LiBorborema

Open libraries

```
library(tidyverse)
library(GGally)
library(widyr)
library(ggraph)
library(igraph)
library(caret) # Machine Learning
library(randomForest) # RF
library(randomForestExplainer) # RF
library(doParallel)
library(gridExtra)
library(ggpubr)
library(sf)
```

1) Reading data

```
#Training data
df <- data.table::fread("E:\\PROJETOS\\LITIO_BORBOREMA\\r\\solo_gama_points.txt") %>%
  mutate(BB_SUBCL_1 = as.factor(BB_SUBCL_1)) %>%
  rename('gamma_k' = 'k_borborem',
         'gamma_eu'= 'eu_borbore',
         'gamma_eth' = 'eth_borbor',
         'gamma_ct' = 'ct_borbore',
         'litho' = 'BB_SUBCL_1') %>%
  filter(gamma_eth > 0) %>%
  mutate(gamma_k = case_when(gamma_k < 0 ~ 0.05*mean(gamma_k, na.rm = TRUE),
                             is.na(gamma_k) ~ 0.05*mean(gamma_k, na.rm = TRUE),
                             TRUE ~ as.double(gamma_k))) %>%
  mutate(gamma_eth = case_when(gamma_eth < 0 ~ 0.05*mean(gamma_eth, na.rm = TRUE),
                               is.na(gamma_eth) ~ 0.05*mean(gamma_eth, na.rm = TRUE),
                               TRUE ~ as.double(gamma_eth))) %>%
  mutate(gamma_eu = case_when(gamma_eu < 0 ~ 0.05*mean(gamma_eu, na.rm = TRUE),
                              # gamma_eu > 20 ~ 30,
                              is.na(gamma_eu) ~ 0.05*mean(gamma_eu, na.rm = TRUE),
                              TRUE ~ as.double(gamma_eu))) %>%
  mutate(gamma_ct = case_when(gamma_ct < 0 ~ 0.05*mean(gamma_ct, na.rm = TRUE),
                              is.na(gamma_ct) ~ 0.05*mean(gamma_ct, na.rm = TRUE),
                              TRUE ~ as.double(gamma_ct))) %>%
  mutate(gamma_k_th = gamma_k/(gamma_eth),
         gamma_u_th = gamma_eu/(gamma_eth),
         gamma_k_eu = gamma_k/(gamma_eu),
         gamma_f_factor = gamma_k*gamma_eu/gamma_eth) %>%
  select(1,3:63,65:68,64)

glimpse(df)
```

```
## Rows: 483
## Columns: 67
## $ FID           <int> 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 1~
## $ x             <dbl> -37.61688, -37.89100, -38.08643, -37.85801, -37.85801, ~
## $ y             <dbl> -9.14987, -9.36639, -9.36344, -9.14286, -9.14286, -9.00~
## $ uf            <chr> "AL", "AL", "AL", "AL", "AL", "AL", "AL", "AL", "AL", "~
## $ num_campo     <chr> "JW-L-3038", "JW-L-3039", "JW-L-3040", "JW-L-3041", "JW~
## $ num_lab       <chr> "FEO319", "FEO320", "FEO321", "FEO322", "FEO323", "FEO2~
## $ ag_ppm        <dbl> 0.050, 0.005, 0.010, 0.005, 0.005, 0.100, 0.020, 0.030,~
## $ al_pct        <dbl> 2.180, 0.860, 0.560, 0.550, 0.505, 1.980, 1.380, 3.770,~
## $ as_ppm        <dbl> 2.00, 2.00, 0.50, 0.50, 0.75, 6.00, 1.00, 2.00, 3.00, 5~
## $ au_ppm        <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0~
## $ b_ppm         <dbl> 5.0, 5.0, 5.0, 5.0, 0.5, 5.0, 5.0, 5.0, 5.0, 5.0, 5.0, ~
## $ ba_ppm        <dbl> 174.0, 98.0, 48.0, 39.0, 35.0, 110.0, 65.0, 101.0, 13.0~
## $ be_ppm        <dbl> 2.40, 0.90, 0.50, 0.40, 0.40, 1.10, 1.10, 0.50, 0.60, 0~
## $ bi_ppm        <dbl> 0.210, 0.110, 0.170, 0.010, 0.055, 0.200, 0.140, 2.090,~
## $ ca_pct        <dbl> 0.250, 0.150, 0.080, 0.060, 0.060, 0.110, 0.190, 0.050,~
## $ cd_ppm        <dbl> 0.005, 0.005, 0.005, 0.005, 0.005, 0.010, 0.030, 0.020,~
## $ ce_ppm        <dbl> 101.07, 29.23, 8.52, 10.72, 10765.00, 67.01, 22.71, 23.~
## $ co_ppm        <dbl> 7.60, 2.90, 1.30, 0.80, 0.55, 4.50, 3.40, 11.20, 1.30, ~
## $ cr_ppm        <dbl> 44, 21, 19, 32, 32, 18, 77, 8, 70, 15, 13, 12, 13, 27, ~
## $ cs_ppm        <dbl> 2.400, 0.830, 1.020, 0.670, 0.635, 2.160, 2.450, 6.520,~
## $ cu_ppm        <dbl> 7.7, 4.7, 3.9, 2.3, 1.5, 12.5, 9.1, 3.6, 5.5, 6.8, 3.5,~
## $ fe_pct        <dbl> 2.25, 0.99, 0.95, 0.55, 0.33, 2.56, 1.56, 3.11, 5.34, 2~
## $ ga_ppm        <dbl> 8.50, 3.70, 2.20, 2.30, 1.95, 8.40, 4.40, 11.70, 20.20,~
## $ ge_ppm        <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0~
## $ hf_ppm        <dbl> 0.070, 0.025, 0.025, 0.025, 0.025, 0.050, 0.025, 0.025,~
## $ hg_ppm        <dbl> 0.030, 0.020, 0.020, 0.030, 0.005, 0.030, 0.020, 0.160,~
## $ in_ppm        <dbl> 0.02, 0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.02, 0.04, 0~
## $ k_pct         <dbl> 0.380, 0.150, 0.170, 0.090, 0.075, 0.230, 0.290, 0.100,~
## $ la_ppm        <dbl> 30.70, 19.20, 5.10, 7.40, 7.45, 33.70, 16.30, 12.30, 2.~
## $ li_ppm        <dbl> 15.0, 7.0, 4.0, 2.0, 1.5, 8.0, 12.0, 13.0, 2.0, 4.0, 0.~
## $ mg_pct        <dbl> 0.220, 0.160, 0.070, 0.030, 0.030, 0.290, 0.250, 0.090,~
## $ mn_ppm        <dbl> 321.0, 135.0, 139.0, 60.0, 34.5, 314.0, 236.0, 542.0, 8~
## $ mo_ppm        <dbl> 0.550, 0.025, 0.025, 0.120, 0.115, 2.220, 1.930, 1.250,~
## $ na_pct        <dbl> 0.050, 0.020, 0.020, 0.010, 0.010, 0.005, 0.040, 0.005,~
## $ nb_ppm        <dbl> 0.460, 0.380, 0.360, 0.080, 0.090, 0.370, 0.300, 0.230,~
## $ ni_ppm        <dbl> 21.80, 10.10, 7.50, 15.80, 16.15, 6.00, 39.30, 3.30, 6.~
## $ p_ppm         <dbl> 1153, 173, 82, 52, 25, 514, 161, 780, 163, 287, 158, 92~
## $ pb_ppm        <dbl> 18.2, 11.5, 5.6, 11.0, 10.4, 12.3, 5.9, 23.2, 6.0, 6.3,~
## $ pd_ppm        <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0~
## $ pt_ppm        <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0~
## $ rb_ppm        <dbl> 39.80, 22.60, 14.60, 8.80, 8.15, 28.70, 37.90, 13.90, 2~
## $ re_ppm        <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0~
## $ s_pct         <dbl> 0.010, 0.005, 0.005, 0.005, 0.005, 0.010, 0.005, 0.030,~
## $ sb_ppm        <dbl> 0.070, 0.025, 0.100, 0.050, 0.025, 0.220, 0.160, 0.140,~
## $ sc_ppm        <dbl> 3.20, 1.50, 0.50, 0.50, 0.45, 4.00, 3.10, 4.90, 6.10, 0~
## $ se_ppm        <dbl> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1.0, 2.0, 0.5, 1.0, ~
## $ sn_ppm        <dbl> 2.00, 1.30, 0.80, 0.80, 0.75, 1.80, 1.20, 1.10, 1.80, 3~
## $ sr_ppm        <dbl> 72.40, 35.30, 15.50, 17.10, 15.95, 17.60, 30.20, 7.60, ~
## $ ta_ppm        <dbl> 0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.025,~
## $ te_ppm        <dbl> 0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.025,~
## $ th_ppm        <dbl> 10.50, 4.00, 2.90, 5.00, 4.85, 12.50, 5.70, 6.30, 9.40,~
## $ ti_pct        <dbl> 0.050, 0.040, 0.030, 0.005, 0.005, 0.040, 0.060, 0.040,~
```

```
## $ u_ppm          <dbl> 1.450, 0.440, 0.440, 0.630, 0.555, 2.380, 0.570, 2.680,~
## $ v_ppm          <dbl> 47, 20, 10, 9, 8, 57, 22, 72, 106, 47, 49, 36, 53, 59, ~
## $ w_ppm          <dbl> 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.30, 0~
## $ y_ppm          <dbl> 7.38, 3.78, 1.61, 1.72, 1675.00, 12.96, 6.34, 4.63, 0.5~
## $ zn_ppm         <dbl> 31.0, 22.0, 11.0, 5.0, 5.5, 35.0, 25.0, 32.0, 10.0, 12.~
## $ zr_ppm         <dbl> 3.00, 0.25, 1.30, 1.10, 0.90, 0.25, 0.25, 0.25, 6.30, 1~
## $ gamma_k        <dbl> 5.110150, 3.912840, 4.013260, 4.375590, 4.375590, 1.351~
## $ gamma_eu       <dbl> 1.410520, 0.546260, 0.875196, 0.621949, 0.621949, 2.027~
## $ gamma_eth      <dbl> 13.692400, 7.164770, 6.775610, 5.909220, 5.909220, 11.8~
## $ gamma_ct       <dbl> 22.25160, 15.00010, 16.07260, 15.70790, 15.70790, 10.55~
## $ gamma_k_th     <dbl> 0.37321068, 0.54612218, 0.59230976, 0.74046823, 0.74046~
## $ gamma_u_th     <dbl> 0.10301481, 0.07624250, 0.12916859, 0.10525060, 0.10525~
## $ gamma_k_eu     <dbl> 3.62288369, 7.16296251, 4.58555558, 7.03528714, 7.03528~
## $ gamma_f_factor <dbl> 0.52642113, 0.29832470, 0.51838713, 0.46053347, 0.46053~
## $ litho          <fct> Metamorfismo regional, Metamorfismo regional, Metamorfi~
```

