

Datos Direccionales.

Proyecto Fin de Máster.

Leyenda Rodríguez, María

9 de junio de 2011

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- En diversos campos surgen problemas estadísticos dónde los datos son dados mediante medidas angulares dando la orientación o ángulos en el plano (datos circulares) o en el espacio (datos esféricos).
- Datos circulares es el caso más simple de esta categoría de datos llamada **datos direccionales**, dónde la única respuesta no es escalar, sino que, es angular o direccional.

Notación

- Las direcciones en el plano pueden ser observadas como vectores en el plano o como puntos en el círculo unidad.
- También se pueden observar, como ángulos o como números complejos.

Luego cada punto x en el círculo unidad puede ser representado por un ángulo θ o por un número complejo unitario z .

$$x = (\cos\theta, \sin\theta)$$
$$z = e^{i\theta} = \cos\theta + i\sin\theta$$

Medidas de localización

- $\bar{\theta}$ es la dirección media de $\theta_1, \dots, \theta_n$. Viene dada por la solución de las ecuaciones

$$\bar{C} = \bar{R} \cos \bar{\theta}$$

$$\bar{S} = \bar{R} \sin \bar{\theta}$$

- dónde $(\bar{C}, \bar{S})$ son las coordenadas cartesianas del centro de masa

$$\bar{C} = \frac{1}{n} \sum_{j=1}^n \cos \theta_j \quad \bar{S} = \frac{1}{n} \sum_{j=1}^n \sin \theta_j$$

- La media de la longitud resultante $\bar{R}$ es dada por

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$$

Medidas de localización

- Notemos que $\bar{\theta}$ no está definida cuando $\bar{R} = 0$
- Cuando $\bar{R} > 0$, $\bar{\theta}$ es dada por

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S}/\bar{C}) & \text{si } \bar{C} \geq 0 \\ \tan^{-1}(\bar{S}/\bar{C}) + \pi & \text{si } \bar{C} < 0 \end{cases}$$

Notemos que en el contexto de estadística circular $\bar{\theta}$ no es la media $(\theta_1 + \dots + \theta_n)/n$

Medidas de concentración y dispersión

- $\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$, entonces si $x_1, \dots, x_n$ son vectores unitarios, $0 \leq \bar{R} \leq 1$.
- Si las direcciones $\theta_1, \dots, \theta_n$ están estrechamente agrupadas luego $\bar{R} = 1$.
- Si $\theta_1, \dots, \theta_n$ están muy dispersas luego $\bar{R}$ será prácticamente 0.

Por tanto, $\bar{R}$ es una medida de concentración del conjunto de datos.

- $\bar{R}$ es invariante bajo rotaciones.
- La longitud resultante $\bar{R}$ es la longitud del vector resultante $x_1 + \dots + x_n$.

Medidas de concentración y dispersión

- La media de longitud resultante $\bar{R}$ es la medida de dispersión más importante.
- A veces emplearemos otras medidas de dispersión de datos circulares, para comparar con datos en línea: la medida más simple es la **varianza circular muestral**

$$V = 1 - \bar{R}$$

- **Desviación circular estándar**

$$v = \sqrt{-2 \log \bar{R}}$$

Distribución circular

- Una distribución circular es una distribución de probabilidad la cual está concentrada en la circunferencia de círculo unidad.
- Las distribuciones circulares son de dos tipos:
 - 1 Discretas- asignan masas de probabilidad solo a un número de direcciones
 - 2 Absolutamente continuas
- Existe la función de densidad $f(\theta)$
 - 1 $f(\theta) \geq 0$
 - 2 $\int_0^{2\pi} f(\theta) d\theta = 1$
 - 3 $f(\theta) = f(\theta + 2\pi)$

Distribución von Mises

Definición

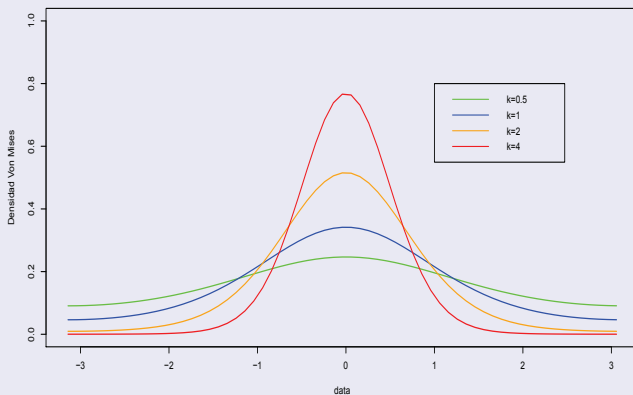
$$g(\theta; \mu, k) = \frac{1}{2\pi I_0(k)} e^{k \cos(\theta - \mu)}$$

- I_0 denota la función de Bessel modificada de primer tipo y orden 0,

$$I_0(k) = \sum_{r=0}^{\infty} \frac{1}{(r!)^2} \left(\frac{k}{2}\right)^{2r}$$

- El parámetro μ es la media de las direcciones
- k es el parámetro de concentración.

Distribución von Mises



Distribución von Mises

Relación $v, \bar{R}, S_n^2$ y parámetro de concentración

- 1 Generamos una muestra $M(\pi, \kappa)$ de tamaño 3000 para distintos valores del parámetro κ (75 valores).
- 2 Para cada una de las muestras calculamos
 - La varianza (1) con respecto a la distancia euclídea (2) y a la distancia definida por la longitud de arco(3),

$$S_n^2(\theta) = \frac{1}{n} \sum_{i=1}^n d^2(\theta_i, \theta) \quad (1)$$

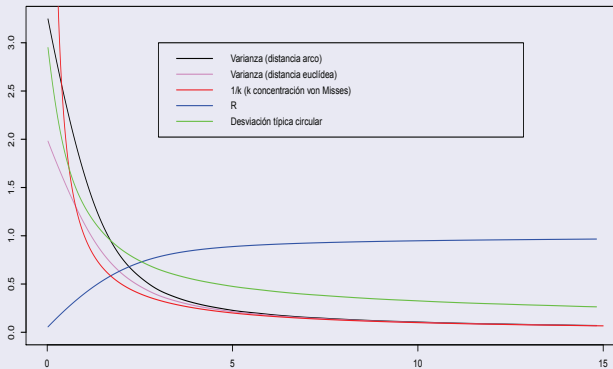
$$\theta = \min_y S_n^2(y)$$

$$d^2(\theta_i, \theta_j) = 2(1 - (\cos(\theta_i - \theta_j))) \quad (2)$$

$$d^2(\theta_i, \theta_j) = \pi - |\pi - |\theta_i - \theta_j|| \quad (3)$$

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} \text{ y } v = \sqrt{-2\log(\bar{R})}.$$

Distribución von Mises



1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- Dada una muestra aleatoria de ángulos $\theta_1, \dots, \theta_n \in [0, 2\pi]$.
- Consideraremos el estimador no paramétrico tipo núcleo de la densidad con núcleo la distribución von Mises

$$\hat{f}(\theta, \nu) = \frac{c_0(\nu)}{n} \sum_{i=1}^n L(\nu \cos(\theta - \theta_i))$$
$$c_0(\nu) = \frac{1}{2\pi I_0(\nu)} \quad L(x) = e^x$$

- Donde $I_r(\nu)$ es la función de Bessel modificada de orden r .
- El parámetro de concentración ν asume el papel del parámetro de suavizado.

- El error asintótico cuadrático medio integrado es

$$AMISE(\nu) = 3\kappa^2 I_2(2\kappa) / \{32\pi\nu^2 I_0(\kappa)^2\}$$

- Minimizando el error AMISE, obtenemos una " Regla plug-in escala von Mises" para el parámetro de suavizado ν basado en la estimación de κ .

