Five Ws on Nonparametric Statistics for Circular Data

J. Ameijeiras[1], Rosa M. Crujeiras[1], M. Oliveira[2] and A. Rodríguez–Casal[1]

[1]Department of Statistics and Operations Research
University of Santiago de Compostela
[2]Inteligencia de Clientes - Abanca

[A tribute to Wences' 'whats']

[A tribute to Wences' 'whats']

What?

[A tribute to Wences' 'whats']

What?

Who did What and When?

[A tribute to Wences' 'whats']

What?

Who did What and When?

What and hoW?

[A tribute to Wences' 'whats']

What?

Who did What and When?

What and hoW?

Why and Where?

[A tribute to Wences' 'whats']

What?

Who did What and When?

What and hoW?

Why and Where?

What else?

Two-dimensional directions or circular data appear in a large variety of disciplines, such as biology, oceanography or meteorology.

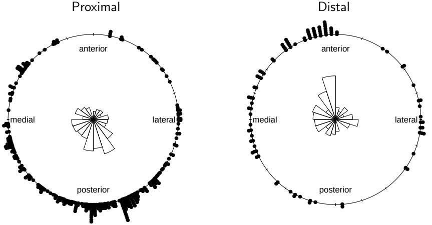
Circular data can be represented as points on the circumference of a unit circle or by one angle in $[0, 2\pi)$.

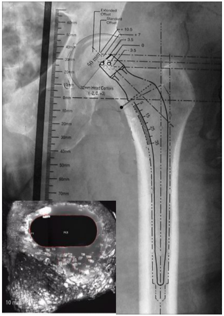
Circular data are periodic, i.e., given $\theta \in [0, 2\pi)$, $\theta = \theta + 2\pi k$, for any integer k.

Techniques for linear data are not valid for circular data.

Example: Cracks in cemented femoral components

- Data*: Angular position of cracks in the cement mantle for proximal and distal regions. Number of cracks in each region: 322 and 99, respectively.



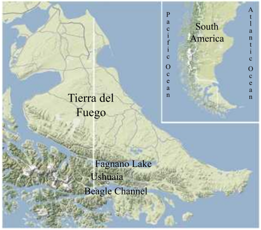


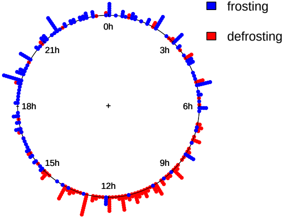
- Is there a preferred direction for cracks?

* Data were provided by Prof. Kenneth A. Mann from Upstate Medical University (New York).

Example: Temperature cycle changes

- Data*: 350 observations corresponding to the hours when the temperature changes from positive to negative and viceversa in periglacial Monte Alvear (Ushuaia, Argentina) from February 2008 to December 2009.



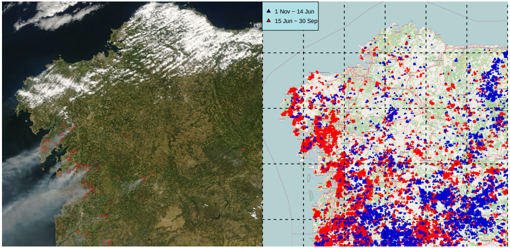


- At what time are frosting/defrosting cycles more likely to happen?

* Data were collected within the Project POL2006-09071 from the Spanish Ministry of Education and Science and provided by Prof. Augusto Pérez–Alberti (USC).

Example: Fires detected by MODIS

- Data*: Time of occurrence and location of the fires detected by MODIS from 10/07/2002 to 9/07/2012.

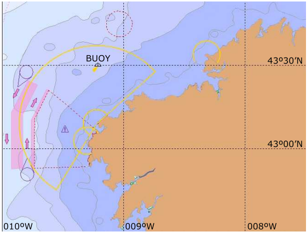


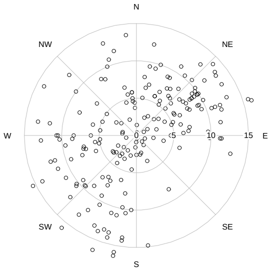
- Aggregating the data at 0.5^o resolution, the objective is find out in which regions the number of fire seasons is greater than the climatological ones.

*Data were collected by Prof. José M. C. Pereira and his group from the University of Lisbon (Portugal).

Example: Wind speed and wind direction

- Data*: hourly observations of wind direction and wind speed in Vilán–Sisargas in winter season (November to February), from 2003 until 2012.