Reading Geophysical Data

```
# Geophysical data
newdata <- readRDS('E:\\PROJETOS\\LITIO_BORBOREMA\\r\\xyz\\borborema_gamma.RDS')

glimpse(newdata)
```

```
## Rows: 401,413
## Columns: 10
## $ x              <dbl> -42.63, -42.57, -42.56, -42.56, -42.55, -42.55, -42.54,~
## $ y              <dbl> -9.00, -9.00, -9.01, -9.00, -9.01, -9.00, -9.00, -9.00,~
## $ gamma_k        <dbl> 0.83, 0.61, 0.58, 0.48, 0.53, 0.54, 0.46, 0.74, 0.22, 0~
## $ gamma_eth      <dbl> 8.71, 7.98, 8.11, 7.87, 8.08, 7.99, 7.70, 7.88, 4.06, 7~
## $ gamma_eu       <dbl> 0.9600000, 1.3000000, 1.4300000, 1.2300000, 1.1900000, ~
## $ gamma_ct       <dbl> 5.14, 4.69, 4.76, 4.43, 4.57, 4.56, 4.20, 4.85, 2.04, 4~
## $ gamma_k_th     <dbl> 0.09529167, 0.07644014, 0.07151576, 0.06099033, 0.06559~
## $ gamma_u_th     <dbl> 0.11021687, 0.16290523, 0.17632335, 0.15628772, 0.14727~
## $ gamma_k_eu     <dbl> 0.8644933, 0.4691947, 0.4055660, 0.3902122, 0.4453407, ~
## $ gamma_f_factor <dbl> 0.09150056, 0.09939612, 0.10229233, 0.07503984, 0.07807~
```

Reading Pegmatite Datasets

```
# Recmin
pegmatites <- foreign::read.dbf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\pegmatitos_li.dbf')
recmin1 <- foreign ::read.dbf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\borborema_PegLi.dbf') #RASTERVALU


glimpse(pegmatites)
```

```
## Rows: 1,226
## Columns: 34
## $ ID_AFLORAM <int> 303195, 303237, 303227, 303182, 303213, 303215, 303216, 410~
## $ ORIGEM     <fct> Oracle, Oracle, Oracle, Oracle, Oracle, Oracle, Oracle, Ora~
## $ METODO_GEO <fct> GPS Manual pós 24/05/2000, GPS Manual pós 24/05/2000, GPS M~
## $ TOPONIMIA  <fct> "GARIMPO IPUEIRAS", "FAZENDA SANTA FÉ", "BERILÓPOLIS", "SOL~
## $ MUNICIPIO  <fct> Milhã, Milhã, Milhã, Solonópole, Solonópole, Solonópole, So~
```

```
## $ UF         <fct> CE, CE, CE, CE, CE, CE, CE, PE, CE, RN, RN, RN, RN, RN, AL,~
## $ GEOLOGO    <fct> Ana Paula Justo, Ana Paula Justo, Ana Paula Justo, Ana Paul~
## $ DATA_CADAS <date> 2010-07-23, 2010-07-19, 2010-07-20, 2010-07-28, 2010-07-21~
## $ TIPO_AFLOR <fct> Mina, Lajedo ou Lajeiro, Mina, Mina, Mina, Mina, Mina, Laje~
## $ DESCRICAO  <fct> "Filão pegmatítico homogêneo, com Qtz-Fsp-Caulim-Ms-Turmver~
## $ NUMERO_CAM <fct> AP-0076, AP-0022, AP-0032, AP-0089, AP-0046, AP-0044, AP-00~
## $ ROCHAS     <fct> "Metaleucogranito, Pegmatito", "Pegmatito", "Metagranitóide~
## $ SUREG      <fct> OUTROS, OUTROS, OUTROS, OUTROS, OUTROS, OUTROS, OUTROS, OUT~
## $ PROJETO    <fct> Geologia da Folha Senador Pompeu, Geologia da Folha Senador~
## $ CODIGO_FOL <fct> SB.24-V-D-VI, SB.24-V-D-VI, SB.24-V-D-VI, SB.24-V-D-VI, SB.~
## $ FOLHA      <fct> Senador Pompeu, Senador Pompeu, Senador Pompeu, Senador Pom~
## $ ID_OCORREN <int> 56966, 57004, 56996, 56953, 56984, 56986, 56987, 66056, 533~
## $ PROVINCIA  <fct> Indeterminada, Indeterminada, Indeterminada, Indeterminada,~
## $ STATUS_ECO <fct> Garimpo, Não explotado, Mina, Garimpo, Garimpo, Garimpo, Ga~
## $ IMPORTANCI <fct> Depósito, Indício, Depósito, Depósito, Depósito, Depósito, ~
## $ LOCALIZACA <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, Open pit, O~
## $ SITUACAO_M <fct> NA, NA, Inativo(a), NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
## $ MOTIVO_INA <fct> NA, NA, Paralisado(a), NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ SITUACAO_G <fct> NA, NA, Inativo(a), NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
## $ MOTIVO_I_1 <fct> NA, NA, Paralisado(a), NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ SUBSTANCIA <fct> "Feldspato, Turmalina", "Água marinha, Berílio, Feldspato",~
## $ ROCHAS_HOS <fct> "Pegmatito", "Pegmatito", "Pegmatito", "Pegmatito", "Pegmat~
## $ ROCHAS_ENC <fct> "Metaleucogranito, Muscovita xisto", "Metagranitóide", "Met~
## $ CLASSES_UT <fct> Gemas, Gemas, Rochas e minerais industriais, Gemas, Gemas, ~
## $ MORFOLOGIA <fct> "Filoneana", "Filoneana", "Filoneana", "Filoneana, Lenticul~
## $ TEXTURAS   <fct> "Pegmatítica", "Pegmatítica, Porfirítica", "Pegmatítica", "~
## $ TIPOS_ALTE <fct> "Turmalinização", NA, NA, "Turmalinização", NA, NA, NA, NA,~
## $ x          <dbl> -39.1381, -39.1082, -39.0981, -39.0607, -39.0304, -39.0224,~
## $ y          <dbl> -5.54275, -5.59631, -5.55298, -5.70445, -5.61417, -5.61922,~
```

Reading Domains and StopWords for further analysis

```
# Domains ----
MC <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_MC.shp',as_tibble = TRUE) %>%
  st_coordinates()
CC <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_CC.shp',as_tibble = TRUE) %>%
  st_coordinates()
JG <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_JG.shp',as_tibble = TRUE) %>%
  st_coordinates()
PS <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_PS.shp',as_tibble = TRUE) %>%
  st_coordinates()
JC <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_JC.shp',as_tibble = TRUE) %>%
  st_coordinates()
ZT <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_ZT.shp',as_tibble = TRUE) %>%
  st_coordinates()
RP <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_RP.shp',as_tibble = TRUE) %>%
  st_coordinates()
PA <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_PA.shp',as_tibble = TRUE) %>%
  st_coordinates()
SE <- sf::read_sf('E:\\PROJETOS\\LITIO_BORBOREMA\\shp\\Domains\\Domain_SE.shp',as_tibble = TRUE) %>%
  st_coordinates()

stop_words_pt <- data.table::fread('C:\\Users\\guilherme.ferreira\\Documents\\STOPWORDS_POTUGUES.txt')
```

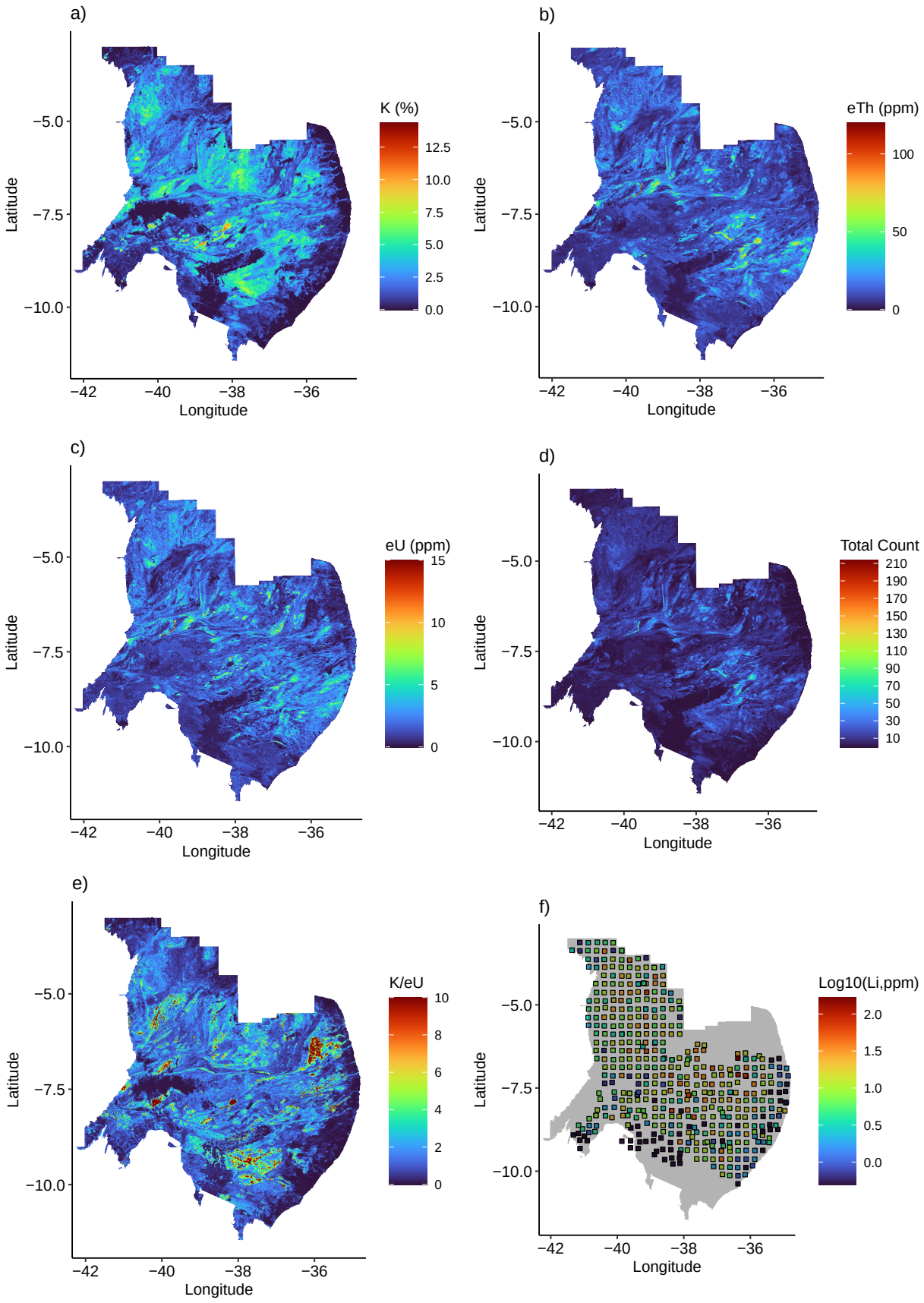
Data Processing

```
# Mapas de Entrada

# Potássio ----
mk <-
   newdata %>%
   mutate(gamma_k = case_when(gamma_k < 0 ~ 0,
                              TRUE ~ gamma_k)) %>%
   ggplot(aes(x,y,fill = gamma_k)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(0,300,2.5)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'K (%)',
        x = 'Longitude',
        y = 'Latitude',
        title = 'a)')


# Tório ----
mth <-
   newdata %>%
   mutate(gamma_eth = case_when(gamma_eth > 150 ~ 150,
                                gamma_eth <= 150 ~ gamma_eth)) %>%
   ggplot(aes(x,y,fill = gamma_eth)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(0,300,50)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'eTh (ppm)',
        x = 'Longitude',
        y = 'Latitude',
        title = 'b)')


# Urânio ----
mu <-
   newdata %>%
   mutate(gamma_eu = case_when(gamma_eu < 0 ~ 0,
```

```
                              gamma_eu > 15 ~ 15,
                              TRUE ~ gamma_eu)) %>%
   ggplot(aes(x,y,fill = gamma_eu)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(-10,300,5)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'eU (ppm)',
        x = 'Longitude',
        y = 'Latitude',
        title = 'c)')


# Total Count ----
mct <-
   newdata %>%
   ggplot(aes(x,y,fill = gamma_ct)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(-10,300,20)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'Total Count',
        x = 'Longitude',
        y = 'Latitude',
        title = 'd)')


# F Factor ----
mf <-
   newdata %>%
   mutate(gamma_f_factor = case_when(gamma_f_factor > 10.1 ~ 10.1,
                                     TRUE ~ gamma_f_factor),
   ) %>%
   ggplot(aes(x,y,fill = gamma_f_factor)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
```

