Modelado de temperaturas extremas en Uruguay

I. Alvarez-Castro   J. Cugliari   N. da Silva   S. De Mello
Agustín Estramil   Manuel Hernández   M. Renom

15 de Noviembre de 2019

orden

Intro:

- proyecto fsd (gente, inst, amigos)
- relevancia y objetivos del proyecto

Datos:

- origen y problemas en los datos
- definicion olas (punto, gralizada), percentiles
- Viz (estatica)

Modelos:

- GEV
- DLM: modelo
- DLM: estimacion

Resultados: mostrar shiny app

1. Introducción

Introducción

Antecedentes y conformación del equipo

- 2015 Jornadas de estadística en la Paloma.
- 2016 Presentación del proyecto I+D de CSIC (no financiado).
- 2017 Fondo sectorial de investigación a partir de datos.($$:-))

Introducción

Instituciones involucradas

- IESTA - Facultad de Ciencias Económicas.
- DCA (DCAOF) - Facultad de Ciencias.
- INUMET

Introducción

Relevancia del problema

- Problema regional de series con datos faltantes.
- Relevancia del estudio de extremos.
- ...

Introducción

Objetivo general

Modelización estadística de las series diarias de temperatura máxima y mínima en varias estaciones meteorológicas en Uruguay.

En base a esos modelos:

- Reconstruir series diarias.
- Analizar eventos de olas de frío y calor.

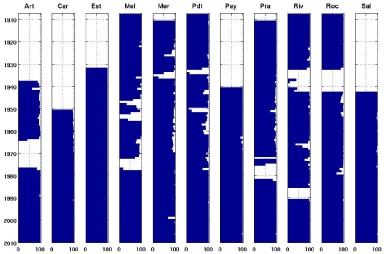
Modelado de temperaturas extremas en Uruguay

Objetivos específicos

- Imputación de datos faltantes.
- Visualización de series diarias.
- Definición y carateterización de olas de extremos.
- Métodos para comparar y evaluar series imputadas.

2. Datos de Temperatura máxima y mínima

Origen y problemas de los datos



Período de análisis: 1950 - 2014

Control de calidad

- Control de calidad RClimdex - umbral 3 STD + Variabilidad interdiaria de Tx y Tn.
- Programa para visualizar los errores detectados por RClimDex.
- Se confeccionó una lista con los datos de todas las estaciones para revisar en el INUMET.

Olas de extremos

Ola de frío: al menos 3 días seguidos con temperaturas mínimas y máximas inferiores a los respectivos percentiles 10 esperados para tales días.

Sucesión de días $t_1, \ldots, t_k$ constituyen una ola de frío de largo k si:

$$\begin{cases} y_{t_i}^n < p_{10_i}^n \\ y_{t_i}^x < p_{10_i}^x \end{cases} \quad \text{para } i = 1, \ldots, k \text{ donde } k \geq 3$$

Se debe verificar al menos 5 de las 6 condiciones anteriores.

Calcular percentiles

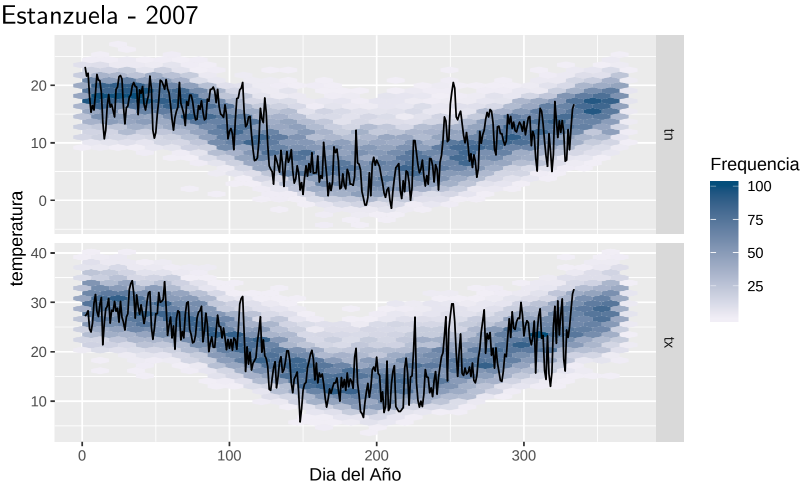
Percentiles para cada dia del año:

$$p^n_{10_t} := \text{ínf}\{y : p(Y^n_t \leq y) \geq 0{,}1\}$$
$$p^x_{10_t} := \text{ínf}\{y : p(Y^x_t \leq y) \geq 0{,}1\}$$