- In the Galician coast, during winter season: is the wind speed influenced by the wind direction?

* Data were downloaded from the Spanish Portuary Authority (www.puertos.es).

When you have a hammer... everything looks like a nail!

The stochastic behaviour of the variables in the previous examples can be characterized by the estimation and deep analysis of a density (e.g. for femoral cracks) or a regression curve (e.g. wind speed and wind direction).

- For regression: see Edu's talk! (just before lunchtime...)

Hall, P., Watson, G.P. and Cabrera, J. (1987)
Kernel density estimation for spherical data.
Biometrika, 74, 751–762.

Bai, Z.D., Rao, C.R. and Zhao, L.C. (1988)
Kernel estimators of density function of directional data.
Journal of Multivariate Analysis, 27, 24–39.

Zhao, L.C. and Wu, C. (2001)
Central limit theorem for integrated square error of kernel estimators for spherical data.
Science in China, Series A, 44, 474–483.

Klemelä, J. (2000)
Estimation of densities and derivatives of densities with directional data.
Journal of Multivariate Analysis, 73, 18–40.

Taylor, C.C. (2008)
Automatic bandwidth selection for circular density estimation.
Computational Statistics and Data Analysis, 76, 705–712.

Di Marzio, M., Panzera A. and Taylor, C.C. (2011)
Kernel density estimation on the torus.
Journal of Statistical Planning and Inference, 141, 2156–2173.

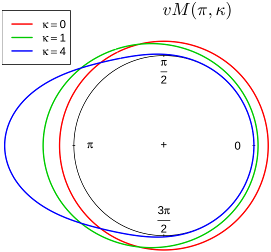
- Oliveira, M., Crujeiras, R.M. and Rodríguez–Casal, A. (2012)
  A plug–in rule for bandwidth selection in circular density estimation.
  Computational Statistics and Data Analysis, 56, 3898–3908.
- Oliveira, M., Crujeiras, R.M. and Rodríguez–Casal, A. (2013)
  Nonparametric circular methods for exploring environmental data.
  Journal of Environmental and Ecological Statistics, 20, 1–17.
- Oliveira, M., Crujeiras, R.M. and Rodríguez–Casal, A. (2014)
  CircSiZer: an exploratory tool for circular data.
  Journal of Environmental and Ecological Statistics, 21, 143–159.
- Oliveira, M., Crujeiras, R.M. and Rodríguez–Casal, A. (2014)
  NPCirc: an R package for nonparametric circular methods.
  Journal of Statistical Software, 61, 1–26.

Parametric circular models

- von Mises, $vM(\mu, \kappa)$:

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \quad 0 \leq \theta < 2\pi$$

  - $\mu \in [0, 2\pi)$ is the mean direction.
  - $\kappa \geq 0$ is the concentration parameter.
  - I_r is the modified Bessel function of order r.

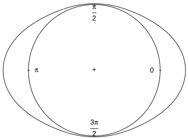


- Other parametric models: Cardioid, wrapped Cauchy, wrapped normal, wrapped skew–normal, etc.

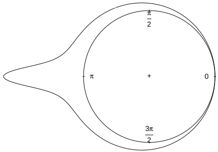
- Parametric mixtures: A finite mixture of M circular distributions, f_m with weights $p_m \geq 0$ for $m = 1, \ldots, M$ and $\sum_{m=1}^{M} p_m = 1$ has density:

$$f(\theta) = \sum_{m=1}^{M} p_m f_m(\theta)$$

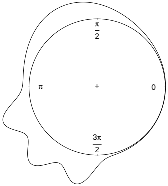
Mixture of two von Mises



Mixture of cardioid and wrapped Cauchy



Mixture of five von Mises



Circular kernel density estimator

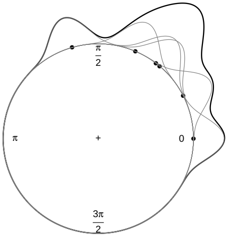
Given a random sample of angles $\Theta_1, \ldots, \Theta_n \in [0, 2\pi)$ from some unknown circular density f, the circular kernel density estimator of f is defined as:

$$\hat{f}(\theta; \nu) = \frac{1}{n} \sum_{i=1}^{n} K_\nu(\theta - \Theta_i)$$

where K_ν is a circular kernel function with concentration parameter $\nu > 0$.

Taking the von Mises density as kernel:

$$\hat{f}(\theta; \nu) = \frac{1}{n 2\pi I_0(\nu)} \sum_{i=1}^{n} e^{\nu \cos(\theta - \Theta_i)}$$



Wences 1st 'what':

Wences 1st 'what': What about the bandwidth?