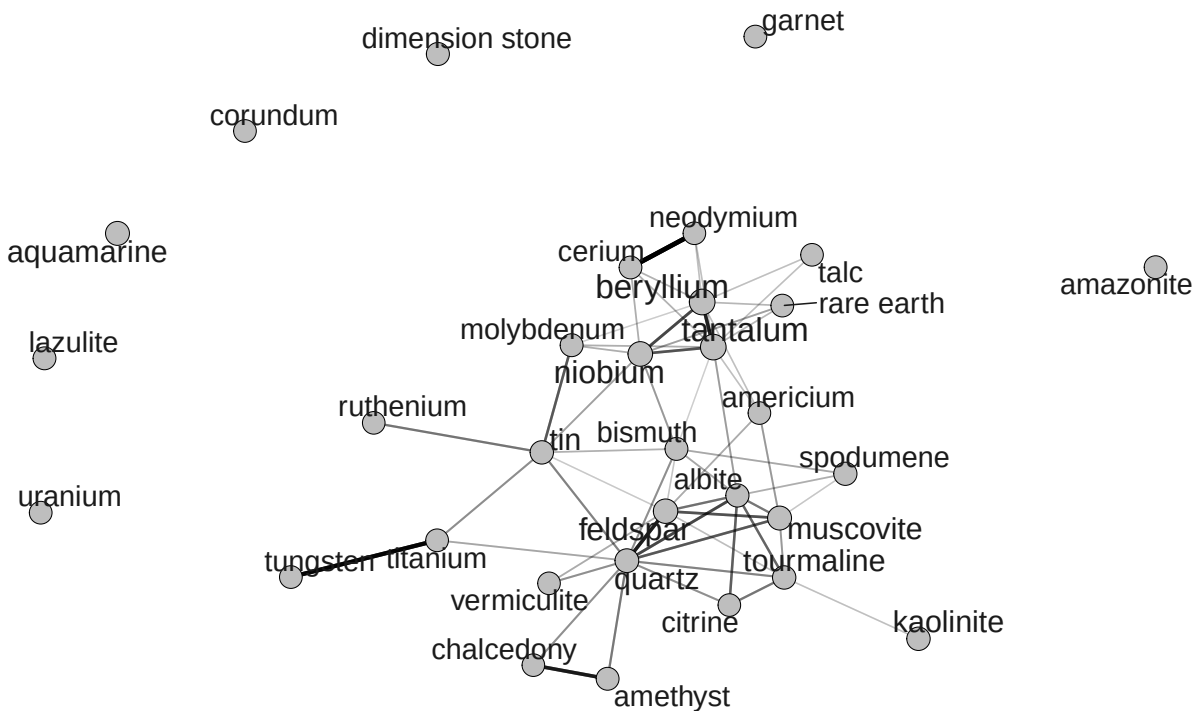
```
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(-10,300,2.5)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'F Factor',
        x = 'Longitude',
        y = 'Latitude')

# K/eU ----
mku <-
   newdata %>%
   mutate(gamma_k_eu = case_when(gamma_k_eu > 10 ~ 10,
                                 gamma_k_eu < 0 ~ 0,
                                 TRUE ~ gamma_k_eu),
   ) %>%
   ggplot(aes(x,y,fill = gamma_k_eu)) +
   geom_raster() +
   coord_equal(xlim = c(-42,-35),
               ylim = c(-11.5,-3),
               expand = TRUE,
               clip = 'on') +
   scale_fill_gradientn(colours = pals::turbo(),
                        guide = guide_colorbar(barheight = 10,
                                               barwidth = 2),
                        breaks=seq(-10,300,2)) +
   ggpubr::theme_pubr() +
   theme(legend.position = 'right') +
   labs(fill = 'K/eU',
        x = 'Longitude',
        y = 'Latitude',
        title = 'e)')

# Soil Geochemistry
mgeq <-
    df %>%
    ggplot(aes(x,y, fill = log10(li_ppm))) +
    geom_point(shape = 22, col = 'black', alpha = 1) +
    geom_raster(inherit.aes = FALSE,
                data = newdata,
                mapping = aes(x,y),
                fill = 'black',
                alpha = .3) +
    coord_equal(xlim = c(-42,-35),
                ylim = c(-11.5,-3),
                expand = TRUE,
                clip = 'on') +
    scale_fill_gradientn(colours = pals::turbo(),
                         guide = guide_colorbar(barheight = 10,
                                                barwidth = 2)) +
    ggpubr::theme_pubr() +
    labs(x = 'Longitude',
```

```
         y = 'Latitude',
         fill = 'Log10(Li,ppm)',
         title = 'f)') +
    scale_size(guide = 'none') +
    theme(legend.position = "right")

grid.arrange(mk,mth,mu,mct,mku,mgeq,
                        layout_matrix = matrix(c(1,2,
                                                 3,4,
                                                 5,6), ncol = 2, byrow = TRUE))
```


a)
K (%)
12.5
10.0
7.5
5.0
2.5
0.0
b)
eTh (ppm)
100
50
0
c)
eU (ppm)
15
10
5
0
d)
Total Count
210
190
170
150
130
110
90
70
50
30
10
e)
K/eU
10
8
6
4
2
0
f)
Log10(Li,ppm)
2.0
1.5
1.0
0.5
0.0
Latitude
−5.0
−7.5
−10.0
Longitude
−42
−40
−38
−36

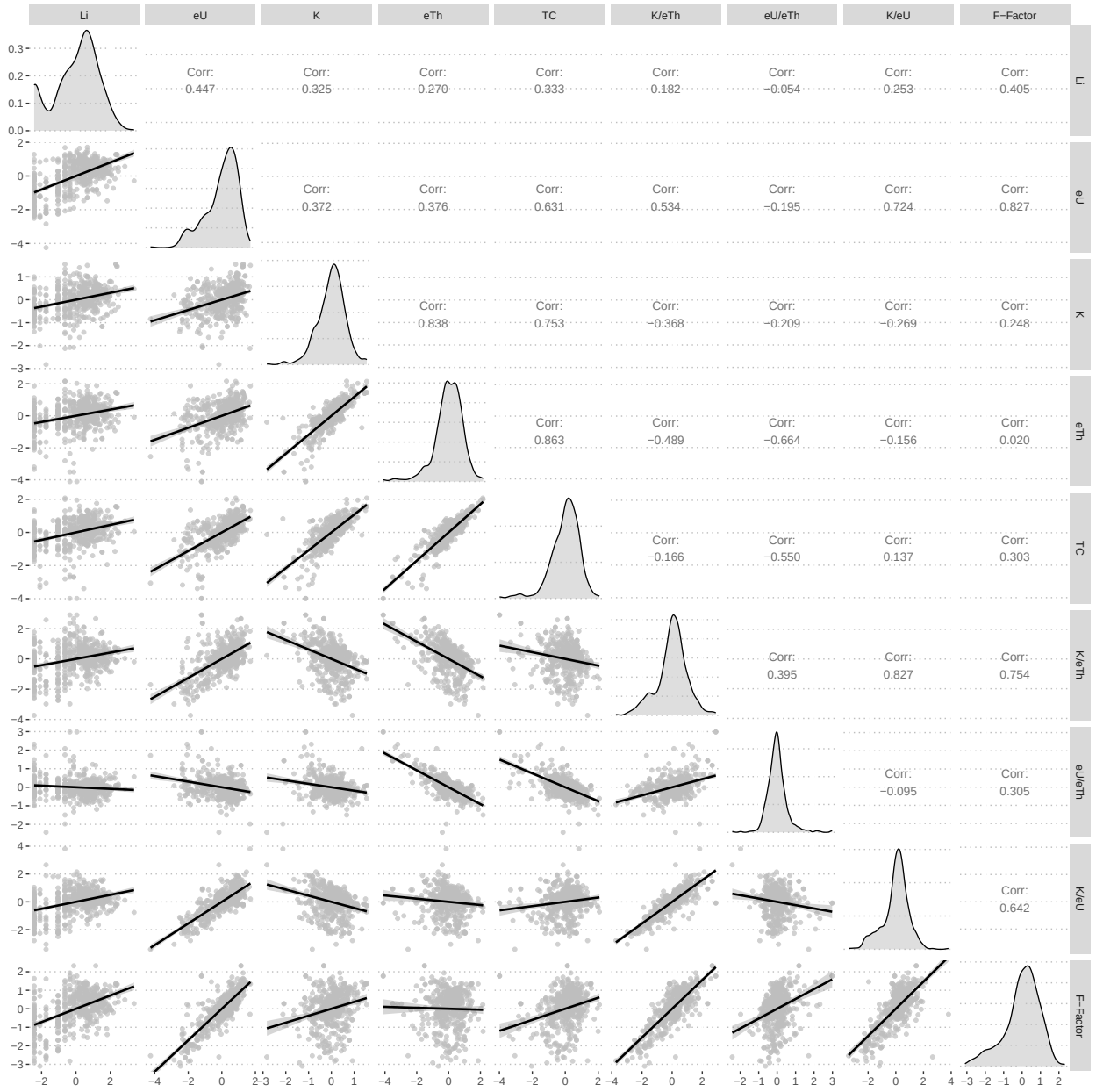
```r
pegmatites %>%
  filter(str_detect(SUBSTANCIA,'Lítio|Espodumênio')) -> t


recmin <-
  pegmatites %>%
  mutate(SUBSTANCIA = as.character(SUBSTANCIA)) %>%
  tidytext::unnest_tokens(output = word, input = SUBSTANCIA) %>%
  anti_join(stop_words_pt, by = c('word' = 'Portugues')) %>%
  filter(!word %in% c('marinha','cristal','ferro','rocha','níquel','rosa','raras','hialino','fumê','ped
  mutate(word = case_when(word == 'água' ~ 'agua-marinha',
                          word == 'terras' ~ 'terras-raras',
                          word == 'ornamental' ~ 'pedra-ornamental',
                          word == 'espodumênio' ~ 'lítio',
                          TRUE ~ as.character(word))) %>%
  mutate(word = case_when(word == 'terras-raras' ~ 'rare earth',
                          word == 'neodímio' ~ 'neodymium',
                          word == 'cério' ~ 'cerium',
                          word == 'granada' ~ 'garnet',
                          word == 'talco' ~ 'talc',
                          word == 'amazonita' ~ 'amazonite',
                          word == 'córindon' ~ 'corundum',
                          word == 'urânio' ~ 'uranium',
                          word == 'agua-marinha' ~ 'aquamarine',
                          word == 'lazulita' ~ 'lazulite',
                          word == 'pegmatito' ~ 'dimension stone',
                          word == 'pedra-ornamental' ~ 'dimension stone',
                          word == 'caulim' ~ 'kaolinite',
                          word == 'amerício' ~ 'americium',
                          word == 'muscovita' ~ 'muscovite',
                          word == 'vermiculita' ~ 'vermiculite',
                          word == 'feldspato' ~ 'feldspar',
                          word == 'citrino' ~ 'citrine',
                          word == 'turmalina' ~ 'tourmaline',
                          word == 'albita' ~ 'albite',
                          word == 'quartzo' ~ 'quartz',
                          word == 'calcedônia' ~ 'chalcedony',
                          word == 'ametista' ~ 'amethyst',
                          word == 'lítio' ~ 'spodumene',
                          word == 'bismuto' ~ 'bismuth',
                          word == 'tântalo' ~ 'tantalum',
                          word == 'nióbio' ~ 'niobium',
                          word == 'berílio' ~ 'beryllium',
                          word == 'estanho' ~ 'tin',
                          word == 'titânio' ~ 'titanium',
                          word == 'tungstênio' ~ 'tungsten',
                          word == 'molibdênio' ~ 'molybdenum',
                          word == 'rutênio' ~ 'ruthenium',
                          word == 'mica' ~ 'muscovite',
                          TRUE ~ as.character(word)))

recmin_freq <-
  recmin %>%
```

```
  count(word, sort = TRUE) %>%
  filter(n > 0)

word_correlations <-
  recmin %>%
  semi_join(recmin_freq, by = 'word') %>%
  widyr::pairwise_cor(item = word, feature = ID_AFLORAM) %>%
  filter(correlation > 0)

set.seed(131)
(p.net <-
    graph_from_data_frame(d = word_correlations,
                          vertices = recmin_freq) %>%
    ggraph(layout = 'fr') +
    geom_edge_link(aes(alpha = correlation,
                       width = correlation)) +
    geom_node_point(aes(size = n),
                    shape = 21,
                    fill = 'grey',
                    col = 'black') +
    geom_node_text(aes(label = name,
                       size = n), repel = TRUE,
                   col = 'grey10') +
    ggpubr::theme_transparent() +
    theme(legend.position = 'none') +
    scale_edge_width(range = c(.5,1.5)) +
    scale_size(range = c(7,8))
)
```