Regla plug-in escala von Mises

$$\nu = [3n\hat{\kappa}^2 I_2(2\hat{\kappa}) \{4\pi^{1/2} I_0(\hat{\kappa})^2\}^{-1}]^{2/5}$$

$$\nu = Cn^{2/5}$$

$$C = [3\hat{\kappa}^2 I_2(2\hat{\kappa}) \{4\pi^{1/2} I_0(\hat{\kappa})^2\}^{-1}]^{2/5}$$

- Introducimos la validación cruzada para minimizar la función pérdida dada por error cuadrático medio y la función de pérdida dada por Kullback-Leiber.
- Consideremos la estimación de la densidad construída dejando fuera el valor θ_j de la muestra.

$$\hat{f}_j(\theta, \nu) = \frac{c_0(\nu)}{n} \sum_{i \neq j} L(\nu \cos(\theta - \theta_i))$$
$$c_0(\nu) = \frac{1}{2\pi I_0(\nu)} \quad L(x) = e^x$$

Funciones de validación cruzada

$$cv_2(\nu) = 2n^{-1} \sum_{i=1}^n \hat{f}_i(\theta_i, \nu) - \int \hat{f}^2(\theta, \nu) d\theta$$

$$cv_{KL} = n^{-1} \sum_{i=1}^n \hat{f}_i(\theta_i, \nu)$$

- Luego $-cv_2(\nu) + \int f^2$ y $-cv_{KL}(\nu) + \int f \log(f)$ son estimaciones insesgadas de la pérdida del error cuadrático $L_2(\nu)$ y la pérdida dada por Kullback-Leiber, respectivamente.

$$v_2 = \operatorname{argmin}_{\nu \geq 0} -cv_2(\nu)$$
$$v_{KL} = \operatorname{argmin}_{\nu \geq 0} -cv_{KL}(\nu)$$

- De estas dos expresiones obtenemos dos posibles valores para el parámetro de suavizado.

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- Dadas n observaciones de una variable direccional explicativa $\theta_1, \dots, \theta_n$ y una variable respuesta lineal $y_1, \dots, y_n$
- Suponemos que $y_i = m(\theta_i) + \epsilon_i$, donde ϵ_i son variables con media cero, independientes e idénticamente distribuidos.

Estimación no paramétrica tipo núcleo de la regresión circular-lineal, $m(\theta)$

$$\hat{m}(\theta; \kappa) = \frac{\sum_{i=1}^n y_i g(\theta - \theta_i, 0, \kappa)}{\sum_{i=1}^n g(\theta - \theta_i, 0, \kappa)}$$

- La elección del grado de suavizado es crucial para la estimación tipo núcleo de la regresión.
- Existen varios métodos para seleccionar el parámetro de suavizado subjetivamente o automáticamente.
- Por ejemplo, cuando el núcleo de suavizado evaluado sobre datos lineales es un núcleo Gausiano, Silverman recomienda un parámetro de suavizado de $0.9\hat{\sigma}n^{1/5}$ y el valor $h_s = 1.06\hat{\sigma}n^{1/5}$ es sugerido por Scott.

- Si θ es distribuída de acuerdo con la distribución von Mises centrada en μ y con parámetro κ , luego

$$\kappa^{-1/2}(\theta - \mu) \rightarrow_D N(0, 1) \quad \kappa \rightarrow \infty$$

- Esto sugiere la elección del parámetro de concentración κ mediante

$$\kappa = \frac{1}{h_s^2}$$

- Otra opción es seleccionar κ por validación cruzada de mínimos cuadrados (LCV).
- Como en el caso lineal, escogemos κ de modo que minimize a la función dónde son las estimaciones sin el dato j -ésimo.

$$CV(\kappa) = n^{-1} \sum_{j=1}^n [y_j - \hat{m}^{-j}(\theta_j; \kappa)]^2$$
$$\hat{m}^{-j}(\theta_j; \kappa) = \frac{\sum_{i \neq j}^n y_i g(\theta_j - \theta_i, 0, \kappa)}{\sum_{i \neq j}^n g(\theta_j - \theta_i, 0, \kappa)}$$

- La estimación tipo núcleo seleccionando el parámetro de suavizado mediante LCV puede ser inconsistente bajo una variedad de circunstancias.
 - para datos discretos con múltiples valores repetidos
 - validación cruzada tiende a sugerir parámetros de suavizado que suavizan muy poco, es decir, κ tiende a ser muy grande.
- Por tanto valores de κ escogidos de acuerdo $\kappa = \frac{1}{h_s^2}$ serán preferidos en este caso.

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- La componente determinística del modelo de regresión estudiado por Downs y Mardia(2002) une a la variable angular dependiente v con la variable angular independiente u mediante

$$\tan \frac{1}{2}(v - \beta) = \omega \tan \frac{1}{2}(u - \alpha)$$
$$v = \beta + 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2}(u - \alpha) \right\}$$

- $\omega \in [-1, 1]$ es el parámetro pendiente.
- $-\pi \leq \alpha, \beta < \pi$ son los parámetros angulares de localización.

- Reemplazamos el ángulo v por θ_t y el ángulo u por θ_{t-1} , $t = 2, \dots, n$. Esta sustitución sugiere un único parámetro de localización, $\alpha = \beta$

$$\tan \frac{1}{2}(\theta_t - \alpha) = \omega \tan \frac{1}{2}(\theta_{t-1} - \alpha)$$
$$\theta_t = \alpha + 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2}(\theta_{t-1} - \alpha) \right\}$$

- Para este modelo de series de tiempo se asume que:

$$\theta_t | \theta_{t-1} \sim M\left(\alpha + 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2}(\theta_{t-1} - \alpha) \right\}, \kappa\right)$$

Modelo de Möbius de series de tiempo

$$\theta_t = \alpha + 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2} (\theta_{t-1} - \alpha) \right\} + \epsilon_t$$

- Dónde $\epsilon_t \sim M(0, \kappa)$
- y la media direccional condicionada, μ_t , de θ_t dado θ_{t-1} viene dada por

$$\mu_t = \alpha + 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2} (\theta_{t-1} - \alpha) \right\}$$

- Estimamos los parámetros α y ω mediante máxima verosimilitud; es decir, maximizando la función

$$l(\alpha, \omega) = \sum_{i=2}^n \cos \left[\theta_t - \alpha - 2 \tan^{-1} \left\{ \omega \tan \frac{1}{2} (\theta_{t-1} - \alpha) \right\} \right]$$

- La maximización la haremos usando la función *nlm* de R pero minimizando la función $-l_C(\alpha, \omega)$ para numerosos conjuntos de datos.

- Una vez que $l(\alpha, \omega)$ ha sido maximizada con respecto a α y ω , una aproximación mediante verosimilitud puede ser usada para obtener una estimación mediante máxima verosimilitud de κ , maximizando

$$l_C(\hat{\alpha}, \hat{\omega}, \kappa) = \text{const.} - (n-1) \log l_0 \kappa + \kappa l_{\hat{\alpha}, \hat{\omega}}$$

respecto de κ .

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- Notemos que vamos a utilizar los datos de dirección de viento de la estación meteorológica A Mourela para llevar a cabo
 - 1 análisis descriptivo circular
 - 2 estimación no paramétrica tipo núcleo de la densidad
 - 3 estimación del modelo de Möbius de series de tiempo
- Utilizaremos la dirección de viento y las concentraciones de SO_2 ; ambas medidas tomadas en cada una de las estaciones de inmisión B1, B2, C9, F2, G2 para realizar la estimación no paramétrica tipo núcleo de la regresión circular-lineal.

- Notemos que tanto la estación metereológica a Mourela como las estaciones de inmisión pertenecen a la Red de Vigilancia de la U.P.T. de As Pontes (Endesa S.A.)
- Todas las estaciones recogen datos minutales pero trabajaremos con las medias horarias con el objetivo de reducir la cantidad de datos.
- Aunque a la hora de presentar los resultados de los datos horarios de viento nos seguimos encontrando con una gran cantidad de datos, 8360.
- Luego, para que los gráficos sean más fáciles de visualizar hemos dividido el año 2010 en cuatro periodos de 2090 datos cada uno.

- Al analizar los datos obtenidos en las estaciones se observó que había valores repetidos tanto en la dirección de viento como en las concentraciones de SO_2 .
- La aparición de medidas repetidas va a ser un problema a la hora de implementar ciertas técnicas, como la validación-cruzada para escoger el parámetro de suavizado.
- Con el fin de solucionar dicho problema, haremos una perturbación en las variables.