Smoothing parameter selection (I)

- Likelihood cross–validation smoothing parameter $\hat{\nu}_{LCV}$ is obtained by maximizing:

$$\mathrm{LCV}(\nu) = \prod_{i=1}^{n} \hat{f}^{-i}(\Theta_i;\nu), \quad \hat{f}^{-i}(\theta;\nu) = \frac{1}{(n-1)(2\pi)I_0(\nu)} \sum_{j\neq i} e^{\nu\cos(\theta-\Theta_j)}$$

- Least squares cross–validation smoothing parameter $\hat{\nu}_{LSCV}$ is obtained by minimizing:

$$\mathrm{LSCV}(\nu) = \int_0^{2\pi} \hat{f}^2(\theta;\nu)d\theta - \frac{2}{n}\sum_{i=1}^{n} \hat{f}^{-i}(\Theta_i;\nu)$$

Hall, P., Watson, G. P. and Cabrera, J. (1987)
Kernel density estimation for spherical data.
Biometrika, 74, 751-762.

Smoothing parameter selection (II)

Consider a global error measurement:

$$\mathrm{MISE}\,(\nu) = \mathbb{E}\left[\int \left(\hat{f}(\theta;\nu) - f(\theta)\right)^2 d\theta\right]$$

For the circular kernel density estimator, the $\mathrm{AMISE}(\nu)$ when $\nu \to \infty$ y $\sqrt{\nu}n^{-1} \to 0$ is given by:

$$\mathrm{AMISE}(\nu) = \frac{1}{16}\left[1 - \frac{I_2(\nu)}{I_0(\nu)}\right]^2 \int_0^{2\pi} \left(f''(\theta)\right)^2 d\theta + \frac{I_0(2\nu)}{2n\pi\left(I_0(\nu)\right)^2}$$

Di Marzio, M., Panzera A. and Taylor, C. C. (2009)
Local polynomial regression for circular predictors.
Statistics & Probability Letters, 79, 2066-2075.

$$\text{AMISE}(\nu) = \frac{1}{16}\left[1 - \frac{I_2(\nu)}{I_0(\nu)}\right]^2 \underbrace{\int_0^{2\pi} \left(f''(\theta)\right)^2 d\theta}_{\text{unknown}} + \frac{I_0(2\nu)}{2n\pi\left(I_0(\nu)\right)^2}$$

- von Mises → Rule of thumb (Taylor, 2008).
- mixture of von Mises → Plug–in rule (Oliveira et al. 2012).

Taylor, C. C. (2008)
Automatic bandwidth selection for circular density estimation.
Computational Statistics and Data Analysis, 52, 3493-3500.

Oliveira, M., Crujeiras, R. M. and Rodríguez–Casal, A. (2012)
A plug–in rule for bandwidth selection in circular density estimation.
Computational Statistics and Data Analysis, 56, 3898–3908.

Plug–in rule

Step 1. Select the number of mixture components M for the reference distribution. For example, by using AIC.

Step 2. Estimate the AMISE as follows:

Step 2.1 Estimate the parameters in the von Mises mixture, $(\mu_m, \kappa_m, \alpha_m)$ for $m = 1, \ldots, M$ by EM.

Step 2.2 Compute the integral $\int_0^{2\pi} (\hat{f}''(\theta))^2 d\theta$.

Step 2.3 Plug–in the quantity $\int_0^{2\pi} (\hat{f}''(\theta))^2 d\theta$ in AMISE to get $\widehat{\mathrm{AMISE}}(\nu)$.

Step 3. Minimize $\widehat{\mathrm{AMISE}}(\nu)$ and obtain $\hat{\nu}_{PI}^{AIC}$.

Wences 2nd 'what':

Wences 2nd 'what': What about the bootstrap?

