

Cantones sin IC

Jimena Murillo

2022-07-04

Paquetes

```
library(keras) # for deep learning  
library(tidyverse) # general utility functions
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.6      v purrr  0.3.4  
## v tibble  3.1.6      v dplyr  1.0.9  
## v tidyr   1.2.0      v stringr 1.4.0  
## v readr   2.1.2      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()    masks stats::lag()
```

```
library(caret) # machine learning utility functions
```

```
## Loading required package: lattice
```

```
##
```

```
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
## lift
```

```
library(tibble)  
library(readr)  
library(ggplot2)  
library(tensorflow)
```

```
##
```

```
## Attaching package: 'tensorflow'
```

```
## The following object is masked from 'package:caret':
```

```
##
```

```
## train
```

```
library(neuralnet)
```

```
##  
## Attaching package: 'neuralnet'  
  
## The following object is masked from 'package:dplyr':  
##  
##      compute
```

Datos

```
load("C:/Users/usuario1/Desktop/CIMPA/Github_CIMPA/PRACTICA_CIMPA/base_cantones.RData")
```

```
basecanton = basecanton %>%
```

```
  dplyr::select(Canton, Year, Month, Nino12SSTA, Nino3SSTA, Nino4SSTA, Nino34SSTA, Nino34SSTA1, Nino34SSTA2
```

```
  arrange(Canton, Year, Month) %>% ungroup() %>% mutate(Month=as.numeric(Month))
```

```
#Funciones
```

```
normalize <- function(x) {  
  return ((x - min(x)) / (max(x) - min(x)))  
}
```

```
denorm <- function(x, base) {  
  return (x*(max(base$RR) - min(base$RR))+min(base$RR))  
}
```

```
metricas <- function(tabla){  
  NRMSE <- mean((tabla$y_pred-tabla$RR)^2)/mean(tabla$RR)  
  return(data.frame(NRMSE))  
}
```

```
basecanton2 = basecanton %>% group_by(basecanton$Canton) %>%  
  mutate_if(is.numeric, normalize)
```

```
## 'mutate_if()' ignored the following grouping variables:  
## * Column 'basecanton$Canton'
```

```
basecanton2 = basecanton2[, -35]
```

```
#Train y test
```

```
data_train = as.data.frame(basecanton2) %>% filter(Year < 1) #PARA ENTRENAR HASTA 2018
```

```

data_test = as.data.frame(basecanton2) %>% filter(Year >= 1)

X_train = data_train[,-ncol(data_train)]
y_train = as.data.frame(data_train[,c("Canton", "RR")])

X_test = as.data.frame(data_test[,-ncol(data_test)])
y_test = as.data.frame(data_test[,c("Canton", "RR")])

Fecha = paste(basecanton$Year, basecanton$Month)
Fecha = Fecha[1:235]

```

Arquitectura del modelo

```

model <- keras_model_sequential()

## Loaded Tensorflow version 2.8.0

# our input layer
model %>%
  layer_simple_rnn(units = 100, input_shape = c(ncol(X_train)-1,1), activation='tanh',
                 kernel_initializer= initializer_constant(0.5),
                 bias_initializer=initializer_zeros()) %>%
  layer_dense(units = 50, activation = "relu")%>%
  layer_dense(units = 50, activation = "relu")%>%
  layer_dense(units = 50, activation = "relu")%>%
  layer_dropout(rate = 0.1)%>%
  layer_dense(units = 25, activation = "relu")%>%
  layer_dense(units = 25, activation = "relu")%>%
  layer_dense(units = 25, activation = "relu")%>%
  layer_dropout(rate = 0.1)%>%
  layer_dense(units = 12, activation = "relu")%>%
  layer_dense(units = 12, activation = "relu")%>%
  layer_dropout(rate = 0.1)%>%
  layer_dense(units = 6, activation = "relu")%>%
  layer_dense(units = 6, activation = "relu")%>%

  layer_dense(units = 1, activation = "sigmoid")

```

Entrenamiento y predicciones

```

Cantones = unique(basecanton$Canton)
Eval.pd = NULL
Eval.tot = NULL

```

```

p1 = list()
p2 = list()

df1 = list()
df2 = list()

Predicciones = matrix(NA, ncol = 2, nrow = 3*length(Cantones))
Index = seq(1,3*length(Cantones), 3)

for (i in 1:length(Cantones)) {

  X_trainc = X_train %>% filter(Canton == Cantones[i])
  X_trainc = as.matrix(X_trainc[,-1])
  y_trainc = y_train %>% filter(Canton == Cantones[i])
  y_trainc = as.matrix(y_trainc[,-1])

  X_testc = X_test %>% filter(Canton == Cantones[i])
  X_testc = as.matrix(X_testc[,-1])
  y_testc = y_test %>% filter(Canton == Cantones[i])
  y_testc = as.matrix(y_testc[,-1])

  X_all = basecanton2 %>% filter(Canton == Cantones[i])
  X_all = as.matrix(X_all[,-c(1,33)])

  base = as.data.frame(basecanton %>% filter(Canton == Cantones[i]) %>% dplyr::select(RR))

  model %>% compile(loss = "mse",
                    optimizer = optimizer_adam(lr = 0.0007),
                    metric = "mean_absolute_error")

  trained_model <- model %>% fit(
    x = X_trainc,
    y = y_trainc,
    batch_size = 18,
    epochs = 100,
    validation_split = 0.1,
    shuffle = F)

  predice = function(x) {
    y_values = (model %>% predict(x))
    result = (y_values*(max(base$RR) - min(base$RR))+min(base$RR))
    return (as.numeric(result))
  }

  Predicciones[Index[i]:(Index[i]+2),1:2] = cbind(Cantones[i], predice(X_testc))
}

```

```

df1[[i]] = as.data.frame(cbind(predice(X_testc), y_testc, Fecha[233:235]))
colnames(df1[[i]]) = c("y_pred", "RR", "Fecha")
df1[[i]]$RR = as.numeric(df1[[i]]$RR)
df1[[i]]$y_pred = as.numeric(df1[[i]]$y_pred)

p1[[i]] = ggplot(df1[[i]], aes(x = Fecha, y = RR, group = 1)) + geom_line(colour = "blue") +
geom_line(aes(x = Fecha, y = y_pred, colour = "red"))+
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank(), axis.text.x = element_text(angle = 45), legend.position = "none")
labs(x = "Fecha", y = "Riesgo Relativo") +
ggtitle(paste("Predicciones 2021 del cantón", Cantones[i], sep = " "))

Eval.pd[i] = as.numeric(metricas(df1[[i]]))

#### VALORES APROXIMADOS ####

## Generar valores ajustados

df2[[i]] = as.data.frame(cbind(predice(X_all), base$RR, Fecha))
colnames(df2[[i]]) = c("y_pred", "RR", "Fecha")
df2[[i]]$RR = as.numeric(df2[[i]]$RR)
df2[[i]]$y_pred = as.numeric(df2[[i]]$y_pred)

everyother1 <- function(x) x[(seq_along(Fecha) + 5)%%12 == 6]

p2[[i]] = ggplot(df2[[i]], aes(x = Fecha, y = RR, group = 1)) + geom_line(colour = "blue") +
geom_line(aes(x = Fecha, y = y_pred, colour = "red"))+
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank(), axis.text.x = element_text(angle = 45), legend.position = "none")
scale_x_discrete(breaks = everyother1) + labs(x = "Fecha", y = "Riesgo Relativo") +
ggtitle(paste("Valores aproximados de training del cantón", Cantones[i], sep = " "))

Eval.tot[i] = as.numeric(metricas(df2[[i]]))

k_clear_session()
}

```

```

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

```

```

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

```

```

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

```



```
## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the 'lr' argument has
## been renamed to 'learning_rate'.
```

Resultados de métricas

```
Metricas = cbind (Eval.pd, Eval.tot)
colnames(Metricas) = c("NMRSE 2021", "NMRSE total")
rownames(Metricas) = Cantones
as.data.frame(Metricas)
```

```
##           NMRSE 2021 NMRSE total
## Alajuela      1.4446560  0.16663460
## Alajuelita    0.2047494  0.07196857
## Atenas        2.2815372  0.52214852
## Cañas         821.8338024  2.74762847
## Carrillo      692.4684731  2.38784369
## Corredores    78.4748838  0.71880550
## Desamparados  0.1310240  0.04601074
## Esparza       221.9349498  0.55024083
## Garabito      3700.0394719  2.49432232
## Golfito       80.1352735  1.12764540
## Guacimo       303.0959877  0.43839840
```

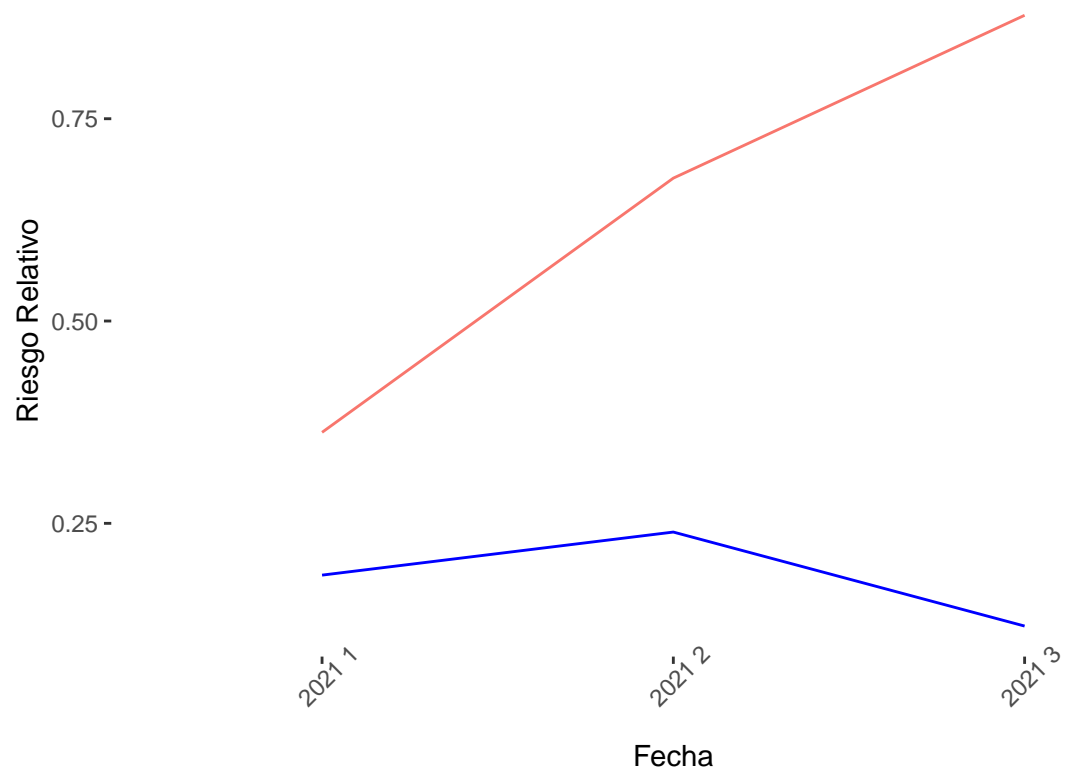
## La Cruz	1010.1042826	4.24121154
## Liberia	82.0557757	0.80820981
## Limon	48.6798393	0.37463552
## Matina	334.5840323	0.69806593
## Montes de Oro	622.6823104	2.29422619
## Nicoya	524.6406187	1.49309352
## Orotina	99.9195521	4.55660260
## Osa	6.6662544	0.61083310
## Parrita	2122.7292768	5.69426665
## Perez Zeledón	0.3282686	0.49367758
## Pococí	81.6122676	0.55843133
## Puntarenas	166.4292687	0.69336345
## Quepos	2901.4815820	3.37209411
## San Jose	0.7877130	0.04893541
## Santa Ana	13.7528285	0.43520149
## SantaCruz	435.5076798	2.20131794
## Sarapiquí	302.2053124	1.49076657
## Siquirres	98.7920736	0.52896842
## Talamanca	120.8983874	1.53538331
## Turrialba	44.3300019	1.26904355
## Upala	Inf	0.29555310

Gráficos

p1

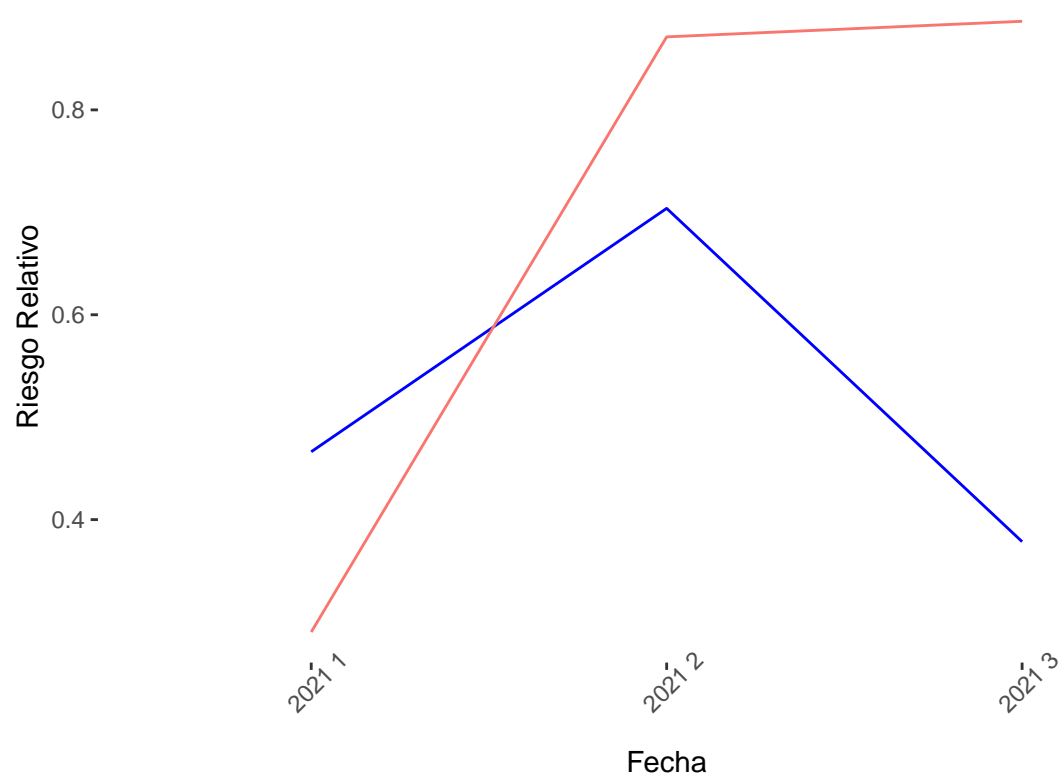
[[1]]

Predicciones 2021 del cantón Alajuela



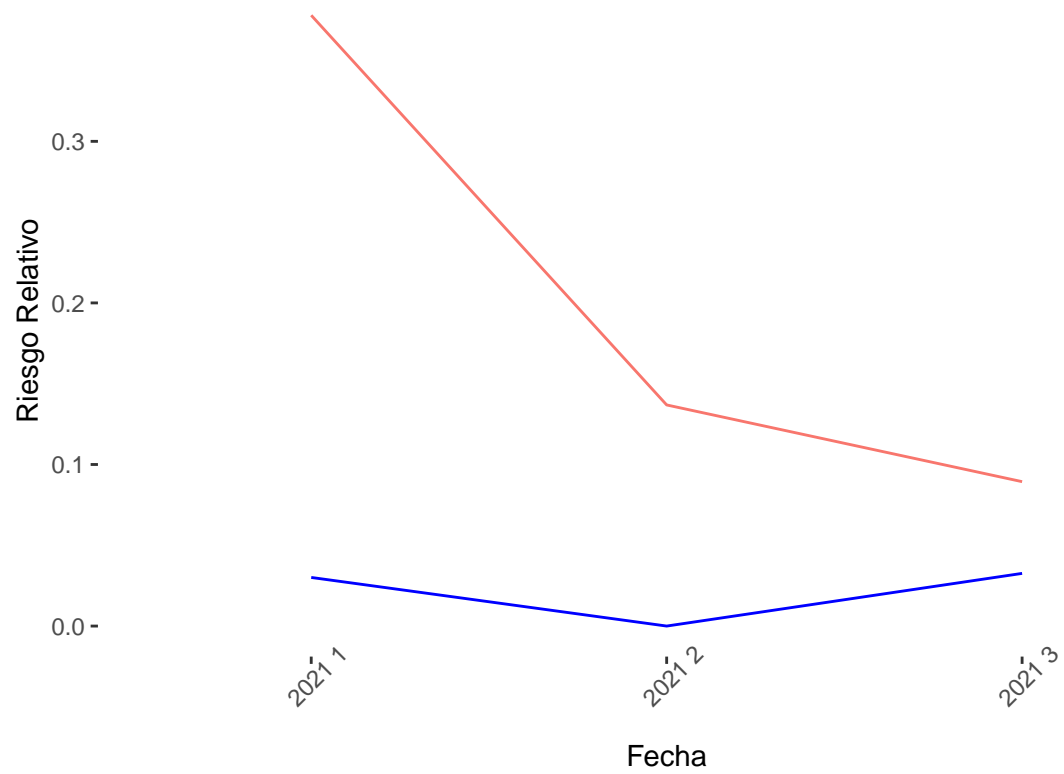
```
##  
## [[2]]
```

Predicciones 2021 del cantón Alajuelita

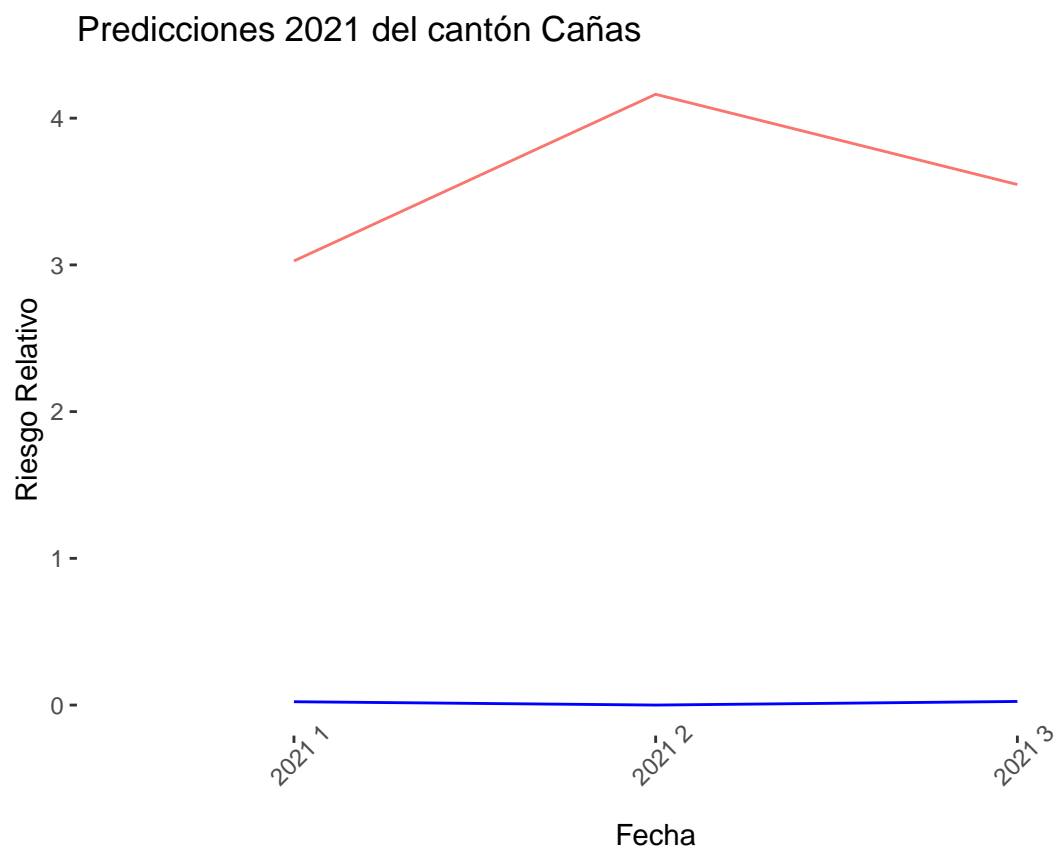


[[3]]