- **Caso lineal:** Azzalini (1981) propone una perturbación que permite una estimación consistente de la función de distribución

$$\begin{aligned}\tilde{X}_i &= X_i + b\epsilon_i, \epsilon_i \sim \text{Epanech}(-\sqrt{5}, \sqrt{5}) \\ b &= 1.03\hat{\sigma}n^{-1/3}\end{aligned}$$

- **Caso circular:** E. Garcia-Portugués et al. (2011)

$$\begin{aligned}\tilde{\theta}_i &= \theta_i + d\epsilon_i, \epsilon_i \sim M(0, 1) \\ b &= n^{-1/5}\end{aligned}$$

	n	Media direccional	Media de longitud resultante
Primer periodo	2090	5.3634	0.2177
Segundo periodo	2090	0.6446	0.2850
Tercer periodo	2090	0.2134	0.1750
Cuarto periodo	2090	5.7374	0.2952

Cuadro: Estudio de los datos de dirección de viento 2010 mediante una medida de localización (media direccional) y una medida de dispersión (media de longitud resultante)

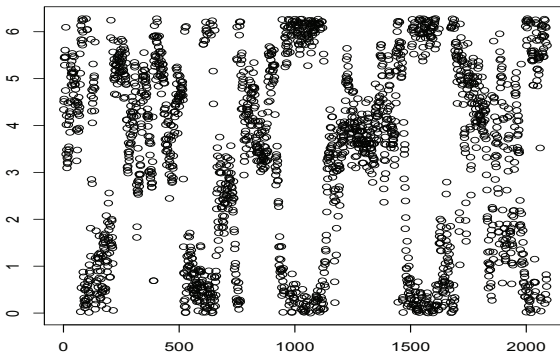
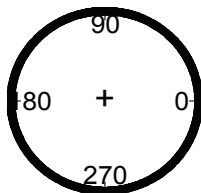
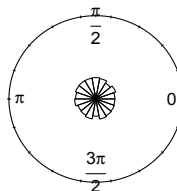


Figura: Gráfico de dispersión de los datos de dirección de viento del primer periodo de 2010

Primer periodo del año 2010



Primer periodo del año 2010



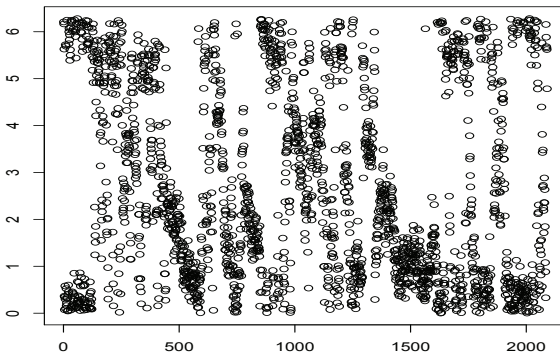
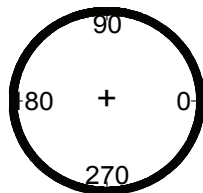
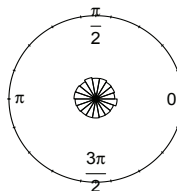


Figura: Gráfico de dispersión de los datos de dirección de viento del primer periodo de 2010

Segundo periodo del año 2010



Segundo periodo del año 2011



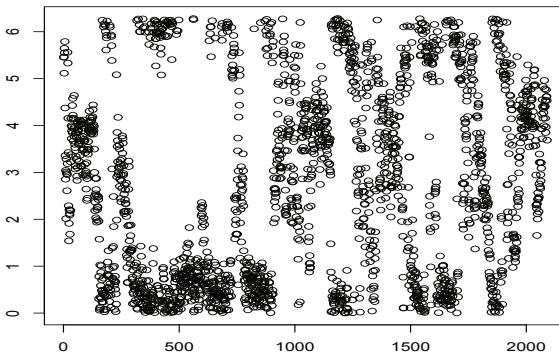
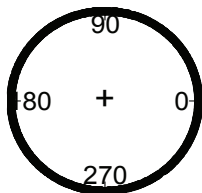
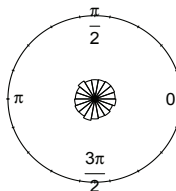


Figura: Gráfico de dispersión de los datos de dirección de viento del tercer periodo de 2010

Tercer periodo del año 2010



Tercer periodo del año 2010



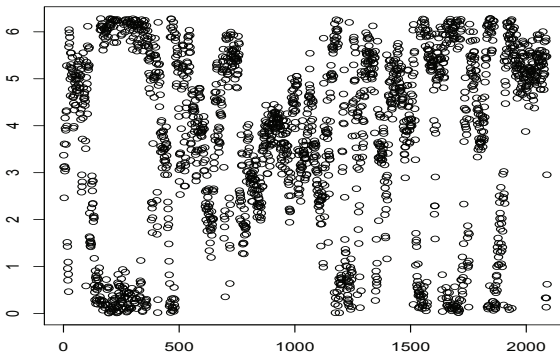
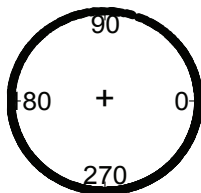
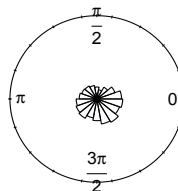


Figura: Gráfico de dispersión de los datos de dirección de viento del cuarto periodo de 2010

Cuarto periodo del año 2010



Cuarto periodo del año 2010



	L2	KL	plug-in
Primer periodo	28.56	26.20	3.10
Segundo periodo	63.57	77.17	4.87
Tercer periodo	41.44	36.17	2.17
Cuarto periodo	35.56	34.31	5.16

Cuadro: Estimaciones del parámetro de suavizado para la función de densidad circular