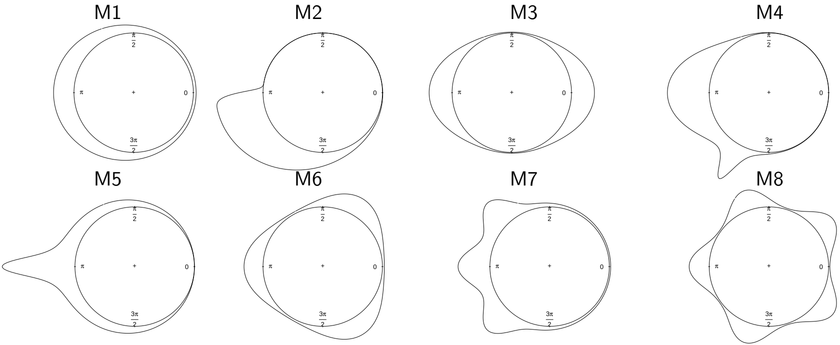
Bootstrap selector

The bootstrap smoothing parameter, $\hat{\nu}_{boot}$, is selected as the value that minimizes the bootstrap MISE:

$$\int_0^{2\pi} \mathbb{E}_B \left[\hat{f}^*(\theta;\nu) - \hat{f}(\theta;\nu) \right]^2 d\theta$$

where $\mathbb{E}_B$ denotes the bootstrap expectation with respect to random samples $\Theta_1^*, \ldots, \Theta_n^*$ generated from $\hat{f}(\theta;\nu)$.

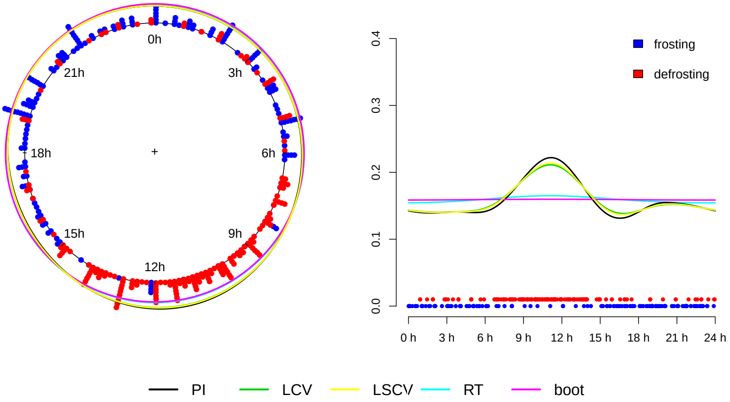
Di Marzio, M., Panzera A. and Taylor, C. C. (2011)
Kernel density estimation on the torus.
Journal of Statistical Planning & Inference, 141, 2156–2173.



$n = 250$	$MISE(\nu_0)$	$MISE(\hat{\nu}_{RT})$	$MISE(\hat{\nu}_{PI})$	$MISE(\hat{\nu}_{LCV})$	$MISE(\hat{\nu}_{boot})$
M1	0.2568	0.3201	0.3499	0.3610	0.3039
M2	1.3422	2.1665	1.6544	1.5842	2.5173
M3	0.5762	10.6753	0.5986	0.5976	0.6038
M4	1.3545	2.0187	1.5816	2.0316	2.7900
M5	1.8929	6.6517	2.2035	3.2325	8.5690
M6	0.6766	6.4797	0.7358	0.7368	1.1453
M7	1.1325	2.9559	1.3480	1.4273	3.2879
M8	1.1141	7.8224	1.1355	1.1473	5.9152

Temperature cycle changes

Data: 350 observations corresponding to the hours when the temperature changes from positive to negative and viceversa.



The idea of CircSiZer method

- CircSiZer is an adaptation to circular data of the original SiZer proposed by Chaudhuri and Marron (1999) for linear data.
- CircSiZer considers nonparametric curve estimates for a wide range of smoothing parameters (ν).
- CircSiZer addresses the question of which features are really there.
- CircSiZer assesses the significance of such features by constructing confidence intervals for the derivative of the smoothed underlying curve at each location $\theta \in [0, 2\pi)$ and scale τ, $f'(\theta;\nu) \equiv \mathbb{E}(\hat{f}'(\theta;\nu))$.

Chaudhuri, P. and Marron, J. S. (1999)
SiZer for exploration of structures in curves.
Journal of the American Statistical Association, 94, 807–823.

Confidence interval

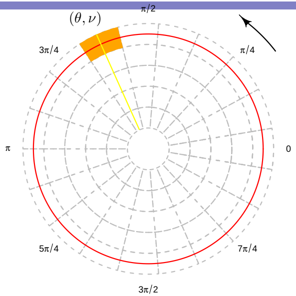
Given a significance level α and for a fixed value of $\nu > 0$ and with $\theta \in [0, 2\pi)$, confidence intervals are of the form

$$\left(\hat{f}'(\theta;\nu) - q^{(1-\alpha/2)} \cdot \widehat{\mathsf{sd}}(\hat{f}'(\theta;\nu)), \hat{f}'(\theta;\nu) - q^{(\alpha/2)} \cdot \widehat{\mathsf{sd}}(\hat{f}'(\theta;\nu))\right)$$

- $\hat{f}'(\theta;\nu)$ is the estimator of the derivative of the curve.
- $q^{(1-\alpha/2)}$ and $q^{(\alpha/2)}$ are appropriate quantiles.
- $\widehat{\mathsf{sd}}(\hat{f}'(\theta;\nu))$ is an estimator of the std of $\hat{f}'(\theta;\nu)$.