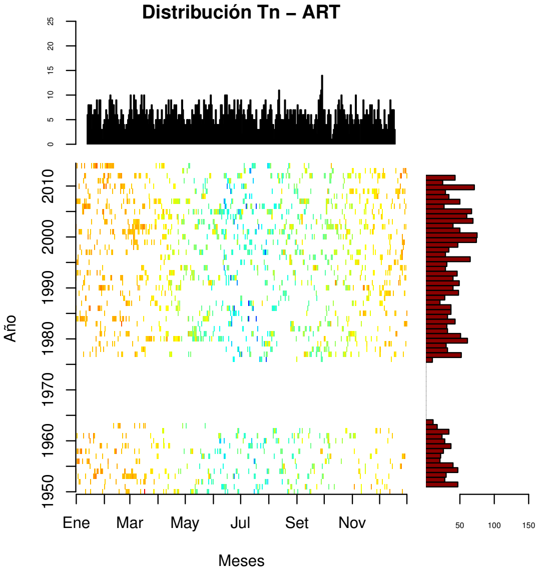
Muestral

- y_{at} es la temperatura de día t en año a
- cada diat t, calcular cuantiles de $\{y_{1t}, \ldots, y_{At}\}$
- obtener $p^n_{10_t}$ como media móvil de ventana 31 (en días)

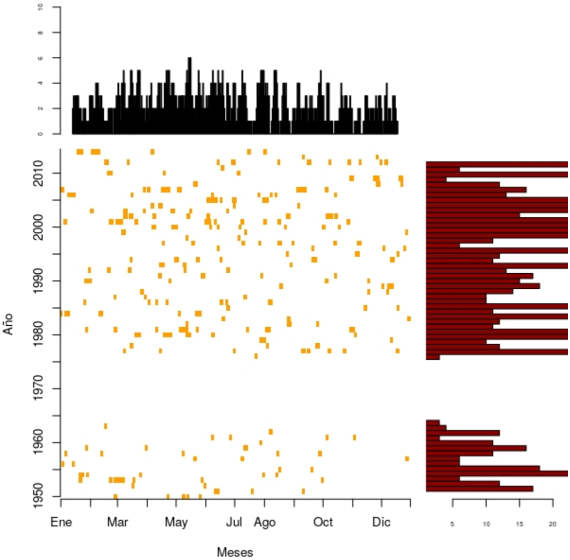
Visualizar olas de temperatura



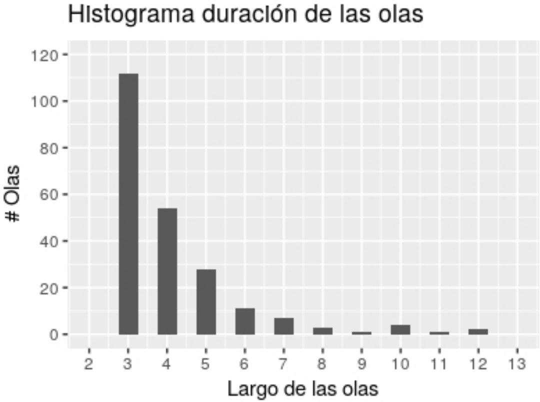
Distribución de datos superando el P90 - ARTIGAS



Olas de calor - ARTIGAS



Olas de calor - ARTIGAS



3. Modelos estadísticos

Modelos estadísticos

- y_{ti}^{n}: temperatura mínima el día t en locación i
- y_{ti}^{x}: temperatura máxima el día t en locación i

Imputar las series diarias:

- Modelos aditivos basados en GEV
- Modelos Bayesianos diámicos
- otros: Regresión + ACP, GAM + estructura dependencia, etc

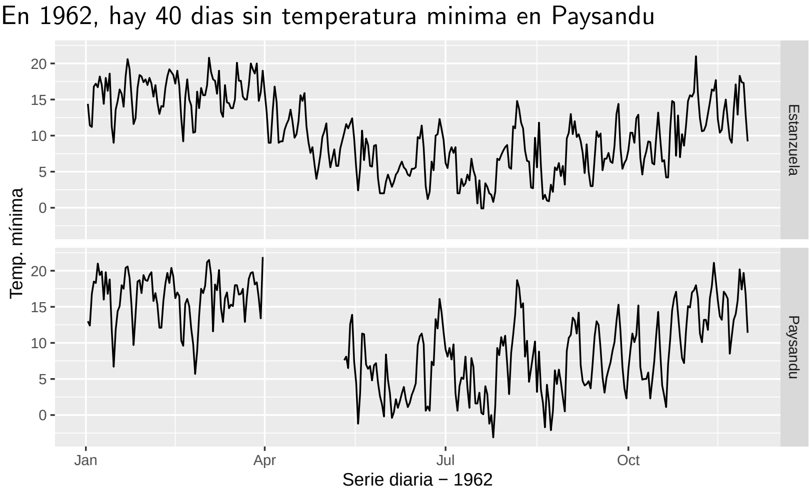
Modelos estadísticos

- y_{ti}^{n}: temperatura mínima el día t en locación i
- y_{ti}^{x}: temperatura máxima el día t en locación i