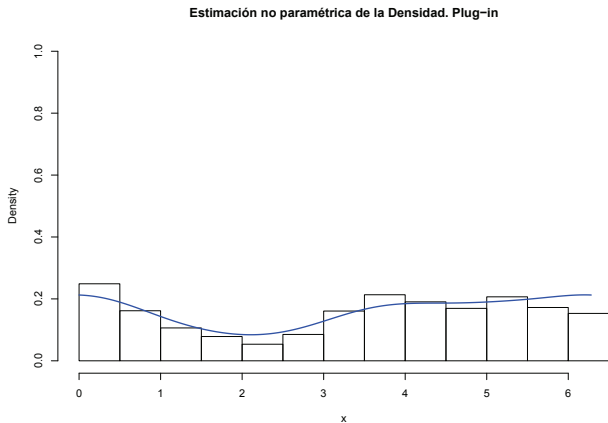


Figura: Estimación plug-in de la densidad del primer periodo de 2010

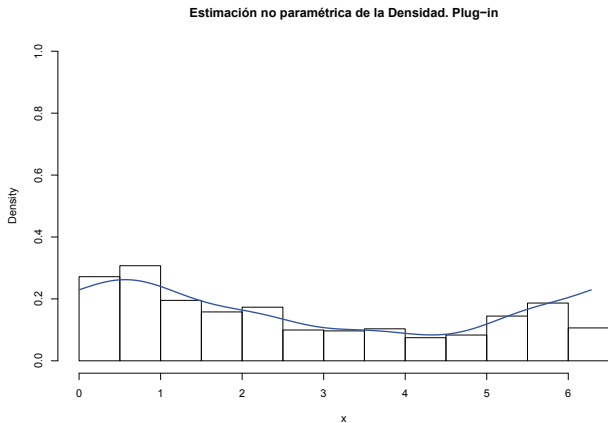


Figura: Estimación plug-in de la densidad del segundo periodo de 2010

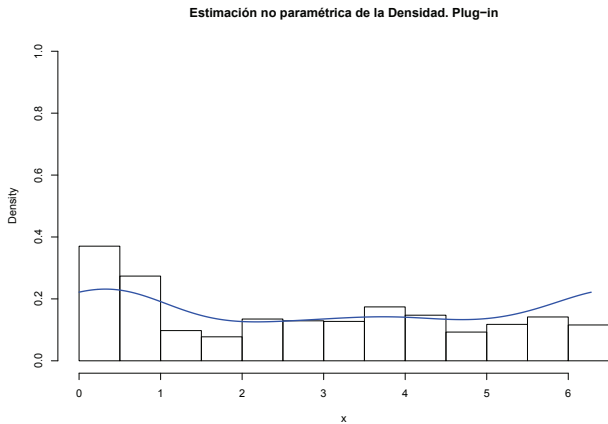


Figura: Estimación plug-in de la densidad del tercer periodo de 2010

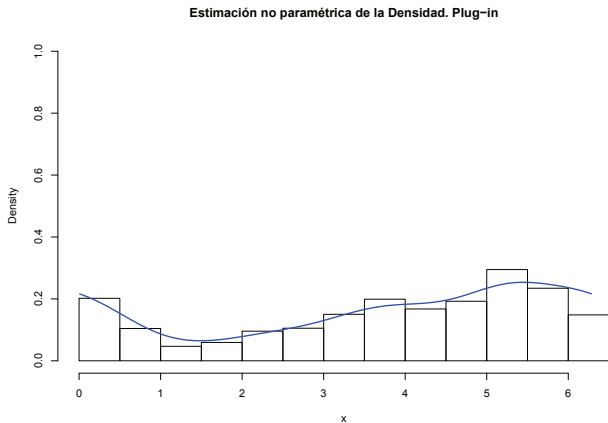


Figura: Estimación plug-in de la densidad del cuarto periodo de 2010

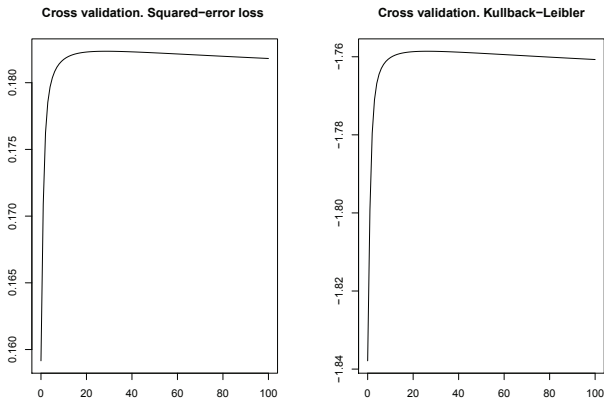


Figura: Validación cruzada del primer periodo de 2010

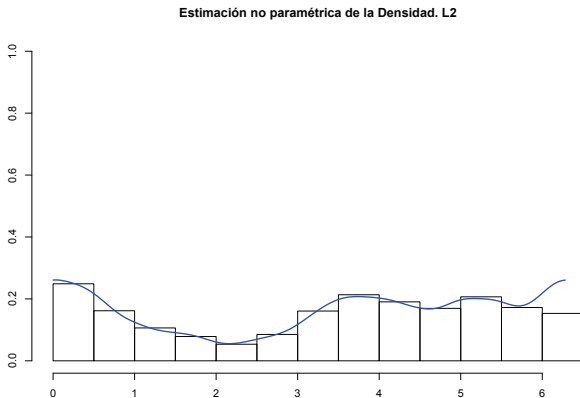


Figura: Estimación L2 de la densidad del primer periodo de 2010

Estimación no paramétrica de la Densidad. Kullback-Leiber

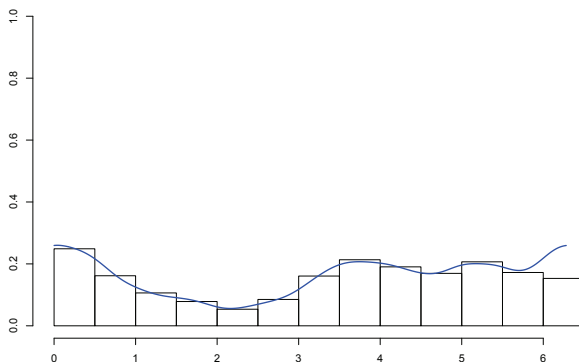


Figura: Estimación KL de la densidad del primer periodo de 2010

Estimación no paramétrica de la Densidad.

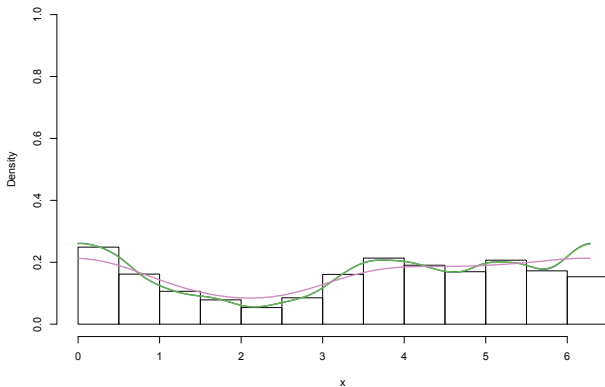


Figura: Estimación no paramétrica de la Densidad. L2 (azul), KL (verde), plug-in (violeta)

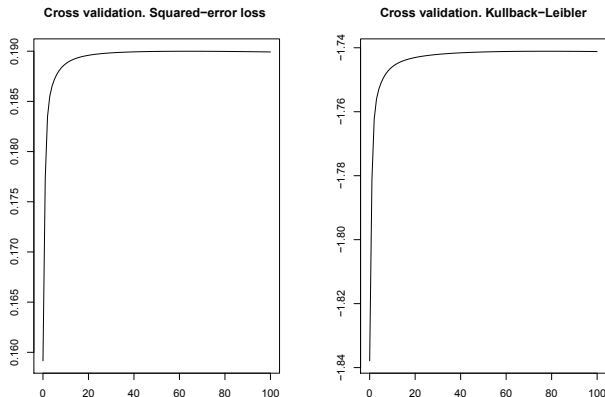


Figura: Validación cruzada del segundo periodo de 2010

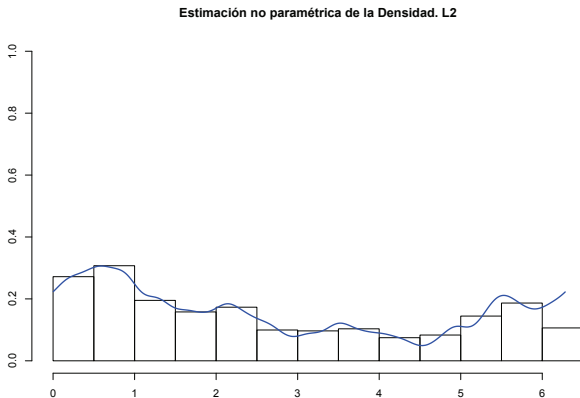


Figura: Estimación L2 de la densidad del segundo periodo de 2010

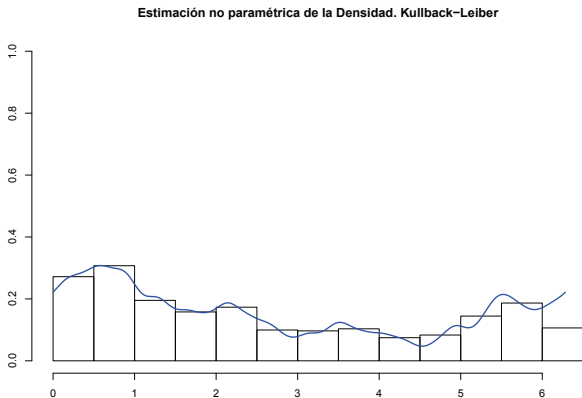


Figura: Estimación KL de la densidad del segundo periodo de 2010

Estimación no paramétrica de la Densidad.

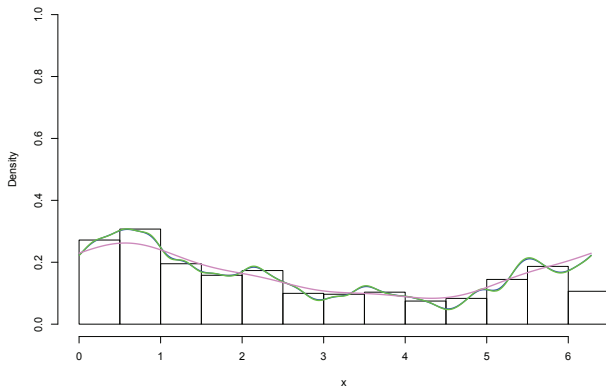


Figura: Estimación no paramétrica de la Densidad. L2 (azul), KL (verde), plug-in (violeta)