Oliveira, M., Crujeiras, R.M. and Rodríguez–Casal, A. (2014)
CircSiZer: an exploratory tool for circular data.
Journal of Environmental and Ecological Statistics, 21, 143–159.

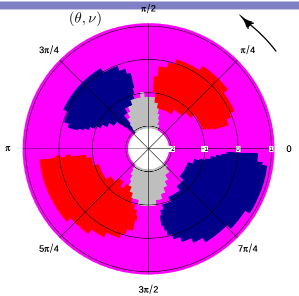
Construction of CircSiZer map



For each pair (θ, ν), with θ varying in $[0, 2\pi)$ and $\nu > 0$:

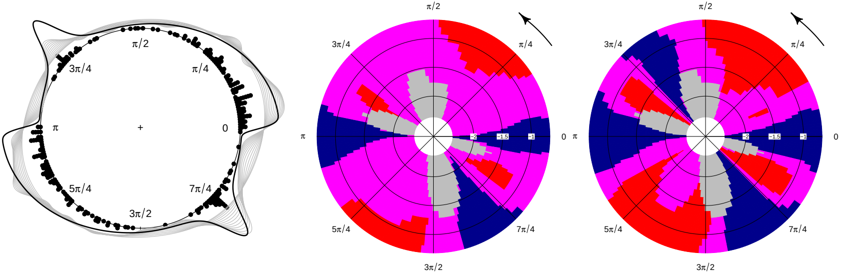
- Compute the confidence interval for $f'(\theta; \nu)$.
- If the interval is
  - above zero $\rightarrow$ the smoothed curve is significantly increasing $\rightarrow$ the location corresponding to the pair (θ, ν) is colored blue.
  - below zero $\rightarrow$ the smoothed curve is significantly decreasing $\rightarrow$ the location corresponding to the pair (θ, ν) is colored red.
  - contains zero $\rightarrow$ the derivative is not sig. dif. from zero $\rightarrow$ the location corresponding to the pair (θ, ν) is colored purple.
  - Location $(\theta, -\log_{10}(\nu))$ is coloured gray if there is not enough data.

Construction of CircSiZer map



For each pair (θ, ν), with θ varying in $[0, 2\pi)$ and $\nu > 0$:

- Compute the confidence interval for $f'(\theta; \nu)$.
- If the interval is
  - above zero $\rightarrow$ the smoothed curve is significantly increasing $\rightarrow$ the location corresponding to the pair (θ, ν) is colored blue.
  - below zero $\rightarrow$ the smoothed curve is significantly decreasing $\rightarrow$ the location corresponding to the pair (θ, ν) is colored red.
  - contains zero $\rightarrow$ the derivative is not sig. dif. from zero $\rightarrow$ the location corresponding to the pair (θ, ν) is colored purple.
  - Location $(\theta, -\log_{10}(\nu))$ is coloured gray if there is not enough data.

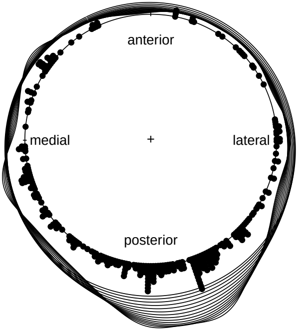


KDEs for a sample with $n = 250$ data. Simultaneous CircSizer map (center) and pointwise CircSizer map (right).

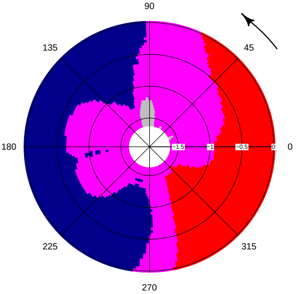
Cracks in cemented femoral components

Is there a preferred direction for cracks in the cement mantle?

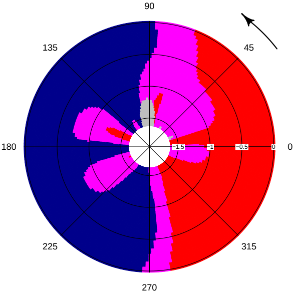
For the proximal region...