Predicciones 2021 del cantón Atenas

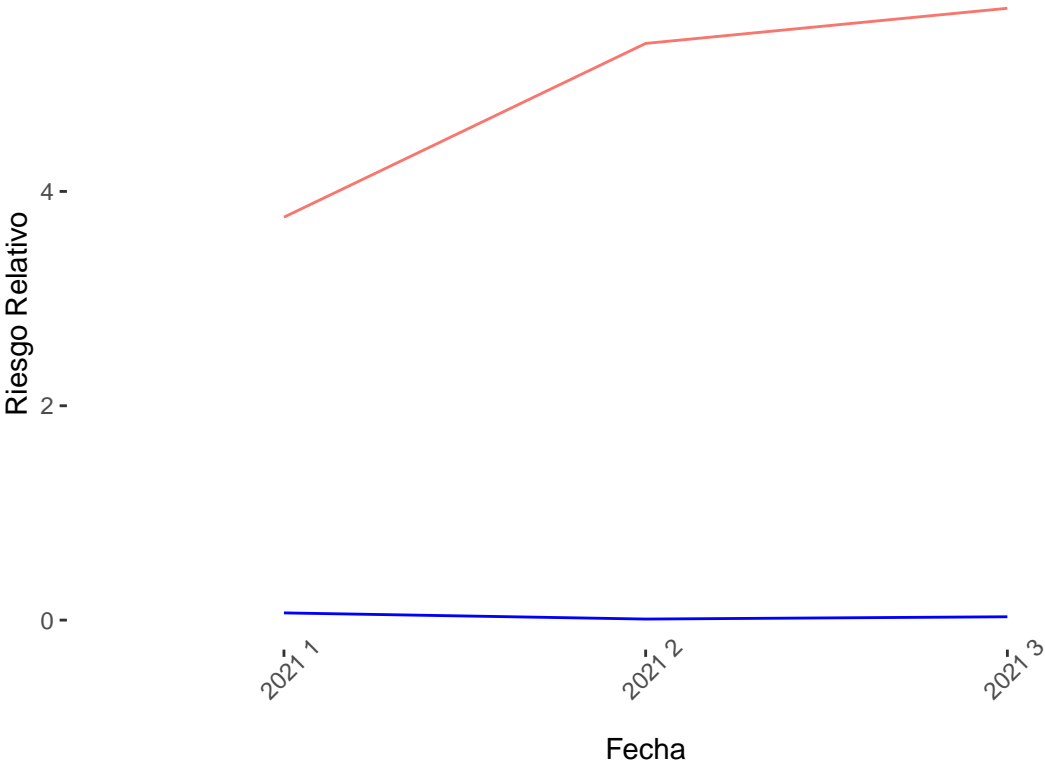


```
##  
## [[4]]
```



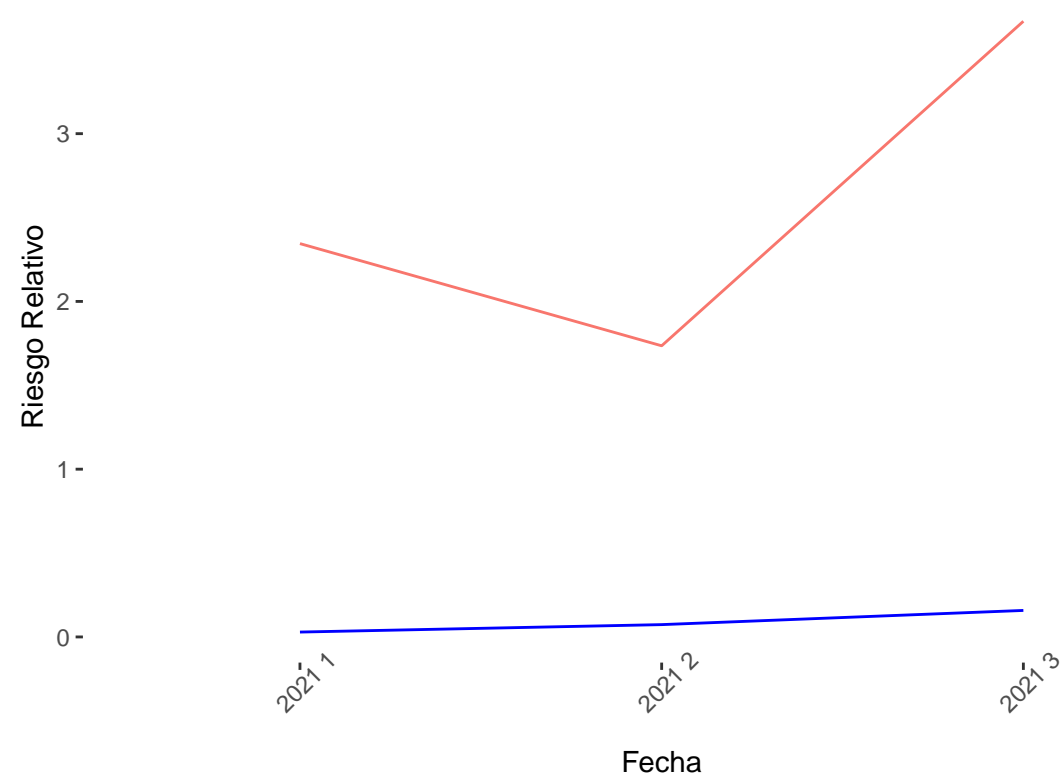
```
##  
## [[5]]
```

Predicciones 2021 del cantón Carrillo



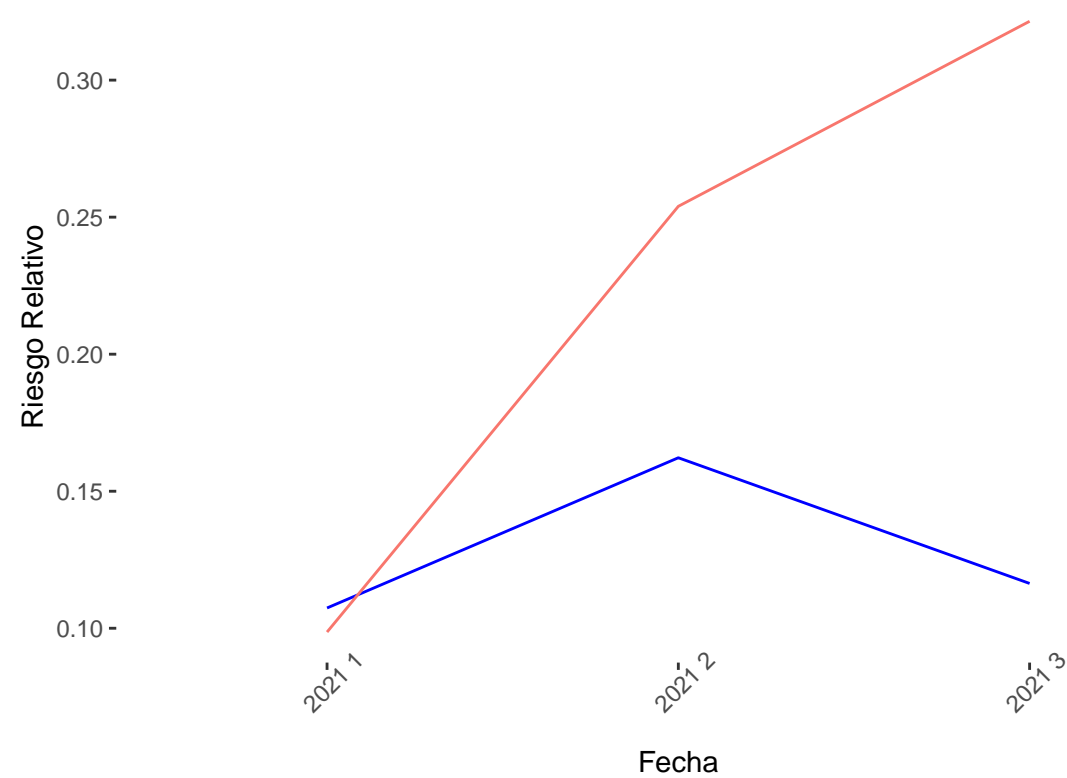
[[6]]

Predicciones 2021 del cantón Corredores



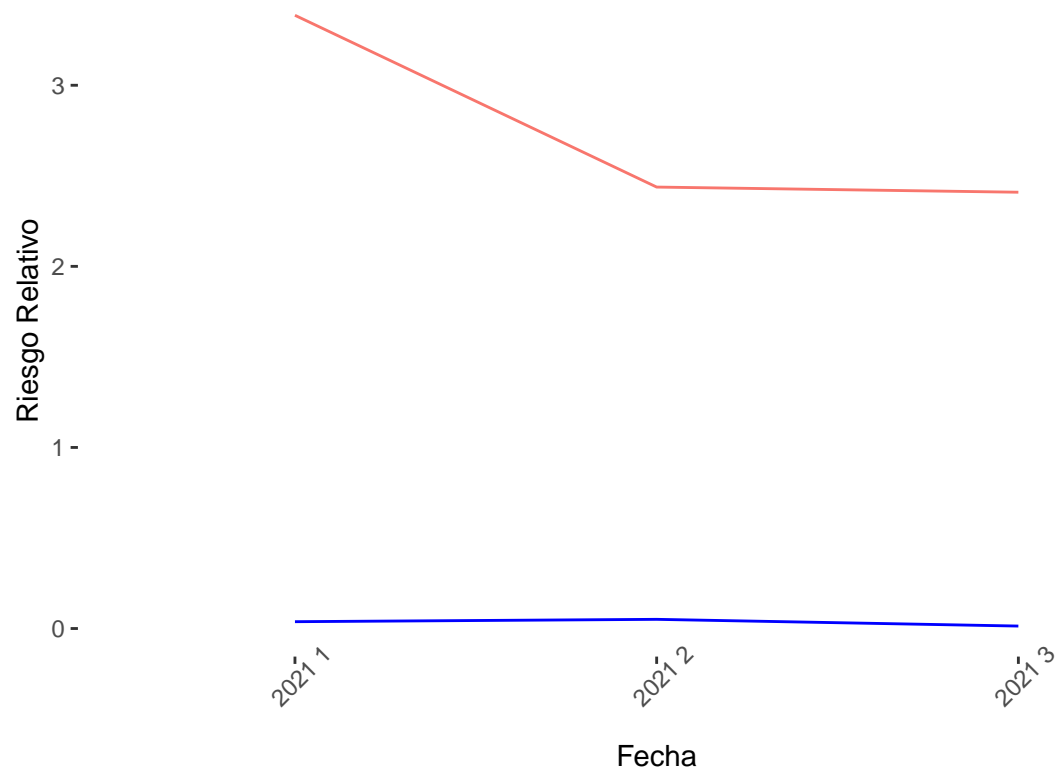
[[7]]

Predicciones 2021 del cantón Desamparados



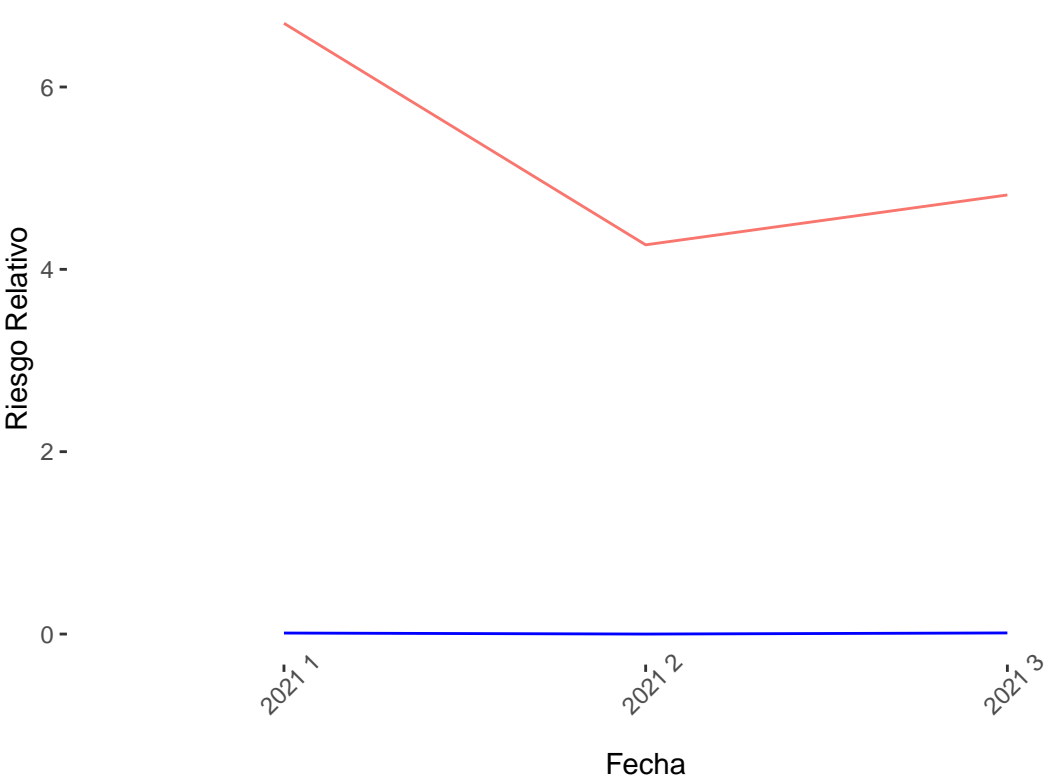
[[8]]

Predicciones 2021 del cantón Esparza



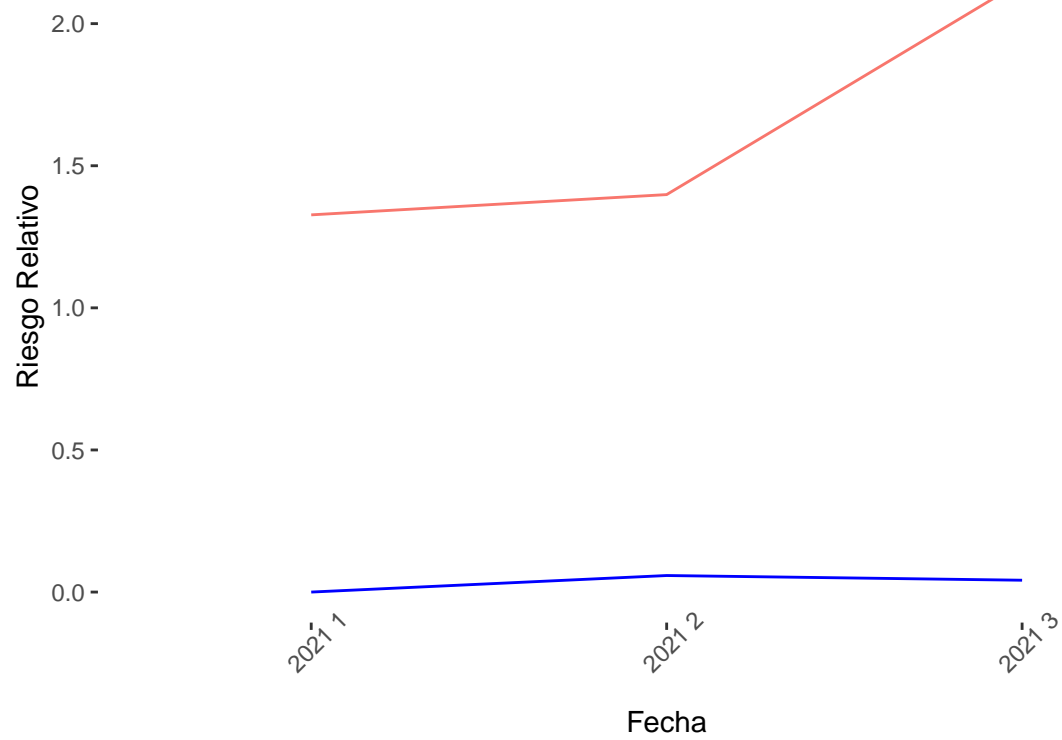
```
##  
## [[9]]
```


Predicciones 2021 del cantón Garabito



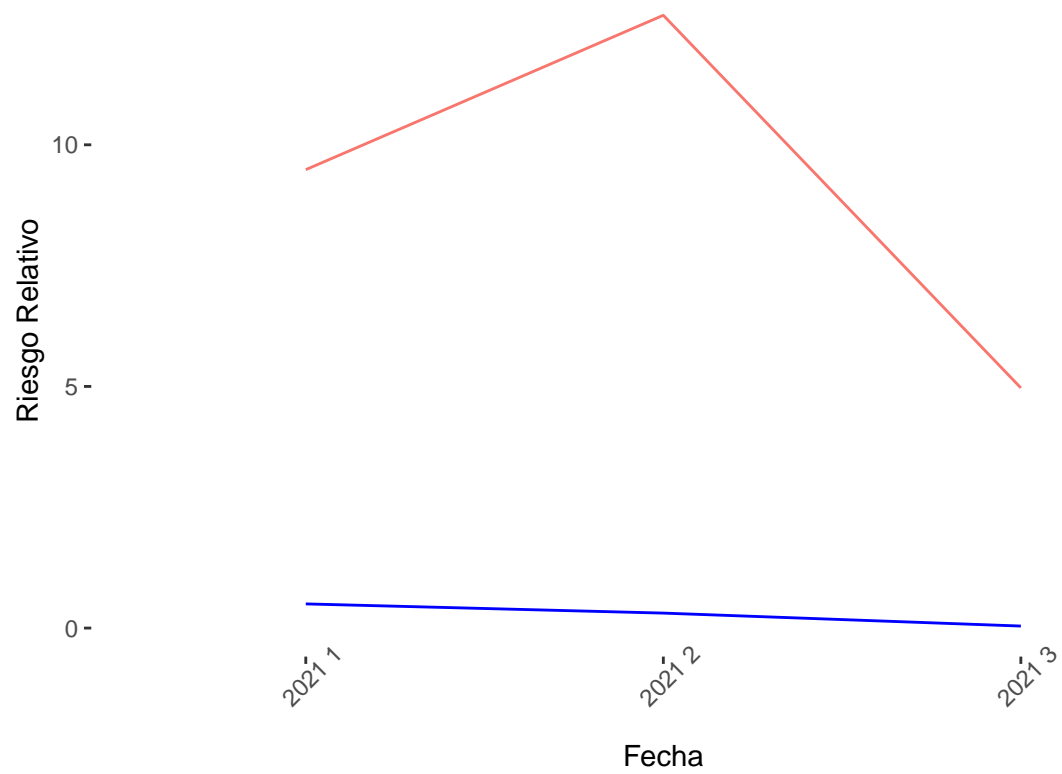
[[10]]

Predicciones 2021 del cantón Golfito



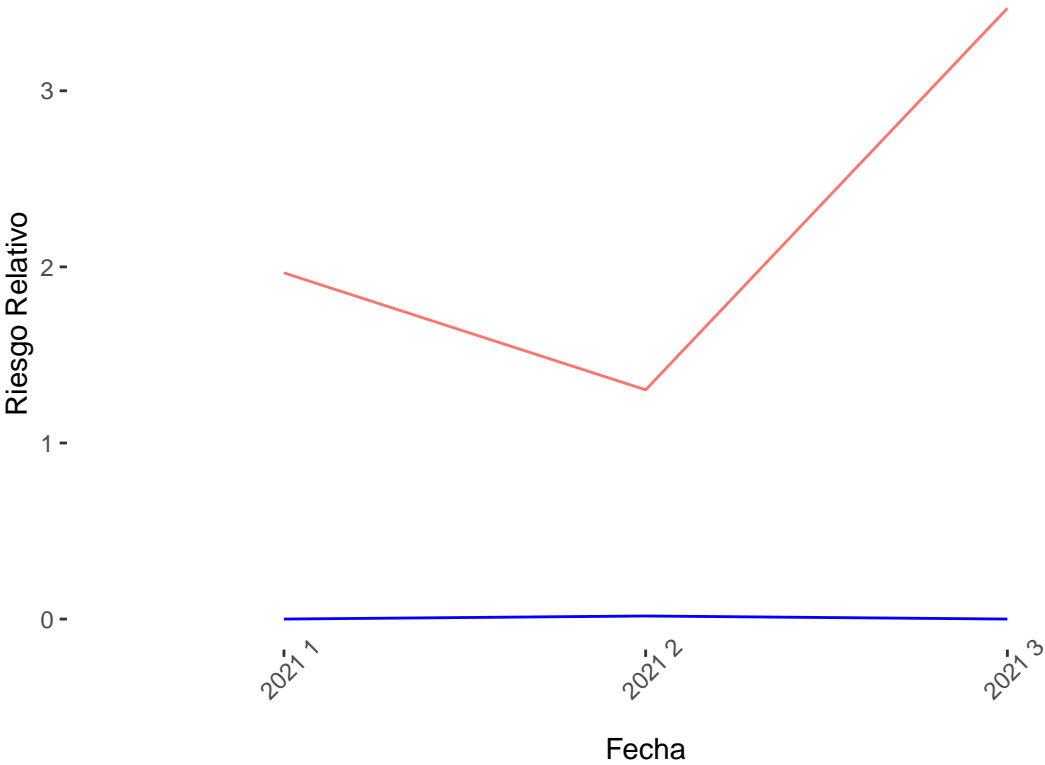
```
##  
## [[11]]
```

Predicciones 2021 del cantón Guacimo



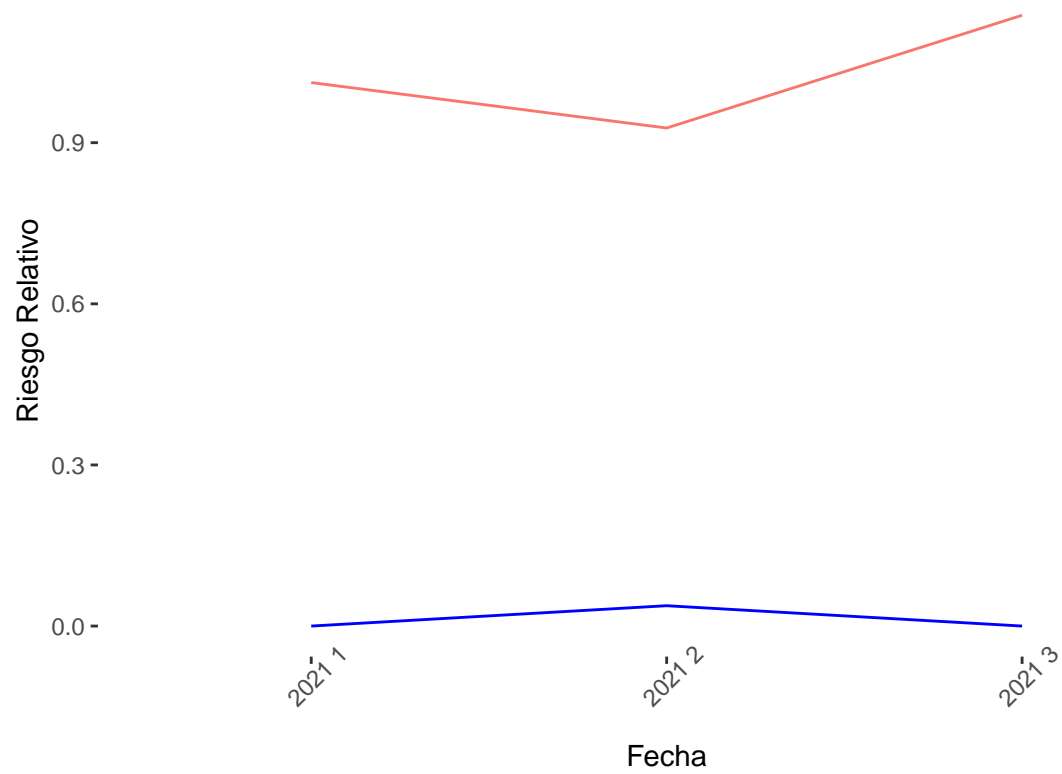
```
##  
## [[12]]
```

Predicciones 2021 del cantón La Cruz



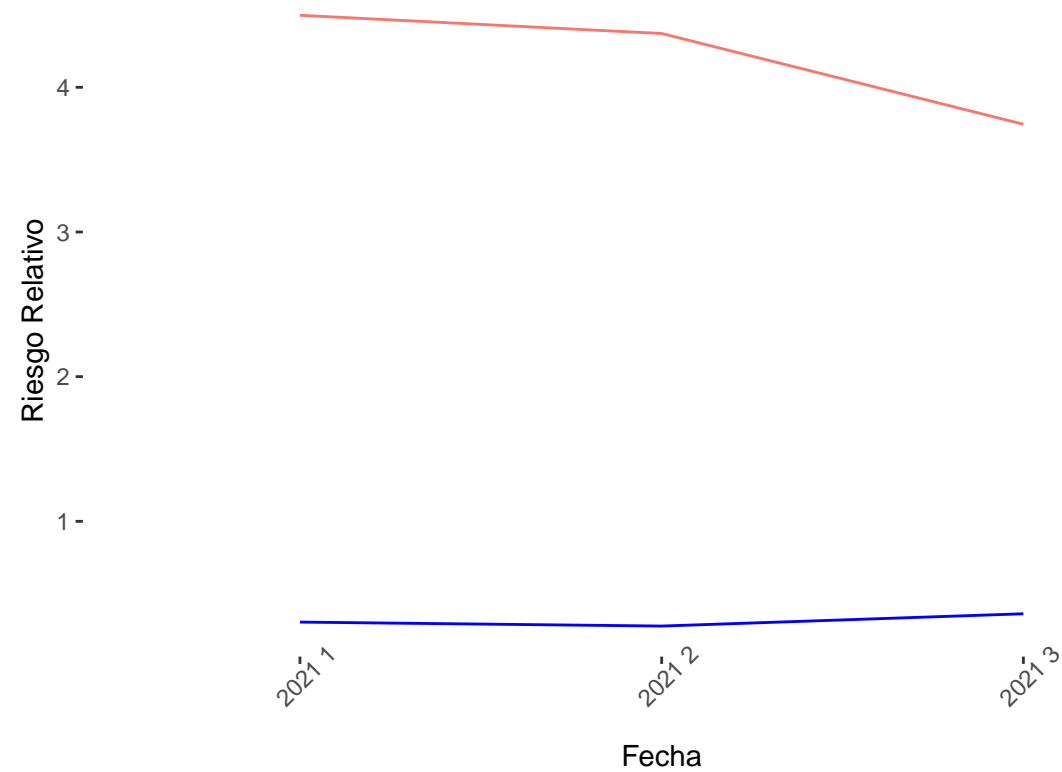
[[13]]