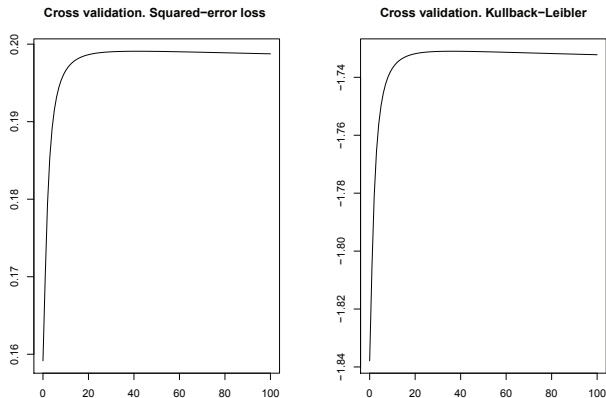


Figura: Validación cruzada del tercer periodo de 2010

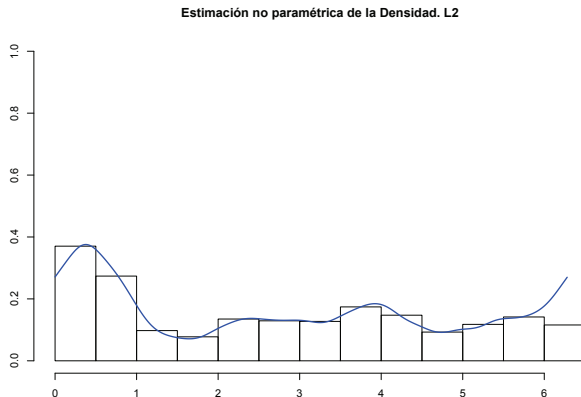


Figura: Estimación L2 de la densidad del tercer periodo de 2010

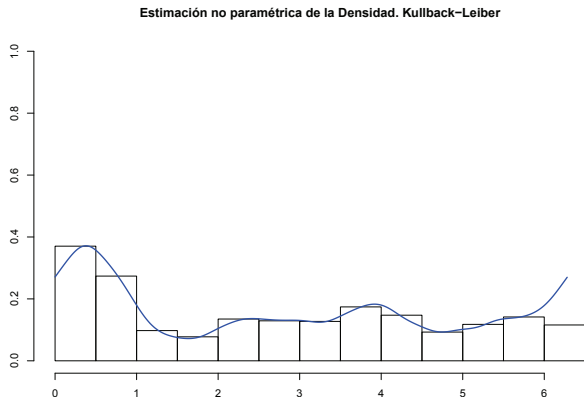


Figura: Estimación KL de la densidad del tercer periodo de 2010

Estimación no paramétrica de la Densidad.

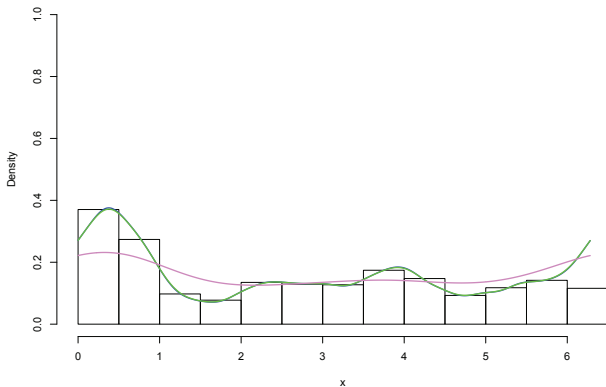


Figura: Estimación no paramétrica de la Densidad. L2 (azul), KL (verde), plug-in (violeta)

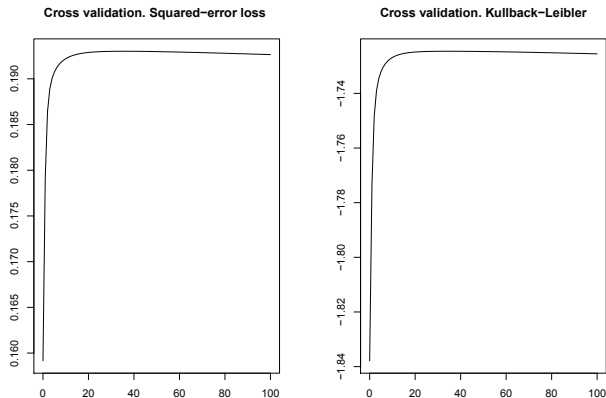


Figura: Validación cruzada del cuarto periodo de 2010

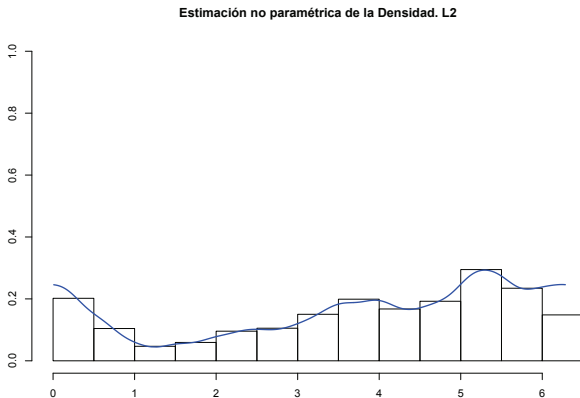


Figura: Estimación L2 de la densidad del cuarto periodo de 2010

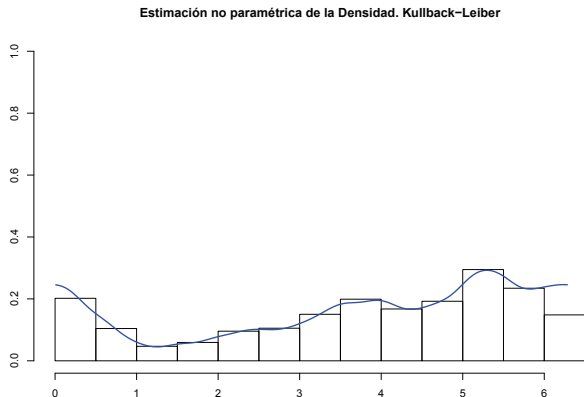


Figura: Estimación KL de la densidad del cuarto periodo de 2010

Estimación no paramétrica de la Densidad.

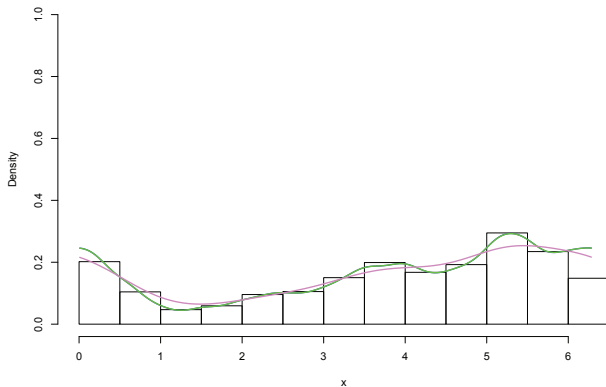


Figura: Estimación no paramétrica de la Densidad. L2 (azul), KL (verde), plug-in (violeta)

	B1	B2	C9	F2	G2
Primer periodo	3.33	2.87	30.28	37.69	26.42
Segundo periodo	13.07	8.83	7.07	10.73	20.13
Tercer periodo	6.61	25.56	2.33	11.97	18.43
Cuarto periodo	0.00	22.63	4.87	11.24	17.74

Cuadro: Estimaciones del parámetro de suavizado obtenidas mediante validación cruzada

	B1	B2	C9	F2	G2
Primer periodo	0.0049	0.0069	0.0074	0.0072	0.0071
Segundo periodo	0.0088	0.0081	0.0097	0.0037	0.0075
Tercer periodo	0.0078	0.0080	0.0087	0.0030	0.0061
Cuarto periodo	0.0045	0.0071	0.0068	0.0062	0.0056

Cuadro: Estimaciones del parámetro de suavizado obtenidas mediante la expresión $\kappa = 1/h_s^2$

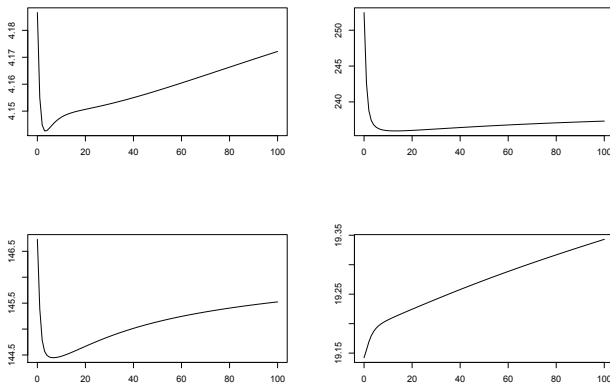


Figura: Validación cruzada en la estación B1

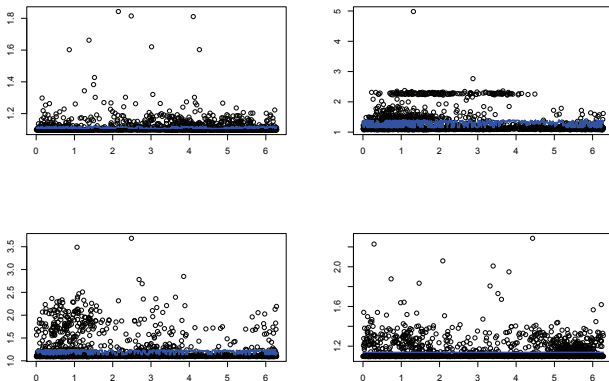


Figura: Estimación de la función de regresión circular-lineal en B1.
Estimación del parámetro de suavizado mediante validación cruzada.