Family of kernel density estimates indexed by the smoothing parameter



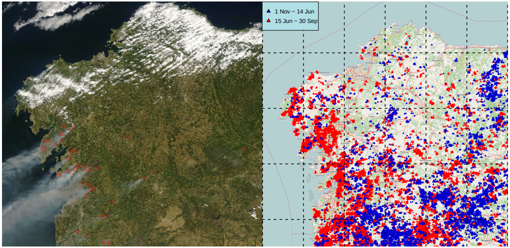
CircSiZer map with simultaneous bootstrap confidence intervals



CircSiZer map with pointwise normal confidence intervals

Remember the example: Fires detected by MODIS

- Data*: Time of occurrence and location of the fires detected by MODIS from 10/07/2002 to 9/07/2012.

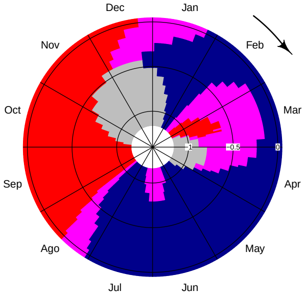


- Aggregating the data at 0.5^o resolution, the objective is find out in which regions the number of fire seasons is greater than the climatological ones.

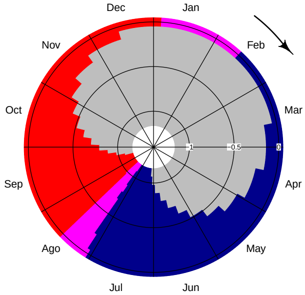
*Data were collected by Prof. José M. C. Pereira and his group from the University of Lisbon (Portugal).

Remember the fires dataset...

CircSiZer maps for two regions in Galicia. Data from July 2002 until July 2012:



Barbanza



Baixo Miño

Wences 3rd 'what':

Wences 3rd 'what': What about doing a test?

I Based on the critical bandwidth

-
-
-
-

II Based on the excess of mass (dip)

-
-
-

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
-
-
-

II Based on the excess of mass (dip)

-
-
-

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
-
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
-
-

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
-
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
-

III Exploratory tools

-
-
-
-

Wences 4th 'what':

Wences 4th 'what': What about calibration?

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
-
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
-

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
-
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

-
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

- Mode tree. Minnote and Scott (1993).
-
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

- Mode tree. Minnote and Scott (1993).
- Mode forest. Minnote *et al.* (1998).
-
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
-
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

- Mode tree. Minnote and Scott (1993).
- Mode forest. Minnote *et al.* (1998).
- SiZer. Chaudhuri and Marron (1999).
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
- Cramér–von Mises. Fisher and Marron (2001).
-

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

- Mode tree. Minnote and Scott (1993).
- Mode forest. Minnote *et al.* (1998).
- SiZer. Chaudhuri and Marron (1999).
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
- Cramér–von Mises. Fisher and Marron (2001).
- Fisher and Marron (2001).

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

III Exploratory tools

- Mode tree. Minnote and Scott (1993).
- Mode forest. Minnote *et al.* (1998).
- SiZer. Chaudhuri and Marron (1999).
-

I Based on the critical bandwidth

- Critical bandwidth. Silverman (1981).
- Calibration of the critical bandwidth. Hall and York (2001).
- Cramér–von Mises. Fisher and Marron (2001).
- Fisher and Marron (2001).

II Based on the excess of mass (dip)

- Dip test. Hartigan and Hartigan (1985).
- Excess of mass. Müller and Sawitzki (1991).
- Calibration of the excess of mass. Cheng and Hall (1998).

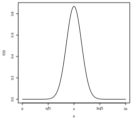
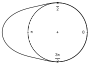
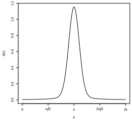
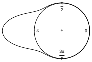
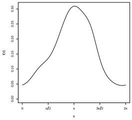
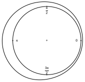
III Exploratory tools

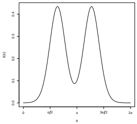
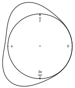
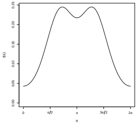
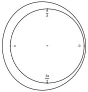
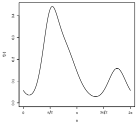
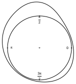
- Mode tree. Minnote and Scott (1993).
- Mode forest. Minnote *et al.* (1998).
- SiZer. Chaudhuri and Marron (1999).
- CircSiZer. Oliveira *et al.* (2012).

