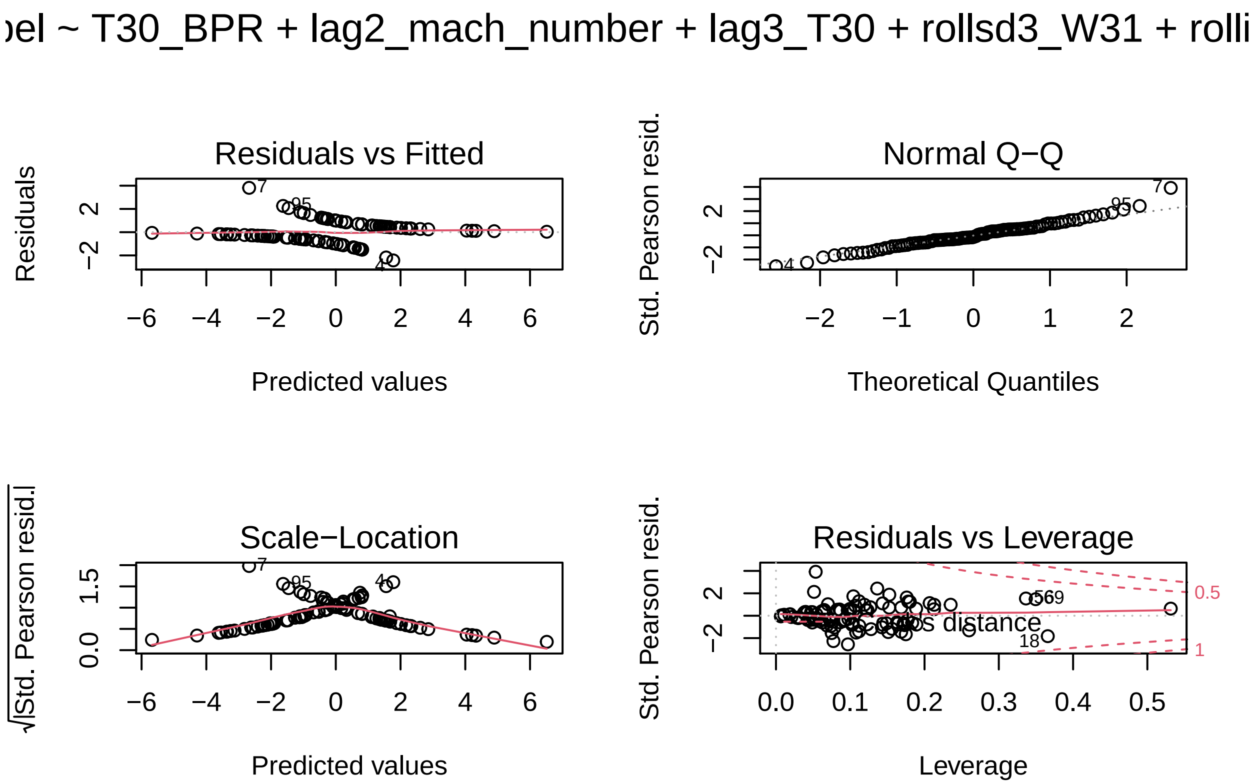
- cummin_T30

- cummax_phi

- cummin_htBleed

5. Representación de los resultados a partir de tablas y gráficas.

```
oldpar <- par(oma = c(0, 0, 3, 0), mfrow = c(2, 2))
plot(model)
```



Vemos el intervalo de confianza de los parametros.

```
confint(model, level = 0.95)
```

```
##                          2.5 %         97.5 %
## (Intercept)       839.08652419 4960.568982273
## T30_BPR            -0.03058919   -0.005583626
## lag2_mach_number 1176.20980745 5977.450565982
## lag3_T30           -0.04353586    0.283654081
## rollsd3_W31      -261.98460973   -3.264000799
## rolliqr3_W31       -2.31786819  265.908040671
## rollsd5_Nf         -0.22533057   51.641153057
## rollsd5_phi       -10.24300577   -0.756712201
## cummin_T30         -0.61570085   -0.036416478
## cummax_phi         -5.82876638   -0.650232964
## cummin_htBleed     -2.88617946   -0.386898867
```

```
pred <-  as.factor(round(predict(model, newdata = df_temp, type="response"),0))

observaciones <- as.factor(if_else(df$label=="bad", 0, 1))

matriz <- confusionMatrix(observaciones, pred)
matriz
```

```
## Confusion Matrix and Statistics
```

```
##
##           Reference
## Prediction  0  1
##          0 39 13
##          1 13 35
##
##                Accuracy : 0.74
##                  95% CI : (0.6427, 0.8226)
##     No Information Rate : 0.52
##     P-Value [Acc > NIR] : 0.000005478
##
##                   Kappa : 0.4792
##
##  Mcnemar's Test P-Value : 1
##
##             Sensitivity : 0.7500
##             Specificity : 0.7292
##          Pos Pred Value : 0.7500
##          Neg Pred Value : 0.7292
##              Prevalence : 0.5200
##          Detection Rate : 0.3900
##    Detection Prevalence : 0.5200
##       Balanced Accuracy : 0.7396
##
##        'Positive' Class : 0
##
```

6. Resolución del problema. A partir de los resultados obtenidos, ¿cuáles son las

conclusiones? ¿Los resultados permiten responder al problema?

Nuestra conclusión es que la información contenida en las variables: lag2_mach_number lag3_T30
rollsd3_W31
rolliqr3_W31
rollsd5_Nf
rollsd5_phi
cummin_T30
cummax_phi
cummin_htBleed

Consigue explicar un 75% sobre cuando una máquina tendrá una vida larga o corta. Nuestro sample size es de 100 máquinas, nuestro modelo es lineal y predecimos sobre el mismo dato usado en el entrenamiento, lo que puede causar overfitting (parcialmente ignorado porque regresión logistica es bastante limitada en generar overfitting).

Este resultado no responde realmente el problema sobre los factores que determinan una vida larga en las máquinas, pero sirve para indicar factores importantes que pueden ser explorados con nuevas variables, modelos más potentes y diferentes "problem framings".

Futuramente continuaremos esta investigación con los otros datasets y tres "problem framings" distintos:

- (1) Como un problema de clasificación para predecir el riesgo de la máquina fallar en su siguiente ciclo;
- (2) Un análisis de supervivencia;

- (3) Un modelos de regresión para predecir el **Remaining Useful Life**