Imputar las series diarias:

- Modelos aditivos basados en GEV
- Modelos Bayesianos diámicos
- otros: Regresión + ACP, GAM + estructura dependencia, etc

Ejemplo Imputación



Modelos aditivos basados en GEV

$$
\begin{aligned}
G(y^x_{ti}) &= \exp\left\{-\left[1+\xi\left(\frac{y^x_{ti}-\mu_{it}}{\sigma_i}\right)\right]_+^{-1/\xi}\right\} \\
\mu_{it} &= s(t) + s(mei) + cl_i
\end{aligned}
$$

- un parámetro de escala por estación, σ_i
- un parámetro de forma, ξ
- imputamos con la locación: $\tilde{y}^x_{ti} = \hat{\mu}_{it}$

Modelo dinámico lineal

$$
\begin{aligned}
Y_t &= F_t\theta_t + v_t & \quad v_t &\sim \mathcal{N}(0, V_t) \\
\theta_t &= G_t\theta_{t-1} + w_t & \quad w_t &\sim \mathcal{N}(0, W_t)
\end{aligned}
$$

$$
\begin{tikzcd}
Y_{t-1} & Y_t & Y_{t+1} \\
\theta_{t-1} \arrow[u] \arrow[r] & \theta_t \arrow[u] \arrow[r] & \theta_{t+1} \arrow[u]
\end{tikzcd}
$$

DLM para temperatura

- temperatura observada: real- ruido
- θ_t entendido como el valor no observable de la temperatura **sin** ruido
- dependencia en θ_t: de otras estaciones y de si misma
- NO toma en cuenta caracter extremo

$$
\begin{array}{ll}
y_t^n = \theta_t + v_t & v_t \sim \mathcal{N}(0, \sigma^2 I_{11}) \\
\theta_t = \theta_{t-1} + w_t & w_t \sim \mathcal{N}(0, W)
\end{array}
$$

Estimación proceso de estados

Inferencia Bayesiana:

$$p(\text{parámetro}|\text{datos}) \propto p(\text{datos}|\text{parámetro}) \times p(\text{parámetro})$$

Inferencia probabilística: concluciones se basan en una distribución de cantidades de interés *condicional* en los datos.

Dos posteriores relevantes:

- Filtro: $p(\theta_t|y_{1:t})$
- Suavizado: $p(\theta_t|y_{1:T})$

Filtro de Kalman: $p(\theta_t|y_{1:t})$

Estimacion recursiva, en base a $p(\theta_{t-1}|y_{1:(t-1)})$:

- Previa:

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

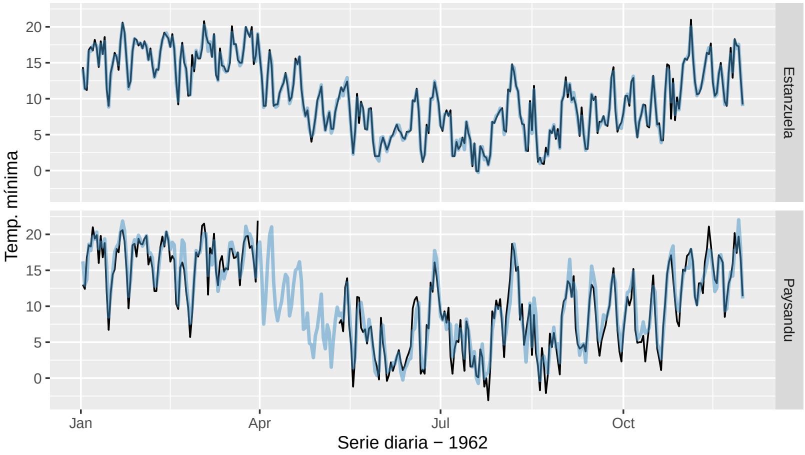
- Posterior Filtrada:

$$p(\theta_t|y_{1:t}) = \frac{p(y_t|\theta_t)p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t})} \propto p(y_t|\theta_t)p(\theta_t|y_{1:t-1})$$

- Previa para estado inicial: $\theta_0 \sim N(m_0, C_0)$
- Soluciones cerradas en el caso *gaussiano lineal*

Ejemplo Imputación

Los metodos para obtener posteriores de θ_t no se ven afectados



Resultados

Aplicacion shiny

Siguientes pasos

- evaluar efecto de la imputacion en las olas estimadas
- caracterizar olas

4. Muchas Gracias !!