GGPAIRS

```r
df %>%
  dplyr::select(li_ppm,gamma_k:gamma_f_factor) %>%
  magrittr::set_colnames(c('Li','eU','K','eTh',
                           'TC','K/eTh','eU/eTh',
                           'K/eU','F-Factor')) %>%
  geoquimica::elem_norm(method = 'clr') %>%
  GGally::ggpairs(lower=list(continuous = GGally::wrap("smooth",
                                                       col = 'grey',
                                                       alpha = .7)),
                  diag = list(continuous = GGally::wrap("densityDiag",
                                                        alpha=0.5,
                                                        fill = 'grey' )),
                  upper = list(continuous = GGally::wrap('cor',
                                                         stars = F,
                                                         method = 'spearman'))
                  ) +
  ggpubr::theme_pubclean()
```



Model Tunning

```
doParallel::registerDoParallel(cores = detectCores()-1)

# Training-Test Split
set.seed(0)
trainset <- df %>%
  select(li_ppm,gamma_k:gamma_f_factor) %>%
  sample_frac(.85)

testset <- df %>%
  select(li_ppm,gamma_k:gamma_f_factor) %>%
  anti_join(trainset)
```

```r
# Data tunning
tg <- data.frame(mtry = seq(2,8, by = 1))
ctrl <- caret::trainControl(method = 'LGOCV',
                            number = 10,
                            # repeats = 3,
                            search = 'grid',
                            allowParallel = TRUE,
                            verboseIter = TRUE,p = .75)

set.seed(0)
tunning <- caret::train(li_ppm ~ ., data = trainset,
                        method = "rf",
                        trControl = ctrl,
                        ntree = 1000,
                        importance = TRUE,
                        tuneGrid = tg, proximity = TRUE)
```

```
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 2 on full training set
```

```r
print(tunning)
```

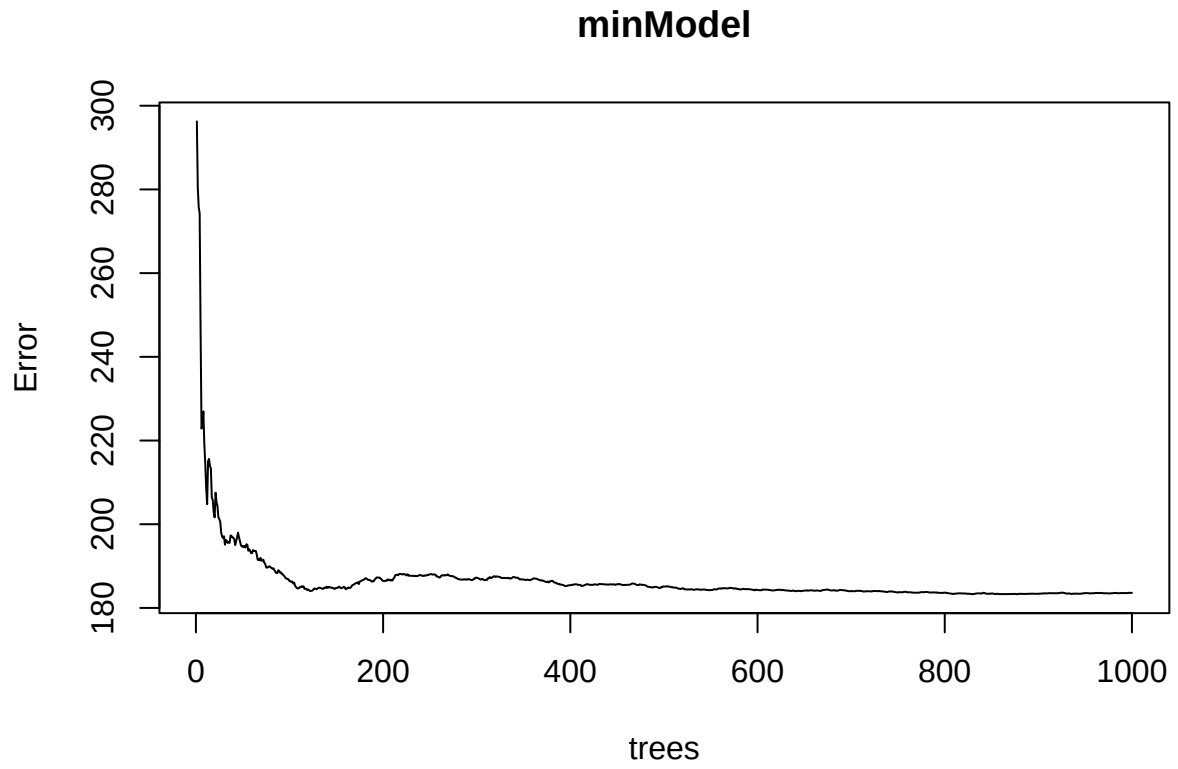
```
## Random Forest
##
## 411 samples
##   8 predictor
##
## No pre-processing
## Resampling: Repeated Train/Test Splits Estimated (10 reps, 75%)
## Summary of sample sizes: 310, 310, 310, 310, 310, 310, ...
## Resampling results across tuning parameters:
##
##   mtry  RMSE      Rsquared   MAE
##   2     13.52949  0.1677051  8.123033
##   3     13.72128  0.1574320  8.268085
##   4     13.88093  0.1467270  8.350370
##   5     14.00980  0.1401867  8.414574
##   6     14.19530  0.1322409  8.525891
##   7     14.32263  0.1247797  8.562173
##   8     14.39980  0.1212411  8.595984
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 2.
```

```r
# Model training

set.seed(0)
minModel <- randomForest(li_ppm ~ .,
                         trainset,
                         proximity = TRUE,
                         ntree = 1000, localImp = TRUE,
```

```
                          mtry = tunning$bestTune[[1,1]])

print(minModel)
```

```
##
## Call:
##  randomForest(formula = li_ppm ~ ., data = trainset, proximity = TRUE,      ntree = 1000, localImp =
##                Type of random forest: regression
##                      Number of trees: 1000
## No. of variables tried at each split: 2
##
##           Mean of squared residuals: 183.6074
##                     % Var explained: 16.57
```

```
plot(minModel)
```