Predicciones 2021 del cantón Liberia

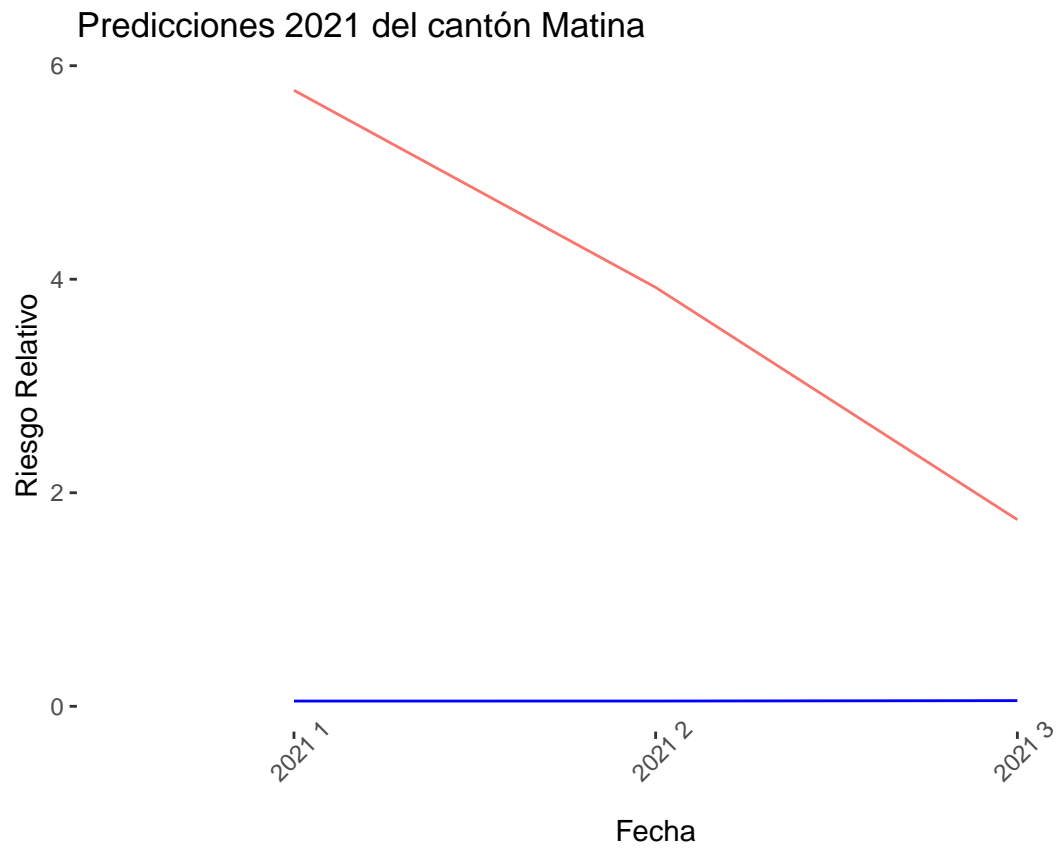


```
##  
## [[14]]
```

Predicciones 2021 del cantón Limon

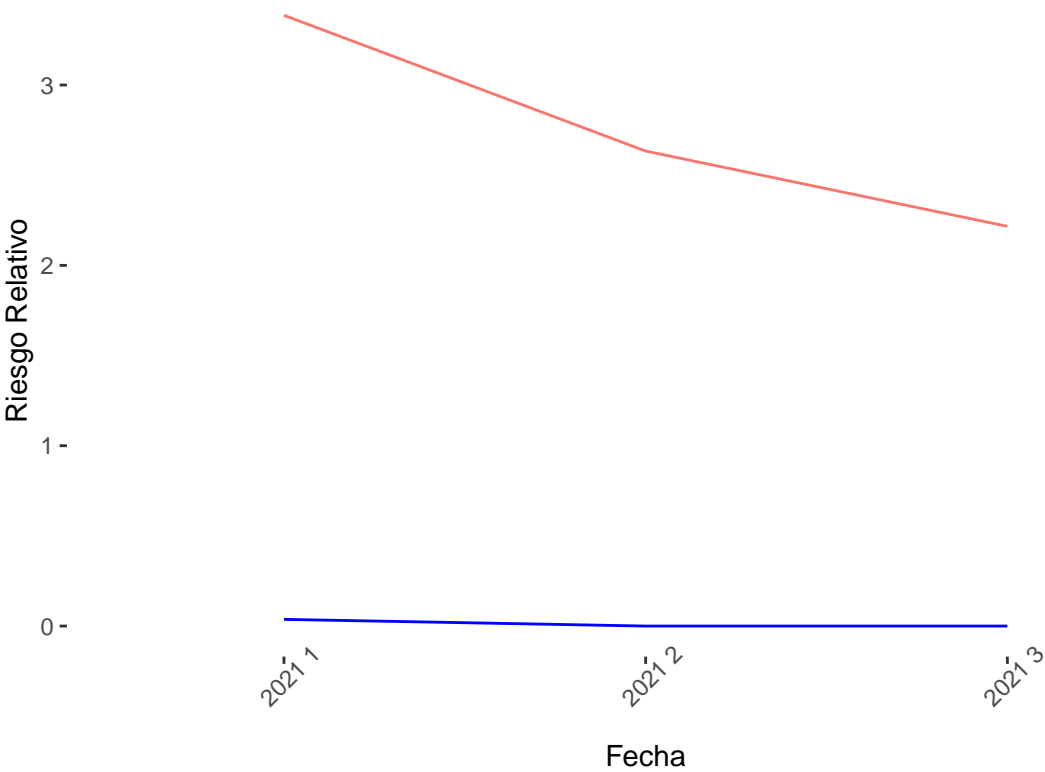


```
##  
## [[15]]
```



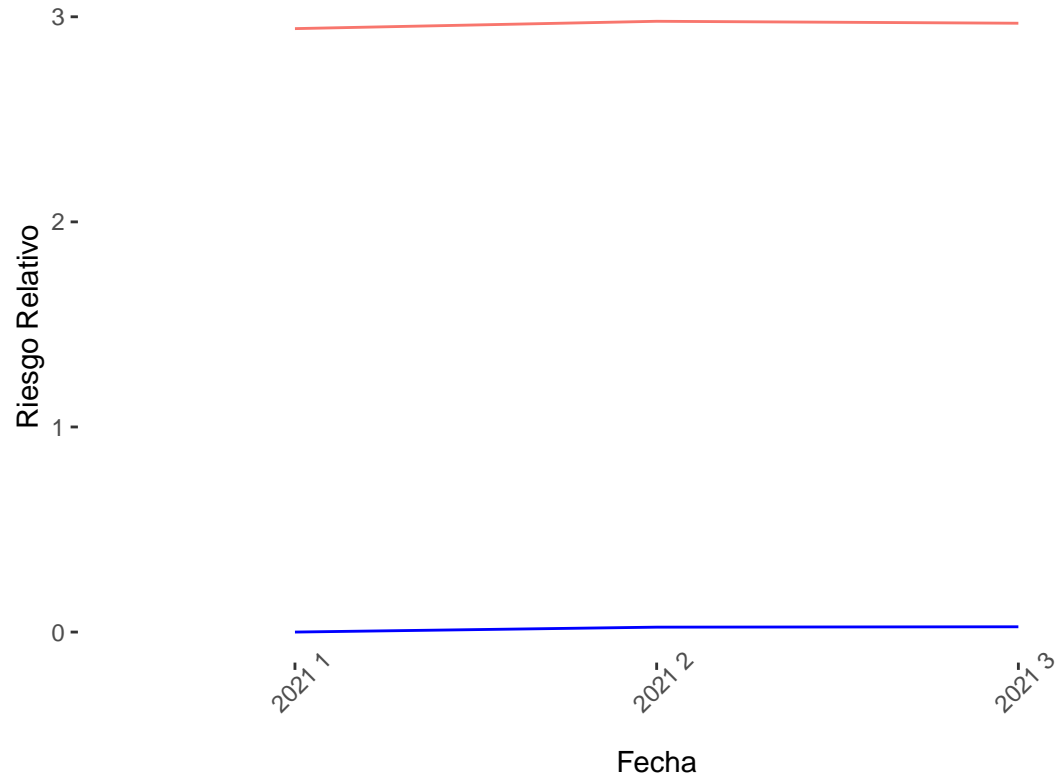
```
##  
## [[16]]
```

Predicciones 2021 del cantón Montes de Oro

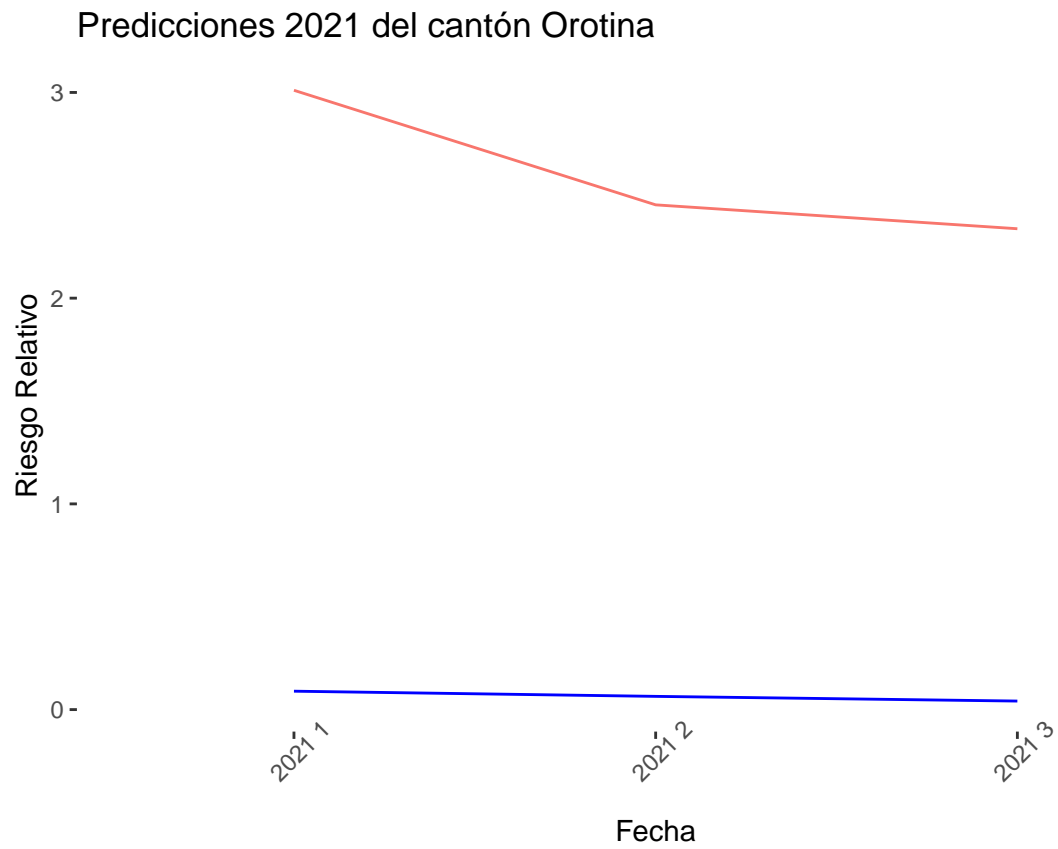


[[17]]

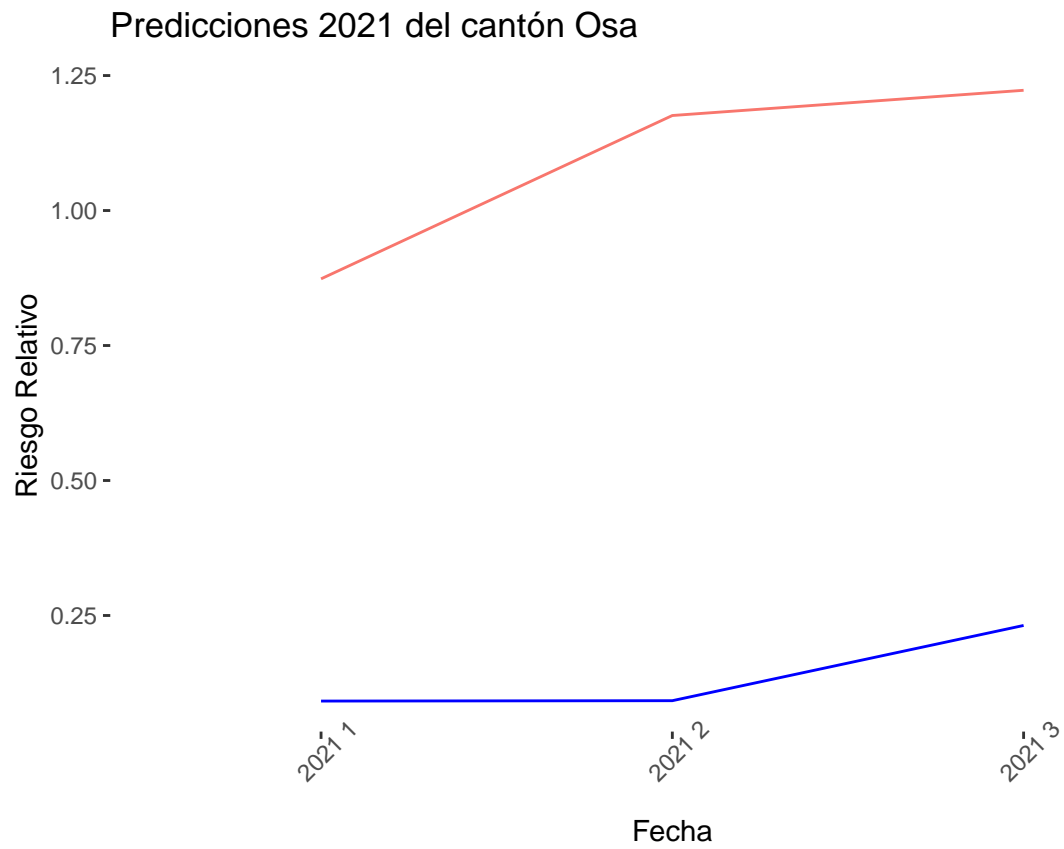
Predicciones 2021 del cantón Nicoya



```
##  
## [[18]]
```

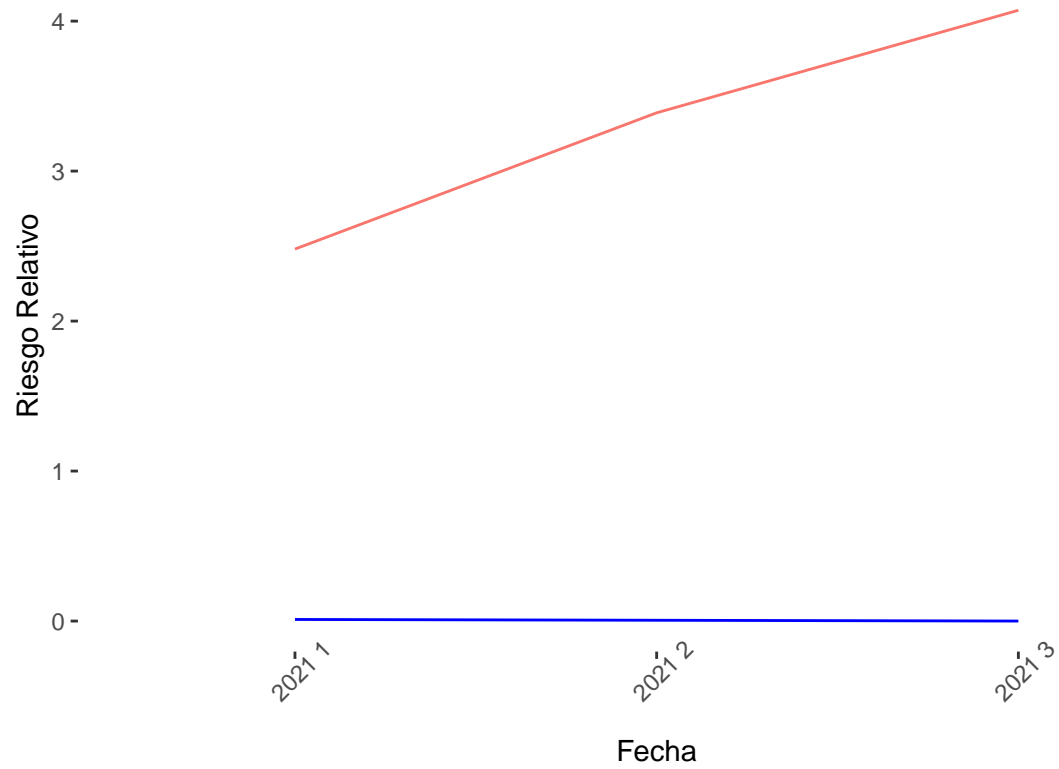


```
##  
## [[19]]
```



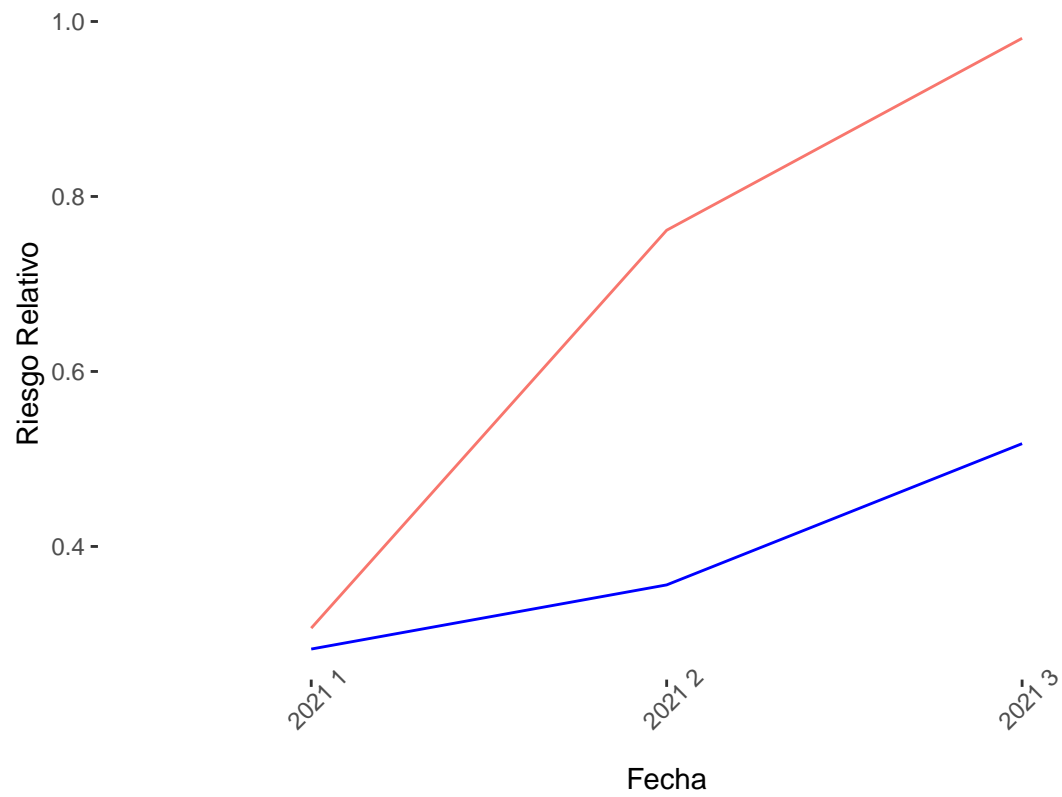
```
##  
## [[20]]
```

Predicciones 2021 del cantón Parrita

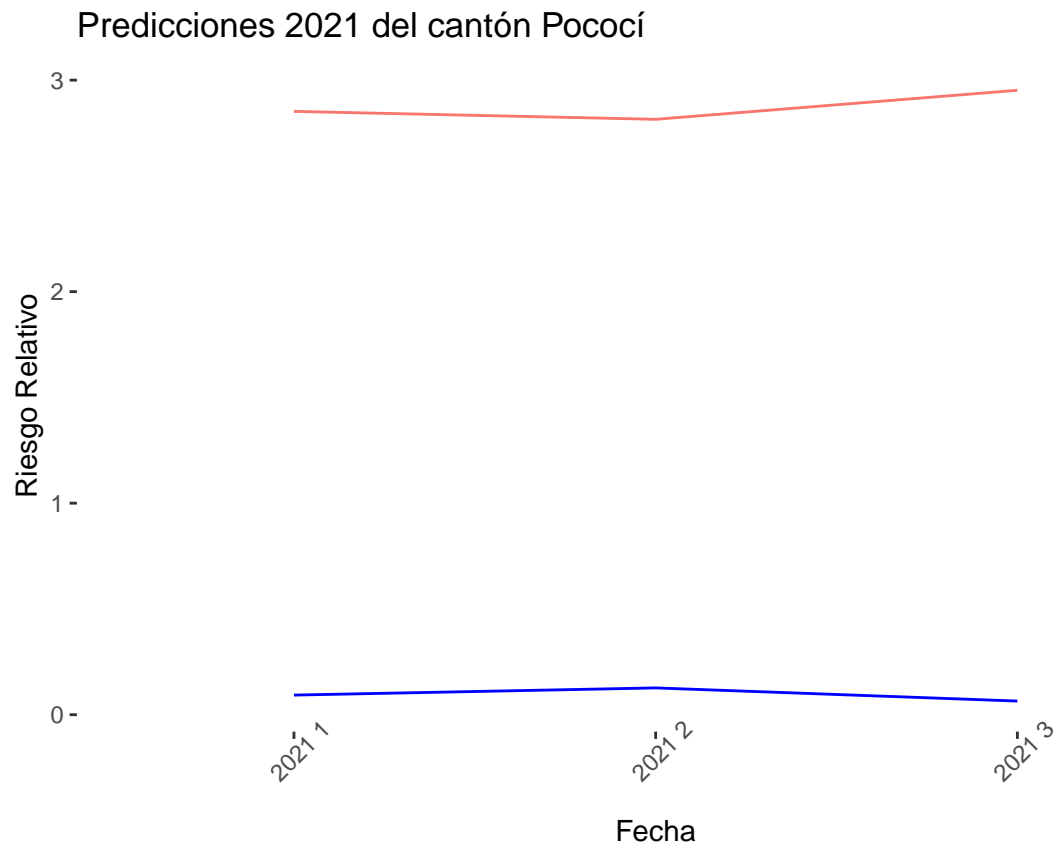


```
##  
## [[21]]
```

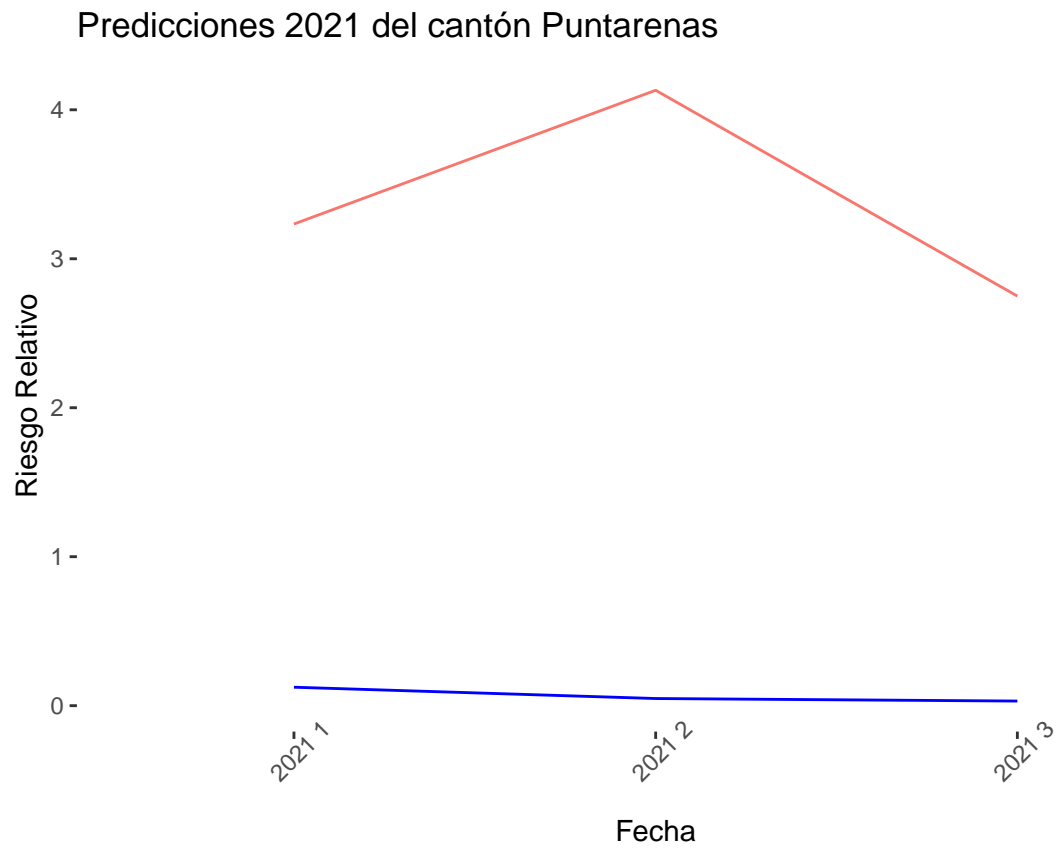
Predicciones 2021 del cantón Perez Zeledón



