

Chirps-validation.R

Check point
1 : df-pol-chirps.RData

§1. Carga los datos

- Pop_col pcp-col.xlsx
- Chirps_df chirps-ts.parquet

§2. Carga datos elevación de datos abiertos

- SRTM_30 SRTM_30.col1.tif

§3. Convierte a objeto sf (Simple features) df with geom col.

- Chirps_point \leftarrow st_as_sf(chirps_df,...)
- Sttns.points \leftarrow st_as_sf(pcp_col,...)

§4. Extrae la altitud de los puntos de las sttns

- SRTM_30.sttns \leftarrow extract(SRTM_30, sttns.points)

El nuevo archivo de estaciones tiene la col.
altitud, chequear dif. respecto al SRTM

Bias	Q1	Q2	Q3	Nan
PBIas	-4.1	4.3	24.6	~7.7% ± 60 y PBIas ± 25
	-4.4%	9.8%	2.6%	Sobreestimación (varias series)

§5. Función para extraer los pxls CHIRPS cercanos

- sttns.points \$r = 0.05 (en concordancia con la resolución espacial de chirps)

- pol_ \leftarrow function(x, r){...}

- pol_list = list()

for (i in 1:nrow(sttns.points)) {

pol_list[[i]] = pol_(sttns.points[i],
sttns.points\$r[i])

5.1 Agrega ID de la sttns

pol_list[[i]]\$ID =

sttns.points\$CodeEstacion[i]

Cada fila del objeto sttns.points es una estación.mes,
la función pol_:

- Crea un polígono alrededor de la coord. \$geometry
- Convierte a un objeto sf y hace joins espaciales st_join
- Calcula distancias st_distance y ordena por distancia
- selecciona el 1er elemento y filtra el objeto Chirps.point

Ajustes

§5

51 col \$ID a c/elemento de pol_list

- [DEPRECATED]
- + Col: \$nrow= 2 pol_list
 - + Vector ID_px comparando \$nrow_ y \$n. chirps en c.e. de pol_list
 - + Cambiar la col: \$r para las estaciones de San Andrés
 - + aplicar nuevamente la función pol_ a los elementos de ID_px

Actualizar

Sttns-IDEAM.R

§1. Concatena los csv

- df_sttns /TS-climatic/ideam

87330 x 3

Obs: sttns-mes

§2. join con diccionario de sttns (CNE-IDEAM)

- CNE-IDEAM CNE-IDEAM.xlsx

4509 x 20

sttns

§3. Aplica criterios de sel de sttns (completitud > 90%)

- df_sttns2 \leftarrow pivot_wider (dplyr::df1, CNE-IDEAM))

501 012 x 20

Sttns_pcp.col-train.parquet

df de 1009 filas x 486 cols, filas son sttns filtradas por criterio
de completitud con excepción de 8 de los del oriente con baja
densidad de sttns x Sup km^2

71 sttns ~7%

y 480 cols meses de
1981-01 a 2020-12 y 6 cols: Cod.est, dpto, n.zona, lat, lon

Seleccionar el max con una función
no permite mantener la col \$geometry
para filtrar en el obj Chirps.point

Checkpoint 2 :

RData

§6. join con stns.points

```
• pol2_list = list()
  for (i in 1:length(stns.points)) {
    # join para traer las cols de stns (obs)
    pol2_list[[i]] = left_join(Pol2.list[[i]], stns.points)
```

§7. Pivotea cols a filas

```
for (i in 1:length(Pol2.list)) {
  # Selecciona cols, pivotea cols a filas y join para
  # Conformar un df con cols ID, date, CHIRPS y Stns.
  pol2_list[[i]] = left_join(
    # geometry-x: CHIRPS
    # geometry-y: Stns
    # geometry: stns
    # geometry.x: CHIRPS
    # geometry.y: stns
    # geometry: stns)
```