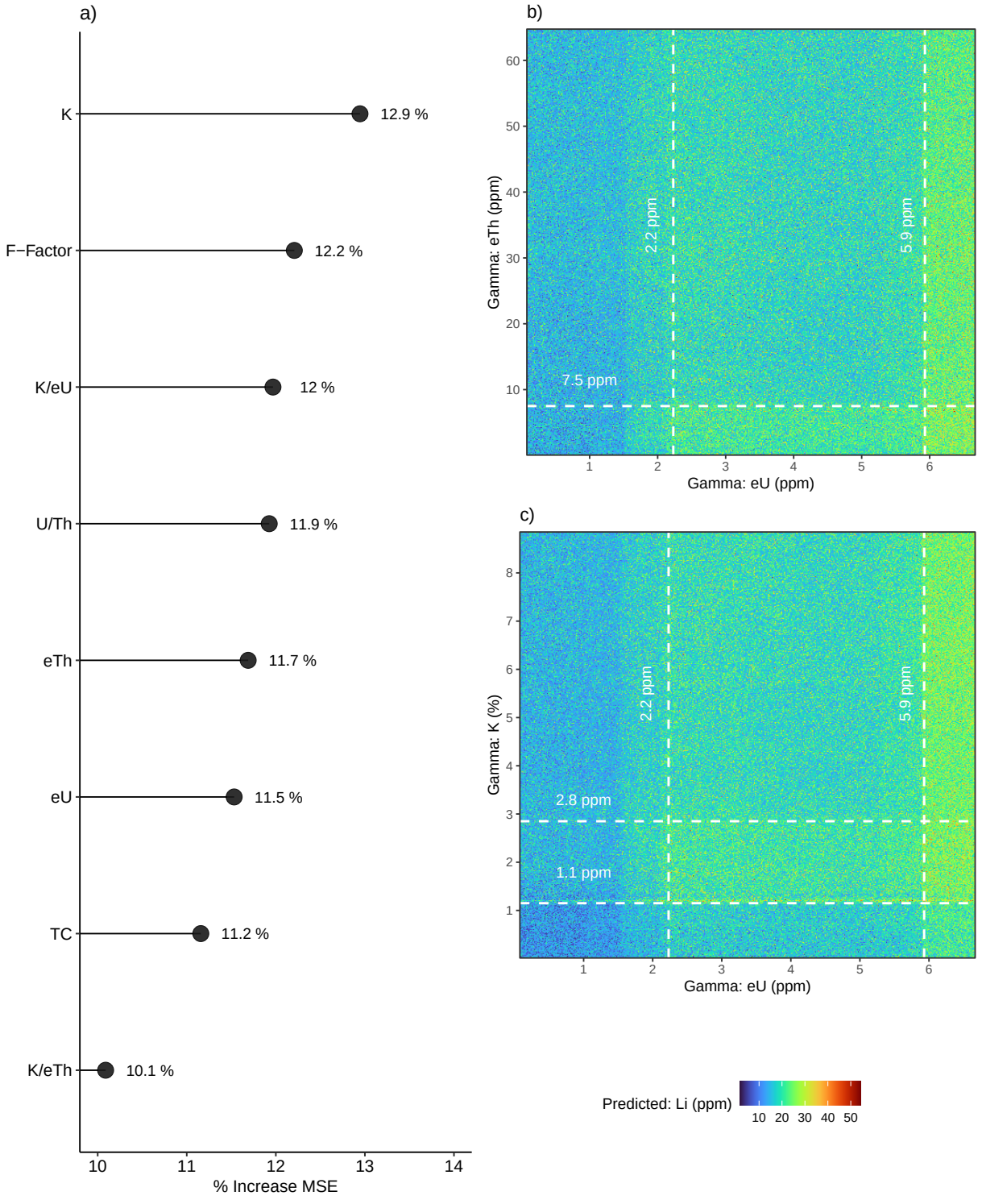



```
p.Imp <-
  varImp(minModel) %>% arrange(desc(Overall)) %>% rownames_to_column() %>%
  mutate(Feature = factor(case_when(rowname == 'gamma_k' ~ 'K',
                                    rowname == 'gamma_u_th' ~ 'U/Th',
                                    rowname == 'gamma_k_eu' ~ 'K/eU',
                                    rowname == 'gamma_eth' ~ 'eTh',
                                    rowname == 'gamma_k_th' ~ 'K/eTh',
                                    rowname == 'gamma_f_factor' ~ 'F-Factor',
```

```
                                    rowname == 'gamma_eu' ~ 'eU',
                                    rowname == 'gamma_ct' ~ 'TC',
                                    TRUE ~ rowname)),levels = c('K',
                                                                'U/Th',
                                                                'K/eU',
                                                                'eTh',
                                                                'K/eTh',
                                                                'F-Factor',
                                                                'eU',
                                                                'TC')) %>%
  ggplot(aes(y = fct_reorder(Feature,Overall), x = Overall)) +
  geom_segment(aes(xend = 0, yend = Feature)) +
  geom_point(size = 5, fill = 'grey10', alpha = .9,
             shape = 21, col = 'black') +
  geom_text(mapping = aes(label = paste(round(Overall,1),'%')),nudge_x = .5) +
  ggpubr::theme_pubr() +
  scale_fill_viridis_c() +
  coord_cartesian(xlim = c(10,14),expand = TRUE,clip = 'on') +
  labs(x = '% Increase MSE',y = '',title = 'a)')

p1 <-
  plot_predict_interaction(forest = minModel,
                           data = trainset,
                           variable1 =  "gamma_eu",
                           variable2 =  "gamma_eth",
                           grid = 500) +
  labs(x = 'Gamma: eU (ppm)',
       y = 'Gamma: eTh (ppm)',
       fill = 'Predicted: Li (ppm)',title = '') +
  coord_cartesian(expand = FALSE,clip = 'off') +
  theme_bw() +
  theme(legend.position = 'bottom',
        axis.text.x = element_blank(),
        axis.title.x = element_blank()) +
  scale_fill_gradientn(colours = pals::turbo()) +
  labs(title = 'b)')

get_only_legend <- function(plot) {
  plot_table <- ggplot_gtable(ggplot_build(plot))
  legend_plot <- which(sapply(plot_table$grobs, function(x) x$name) == "guide-box")
  legend <- plot_table$grobs[[legend_plot]]
  return(legend)
}

legend <- get_only_legend(p1)

p1 <-
    plot_predict_interaction(forest = minModel,
                             data = trainset,
                             variable1 =  "gamma_eu",
                             variable2 =  "gamma_eth",
                             grid = 500) +
    geom_hline(yintercept = 7.5,
```

```
              linetype = 'dashed',
              col = 'white',
              size = .8) +
    annotate(geom = 'text',
             label = '7.5 ppm',
             x = 1,
             y = 11.5,
             col = 'white') +
    geom_vline(xintercept = 2.23,
              linetype = 'dashed',
              col = 'white',
              size = .8) +
    annotate(geom = 'text',
             label = '2.2 ppm',
             x = 1.9,
             y = 35,
             angle = 90,
             col = 'white') +
    geom_vline(xintercept = 5.93,
              linetype = 'dashed',
              col = 'white',
              size = .8) +
    annotate(geom = 'text',
             label = '5.9 ppm',
             x = 5.65,
             y = 35,
             angle = 90,
             col = 'white') +
    labs(x = 'Gamma: eU (ppm)',
         y = 'Gamma: eTh (ppm)',
         fill = 'Predicted: Li (ppm)',title = '') +
    coord_cartesian(expand = FALSE,clip = 'off') +
    scale_x_continuous(breaks = seq(0,7,1)) +
    scale_y_continuous(breaks = seq(0,70,10)) +
    theme_bw() +
    theme(legend.position = 'none') +
    scale_fill_gradientn(colours = pals::turbo()) +
    labs(title = 'b)')


p2 <-
  plot_predict_interaction(forest = minModel,
                           data = trainset,
                           variable1 =  "gamma_eu",
                           variable2 =  "gamma_k",
                           grid = 500) +
  geom_hline(yintercept = 1.15,
             linetype = 'dashed',
             col = 'white',
             size = .8) +
  annotate(geom = 'text',
           label = '1.1 ppm',
           x = 1,
```

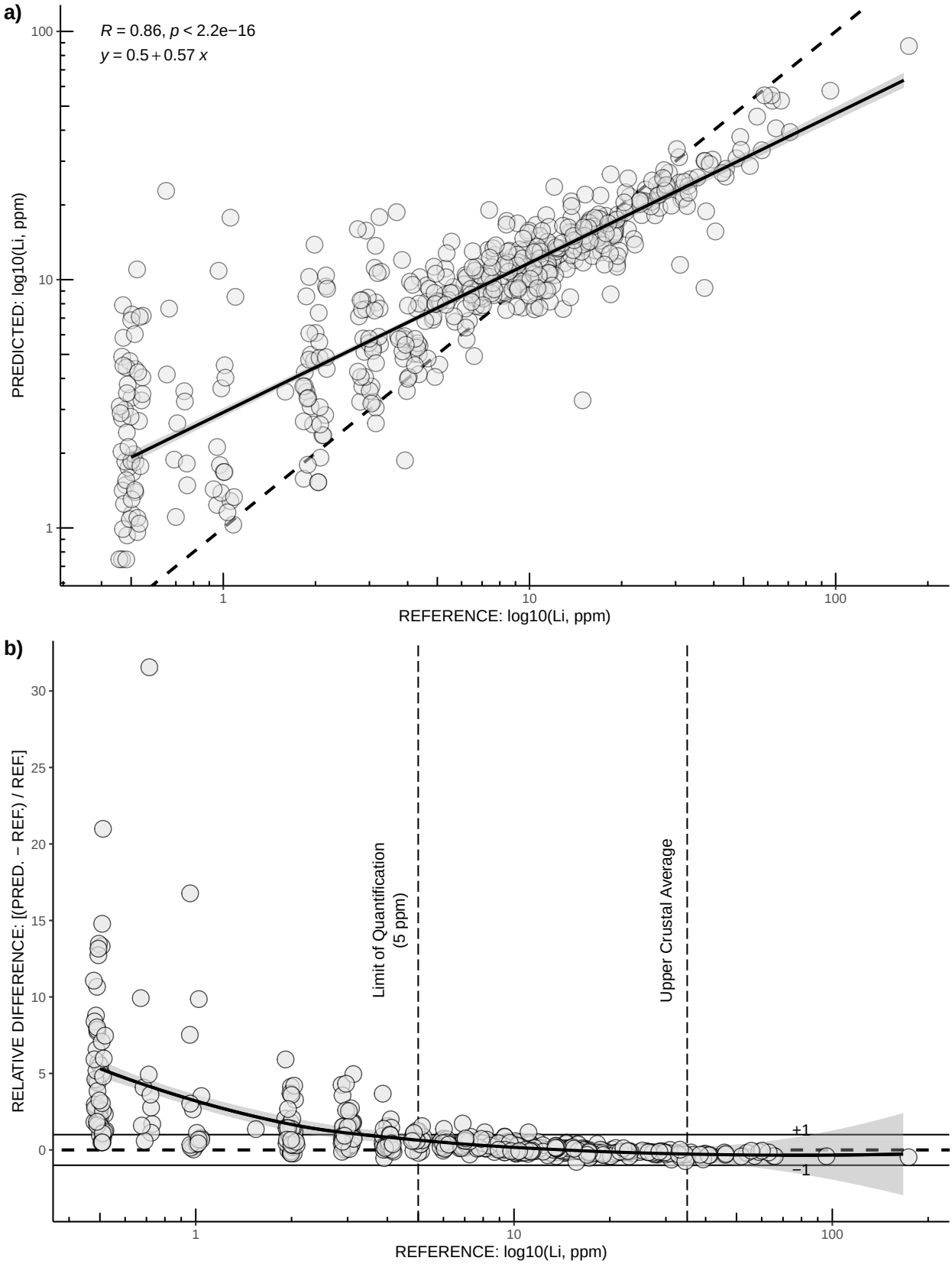
```
           y = 1.8,
           col = 'white') +
    geom_hline(yintercept = 2.85,
               linetype = 'dashed',
               col = 'white',
               size = .8) +
    annotate(geom = 'text',
             label = '2.8 ppm',
             x = 1,
             y = 3.3,
             col = 'white') +
  geom_vline(xintercept = 2.23,
             linetype = 'dashed',
             col = 'white',
             size = .8) +
  annotate(geom = 'text',
           label = '2.2 ppm',
           x = 1.9,
           y = 5.5,
           angle = 90,
           col = 'white') +
  geom_vline(xintercept = 5.93,
             linetype = 'dashed',
             col = 'white',
             size = .8) +
  annotate(geom = 'text',
           label = '5.9 ppm',
           x = 5.65,
           y = 5.5,
           angle = 90,
           col = 'white') +
  labs(x = 'Gamma: eU (ppm)',
       y = 'Gamma: K (%)',
       fill = 'Predicted: Li (ppm)',title = '') +
  coord_cartesian(expand = FALSE,clip = 'off') +
  scale_x_continuous(breaks = seq(0,7,1)) +
  scale_y_continuous(breaks = seq(0,8,1)) +
  theme_bw() +
  theme(legend.position = 'none') +
  scale_fill_gradientn(colours = pals::turbo()) +
  labs(title = 'c)')


gridExtra::grid.arrange(p.Imp,p1,p2,legend,
                        layout_matrix = matrix(c(1,2,
                                                 1,3,
                                                 1,4), ncol = 2, byrow = TRUE), heights = c(5,5,2))
```




```r
# Predição
predOut <- predict(minModel,newdata =  df)

df1 <- df %>%
  bind_cols(as_tibble(predOut)) %>%
  rename(li_ppm_pred = 'value')
```

```r
set.seed(1123)

p0 <-
  df1 %>%
  ggplot(aes(li_ppm, li_ppm_pred)) +
  # geom_smooth(method = MASS::rlm) +
  labs(x = 'REFERENCE: log10(Li, ppm)',
       y = 'PREDICTED: log10(Li, ppm)') +
  geom_abline(intercept = 0,slope = 1,
              linetype = 2, size = 1,
              col = 'black') +
  geom_jitter(width = 0.04,height = 0,
              shape = 21, size = 5, alpha = .5,
              fill = 'grey90'
  ) +
  geom_smooth(method = MASS::rlm,
              col = 'black') +
  ggpubr::stat_cor(label.x = -0.4,
                   label.y = 2) +
  ggpubr::stat_regline_equation(label.x = -0.4,
                                label.y = 1.9) +
  # coord_equal() +
  scale_y_continuous(trans = 'log10') +
  scale_x_continuous(trans = 'log10') +
  theme_classic() +
  annotation_logticks()

p01 <-
    df1 %>%
    ggplot(aes(li_ppm,
               (li_ppm_pred - li_ppm)/li_ppm)
    ) +
    geom_hline(yintercept = 1) +
    geom_hline(yintercept = -1) +
    geom_vline(xintercept = 35,
               linetype = 5) +
    geom_vline(xintercept = 5,
               linetype = 5) +
    geom_abline(intercept = 0,slope = 0,
                linetype = 2, size = 1,
                col = 'black') +
    geom_smooth(col = 'black', se = TRUE) +
    geom_jitter(width = 0.02,height = 0,
                shape = 21, size = 5, alpha = .7,
                fill = 'grey90'
    ) +
    geom_smooth(col = 'black', se = FALSE) +
    scale_x_continuous(
      # limits = c(1,180),
      trans = 'log10'
    ) +
    theme_classic() +
    annotation_logticks(sides = 'b') +
```

```
    labs(x = 'REFERENCE: log10(Li, ppm)',
         y = 'RELATIVE DIFFERENCE: [(PRED. - REF.) / REF.]') +
    annotate(geom = 'text',label = '+1',
             x = 80, y = 1.3) +
    annotate(geom = 'text',label = '-1',
             x = 80, y = -1.3) +
    annotate(geom = 'text',label = 'Limit of Quantification',
             x = 3.75,y = 15,angle = 90) +
    annotate(geom = 'text',label = '(5 ppm)',
             x = 4.35,y = 15,angle = 90) +
    annotate(geom = 'text',label = 'Upper Crustal Average',
             x = 35-5,y = 15,angle = 90) +
    scale_y_continuous(breaks = seq(-5,40,5))

ggpubr::ggarrange(p0,p01,
                  ncol = 1,hjust = 0,
                  common.legend = FALSE,
                  labels = c('a)','b)'),
                  widths = c(2, 1))
```