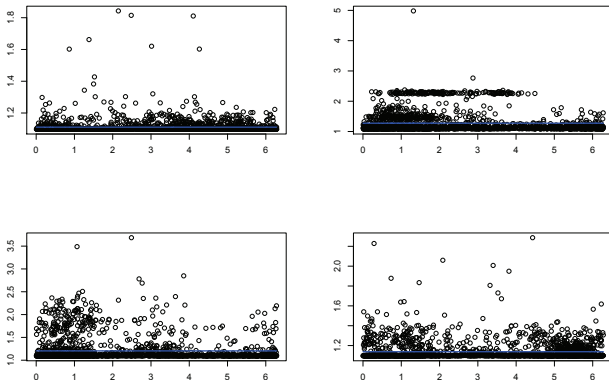


Figura: Estimación de la función de regresión circular-lineal en B1.
Estimación del parámetro de suavizado mediante $1/h_s^2$.

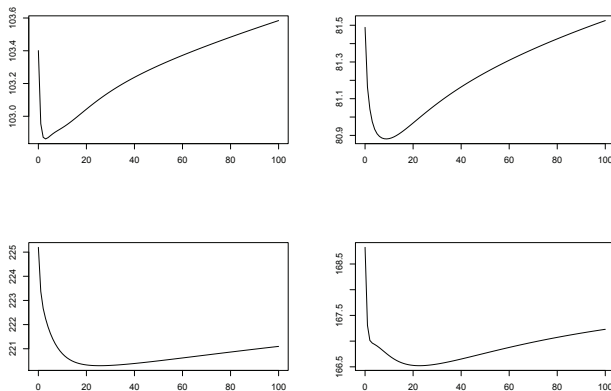


Figura: Validación cruzada en la estación B2

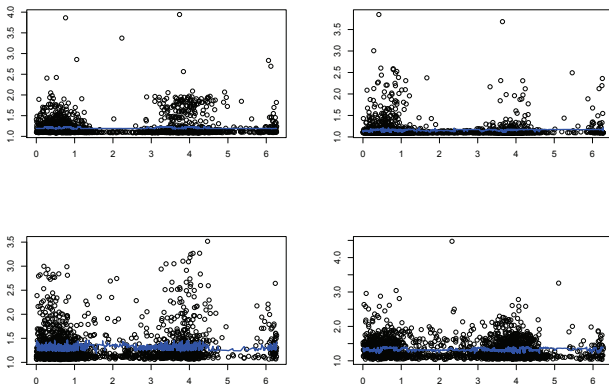


Figura: Estimación de la función de regresión circular-lineal en B2.
Estimación del parámetro de suavizado mediante validación cruzada.

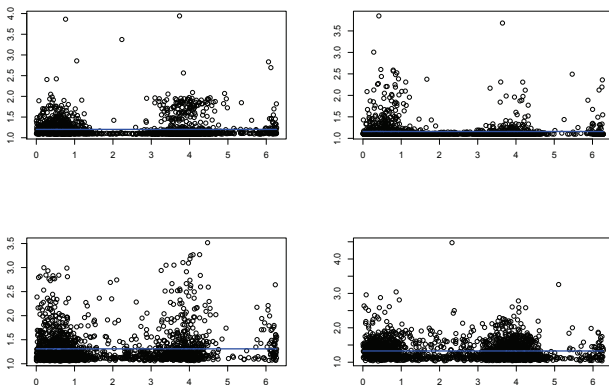


Figura: Estimación de la función de regresión circular-lineal en B2.
Estimación del parámetro de suavizado mediante $1/h_s^2$.

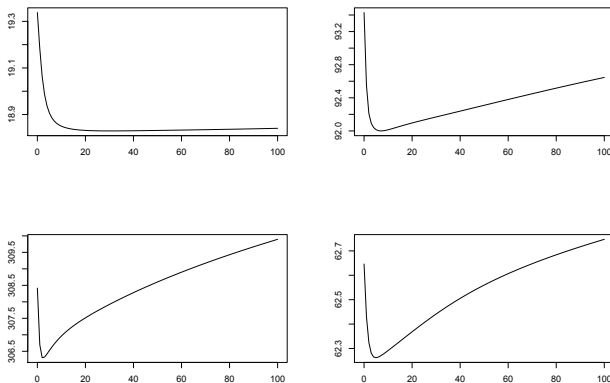
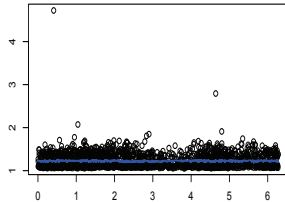
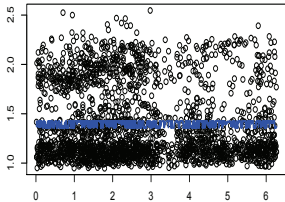
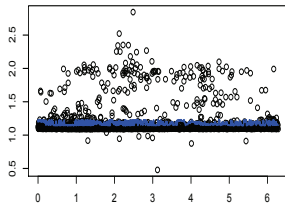
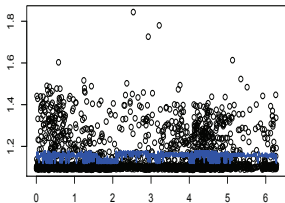


Figura: Validación cruzada en la estación C9



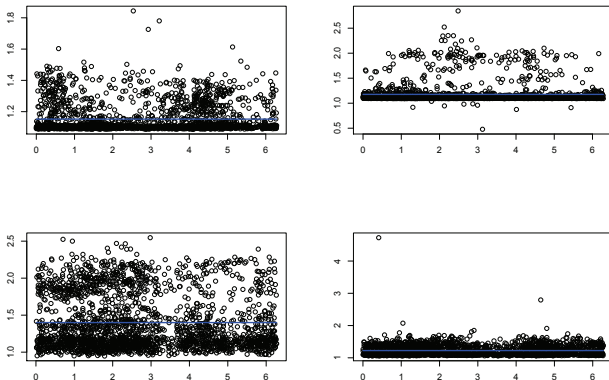
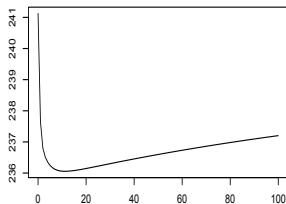
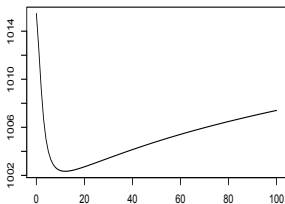
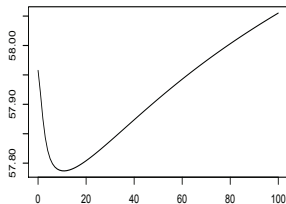
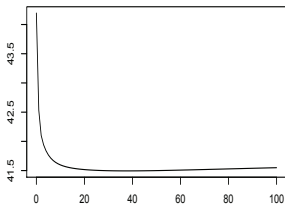
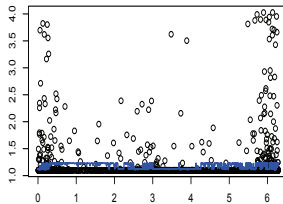
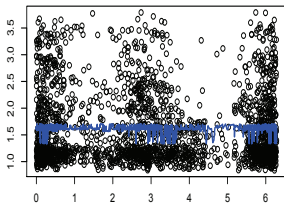
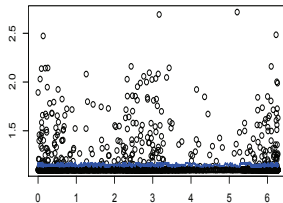
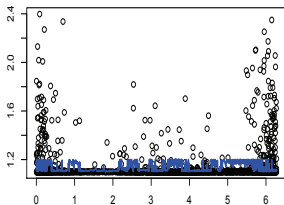


Figura: Estimación de la función de regresión circular-lineal en C9.
Estimación del parámetro de suavizado mediante $1/h_s^2$.