CHIRPS hasta 2022, mientras stns (train) hasta 2020-12

TS original: stns

```
# obtener el orden del modelo AR
# el orden será el parámetro de la prueba ADF
# diff para convertir TS a estacionaria
# y poder ajustar "correctamente" modelo AR
# Sin diff, dado que se parte
# de lo que se quiere fijar;
# TS climáticas son no-estacionarias
# debido a la comp. estacional.
```

for (i in 1:length(Pol2.list)) {

```
m_order[i] <- tryCatch(
  {ar.obj <- ar(diff(Pol2.list[[i]])$stns),
   method = "mle")
  ar.obj$order},
  error = f[i] ~ i
)
k = k + 1
}
```

m_order[k]

ar(Pol2.list[[i]]\$stns,
method = "mle")

tal vez por
TS no-estacionaria → se podría
diff doble

gráfico stns con Error al ajustar modelo AR

eTSA::adf.test # tipo regresión: type 1: No drift, No trend
for (i in 1:length(Pol2.list)) {

l_ADF = eTSA::adf.test(Pol2.list[[i]]\$stns,
 flag = m_order[i], output = FALSE)

df_ADF\$log_t1[i] = which(l_ADF\$type1\$p.value == max(l_ADF\$type1\$p.value))

df_ADF\$p_value_t1[i] = max(l_ADF\$type1\$p.value))

Urca::Ur.df → si por qué es de interés? es el valor crítico? checar documentación

diff estacional

Calcula el periodo del ciclo estacional a partir de la f. densidad espectral

df_pgram = data.frame(freq = rep(0, len), period = rep(0, len))

for (i in 1:length(Pol2.list)) {

pgram = spectrum(Pol2.list[[i]]\$stns, log = "no")

f_max = which.max(Periodograma\$spec)

df_pgram\$p.freq[i] = pgram\$freq[f_max]

df_pgram\$period[i] = round(1/pgram\$freq[f_max])

gráfico muestra stns con periodo 4 y 6

Tedundante
Calculo más
2 veces.

desestacionaliza la TS stns

```
d_stns = list()
for (i in 1:length(Pol2.list)) {
  d_stns[[i]] = diff(Pol2.list[[i]]$stns,
    lag = df.pgram$period[[i]]) }
```

Crea Col d_stns con serie desestacionalizada

```
for (i in 1:length(Pol2.list)) {
  Pol2.list[[i]]$d_stns = c(rep(NA, df.pgram$period[[i]]),
    diff(Pol2.list[[i]]$stns,
      lag = df.pgram$period[[i]])) }
```

```
for (i in 1:length(Pol2.list)) {
```

M_order[3][k] <- tryCatch(

far.obj <- ar (diff(Pol2.list[[i]]\$d_stns),
 method = 'mlc')

ar.obj\$order},
 error = f(e){-1}

)
k = k+1

M_order[4][k]

ar (Pol2.list[[i]]\$d_stns,
 method = 'mlc')

arroja error
Por los NA's?
sólo para obtener el orden
crea otra lista d_stns?

eTSA::adf.test

urca :: ur.df

l. 429 - 628

f.get.AR_order(TS, len, diff) { ... }

return(m_order)

m_order # <- f.get.AR_order(...)

l.b.

descomposición STL base

Obtiene Orden del modelo AR

eTSA::adf.test

urca :: ur.df

descomposición STL feast

Obtiene Orden del modelo AR

eTSA::adf.test

urca :: ur.df

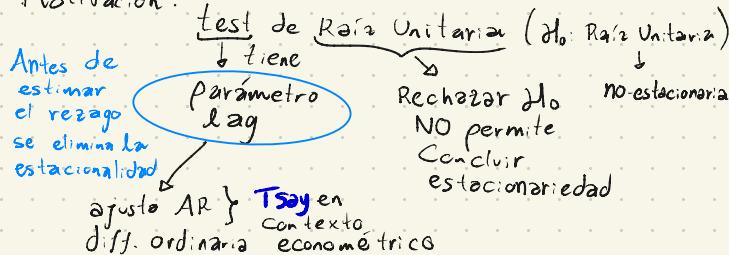
○ Chirps_Validation.R

SB. Evalua la estacionariedad de las series

Si hay una estacionalidad determinística el proceso puede ser no-estacionario
Pero la prueba ADF

Concluye estacionario

Motivación:

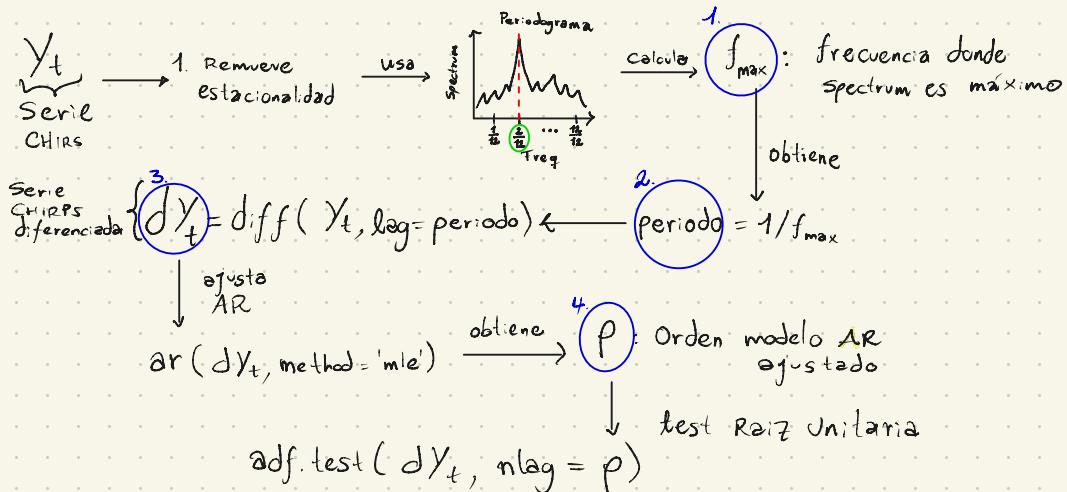


estructuras autocorrelación temporal → No cumple supuesto independencia

↓
No cumple supuesto i.i.d.

↓
Puede llevar a relaciones espurias

Uso correcto corr.



Rechaza H_0 :
Si $P\text{-value} < 0.05$

argumentar
Si afirmar que es estacionaria
a) no tener Raiz Unitaria
i) Cuál es la ventaja?
ii) mantiene propiedades estadísticas?
iii) no requiere diff. ordinaria para ajustar modelo ARIMA

$f(dY_t, dZ_t)$ →

las métricas se aprox. más al valor real del parámetro y NO Sobreestima/Subestima

Statistical methods for climate scientist

5.5 Intro. to stochastic Process

5.2 Stochastic Processes. (Pg. 102)

most climate TS are nonstationary, because they depend on calendar day and time. Winter and summer temp have different dist.

Unit Root Test in Time Series

5.6 Dickey-Fuller and related Tests

6.3 DF tests for a Unit Root (Pg. 205)

o Chirps-validation.R

§3 diff estacional

Motivación: Al aplicar la prueba ADF a la TS original, luego de ajustar un modelo AR (Tsay, 2014) y obtener el orden de este modelo para usarlo como arg $\log d$ dado que la prueba ADF es sensible al orden, se concluye que No rechaza H_0 , luego hay raíz unitaria por tanto la TS es no-estacionaria, lo que lleva a que las métricas no se aprox. al valor real y puedan ser subestimadas/ Sobreestimadas debido a las relaciones espurias (uso correcto de la cor...)

Uno de los motivos por los que la serie es no-estacionaria es debido a la componente estacional determinística, por lo cual en el contexto de

In the area of time series called Spectral analysis

trend stationarity

¿Cuál es la relación entre la prueba de raíz Unitaria y la estacionariedad?

yo

Por qué la teoría

Spectral analysis for Stationary time series?
yo Y si es no-estacionaria?

El periodograma puede verse como una herramienta para la detección de posibles ciclos (ocultos) deterministas en una serie temporal.

una forma de eliminar la comp. estacional es usando diferencias estacionales $(1 - B^s)^D$, donde s es el periodo del ciclo estacional.

Para encontrar s y proceder a eliminar la comp. estacional;

1. Se determina el valor de la freq. donde se maximiza el periodograma de la serie, un valor alto de periodograma para una f_j está asociado al Ciclo estacional que se quiere eliminar de la TS.

2. Calcula el periodo $s = 1/f_{\max I(f)}$

Por qué el máx del Periodograma?

yo

Fourier Analysis of Time Series

5.5 The Fast Fourier Transform

base teórica de R stats::specgram