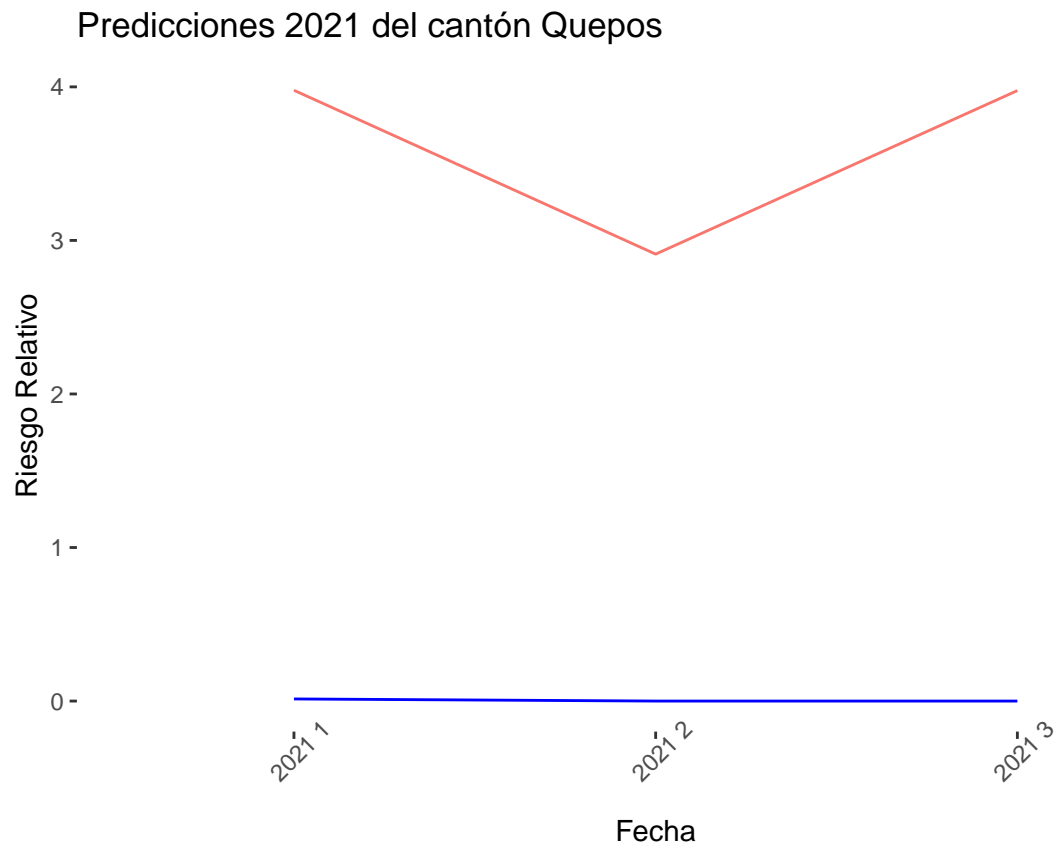
[[22]]



```
##  
## [[23]]
```

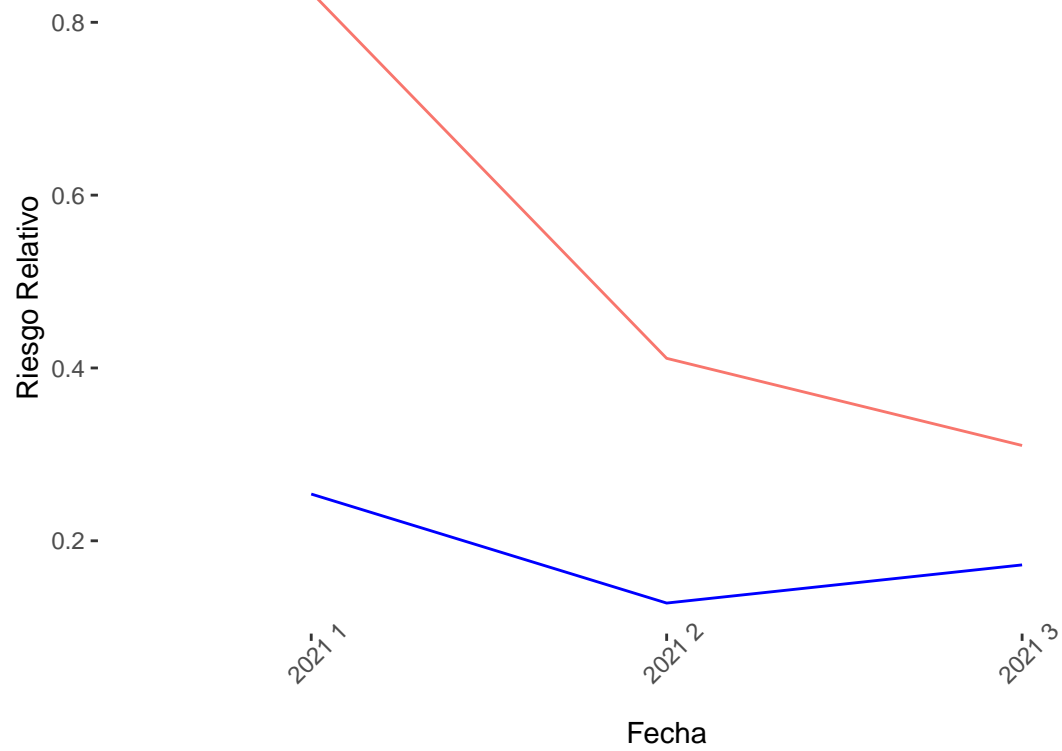


```
##  
## [[24]]
```



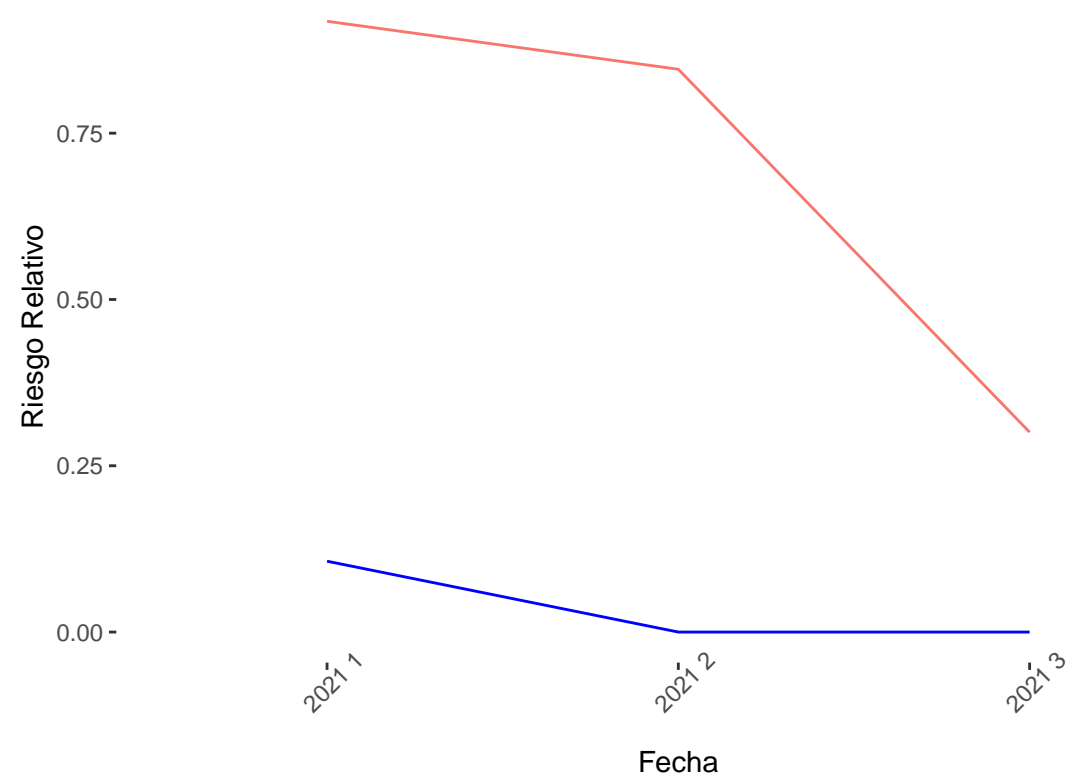
```
##  
## [[25]]
```


Predicciones 2021 del cantón San Jose



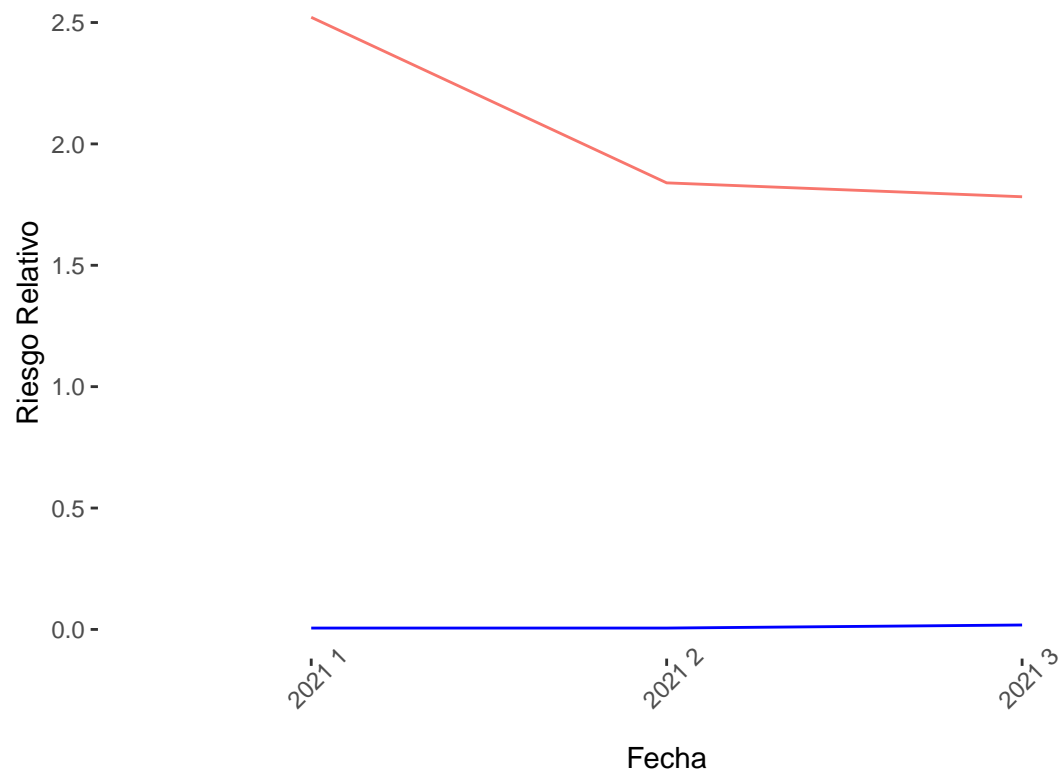
```
##  
## [[26]]
```

Predicciones 2021 del cantón Santa Ana



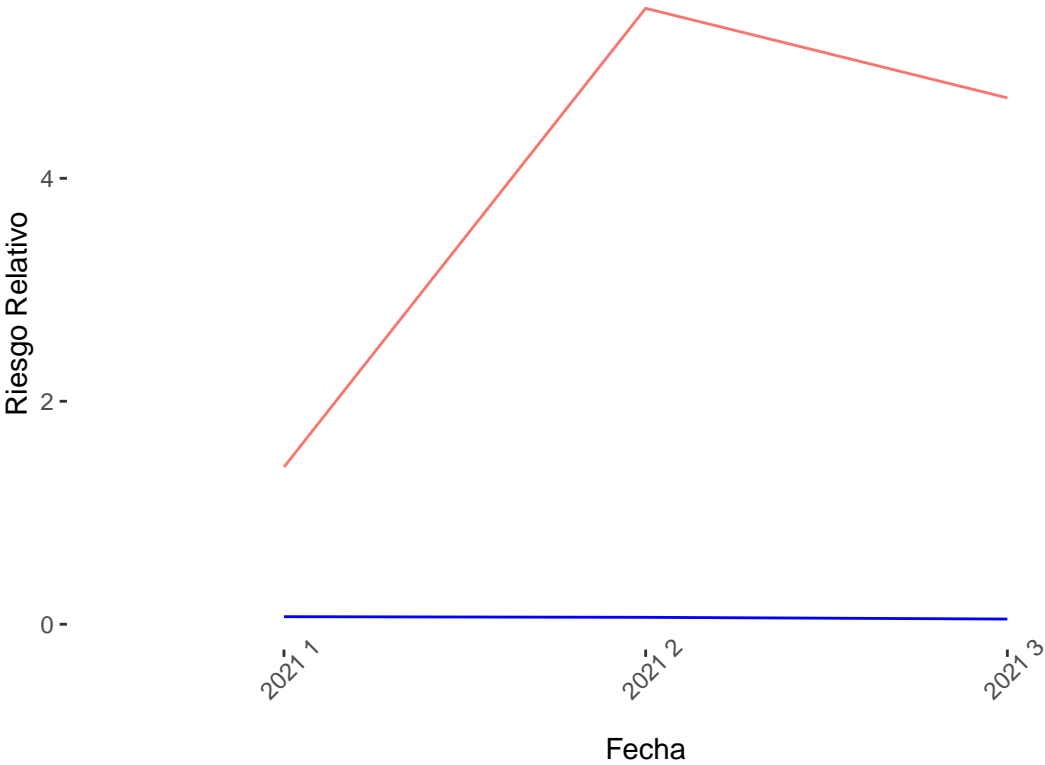
[[27]]

Predicciones 2021 del cantón SantaCruz

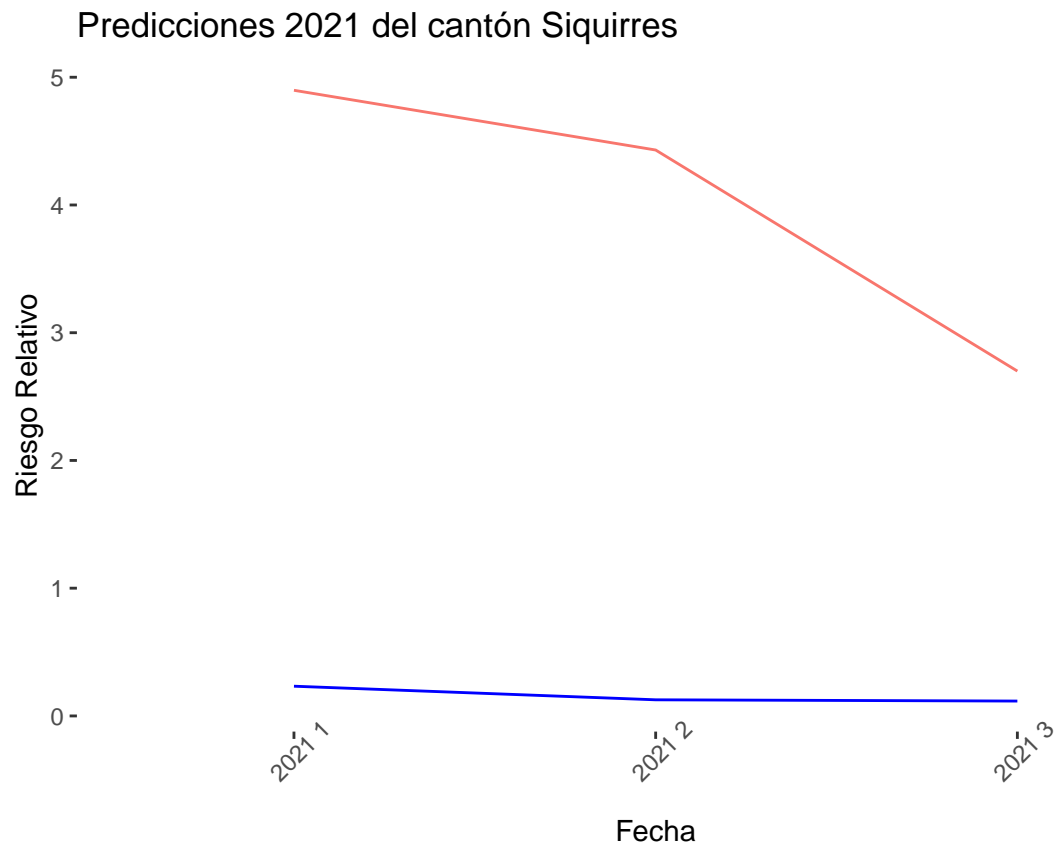


[[28]]

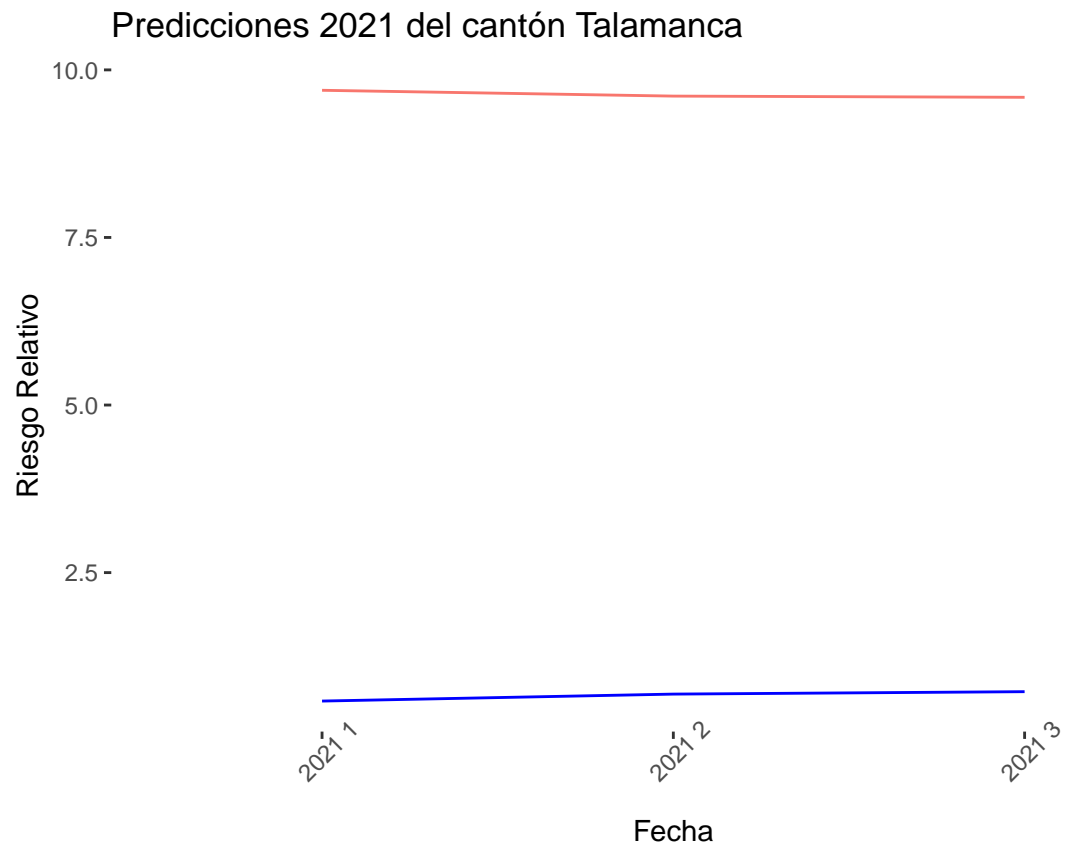
Predicciones 2021 del cantón Sarapiquí



[[29]]

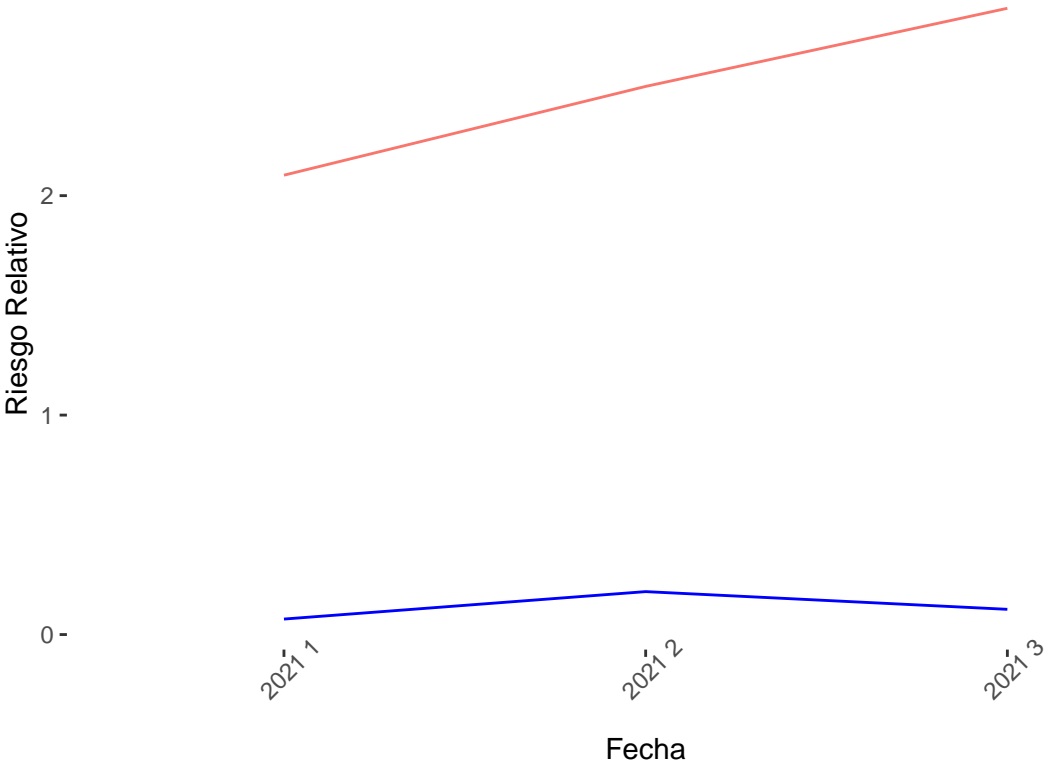


[[30]]



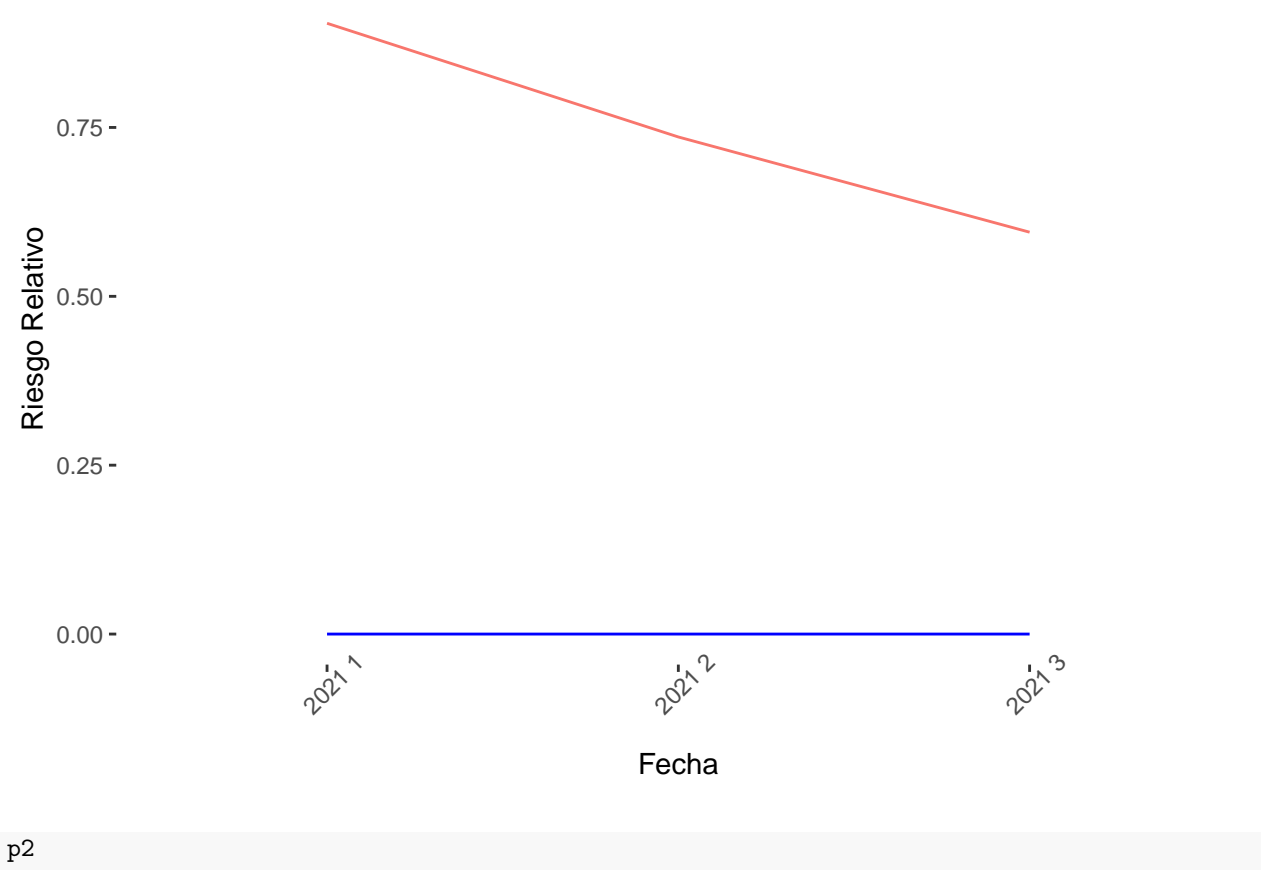
```
##  
## [[31]]
```

Predicciones 2021 del cantón Turrialba



[[32]]