a)
R = 0.86, p < 2.2e−16
y = 0.5 + 0.57 x
PREDICTED: log10(Li, ppm)
REFERENCE: log10(Li, ppm)
b)
RELATIVE DIFFERENCE: [(PRED. − REF.) / REF.]
Limit of Quantification (5 ppm)
Upper Crustal Average
+1
−1
REFERENCE: log10(Li, ppm)

```r
li_mapa <- predict(minModel,newdata =  newdata)

mapa_output <- newdata %>%
  bind_cols(as_tibble(li_mapa)) %>%
  rename(li_ppm_pred = 'value')
print(mapa_output)
```

```
##              x     y gamma_k gamma_eth gamma_eu  gamma_ct  gamma_k_th
##          <num> <num>   <num>     <num>    <num>     <num>       <num>
##      1: -42.63 -9.00    0.83      8.71     0.96 5.1400000 0.095291673
##      2: -42.57 -9.00    0.61      7.98     1.30 4.6900000 0.076440145
##      3: -42.56 -9.01    0.58      8.11     1.43 4.7600000 0.071515764
##      4: -42.56 -9.00    0.48      7.87     1.23 4.4300000 0.060990330
##      5: -42.55 -9.01    0.53      8.08     1.19 4.5700000 0.065593248
##     ---
## 401409: -34.80 -7.30    0.03      2.21     0.71 0.4725986 0.013574046
## 401410: -34.80 -7.28    0.16      5.19     1.17 0.4725986 0.030827922
## 401411: -34.80 -7.20    0.07      4.26     0.92 0.4725986 0.016431539
## 401412: -34.80 -7.19    0.03      3.55     0.14 0.4725986 0.008450466
## 401413: -34.80 -7.18    0.08      2.77     0.17 0.4725986 0.028879824
##         gamma_u_th gamma_k_eu gamma_f_factor li_ppm_pred
##              <num>      <num>          <num>       <num>
##      1: 0.11021687 0.86449328    0.091500558   12.413270
##      2: 0.16290523 0.46919468    0.099396124    7.139073
##      3: 0.17632335 0.40556604    0.102292328    5.210732
##      4: 0.15628772 0.39021218    0.075039836    3.654503
##      5: 0.14727541 0.44534073    0.078077253    6.156733
##     ---
## 401409: 0.32125243 0.04224757    0.009671060    1.963503
## 401410: 0.22542918 0.13674045    0.036094297    2.351990
## 401411: 0.21595737 0.07607869    0.015140257    2.302193
## 401412: 0.03943551 0.21413276    0.001187857    2.451032
## 401413: 0.06136963 0.47031158    0.004918599    5.012990
```

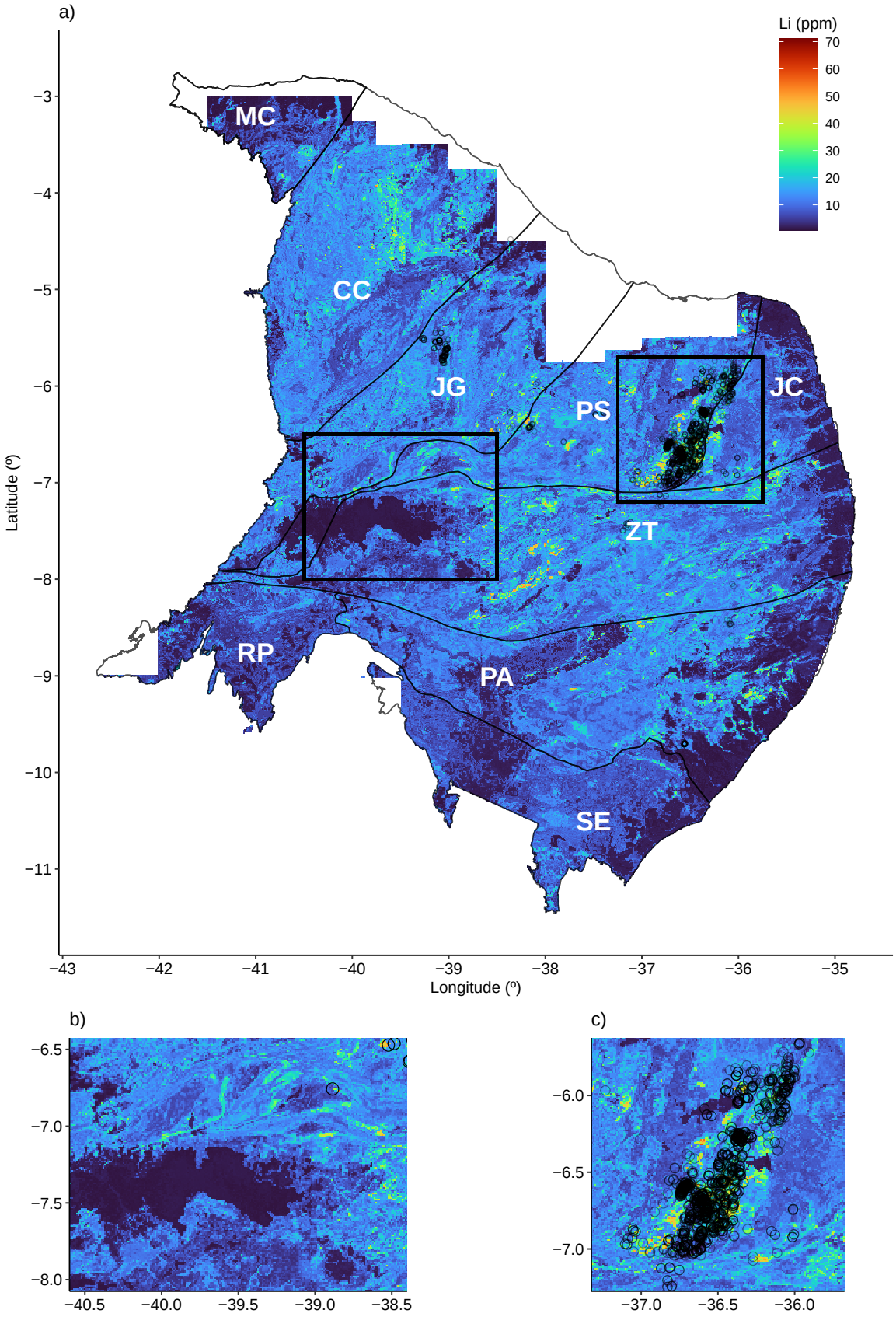
```r
p4 <-
    mapa_output %>%
    ggplot(aes(x,y,fill = li_ppm_pred)) +
    geom_tile() +
    scale_fill_gradientn(colours = pals::turbo(),
                         guide = guide_colorbar(barheight = 10,
                                                barwidth = 2),
                         breaks=seq(0,270,10)) +
    coord_equal() +
    geom_point(inherit.aes = FALSE,
               data = recmin %>%
                 sample_frac(.6),
               mapping = aes(x,y),
               shape = 21,
               fill = NA,
               col = 'black',
               alpha = .3) +
    ggpubr::theme_pubr() +
    scale_x_continuous(breaks = seq(-50,-30,1),
```

```
                    minor_breaks = seq(-50,-30,0.5)) +
  scale_y_continuous(breaks = seq(-20,0,1),
                     minor_breaks = seq(-20,0,0.5)) +
  theme(legend.position = 'inside',
        legend.position.inside = c(0.9,0.9)) +
  geom_path(inherit.aes = FALSE,
            data = MC,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'MC',
           x = -41,
           y = -3.2,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = MC,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'CC',
           x = -40,
           y = -5,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = CC,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'JG',
           x = -39,
           y = -6,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = JG,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'PS',
           x = -37.5,
           y = -6.25,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = PS,
            aes(X,Y),
            alpha = .7) +
```

```
  annotate(geom = 'text',
           label = 'JC',
           x = -35.5,
           y = -6,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = JC,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'ZT',
           x = -37,
           y = -7.5,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = ZT,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'PA',
           x = -38.5,
           y = -9,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = PA,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'RP',
           x = -41,
           y = -8.75,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = RP,
            aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'SE',
           x = -37.5,
           y = -10.5,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
            data = SE,
```