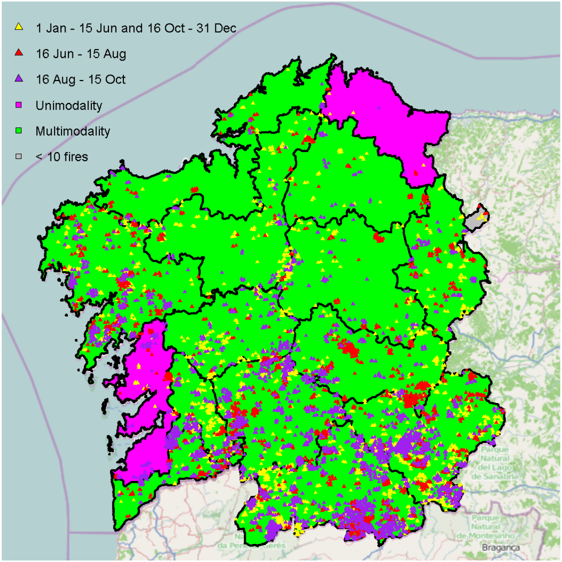
Our proposal, in a nutshell: f unkown density with j modes.

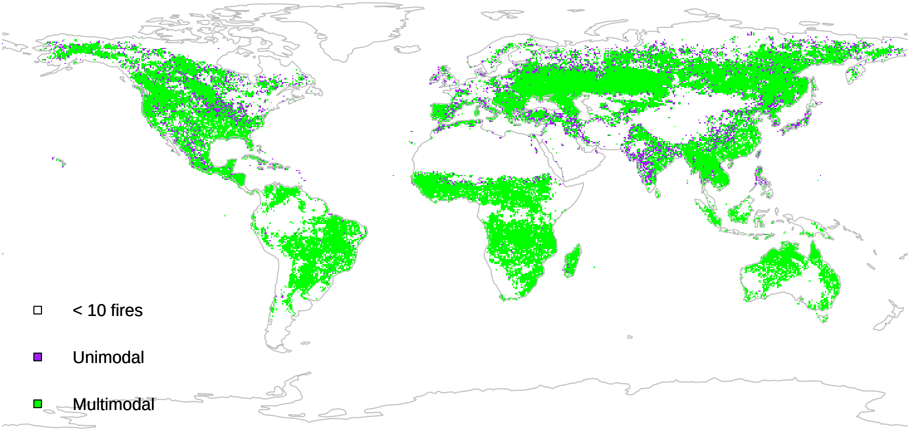
$$H_0: \ j = 1 \quad \text{vs.} \quad H_a: \ j > 1$$

- Consider the excess mass statistic. (Independent of unknowns except for a factor which includes f and f'' at modes and antimodes).
- Estimate the unknowns using kernel estimators with critical bandwidth (for f).
- Calibrate by Bootstrap, with generated samples from a (modified version of) kernel estimator with critical bandwidth.

		α	0.01	0.05	0.10
			Fisher and Marron		
		$n = 50$	0(0)	0.022(0.013)	0.068(0.022)
		$n = 200$	0(0)	0.012(0.010)	0.058(0.020)
		$n = 1000$	0.004(0.006)	0.008(0.008)	0.032(0.015)
			Excess mass		
		$n = 50$	0.012(0.010)	0.054(0.020)	0.094(0.026)
		$n = 200$	0.008(0.008)	0.038(0.017)	0.092(0.025)
		$n = 1000$	0.006(0.007)	0.040(0.017)	0.086(0.025)
			Fisher and Marron		
		$n = 50$	0.370(0.042)	0.520(0.044)	0.598(0.043)
		$n = 200$	0.648(0.042)	0.776(0.037)	0.848(0.031)
		$n = 1000$	0.498(0.044)	0.648(0.042)	0.728(0.039)
			Excess mass		
		$n = 50$	0.008(0.008)	0.044(0.018)	0.104(0.027)
		$n = 200$	0.014(0.010)	0.036(0.016)	0.076(0.023)
		$n = 1000$	0.012(0.010)	0.040(0.017)	0.074(0.023)
			Fisher and Marron		
		$n = 50$	0.004(0.006)	0.030(0.015)	0.062(0.021)
		$n = 200$	0.004(0.006)	0.032(0.015)	0.052(0.019)
		$n = 1000$	0.002(0.004)	0.018(0.012)	0.048(0.019)
			Excess mass		
		$n = 50$	0(0)	0.024(0.013)	0.044(0.018)
		$n = 200$	0.006(0.007)	0.044(0.018)	0.100(0.026)
		$n = 1000$	0.010(0.009)	0.048(0.019)	0.100(0.026)