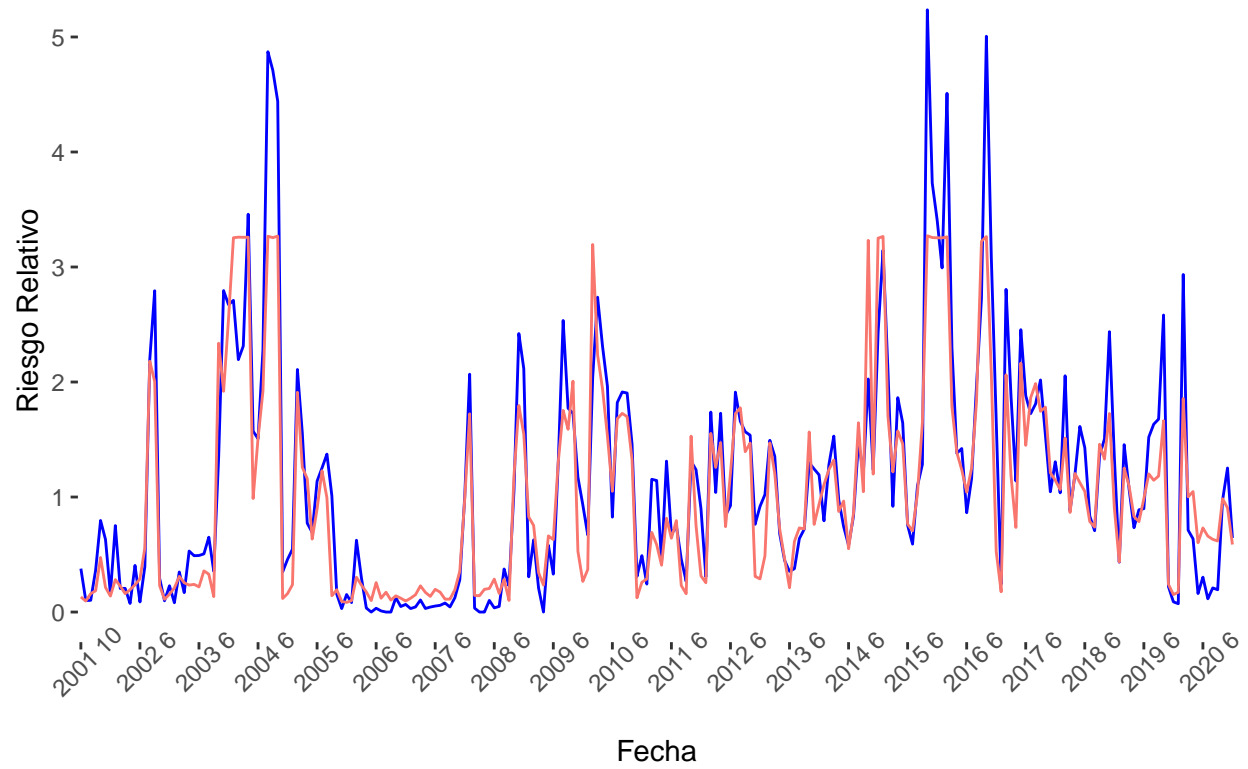
Predicciones 2021 del cantón Upala



p2

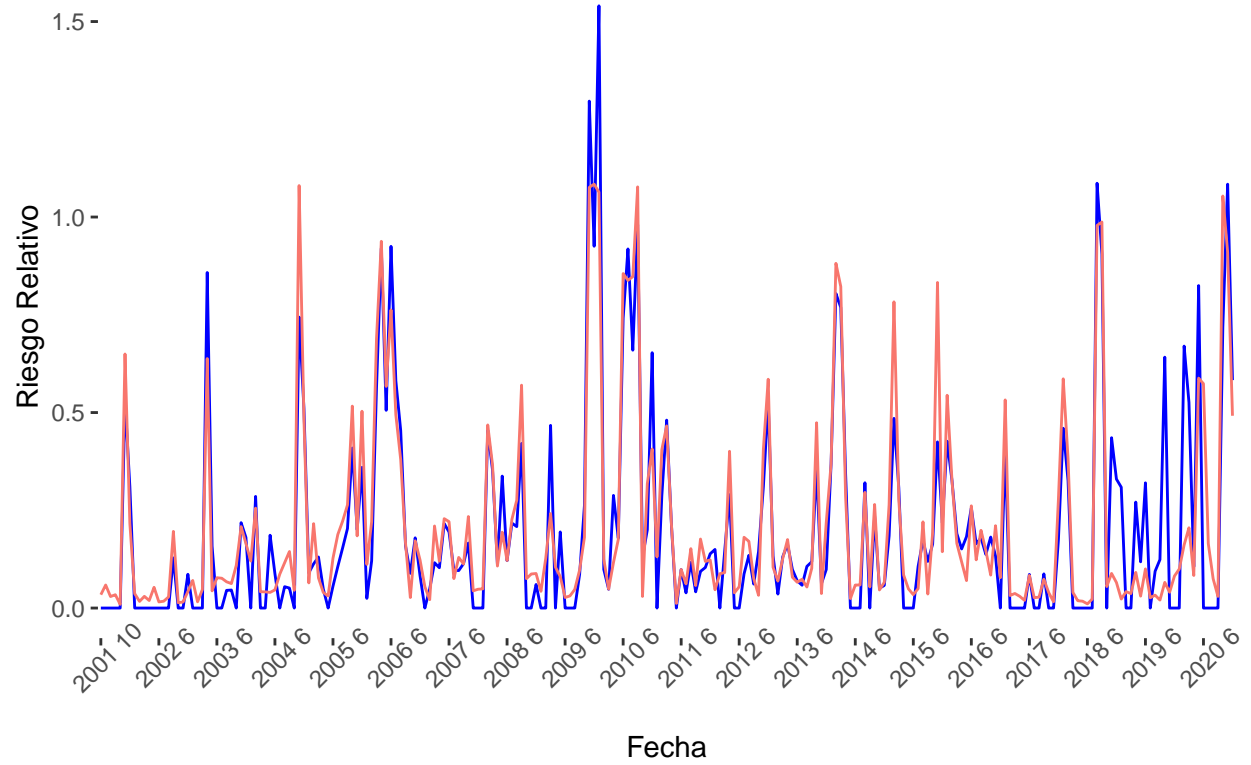
[[1]]

Valores aproximados de training del cantón Alajuela

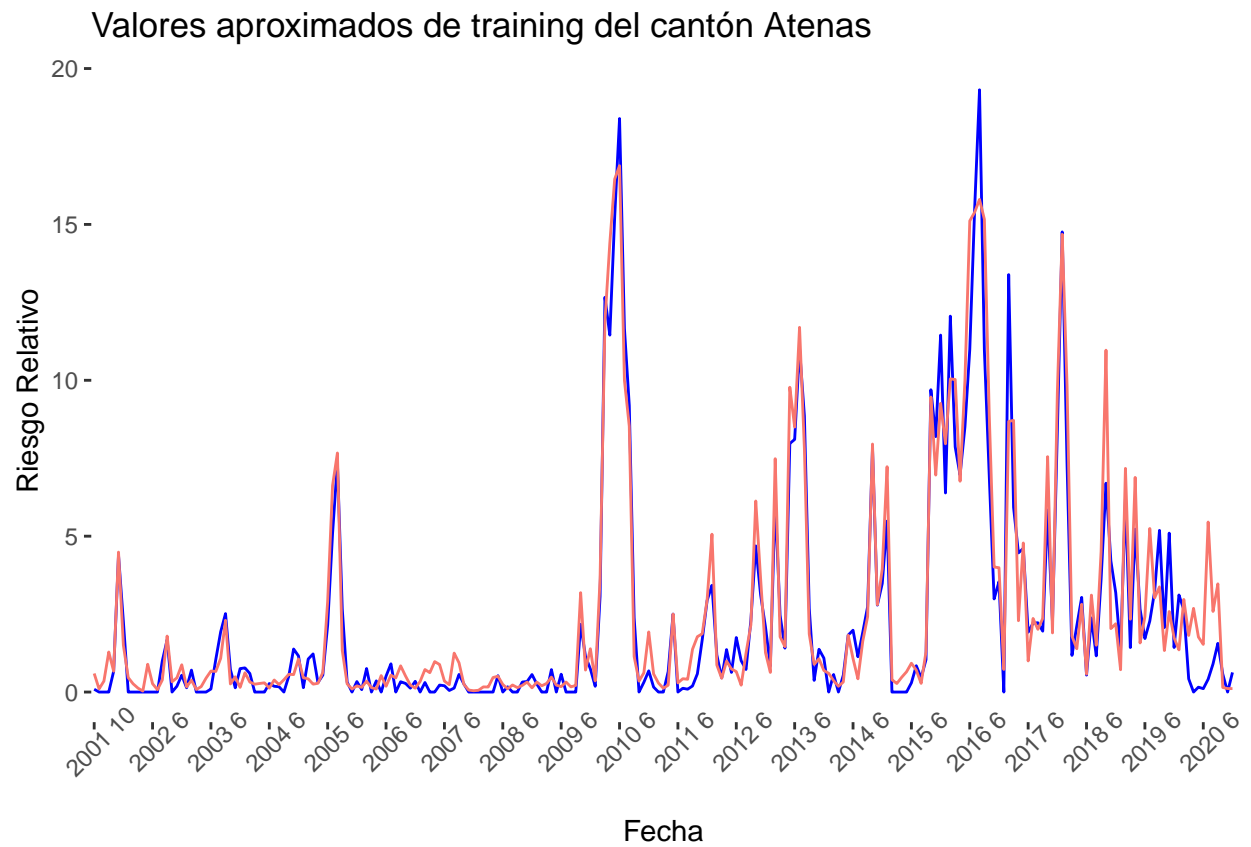


```
##  
## [[2]]
```

Valores aproximados de training del cantón Alajuelita

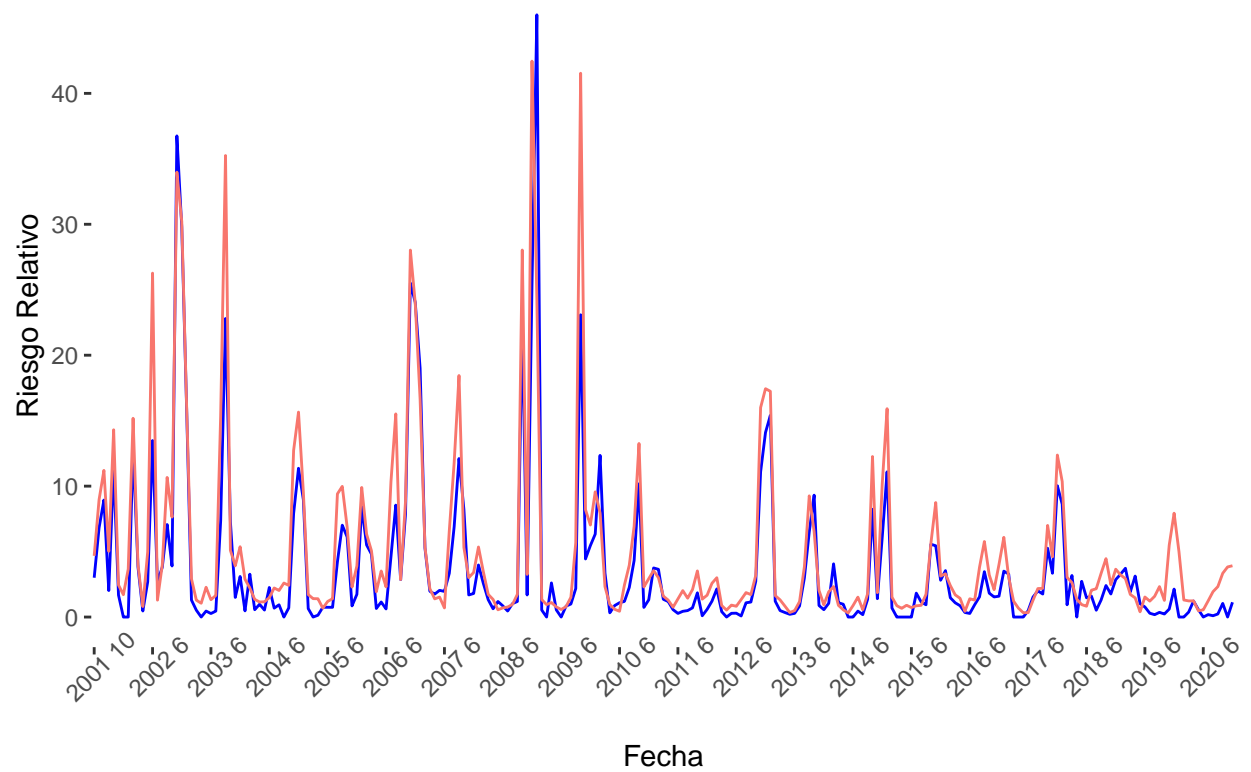


```
##  
## [[3]]
```



```
##  
## [[4]]
```

Valores aproximados de training del cantón Cañas

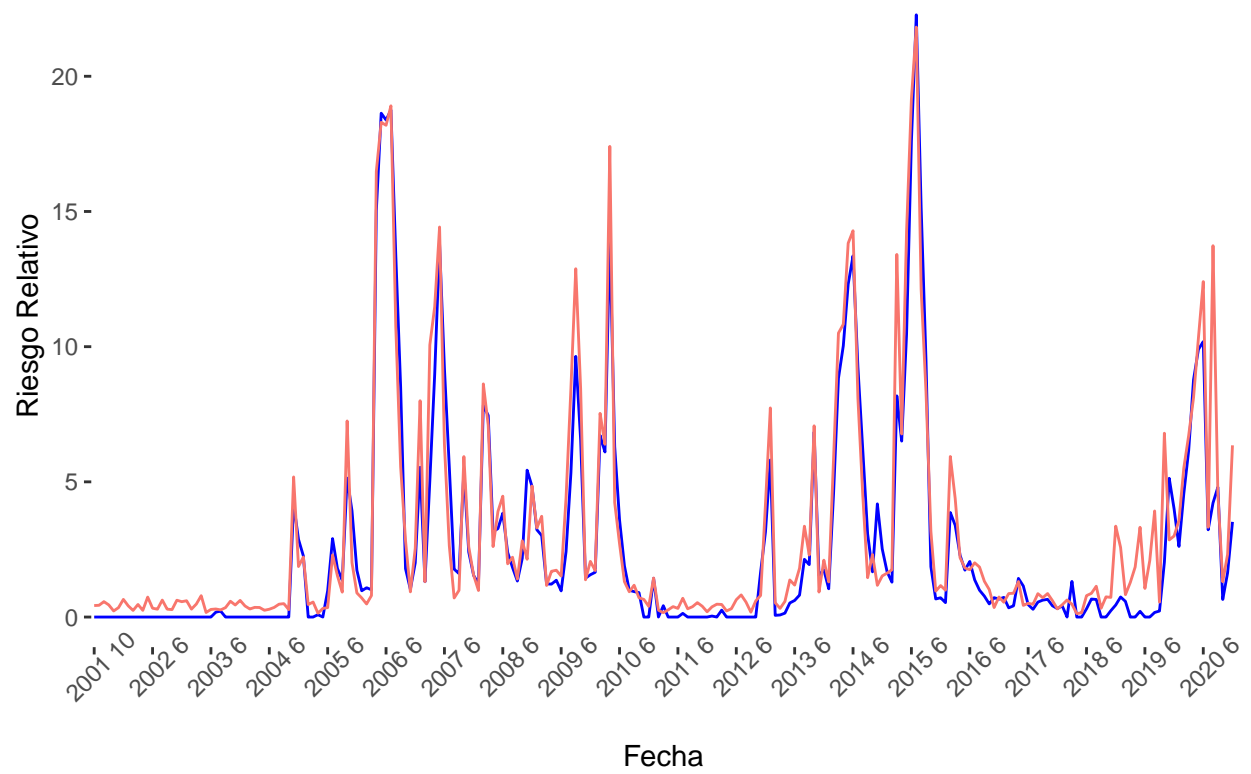


```
##  
## [[5]]
```



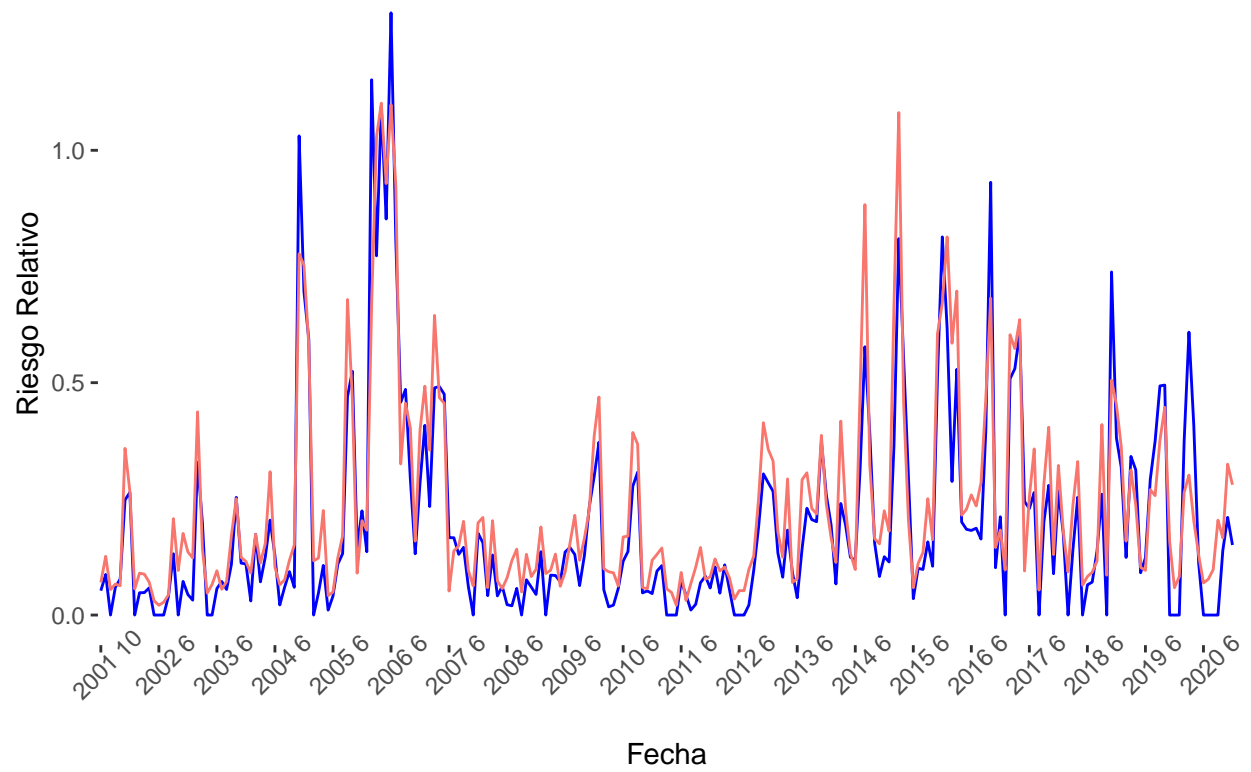
```
##  
## [[6]]
```

Valores aproximados de training del cantón Corredores



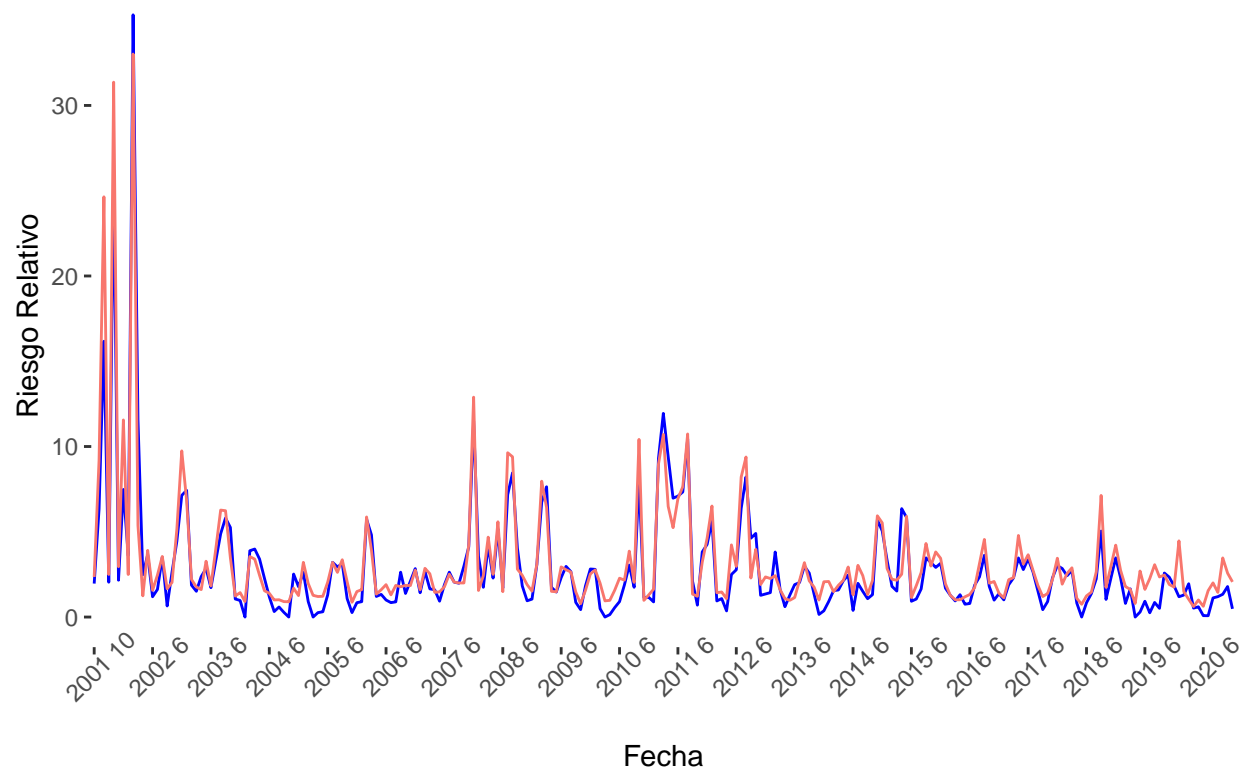
[[7]]