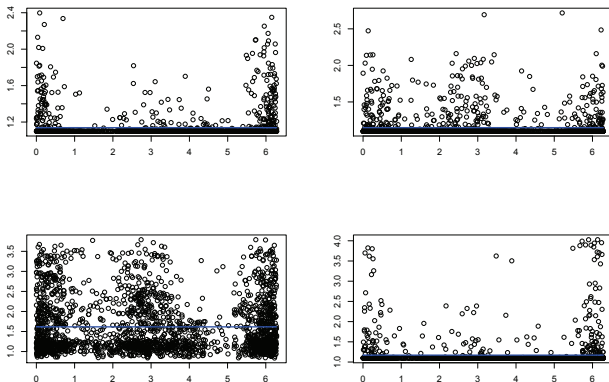


Figura: Estimación de la función de regresión circular-lineal en F2.
Estimación del parámetro de suavizado mediante $1/h_s^2$.

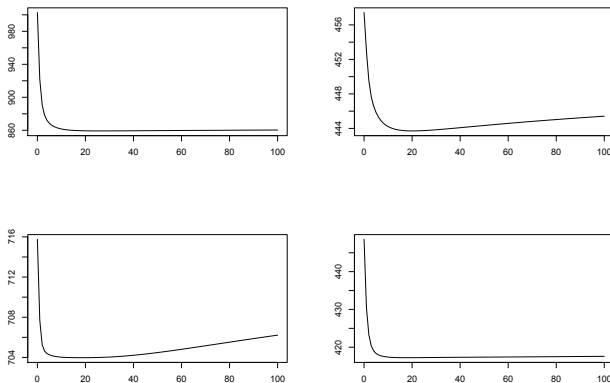


Figura: Validación cruzada en la estación G2

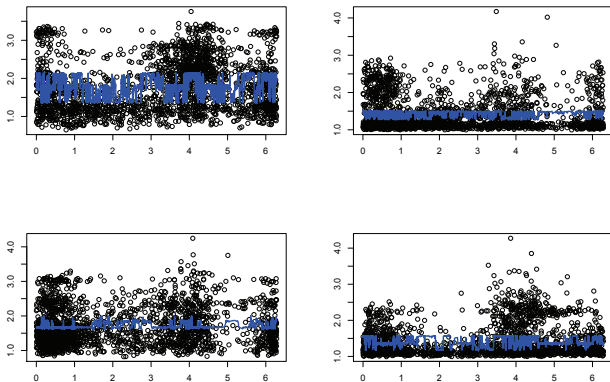


Figura: Estimación de la función de regresión circular-lineal en G2.
Estimación del parámetro de suavizado mediante validación cruzada.

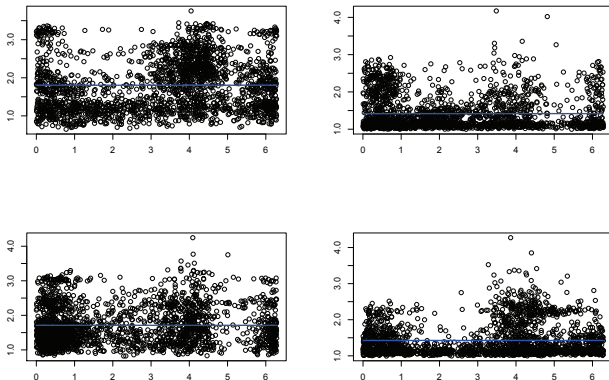


Figura: Estimación de la función de regresión circular-lineal en G2.
Estimación del parámetro de suavizado mediante $1/h_s^2$.

	α	ω	κ
Primer periodo	-0.7492	0.9634	4.12
Segundo periodo	0.5133	0.9296	4.09
Tercer periodo	-0.1431	0.9644	4.11
Cuarto periodo	-1.3190	0.9382	4.11

Cuadro: Estimaciones de los parámetros que están involucrados en la otención del modelo de Möbius de series de tiempo

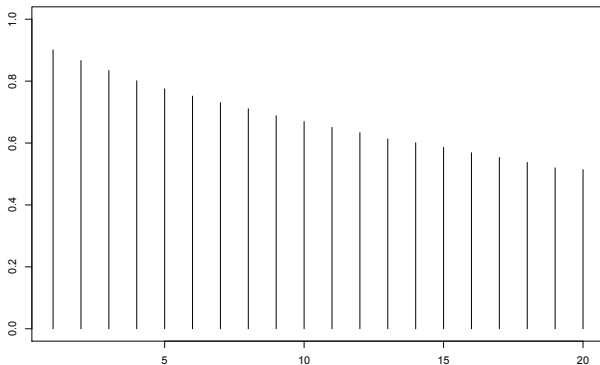


Figura: Autocorrelaciones circulares para dirección de viento del primer periodo de 2010

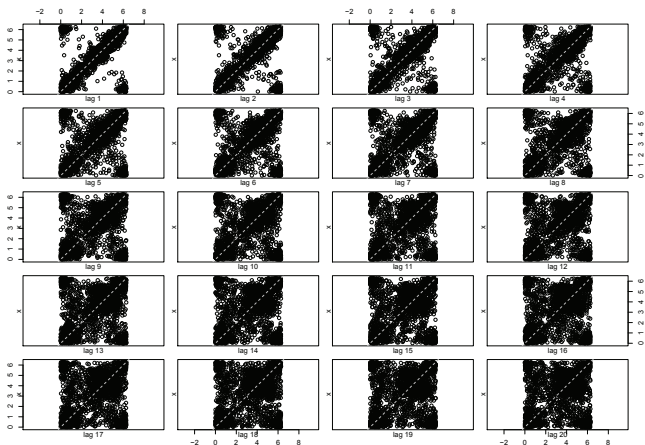


Figura: Representación de la serie de tiempo de la dirección del viento del primer periodo de 2010 frente a sus retardos.

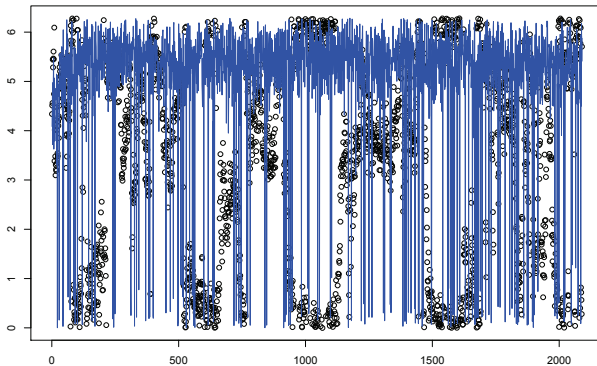


Figura: Serie de tiempo de la dirección de viento del primer periodo de 2010

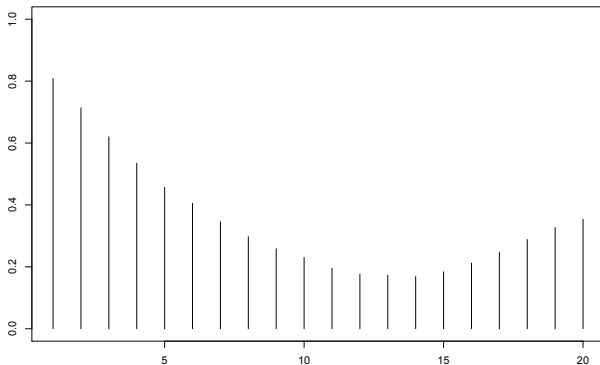


Figura: Autocorrelaciones circulares para dirección de viento del segundo periodo de 2010

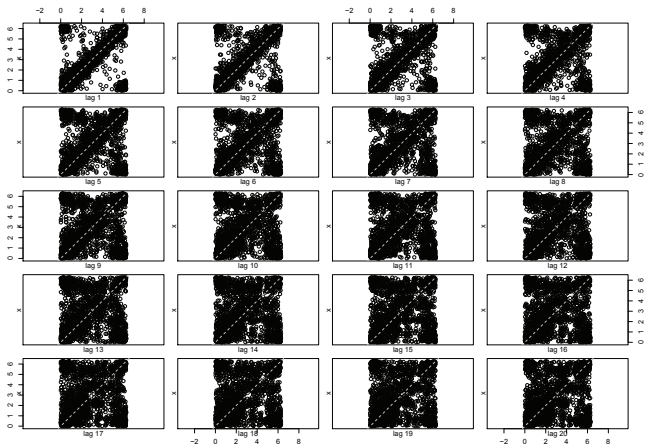


Figura: Representación de la serie de tiempo de la dirección del viento del segundo periodo de 2010 frente a sus retardos.