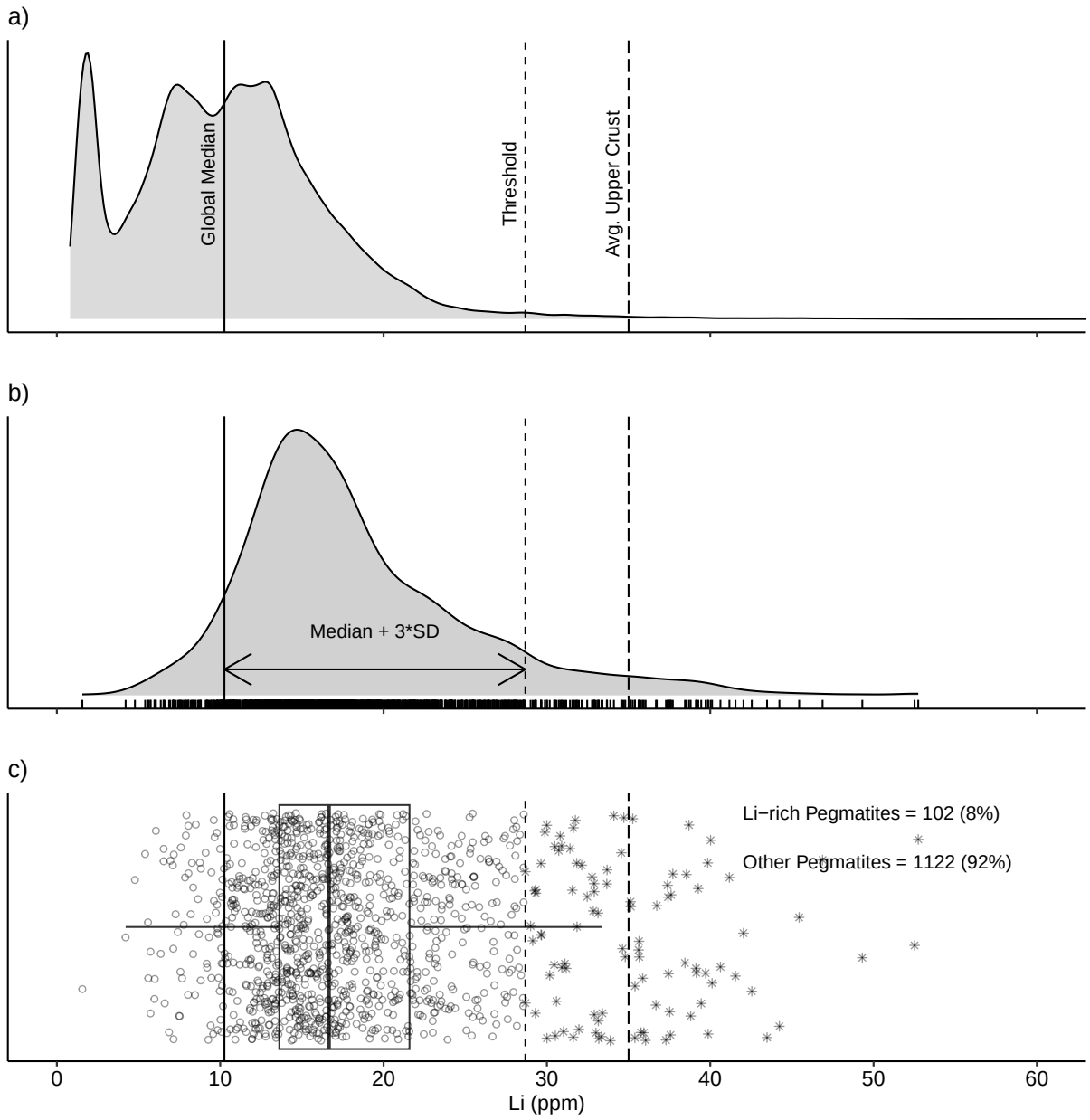
```
              aes(X,Y),
              alpha = .7) +
    labs(title = 'a)',
         fill = 'Li (ppm)',
         x = 'Longitude (º)',
         y = 'Latitude (º)') +
    geom_rect(xmin = -37.25,
              xmax = -35.75,
              ymin = -7.2,
              ymax = -5.7,
              fill = NA,
              col = 'black',
              size = 1) +
    geom_rect(xmin = -38.5,
              xmax = -40.5,
              ymin = -8,
              ymax = -6.5,
              fill = NA,
              col = 'black',
              size = 1)


# Seridó
set.seed(0)
p5 <-
    mapa_output %>%
    ggplot(aes(x,y,fill = li_ppm_pred)) +
    geom_tile() +
    scale_fill_gradientn(colours = pals::turbo()) +
    coord_equal(xlim = c(-37.25,-35.75),
                ylim = c(-7.2,-5.7)) +
    geom_point(inherit.aes = FALSE,
               data = recmin %>%
                 sample_frac(.7),
               mapping = aes(x,y),
               shape = 21,
               fill = NA,
               col = 'black',
               alpha = .4,
               size = 3) +
    scale_x_continuous(breaks = seq(-50,-30,0.5)
                       ) +
    scale_y_continuous(breaks = seq(-20,0,0.5),
                       ) +
    ggpubr::theme_pubr() +
    theme(legend.position = 'none') +
    labs(title = 'c)',
         fill = 'Li (ppm)',
         x = '',
         y = ''
    )
```

```r
# Araripe
p6 <-
    mapa_output %>%
    ggplot(aes(x,y,fill = li_ppm_pred)) +
    geom_tile() +
    scale_fill_gradientn(colours = pals::turbo()) +
    coord_equal(xlim = c(-40.5,-38.5),
                ylim = c(-8,-6.5)) +
    scale_x_continuous(breaks = seq(-50,-30,0.5)
    ) +
    scale_y_continuous(breaks = seq(-20,0,0.5),
    ) +
    geom_point(inherit.aes = FALSE,
               data = recmin %>%
                 sample_frac(1),
               mapping = aes(x,y),
               shape = 21,
               fill = NA,
               col = 'black',
               alpha = 1,
               size = 4) +
    ggpubr::theme_pubr() +
    theme(legend.position = 'none') +
    labs(title = 'b)',
         fill = 'Li (ppm)',
         x = '',
         y = ''
    )


grid.arrange(p4,p6,p5,
             layout_matrix = matrix(c(1,1,
                                      1,1,
                                      1,1,
                                      2,3), ncol = 2, byrow = TRUE))
```


a)
Li (ppm)
70
60
50
40
30
20
10
MC
CC
JG
PS
JC
ZT
RP
PA
SE
Latitude (º)
Longitude (º)
−3
−4
−5
−6
−7
−8
−9
−10
−11
−43
−42
−41
−40
−39
−38
−37
−36
−35
b)
−6.5
−7.0
−7.5
−8.0
−40.5
−40.0
−39.5
−39.0
−38.5
c)
−6.0
−6.5
−7.0
−37.0
−36.5
−36.0

```r
mu <- median(mapa_output$li_ppm_pred)
sd <- sd(mapa_output$li_ppm_pred)

p7 <-
    mapa_output %>%
    ggplot(aes(li_ppm_pred)) +
    geom_density(fill = 'grey80',
                 alpha = .7) +
    labs(title = 'a)',x = '',y = '') +
    ggpubr::theme_pubr() +
    coord_cartesian(xlim = c(0,60)) +
    labs(title = 'a)',x = '',y = '') +
    annotate(geom = 'text',label = 'Global Median',
             x = mu-1,y = .04,angle = 90) +
    annotate(geom = 'text',label = 'Avg. Upper Crust',
             x = (34),y = .040, angle = 90) +
    geom_vline(aes(xintercept = 35), linetype = 'longdash') +
    annotate(geom = 'text',label = 'Threshold',
             x = mu+3*sd-1,y = .04,angle = 90) +
    geom_vline(aes(xintercept = mu)) +
    geom_vline(aes(xintercept = mu+3*sd),
               linetype = 'dashed') +
    theme(axis.text.y = element_blank(),
          axis.ticks.y = element_blank(),
          axis.text.x = element_blank())
p8 <-
    recmin1 %>%
    rename(li_pred = RASTERVALU) %>%
    ggplot() +
    geom_density(aes(li_pred), fill = 'grey', alpha = .7) +
    geom_vline(aes(xintercept = mu)) +
    geom_vline(aes(xintercept = mu+3*sd),
               linetype = 'dashed') +
    geom_vline(aes(xintercept = 35), linetype = 'longdash') +
    annotate(geom = 'segment',
             x = mu, y = .008, xend = mu+3*sd, yend = .008,
             arrow = arrow(ends = "both")) +
    geom_rug(aes(li_pred)) +
    annotate(geom = 'text',label = 'Median + 3*SD',
             x = ((mu+(mu+3*sd))/2),y = .020) +
    ggpubr::theme_pubr() +
    coord_cartesian(xlim = c(0,60)) +
    labs(title = 'b)',x = '',y = '') +
    theme(axis.text.y = element_blank(),
          axis.ticks.y = element_blank(),
          axis.text.x = element_blank())

set.seed(42)
p9 <-
    recmin1 %>%
    rename(li_pred = RASTERVALU) %>%
    mutate(symbol = case_when(li_pred > (mu+3*sd) ~ 'OUTLIER',
                              TRUE ~ 'REGULAR')) %>%
```

```
    ggplot(aes(x = li_pred)) +
    geom_boxplot(outlier.shape = NA) +
    geom_jitter(aes(y = 0, shape = symbol),width = 0,height = 0.35,
                fill = NA, col = 'black', alpha = .4) +
    geom_vline(aes(xintercept = mu)) +
    geom_vline(aes(xintercept = mu+3*sd),
               linetype = 'dashed') +
    geom_vline(aes(xintercept = 35), linetype = 'longdash') +
    ggpubr::theme_pubr() +
    coord_cartesian(xlim = c(0,60)) +
    labs(title = 'c)',x = 'Li (ppm)', y ='') +
    scale_shape_manual(values = c(8,21)) +
    scale_x_continuous(breaks = seq(0,60,by = 10)) +
    annotate(geom = 'text',label = 'Li-rich Pegmatites = 102 (8%)',
             x = 42,y = 0.35,
             hjust = 0) +
    annotate(geom = 'text',label = 'Other Pegmatites = 1122 (92%)',
             x = 42,y = 0.2,
             hjust = 0) +
    theme(axis.text.y = element_blank(),
          axis.ticks.y = element_blank(),
          legend.position = 'none')

(p.density <-grid.arrange(p7,p8,p9, ncol = #,
                          # heights = c(6,6,6)
                          )
)
```




```
## TableGrob (3 x 1) "arrange": 3 grobs
##   z     cells    name           grob
## 1 1 (1-1,1-1) arrange gtable[layout]
## 2 2 (2-2,1-1) arrange gtable[layout]
## 3 3 (3-3,1-1) arrange gtable[layout]
```

```
recmin1 %>%
  rename(li_pred = RASTERVALU) %>%
  mutate(symbol = case_when(li_pred > (mu+3*sd) ~ 'OUTLIER',
                            TRUE ~ 'REGULAR')) %>%
  select(symbol) %>%
  group_by(symbol) %>%
  summarize(n = n())
```

```
## # A tibble: 2 x 2
##   symbol      n
##   <chr>   <int>
## 1 OUTLIER   102
## 2 REGULAR  1122
```

```
recmin1 %>%
  rename(li_pred = RASTERVALU) %>%
  arrange(desc(li_pred)) %>%
  mutate(SUBSTANCIA = tolower(SUBSTANCIA)) %>%
  mutate(Substances = str_replace_all(SUBSTANCIA,
                                      c('terras-raras' = 'Rare Earth',
                                        'neodímio' = 'Neodymium',
                                        'cério' = 'Cerium',
                                        'granada' = 'Garnet',
                                        'talco' = 'Talc',
                                        'amazonita' = 'Amazonite',
                                        'córindon' = 'Corundum',
                                        'urânio' = 'Uranium',
                                        'água-marinha' = 'Aquamarine',
                                        'água marinha' = 'Aquamarine',
                                        'lazulita' = 'Lazulite',
                                        'pegmatito' = 'Dimension stone',
                                        'pedra-ornamental' = 'Dimension stone',
                                        'caulim' = 'Kaolinite',
                                        'amerício' = 'Americium',
                                        'muscovita' = 'Muscovite',
                                        'vermiculita' = 'Vermiculite',
                                        'feldspato' = 'Feldspar',
                                        'citrino' = 'Citrine',
                                        'turmalina' = 'Tourmaline',
                                        'albita' = 'Albite',
                                        'quartzo' = 'Quartz',
                                        'quartzo rosa' = 'Pink quartz',
                                        'calcedônia' = 'Chalcedony',
                                        'ametista' = 'Amethyst',
                                        'lítio' = 'Spodumene',
                                        'bismuto' = 'Bismuth',
                                        'tântalo' = 'Tantalum',
                                        'nióbio' = 'Niobium',
                                        'berílio' = 'Beryllium',
                                        'estanho' = 'Tin',
                                        'titânio' = 'Titanium',
                                        'tungstênio' = 'Tungsten',
                                        'molibdênio' = 'Molybdenum',
                                        'rutênio' = 'Ruthenium',
                                        'mica' = 'Muscovite',
                                        'rocha ornamental' = 'Dimension stone'))) %>%
  select(TOPONIMIA,x,y, Substances, li_pred) %>%
  filter(li_pred > 35) %>%
  rename('Pegmatite Name' = 'TOPONIMIA',
         'Longitude' = 'x',
         'Latitude' = 'y',
         'Known substances' = 'Substances',
```

```
        'Li (ppm) in soil' = 'li_pred') %>%
  head(40)
```