Valores aproximados de training del cantón Desamparados



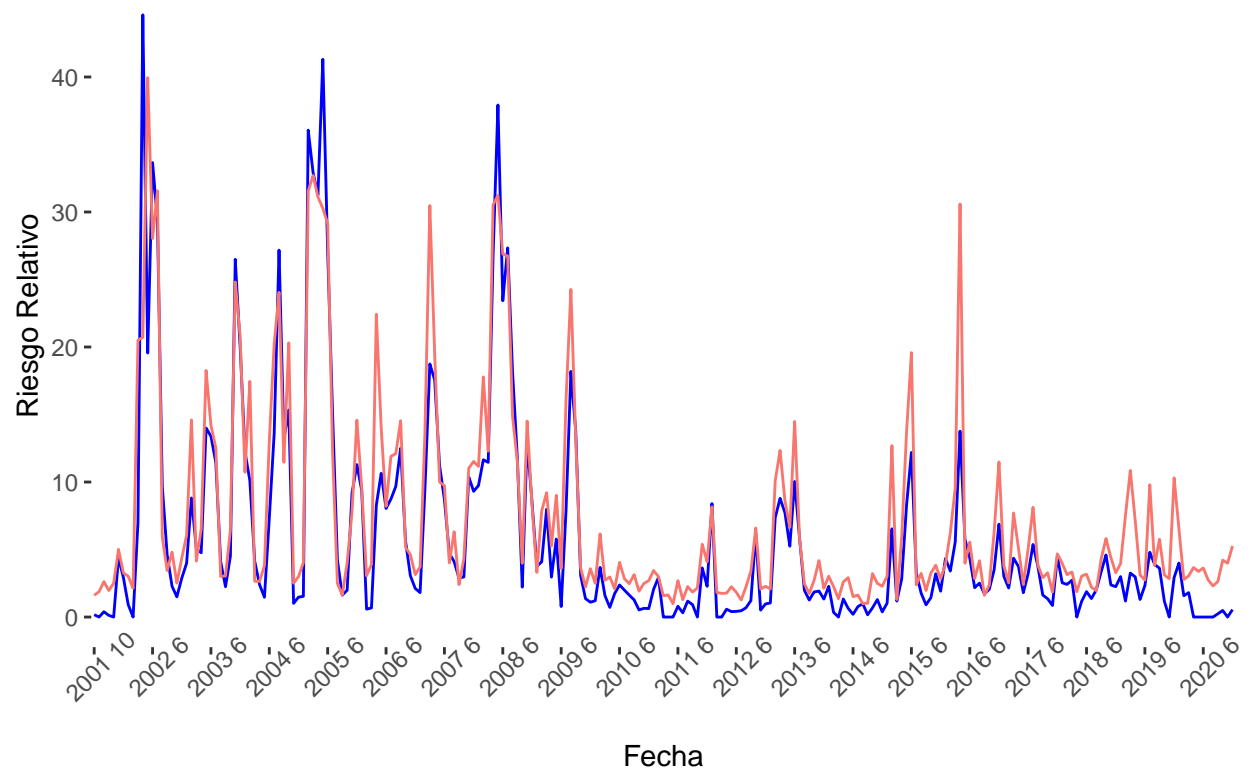
```
##  
## [[8]]
```

Valores aproximados de training del cantón Esparza



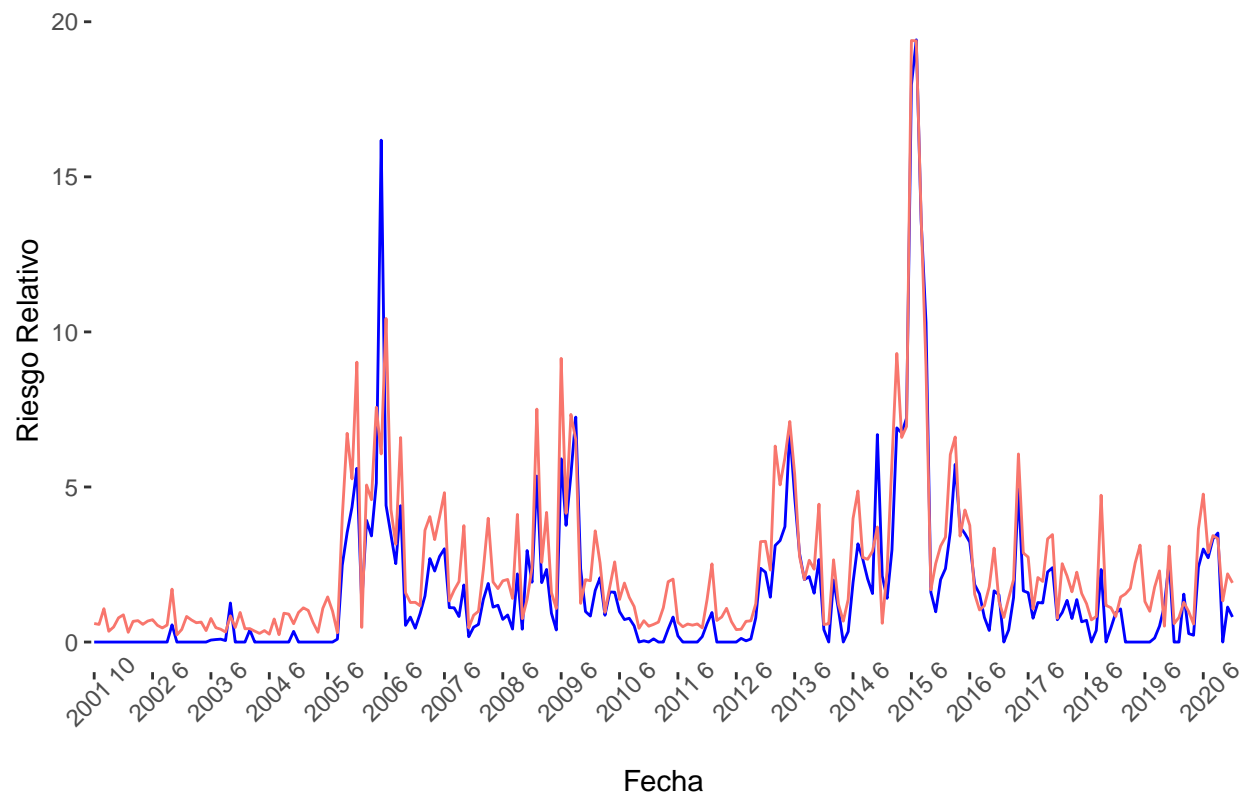
```
##  
## [[9]]
```


Valores aproximados de training del cantón Garabito



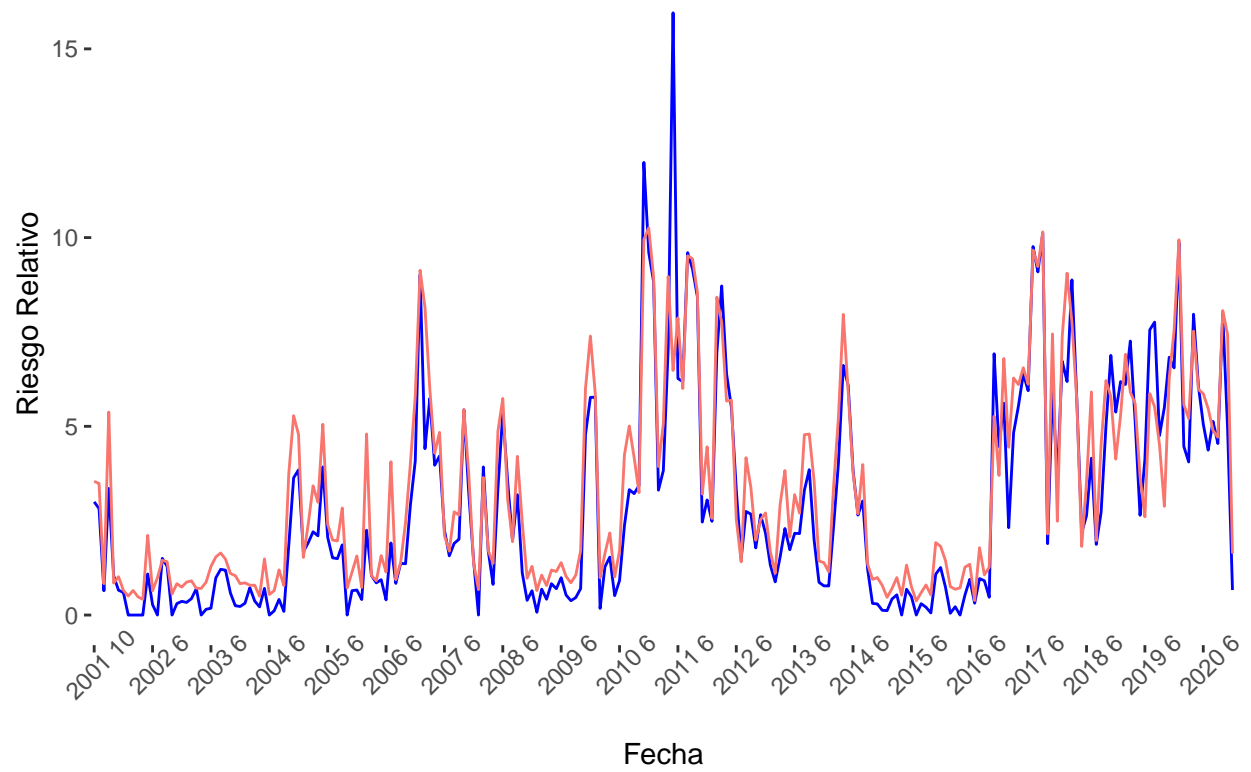
```
##  
## [[10]]
```

Valores aproximados de training del cantón Golfito



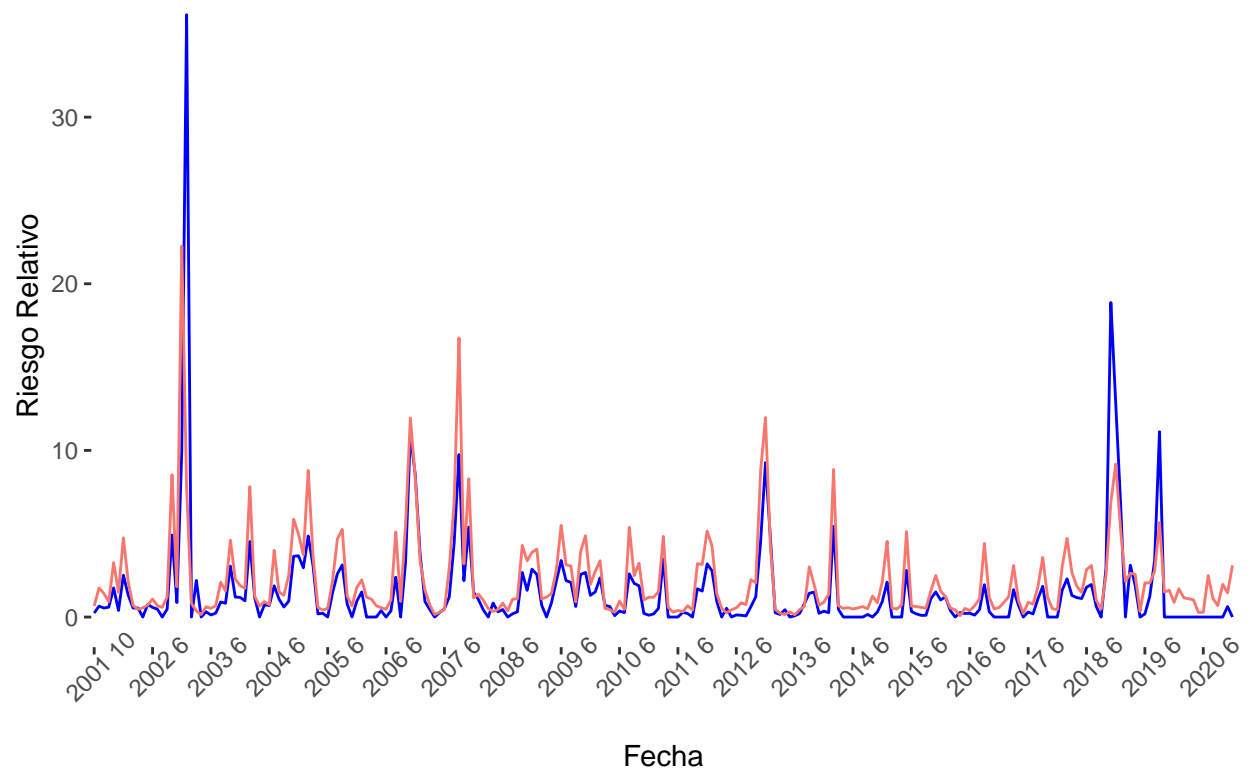
```
##  
## [[11]]
```

Valores aproximados de training del cantón Guacimo



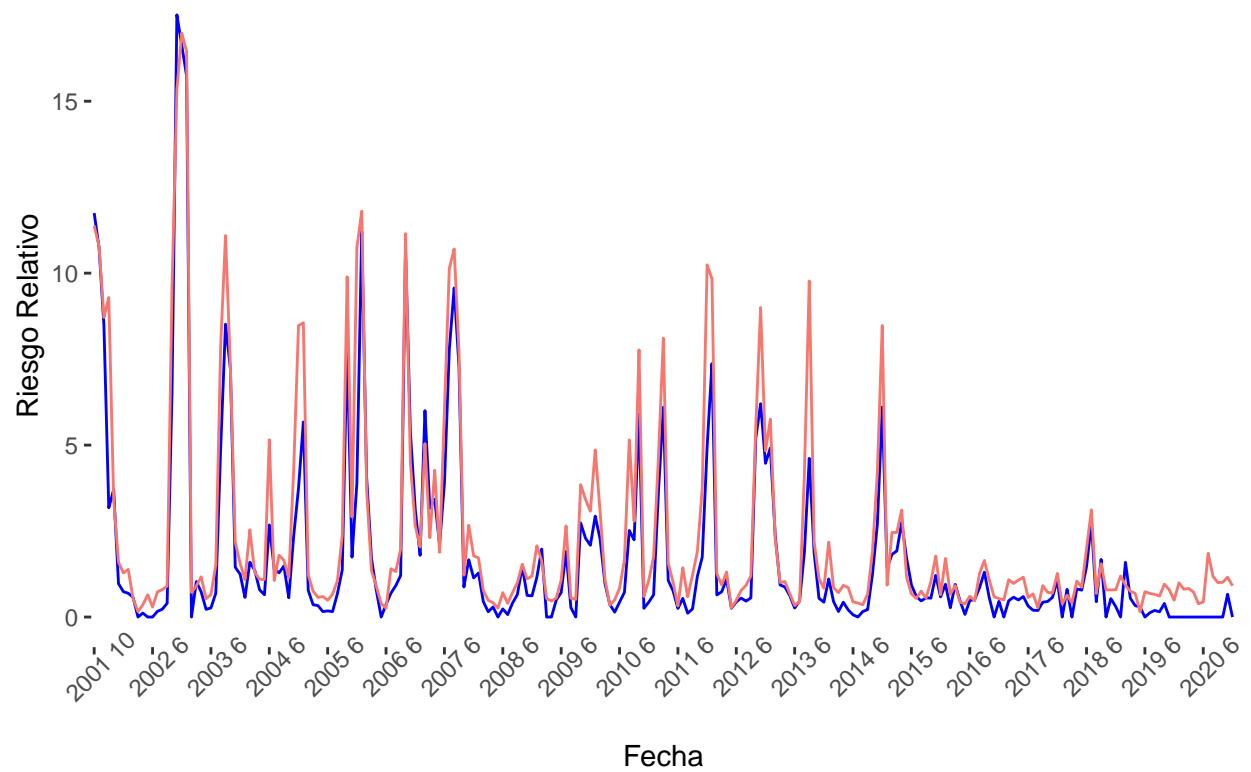
```
##  
## [[12]]
```

Valores aproximados de training del cantón La Cruz



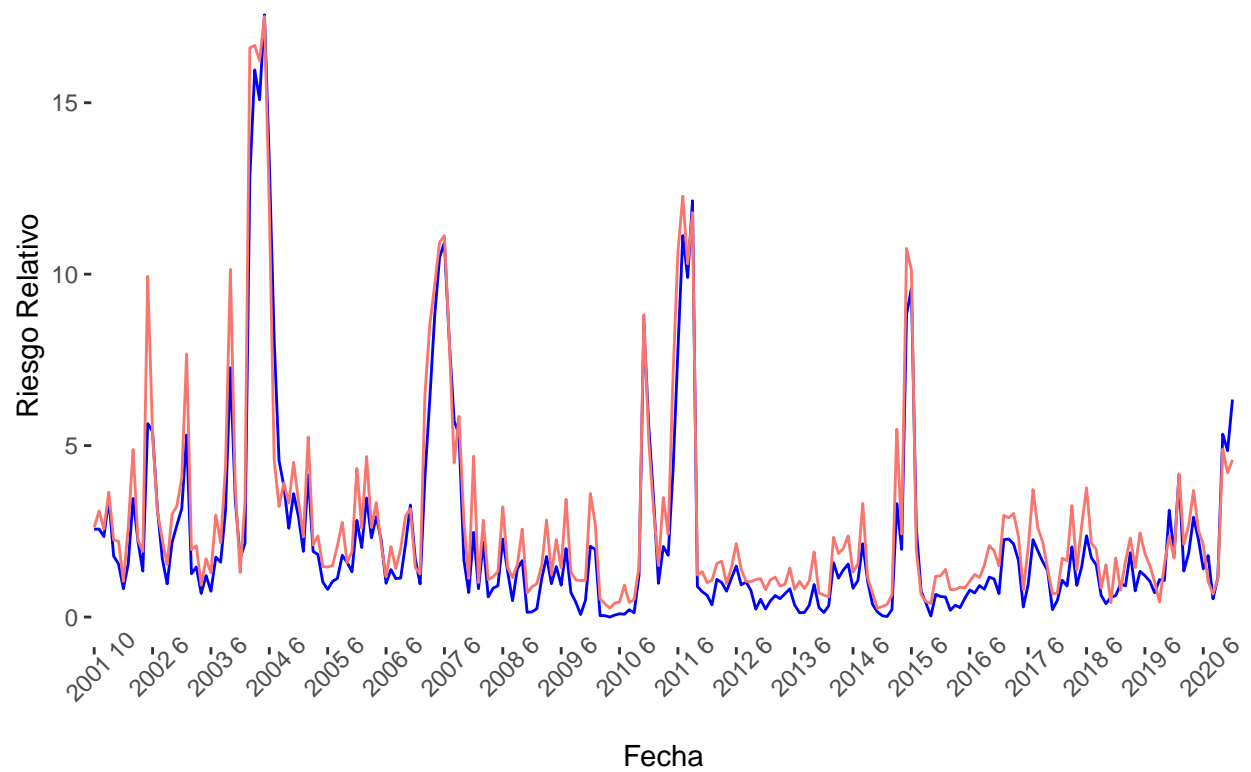
```
##  
## [[13]]
```

Valores aproximados de training del cantón Liberia



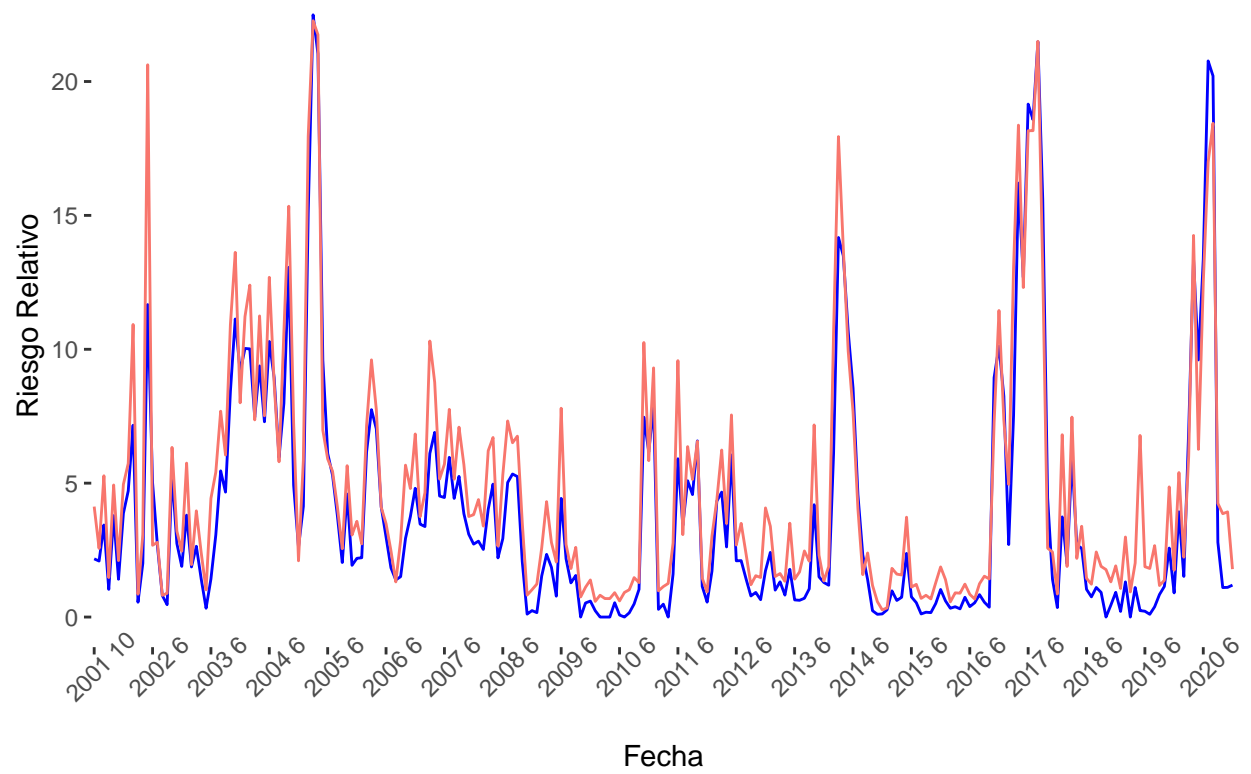
[[14]]