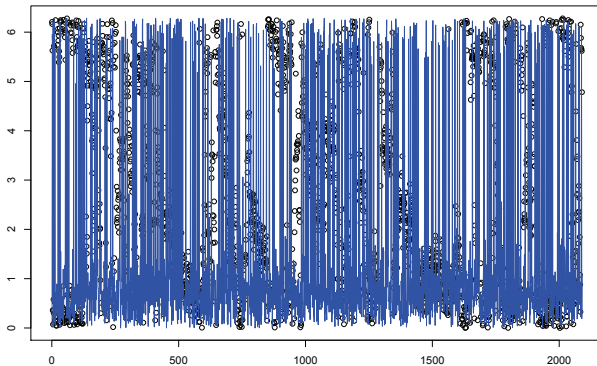


Figura: Serie de tiempo de la dirección de viento del segundo periodo de 2010

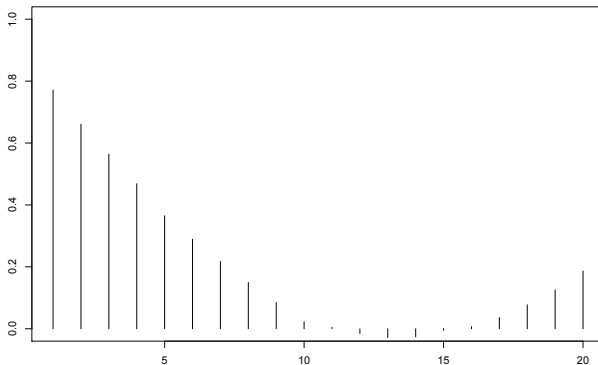


Figura: Autocorrelaciones circulares para dirección de viento del tercer periodo de 2010

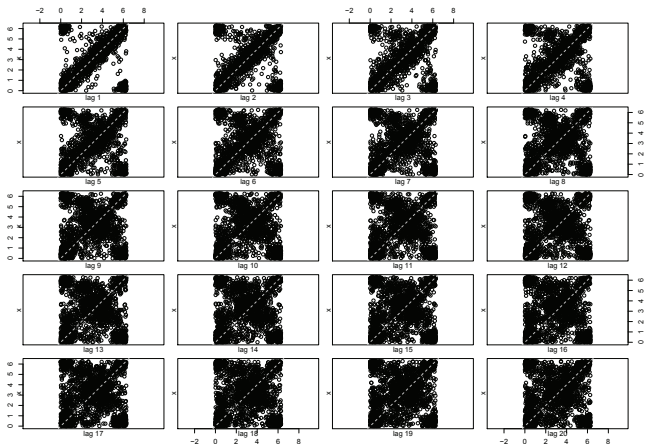


Figura: Representación de la serie de tiempo de la dirección del viento del tercer periodo de 2010 frente a sus retardos.

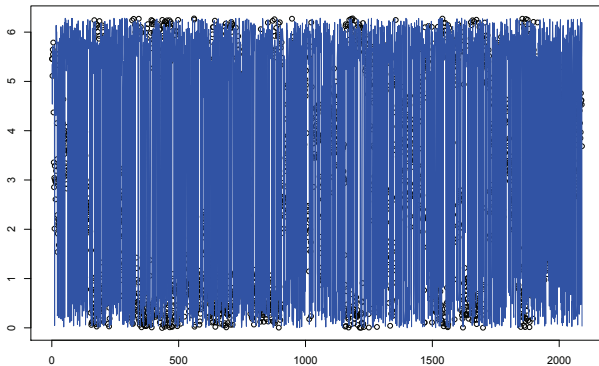


Figura: Serie de tiempo de la dirección de viento del tercer periodo de 2010

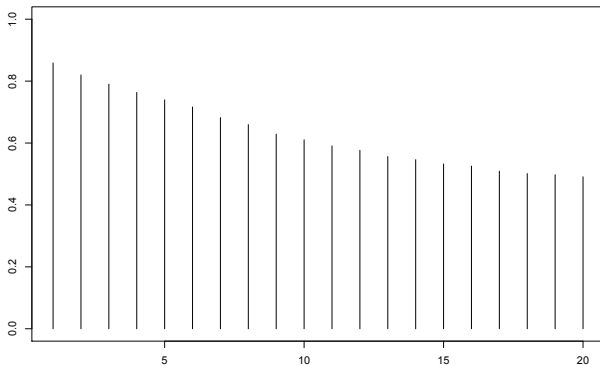


Figura: Autocorrelaciones circulares para dirección de viento del cuarto periodo de 2010

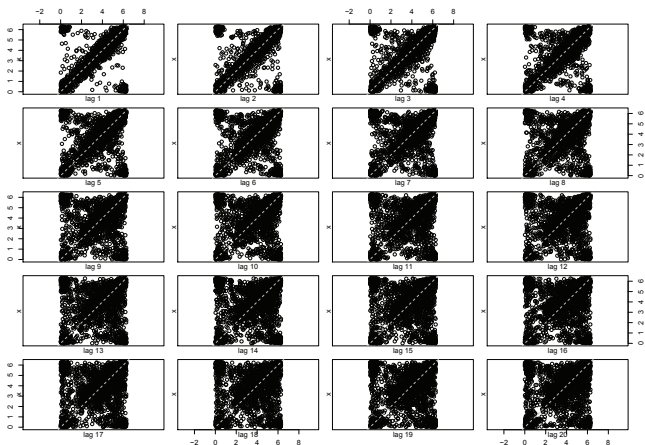


Figura: Representación de la serie de tiempo de la dirección del viento del cuarto periodo de 2010 frente a sus retardos.

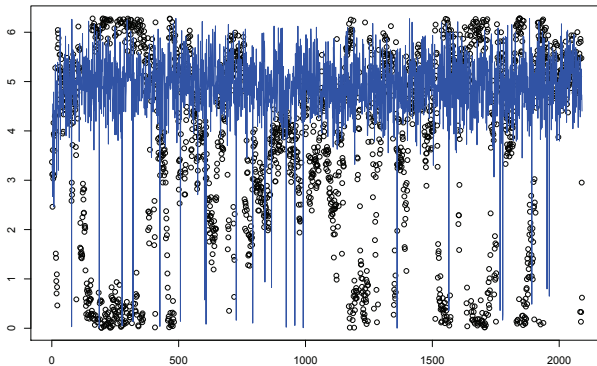


Figura: Serie de tiempo de la dirección de viento del cuarto periodo de 2010

1 Introducción

2 Función de densidad circular

- Estimación no paramétrica tipo núcleo de la densidad circular
- Selección del parámetro de suavizado. Plug-in
- Selección del parámetro de suavizado. Validación cruzada

3 Función de regresión circular-lineal.

- Selección del parámetro de suavizado

4 El modelo de Möbius de series de tiempo

- Estimación de los parámetros. Máxima verosimilitud

5 Aplicación

- Datos perturbados
- Análisis de los datos de dirección de viento en A Mourela
- Estimación tipo núcleo de la función de densidad circular
- Estimación tipo núcleo de la función de regresión circular-lineal
- Estimación del modelo de Möbius de series de tiempo

6 Bibliografía

- [1] A. Azzalini. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, 68:326–328, 1981.
- [2] Panzera A. Di Marzio, M. and C.C. Taylor. Local polynomial regression for circular predictors. *Statistics and Probability Letters*, 79:2066–2075, 2009.
- [3] T. D. Downs and K. V. Mardia. Circular regression. *Biometrika*, 89:683–698, 2002.
- [4] Crujeiras R.M. González Manteiga W. García Portugués, E. Exploring with direction and so2 concentration by circular-linear density estimation. Technical report, Universidade de Santiago de Compostela, Departamento de Estatística e Investigación Operativa., 2011.
- [5] X. Haiyong and F. P. Schoenberg. Kernel regression of directional data with application to wind and wildfire data in los angeles country, california. Department of Statistics, University of California, Los Angeles, CA 90095-1554, USA.

- [6] Watson G.S. Hall, P. and J. Cabrera. Kernel density estimation with spherical data. *Biometrika*, 74:751–762, 1987.
- [7] G. Hughes. *Multivariate and Times Series Models for Circular Data with Applications to Protein Conformational Angles*. PhD thesis, The University of Leeds, Department of Statistics., 2007.
- [8] K.V. Mardia and P.E Jupp. *Directional Statistics*. Wiley; New York, 2000.
- [9] C.C. Taylor. Automatic bandwidth selection for circular density estimation. *Computational Statistics and Data Analysis*, 52:3493–3500, 2008.

Datos Direccionales.

Proyecto Fin de Máster.

Leyenda Rodríguez, María

9 de junio de 2011