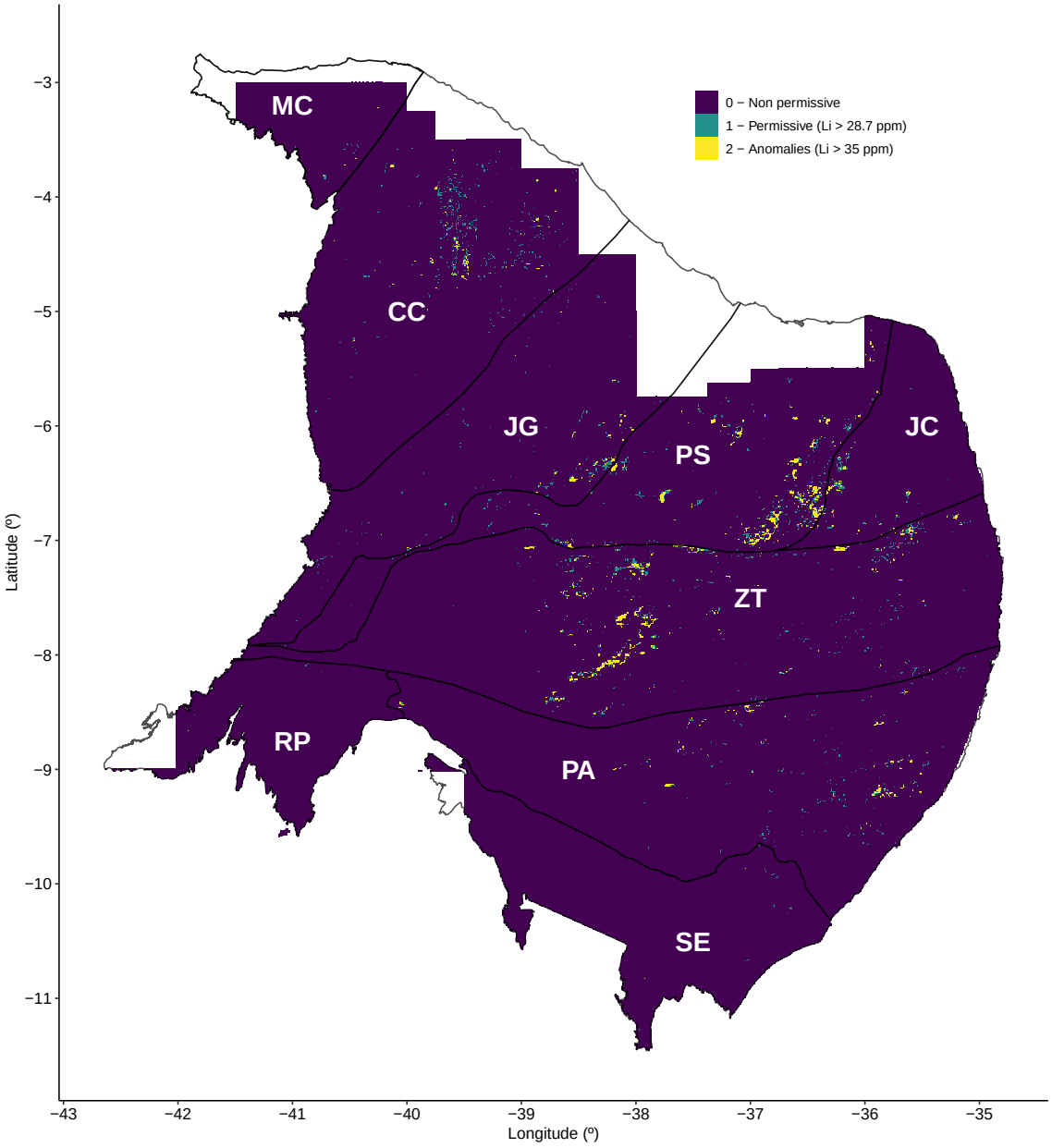
```
##            Pegmatite Name Longitude Latitude
## 1             Alto Grande  -36.4903 -6.58403
## 2          Alto Belo Açude  -36.4603 -6.57375
## 3           Alto Primavera  -36.4789 -6.65042
## 4          Cabeça de Velho  -36.4856 -6.65319
## 5                 Formigão  -36.4920 -6.64208
## 6             Alto Ovelhas  -36.4845 -6.64569
## 7   Mina de Chico Cândido  -36.4014 -6.72181
## 8   Alto do Trigueiro III  -36.5453 -6.65708
## 9            Alto da Serra  -36.4817 -6.62180
## 10    Alto do Trigueiro I  -36.5361 -6.66125
## 11  Alto Serrote do Mariz  -36.4003 -6.73180
## 12   Alto do Trigueiro II  -36.5339 -6.66292
## 13       Olho D'Água Sêco  -36.5686 -6.62958
## 14    Alto Olho D'Aguinha  -36.5675 -6.62625
## 15     Alto Pedra Racháda  -36.4706 -6.70042
## 16           Alto Zè Preto  -36.4786 -6.54208
## 17      Alto da Corujinha  -36.5339 -6.67430
## 18   Alto Malhada Redonda  -36.3964 -6.74431
## 19                Formigão  -36.4900 -6.63930
## 20        Alto Saco Largo  -36.5528 -6.66042
## 21           Alto Pedrinha  -36.4559 -6.68930
## 22           Alto da Chápa  -36.4614 -6.65597
## 23             Moça Bonita  -36.4281 -6.64680
## 24     Alto do Garrote II  -36.4770 -6.54375
## 25         Serra Cabeluda  -37.0392 -7.00458
## 26             Alto Túnica  -36.4636 -6.56875
## 27     Alto Belo Açude II  -36.4547 -6.57097
## 28        Alto Tanquinhos  -36.3892 -6.59486
## 29 Alto Malhada Escondida  -36.4447 -6.64292
## 30            Lagoa I e II  -38.1092 -6.35736
## 31          Alto Corrixauá  -36.4503 -6.68847
## 32           Garimpo Lagoa  -38.1131 -6.36014
## 33             Alto Cágado  -36.4309 -6.52541
## 34       Alto da Serrinha  -36.4636 -6.56708
## 35             Sítio Patos  -36.4871 -6.71500
## 36      Alto Pedra Branca  -36.4114 -6.49208
## 37    Serrote da Serraria  -36.4781 -6.69014
## 38         Alto da Marinha  -36.4595 -6.55875
## 39             Alto Damião  -36.4039 -6.48514
## 40       Alto do Tibiri I  -36.5222 -6.79181
##                             Known substances Li (ppm) in soil
## 1              Beryllium, Niobium, Tantalum          52.7342
## 2   Beryllium, Muscovite, Niobium, Tantalum          52.5080
## 3                        Beryllium, Tantalum          49.3106
## 4              Beryllium, Niobium, Tantalum          46.8765
## 5                        Beryllium, Tantalum          45.4437
## 6                                  Tantalum          44.2338
## 7              Beryllium, Niobium, Tantalum          43.4790
## 8                        Beryllium, Tantalum          42.5447
```

```
## 9                          Niobium, Tantalum          42.0337
## 10             Beryllium, Niobium, Tantalum          41.5447
## 11                         Niobium, Tantalum          41.1723
## 12             Beryllium, Niobium, Tantalum          40.6238
## 13                       Beryllium, Niobium          40.1125
## 14             Beryllium, Niobium, Tantalum          40.0199
## 15             Beryllium, Niobium, Tantalum          39.8710
## 16                       Beryllium, Tantalum          39.8474
## 17                       Beryllium, Tantalum          39.7138
## 18                       Beryllium, Tantalum          39.4474
## 19                         Niobium, Tantalum          39.2645
## 20                        Feldspar, Tantalum          39.1463
## 21                       Beryllium, Tantalum          39.1264
## 22             Beryllium, Niobium, Tantalum          38.8015
## 23             Beryllium, Niobium, Tantalum          38.6984
## 24            Beryllium, Feldspar, Tantalum          38.5401
## 25                                 Beryllium          38.4528
## 26             Beryllium, Niobium, Tantalum          37.7086
## 27            Beryllium, Muscovite, Niobium          37.6016
## 28             Beryllium, Niobium, Tantalum          37.5658
## 29                         Niobium, Tantalum          37.5192
## 30                                  Feldspar          37.4447
## 31                                       Tin          37.4160
## 32                               Aquamarine          37.3846
## 33                                 Beryllium          37.2887
## 34                       Beryllium, Niobium          36.7249
## 35                          Dimension stone          36.6757
## 36             Beryllium, Niobium, Tantalum          36.0432
## 37                         Niobium, Tantalum          35.9354
## 38                       Beryllium, Tantalum          35.8649
## 39        Beryllium, Tin, Niobium, Tantalum          35.7638
## 40 Beryllium, Feldspar, Muscovite, Tantalum          35.6542
```

```
mapa_output %>%
    mutate(li_bin = case_when(li_ppm_pred >= 35 ~ "2 - Anomalies (Li > 35 ppm)",
                              li_ppm_pred > 28.7 & li_ppm_pred < 35 ~ '1 - Permissive (Li > 28.7 ppm)',
                              li_ppm_pred <= 28.7 ~ "0 - Non permissive")) %>%
    ggplot(aes(x,y,fill = li_bin)) +
    geom_tile() +
    geom_path(inherit.aes = FALSE,
                data = MC,
                aes(X,Y),
             alpha = .7) +
    annotate(geom = 'text',
             label = 'MC',
             x = -41,
             y = -3.2,
             color = 'white',
             fontface = 'bold',
             size = 7) +
    geom_path(inherit.aes = FALSE,
                data = MC,
                aes(X,Y),
             alpha = .7) +
```

```
  annotate(geom = 'text',
           label = 'CC',
           x = -40,
           y = -5,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
              data = CC,
              aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'JG',
           x = -39,
           y = -6,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
              data = JG,
              aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'PS',
           x = -37.5,
           y = -6.25,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
              data = PS,
              aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'JC',
           x = -35.5,
           y = -6,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
              data = JC,
              aes(X,Y),
            alpha = .7) +
  annotate(geom = 'text',
           label = 'ZT',
           x = -37,
           y = -7.5,
           color = 'white',
           fontface = 'bold',
           size = 7) +
  geom_path(inherit.aes = FALSE,
              data = ZT,
```

```
              aes(X,Y),
           alpha = .7) +
annotate(geom = 'text',
         label = 'PA',
         x = -38.5,
         y = -9,
         color = 'white',
         fontface = 'bold',
         size = 7) +
geom_path(inherit.aes = FALSE,
              data = PA,
              aes(X,Y),
           alpha = .7) +
annotate(geom = 'text',
         label = 'RP',
         x = -41,
         y = -8.75,
         color = 'white',
         fontface = 'bold',
         size = 7) +
geom_path(inherit.aes = FALSE,
              data = RP,
              aes(X,Y),
           alpha = .7) +
annotate(geom = 'text',
         label = 'SE',
         x = -37.5,
         y = -10.5,
         color = 'white',
         fontface = 'bold',
         size = 7) +
geom_path(inherit.aes = FALSE,
              data = SE,
              aes(X,Y),
           alpha = .7) +
scale_fill_viridis_d() +
coord_equal() +
scale_x_continuous(breaks = seq(-50,-30,1)
) +
scale_y_continuous(breaks = seq(-20,0,1),
) +
ggpubr::theme_pubr() +
theme(legend.position = 'inside',
      legend.position.inside = c(0.75,0.9)) +
labs(fill = '',
     x = 'Longitude (º)',
     y = 'Latitude (º)')
```


0 − Non permissive
1 − Permissive (Li > 28.7 ppm)
2 − Anomalies (Li > 35 ppm)
MC
CC
JG
PS
JC
ZT
RP
PA
SE
Latitude (°)
−3
−4
−5
−6
−7
−8
−9
−10
−11
−43
−42
−41
−40
−39
−38
−37
−36
−35
Longitude (°)