Valores aproximados de training del cantón Limon



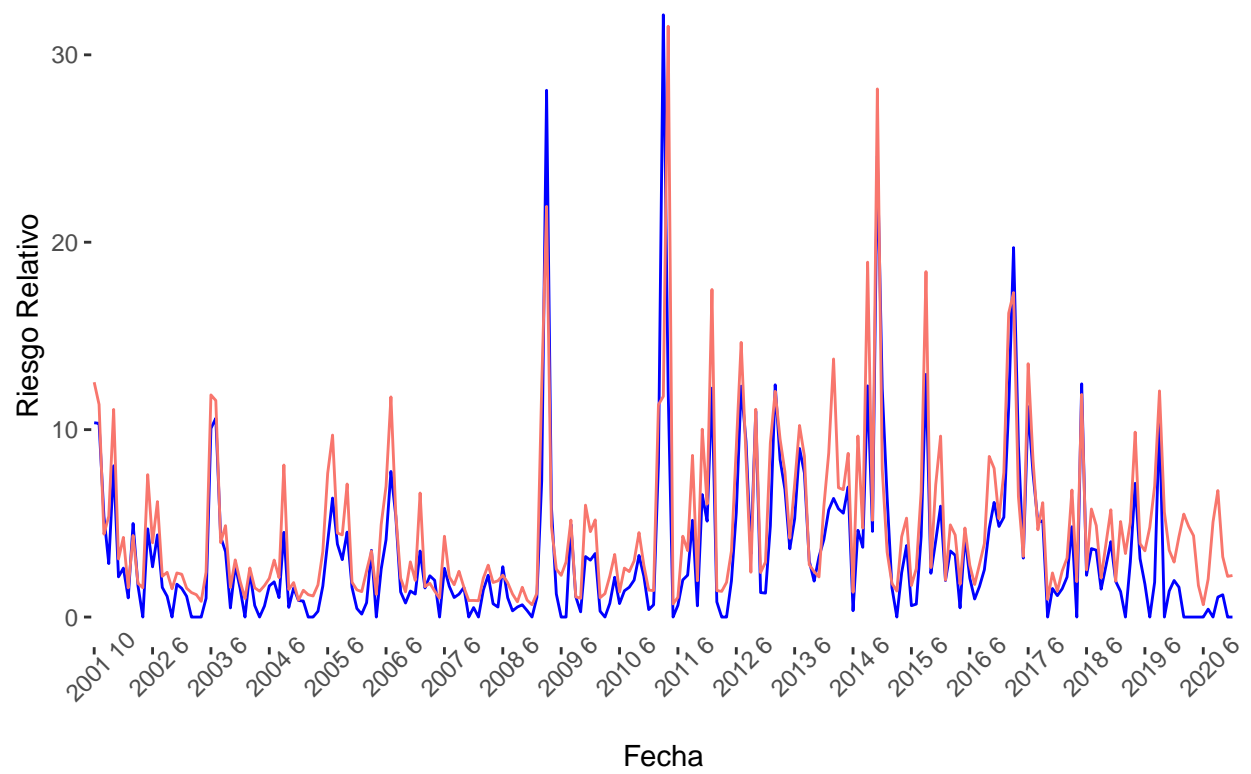
```
##  
## [[15]]
```

Valores aproximados de training del cantón Matina



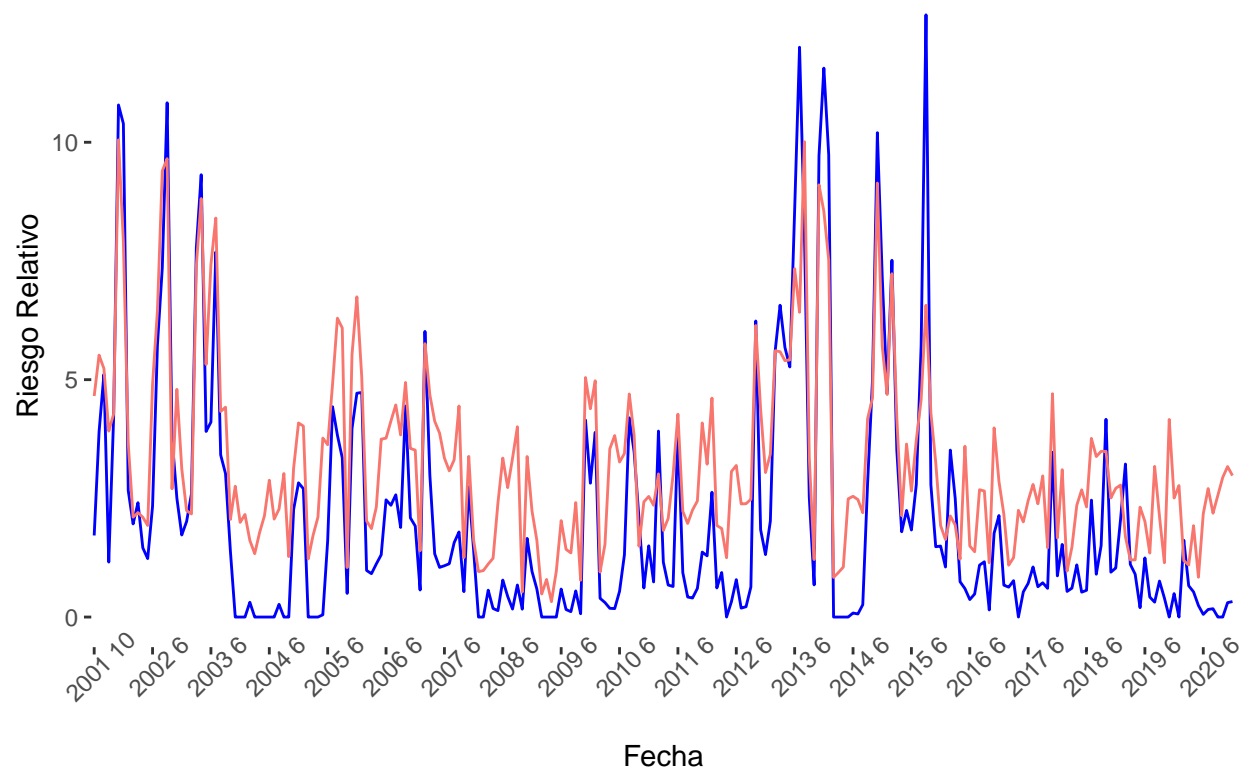
```
##  
## [[16]]
```

Valores aproximados de training del cantón Montes de Oro



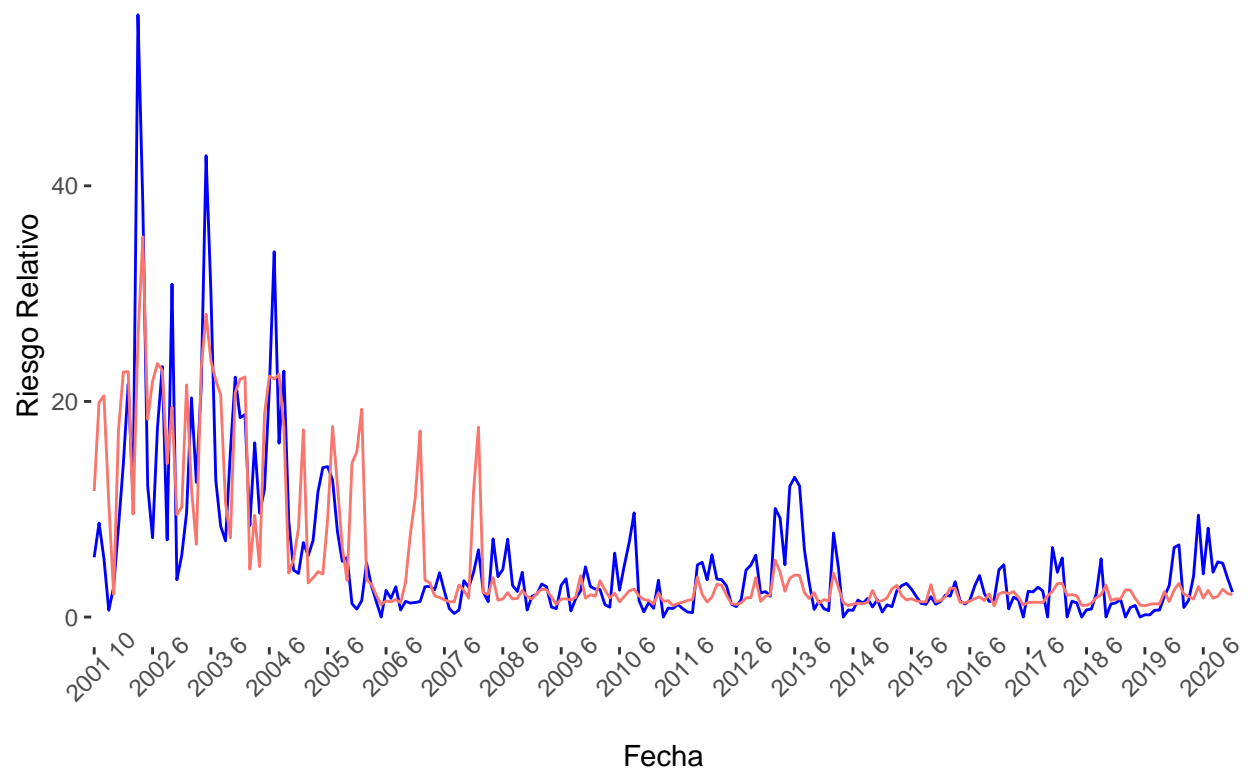
```
##  
## [[17]]
```


Valores aproximados de training del cantón Nicoya



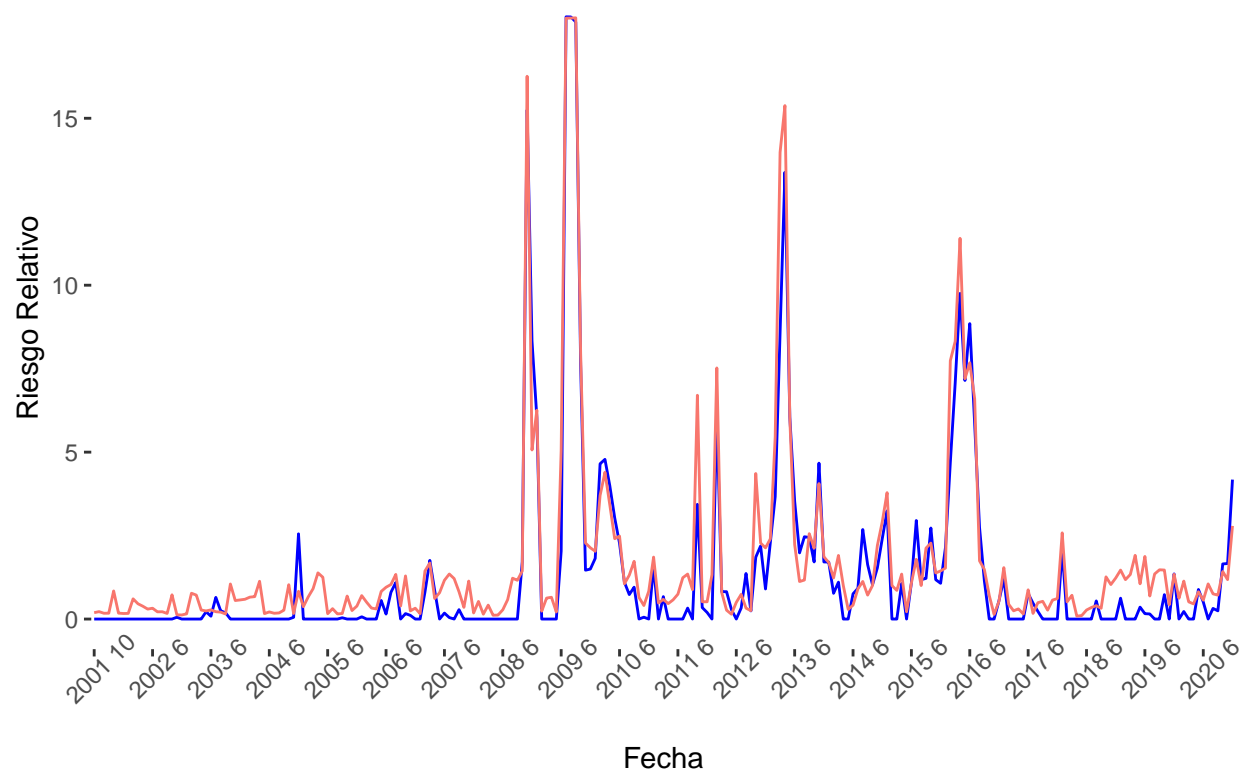
```
##  
## [[18]]
```

Valores aproximados de training del cantón Orotina



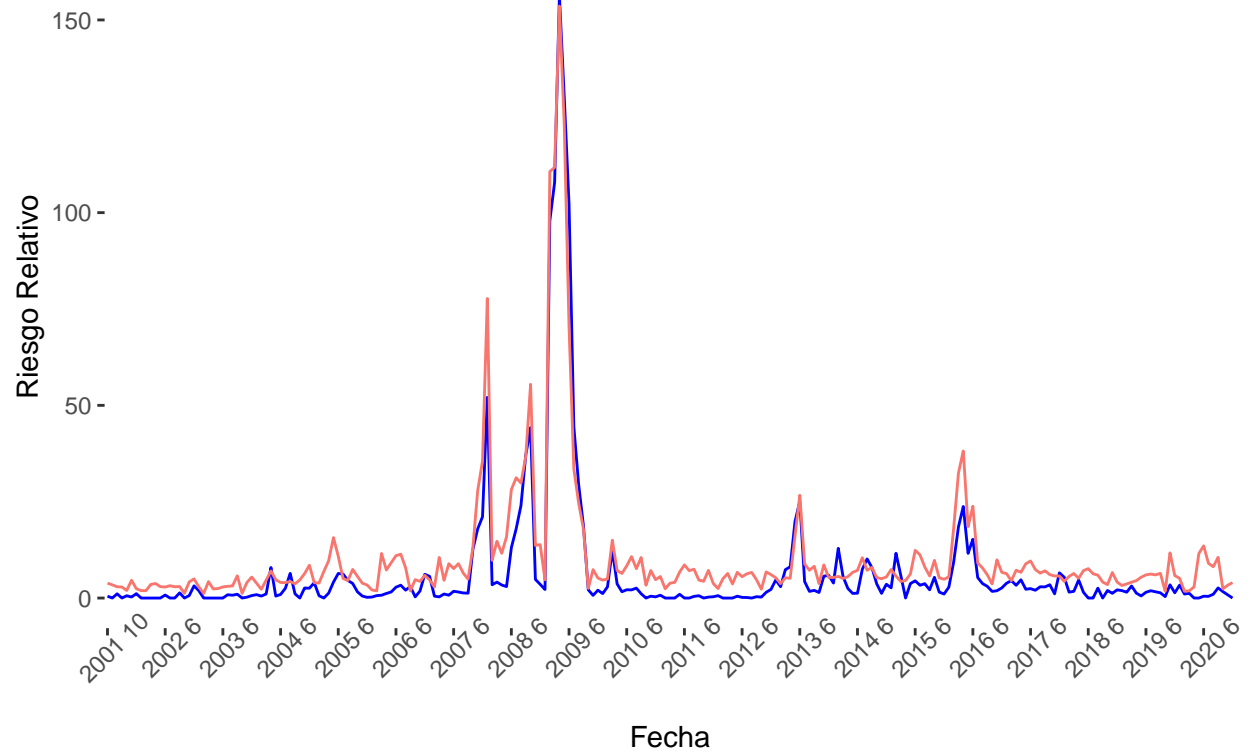
```
##  
## [[19]]
```

Valores aproximados de training del cantón Osa



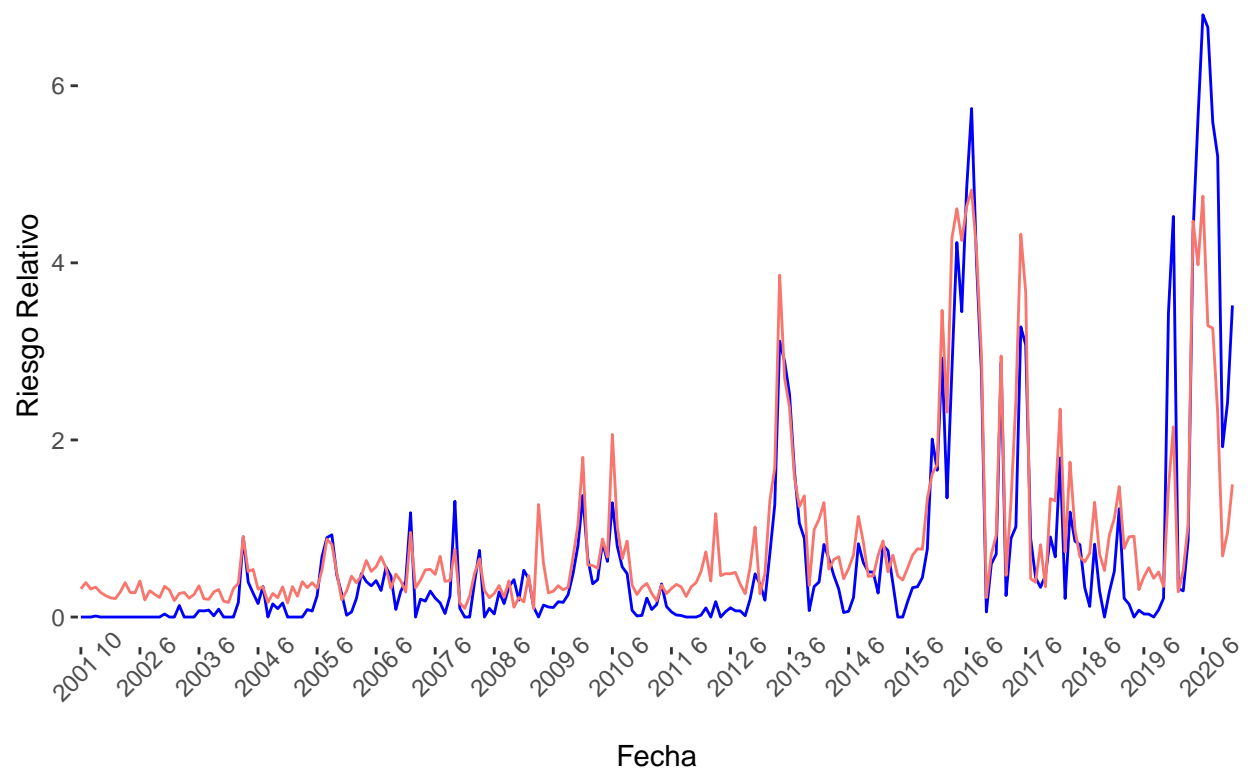
[[20]]

Valores aproximados de training del cantón Parrita



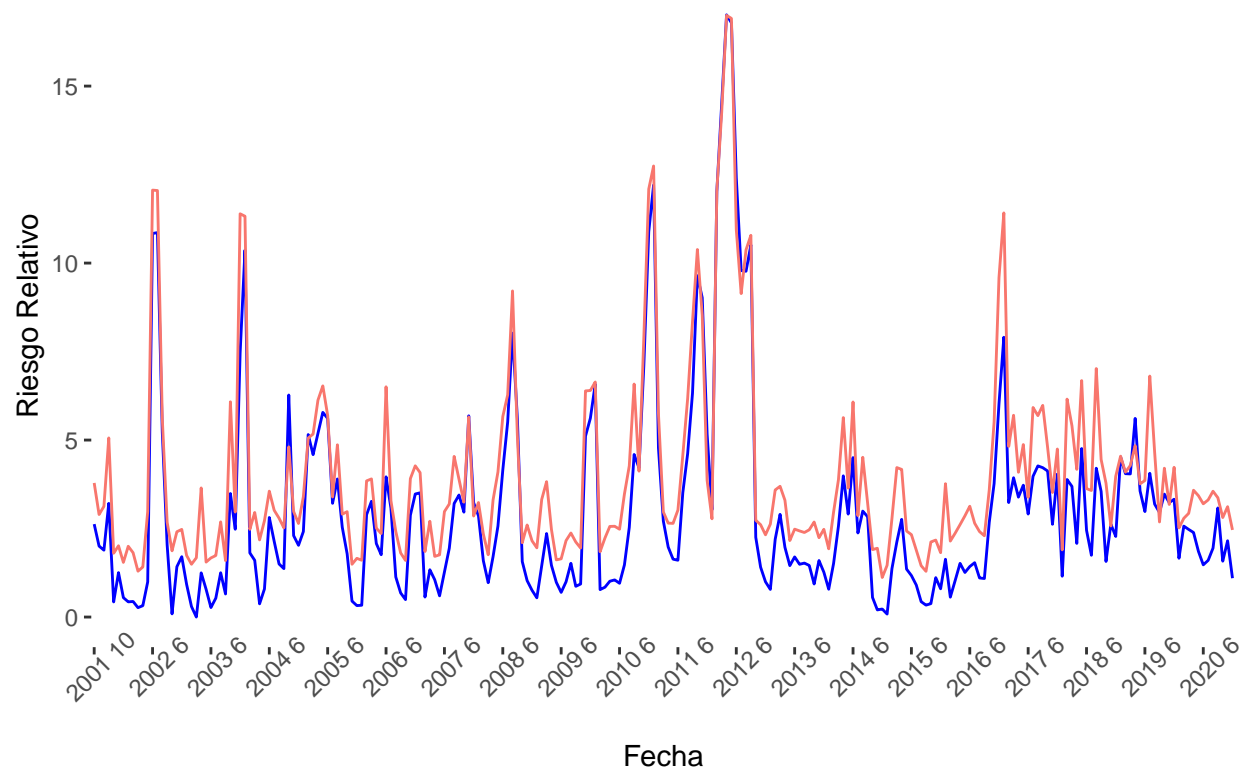
```
##  
## [[21]]
```

Valores aproximados de training del cantón Perez Zeledón



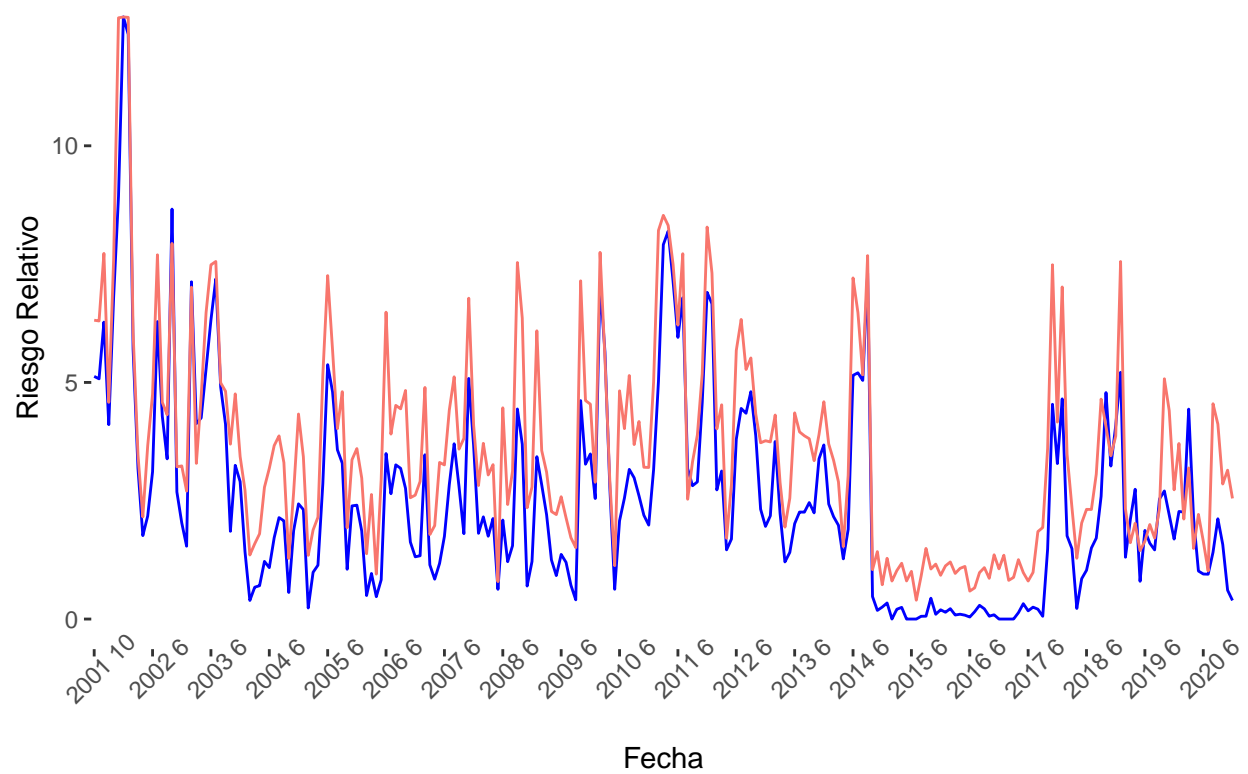
[[22]]