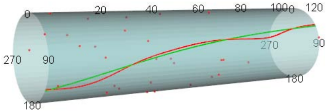
		α	0.01	0.05	0.10
			Fisher and Marron		
		$n = 50$	0.782(0.036)	0.920(0.024)	0.958(0.018)
		$n = 100$	0.978(0.013)	0.996(0.006)	0.998(0.004)
		$n = 200$	1(0)	1(0)	1(0)
			Excess mass		
		$n = 50$	0.534(0.044)	0.758(0.038)	0.854(0.031)
		$n = 100$	0.914(0.025)	0.968(0.015)	0.984(0.011)
		$n = 200$	0.996(0.006)	1(0)	1(0)
			Fisher and Marron		
		$n = 50$	0.014(0.010)	0.070(0.022)	0.130(0.029)
		$n = 100$	0.038(0.017)	0.124(0.029)	0.196(0.035)
		$n = 200$	0.066(0.022)	0.170(0.033)	0.246(0.038)
			Excess mass		
		$n = 50$	0.022(0.013)	0.072(0.023)	0.140(0.030)
		$n = 100$	0.032(0.015)	0.084(0.024)	0.156(0.032)
		$n = 200$	0.044(0.018)	0.102(0.027)	0.204(0.035)
			Fisher and Marron		
		$n = 50$	0.476(0.044)	0.730(0.039)	0.840(0.032)
		$n = 100$	0.828(0.033)	0.952(0.019)	0.976(0.013)
		$n = 200$	0.990(0.009)	0.998(0.004)	1(0)
			Excess mass		
		$n = 50$	0.044(0.018)	0.164(0.032)	0.284(0.040)
		$n = 100$	0.208(0.036)	0.438(0.043)	0.554(0.044)
		$n = 200$	0.578(0.043)	0.796(0.035)	0.870(0.029)

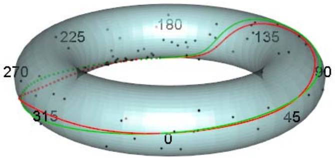

1 Jan - 15 Jun and 16 Oct - 31 Dec
16 Jun - 15 Aug
16 Aug - 15 Oct
Unimodality
Multimodality
< 10 fires
Parque Natural del Lago de Sanabria
Parque Natural de Montesinho
Bragança


< 10 fires
Unimodal
Multimodal

When a circular meets a linear...

and when a circular meets another circular...





García–Portugués, E., Crujeiras, R.M. and González–Manteiga, W. (2013)
Exploring wind direction and SO2 concentration by circular–linear density estimation.
Stochastic Environmental Research and Risk Assessment, 27, 1055–1067.

García–Portugués, E., Crujeiras, R.M. and González–Manteiga, W. (2013)
Kernel density estimation for directional–linear data.
Journal of Multivariate Analysis, 121, 152–175.

García–Portugués, E. (2014)
Exact risk improvement of bandwidth selectors for kernel density estimation with directional data.
Electronic Journal of Statistics, 7, 1655–1685.

García–Portugués, E., Barros, A.M.G., Crujeiras, R.M., González–Manteiga, W. and Pereira, J. (2014)
A test for directional–linear independence, with applications to wildfire orientation and size.
Stochastic Environmental Research and Risk Assessment, 28, 1261–1275.

García–Portugués, E., Crujeiras, R.M. and González–Manteiga, W. (2015)
Central limit theorems for directional and linear data with applications.
Statistica Sinica, 25, 1207–1229.

Summarizing...

Summarizing...

- We've had some fun with kernel methods and circular data.

Summarizing...

- We've had some fun with kernel methods and circular data.
- We've selected bandwidths!

Summarizing...

- We've had some fun with kernel methods and circular data.
- We've selected bandwidths!
- We've run Bootstraps!!!

Summarizing...

- We've had some fun with kernel methods and circular data.
- We've selected bandwidths!
- We've run Bootstraps!!!
- We're even working on a test!!!!

Summarizing...

- We've had some fun with kernel methods and circular data.
- We've selected bandwidths!
- We've run Bootstraps!!!
- We're even working on a test!!!!
- ... but there are many questions to answer...

Summarizing...

- We've had some fun with kernel methods and circular data.
- We've selected bandwidths!
- We've run Bootstraps!!!
- We're even working on a test!!!!
- ... but there are many questions to answer...

Thanks for your attention!