Valores aproximados de training del cantón Pococí



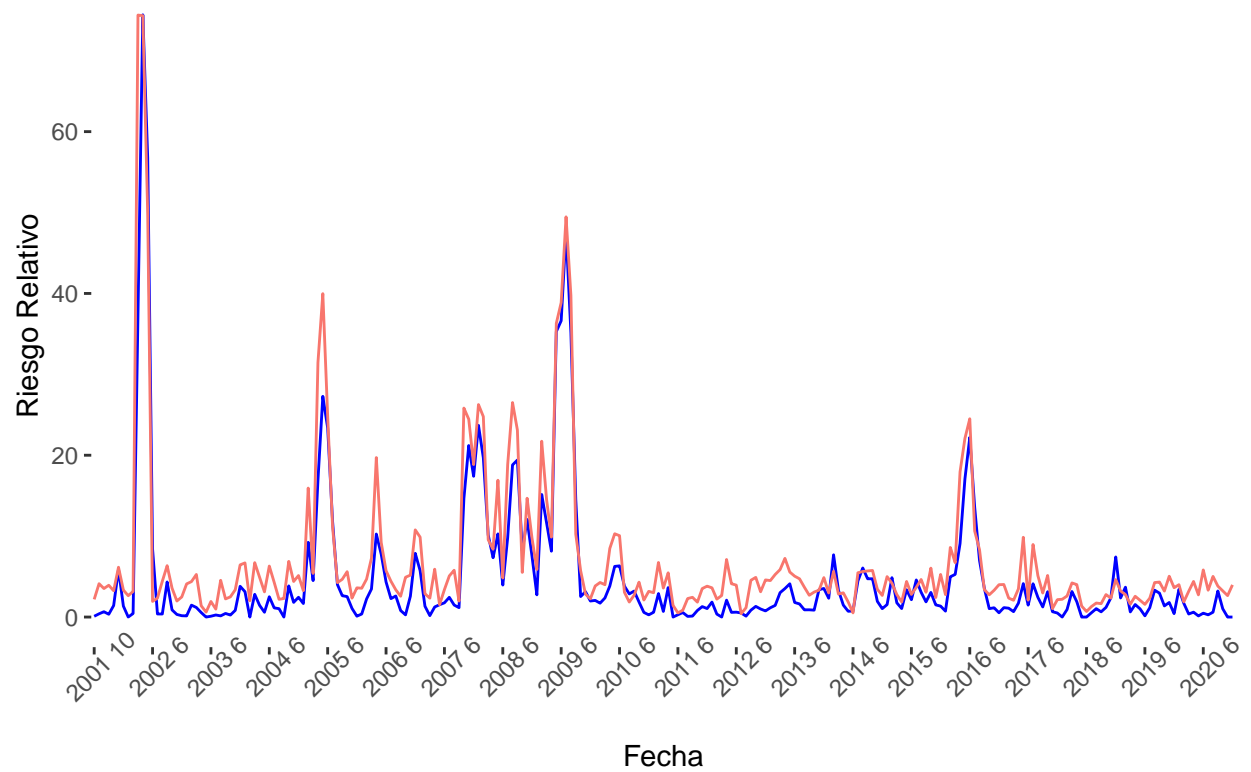
```
##  
## [[23]]
```

Valores aproximados de training del cantón Puntarenas



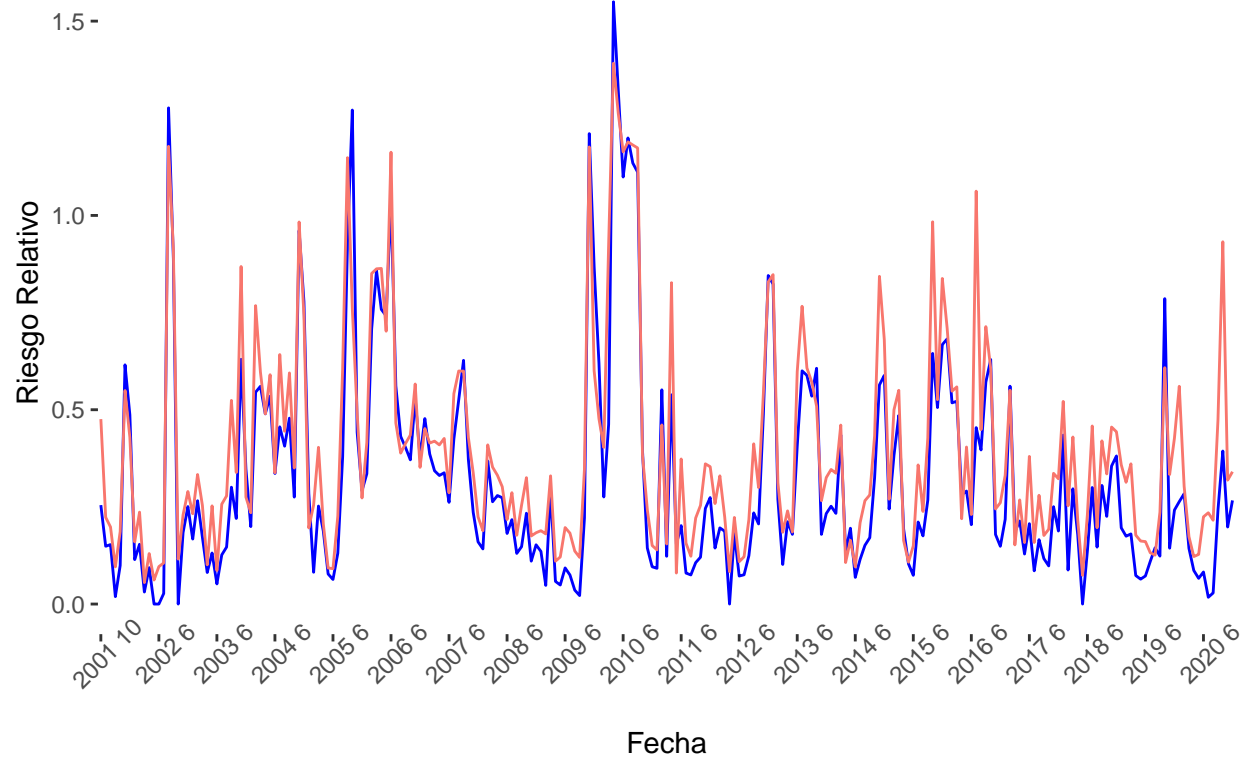
[[24]]

Valores aproximados de training del cantón Quepos



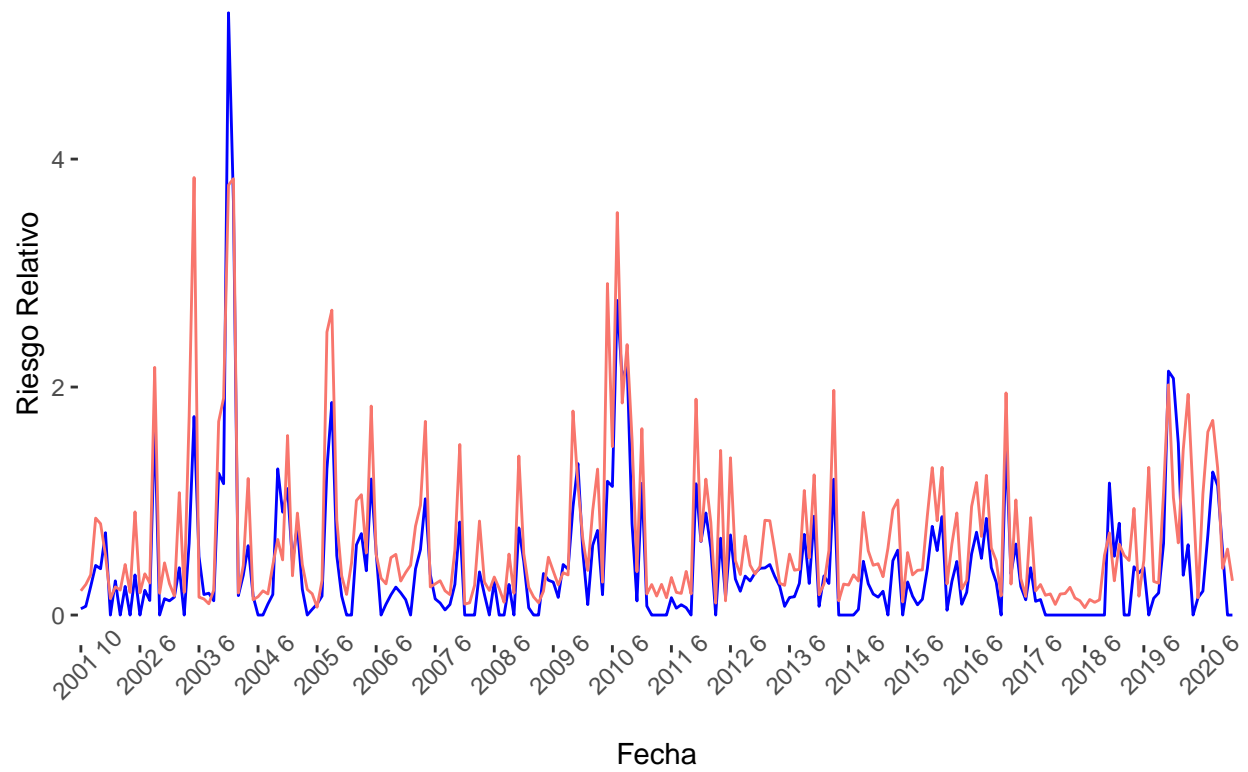
```
##  
## [[25]]
```


Valores aproximados de training del cantón San Jose



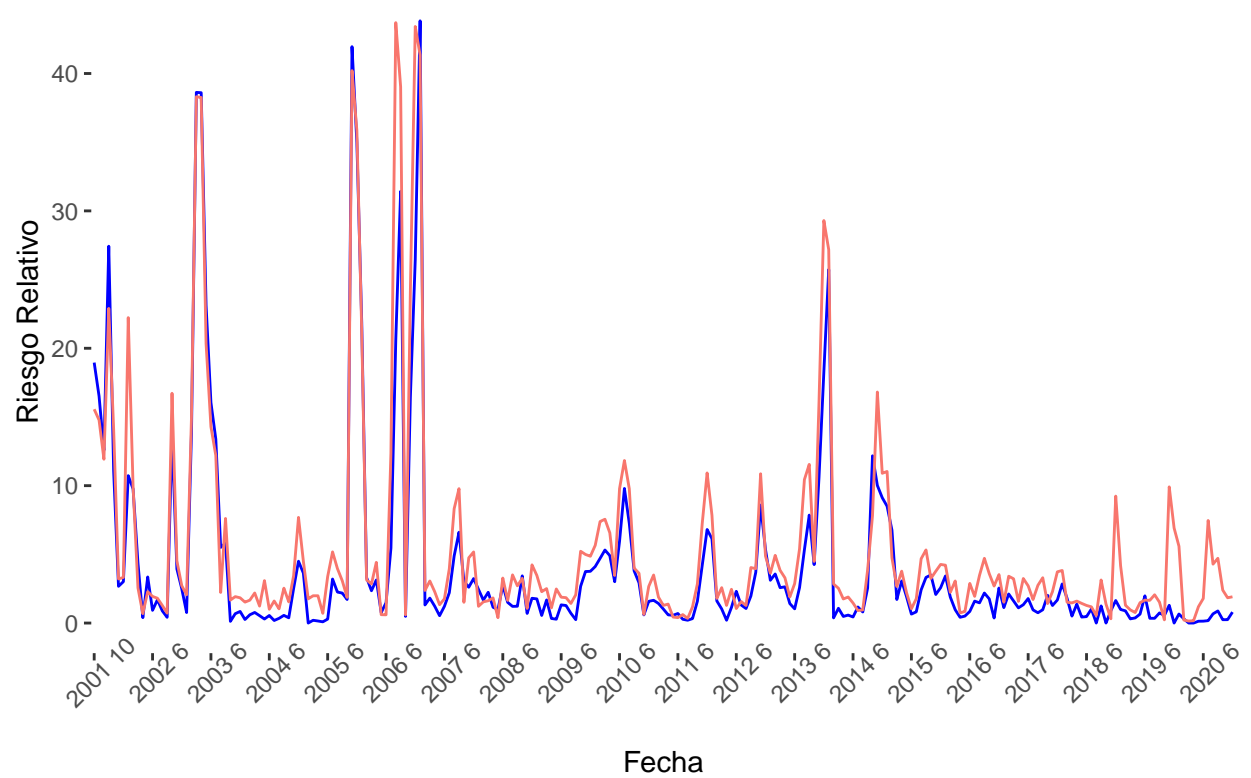
```
##  
## [[26]]
```

Valores aproximados de training del cantón Santa Ana



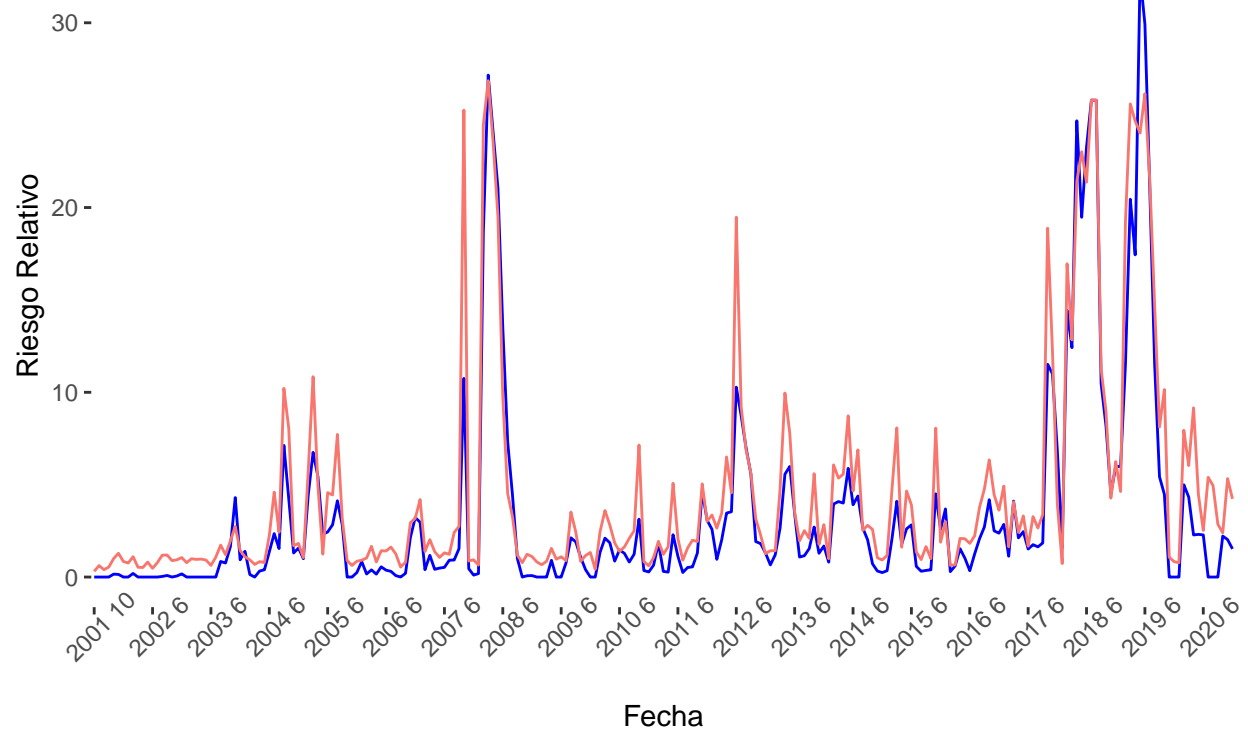
```
##  
## [[27]]
```

Valores aproximados de training del cantón SantaCruz



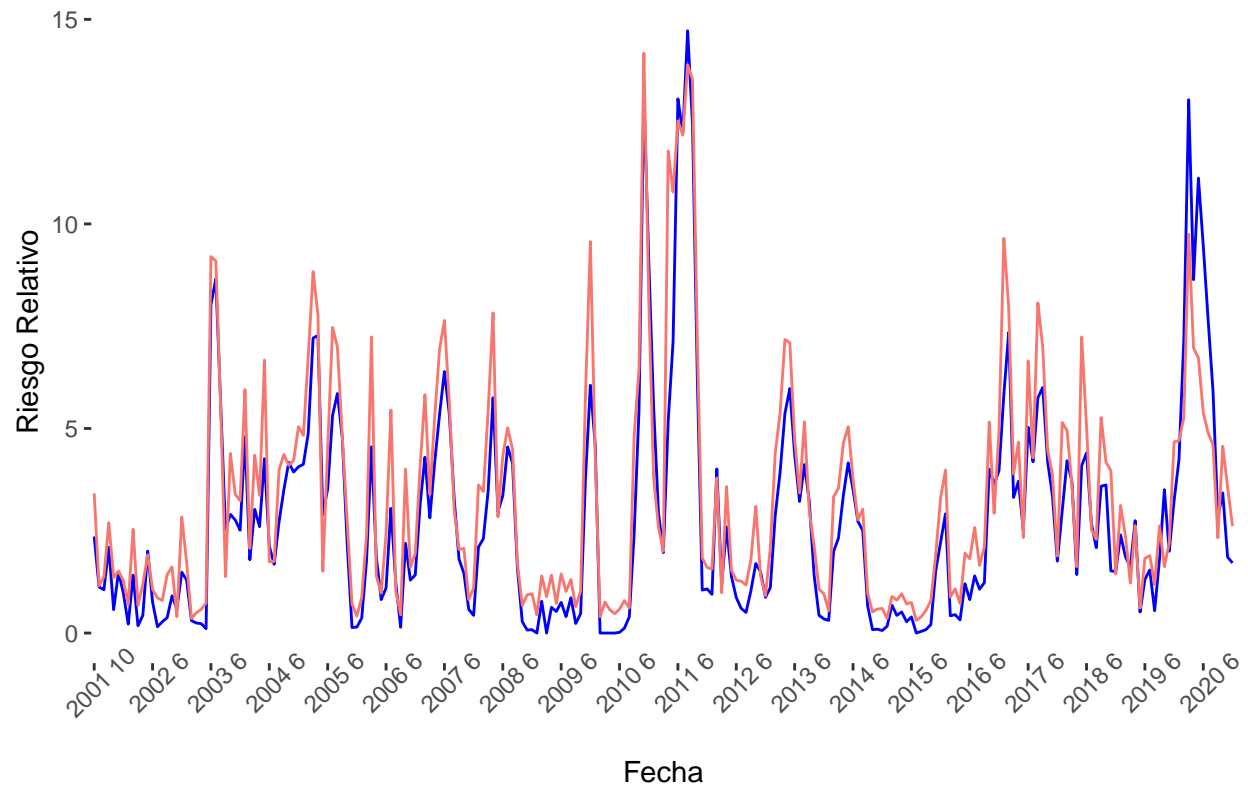
```
##  
## [[28]]
```

Valores aproximados de training del cantón Sarapiquí



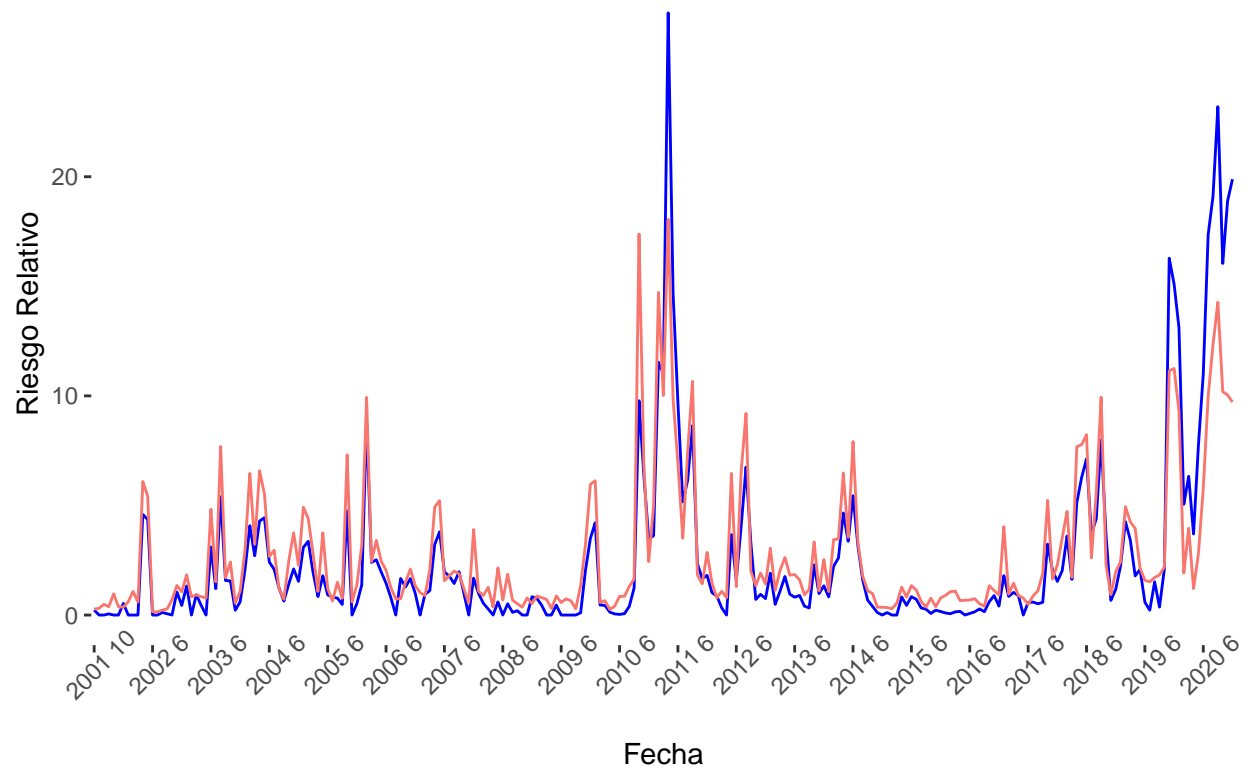
[[29]]

Valores aproximados de training del cantón Siquirres



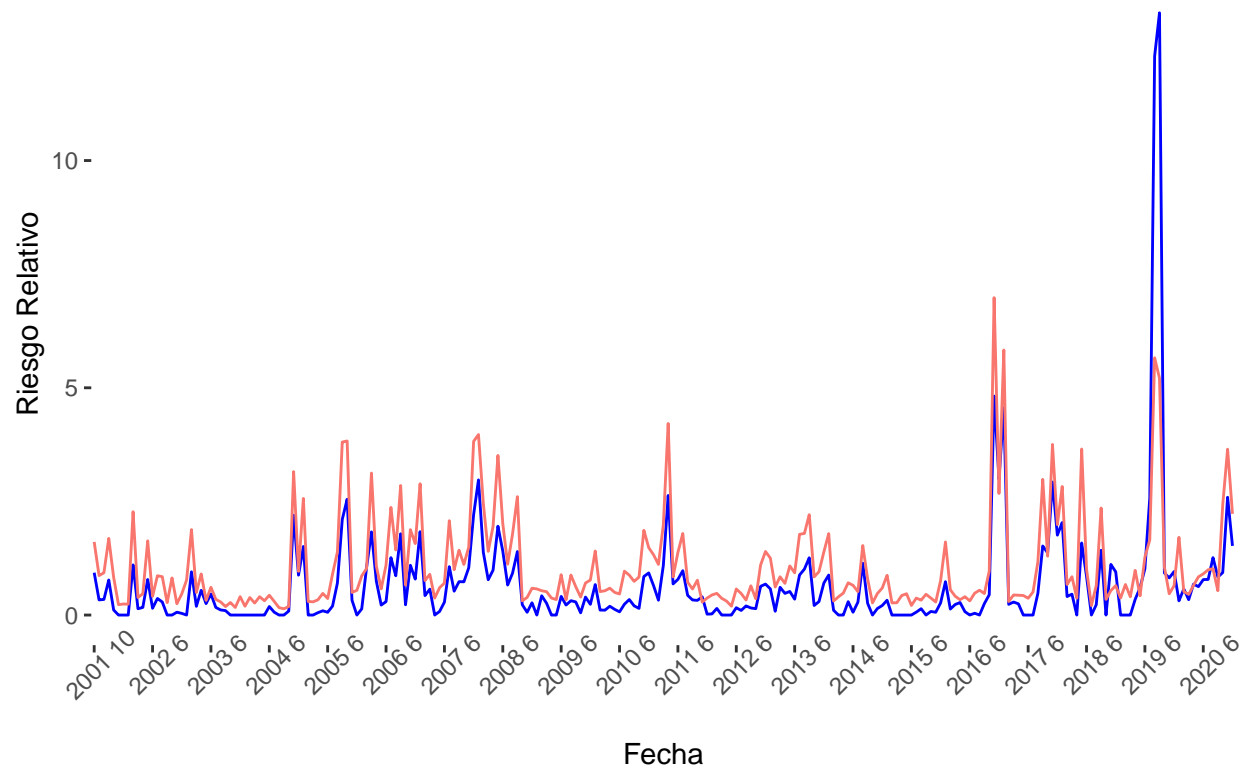
[[30]]

Valores aproximados de training del cantón Talamanca



```
##  
## [[31]]
```

Valores aproximados de training del cantón Turrialba



```
##  
## [[32]]
```

Valores aproximados de training del cantón Upala

