A versatile trivariate wrapped Cauchy copula with applications to toroidal and cylindrical data

Christophe Ley

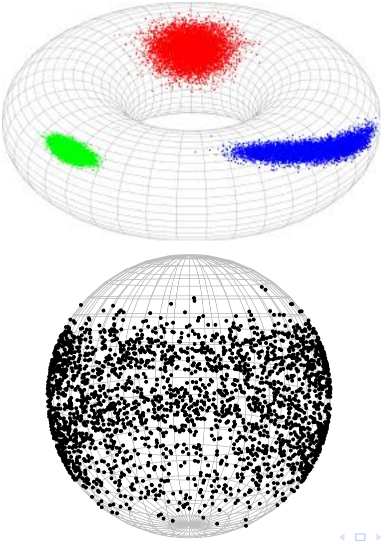
University of Luxembourg

joint work with Shogo Kato (Institute of Statistical Mathematics), Sophia Loizidou (University of Luxembourg) and Kanti Mardia (Leeds University)

StatMathAppli2025

September 4 2025

Directional data



Plan

Models for angular data

The literature

- is full of probability distributions for circular data (1 angle)

Models for angular data

The literature

- is full of probability distributions for circular data (1 angle)
- The literature also contains several models for toroidal data (2 angles)

Models for angular data

The literature

- is full of probability distributions for circular data (1 angle)
- The literature also contains several models for toroidal data (2 angles)
- The literature also contains several models for cylindrical data (1 angle and 1 linear component)

Models for angular data

The literature

- is full of probability distributions for circular data (1 angle)
- The literature also contains several models for toroidal data (2 angles)
- The literature also contains several models for cylindrical data (1 angle and 1 linear component)

But scarcity for data consisting of 3 angles, 2 angles and 1 linear component, 1 angle and 2 linear components

Motivating examples

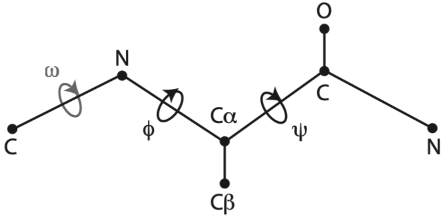


Fig. 1. The ϕ, ψ angular degrees of freedom in one residue of the protein backbone. The ω dihedral angle can be assumed to be fixed at 180° (*trans*) or 0° (*cis*).

Pairs of dihedral angles ϕ and ψ are data on the torus. However, there is a third, not-modelled variable: the torsion angle ω!

Come back for the talk tomorrow by Sophia Loizidou!

Motivating examples

Many environmental agencies are collecting data on wave heights and directions in order to identify sea regimes. This is not enough: wind direction should be added.

Motivating examples

Many environmental agencies are collecting data on wave heights and directions in order to identify sea regimes. This is not enough: wind direction should be added.

Recent Environmetrics paper: "In wintertime, relevant events in the Adriatic Sea are typically generated by the southeastern Sirocco wind and the northern Bora wind"

Our solution

We propose the trivariate wrapped Cauchy copula for angular as well as linear components!

Plan

Our new trivariate copula model

We propose the density

$$c(u_1, u_2, u_3) = c_2\left[c_1+2\left\{\rho_{12}\cos(u_1-u_2)+\rho_{13}\cos(u_1-u_3)+\rho_{23}\cos(u_2-u_3)\right\}\right]^{-1}, \quad (1)$$

with $0 \leq u_1, u_2, u_3 < 2\pi$ and where $\rho_{12}, \rho_{13}, \rho_{23} \in \mathbb{R} \setminus \{0\}$, $\rho_{12}\rho_{13}\rho_{23} > 0$,

$$c_1 = \frac{\rho_{12}\rho_{13}}{\rho_{23}} + \frac{\rho_{12}\rho_{23}}{\rho_{13}} + \frac{\rho_{13}\rho_{23}}{\rho_{12}}, \quad (2)$$

and

$$c_2 = \frac{1}{(2\pi)^3}\left\{\left(\frac{\rho_{12}\rho_{13}}{\rho_{23}}\right)^2 + \left(\frac{\rho_{12}\rho_{23}}{\rho_{13}}\right)^2 + \left(\frac{\rho_{13}\rho_{23}}{\rho_{12}}\right)^2 - 2\rho_{12}^2 - 2\rho_{13}^2 - 2\rho_{23}^2\right\}^{1/2}. \quad (3)$$

Moreover, the parameters satisfy

$$|\rho_{jk}| < |\rho_{ij}\rho_{ik}|/(|\rho_{ij}| + |\rho_{ik}|),$$

where $\rho_{ji} = \rho_{ij}$ for $1 \leq i < j \leq 3$ for a permutation of $(1, 2, 3)$, (i, j, k).

Identifiability

Proposition

Assume that the parameter space of the family (1) is constrained to the condition $\rho_{12}\rho_{13}\rho_{23} = \beta(> 0)$. With this restriction, the family of distributions (1) is identifiable.

- One-dimensional marginal densities are uniform on the circle
- Conditional as well as two-dimensional marginal densities are well-known densities on the torus and circle

Simple random variate generation mechanism

Theorem

The following algorithm generates random variates from the distribution (1) without rejection.

1. *Generate uniform* $(0,1)$ *random variates* ω_1, ω_2 *and* ω_3*.*
2. *Compute*

$$u_1 = 2\pi\omega_1, \quad u_2 = u_1 + \eta_{12} + 2\arctan\left[\left(\frac{1-\delta_{12}}{1+\delta_{12}}\right)\tan\{\pi(\omega_2 - 0.5)\}\right],$$

$$u_3 = \eta_{3|12} + 2\arctan\left[\left(\frac{1-\delta_{3|12}}{1+\delta_{3|12}}\right)\tan\{\pi(\omega_3 - 0.5)\}\right],$$

   where (η_{12}, δ_{12}) *and* $(\eta_{3|12}, \delta_{3|12})$ *have simple forms.*
3. *Record* (u_1, u_2, u_3) *as the random variate from the distribution (1).*

Unimodality, modes and antimodes

Theorem

For $\rho_{jk} > 0$ and $\rho_{j\ell}, \rho_{k\ell} < 0$, the modes of the density (1) are given by

$$\begin{array}{lll} \text{(i)} & u_j = u_k = u_\ell & \text{if } |\rho_{jk}| < |\rho_{j\ell}\rho_{k\ell}|/(|\rho_{j\ell}| + |\rho_{k\ell}|), \\ \text{(ii)} & u_j = u_k + \pi = u_\ell + \pi & \text{if } |\rho_{j\ell}| < |\rho_{jk}\rho_{k\ell}|/(|\rho_{jk}| + |\rho_{k\ell}|), \\ \text{(iii)} & u_j = u_k + \pi = u_\ell & \text{if } |\rho_{k\ell}| < |\rho_{jk}\rho_{j\ell}|/(|\rho_{jk}| + |\rho_{j\ell}|), \end{array} \quad (4)$$

and the antimodes of the density (1) are given by

$$u_j = u_k = u_\ell + \pi. \quad (5)$$

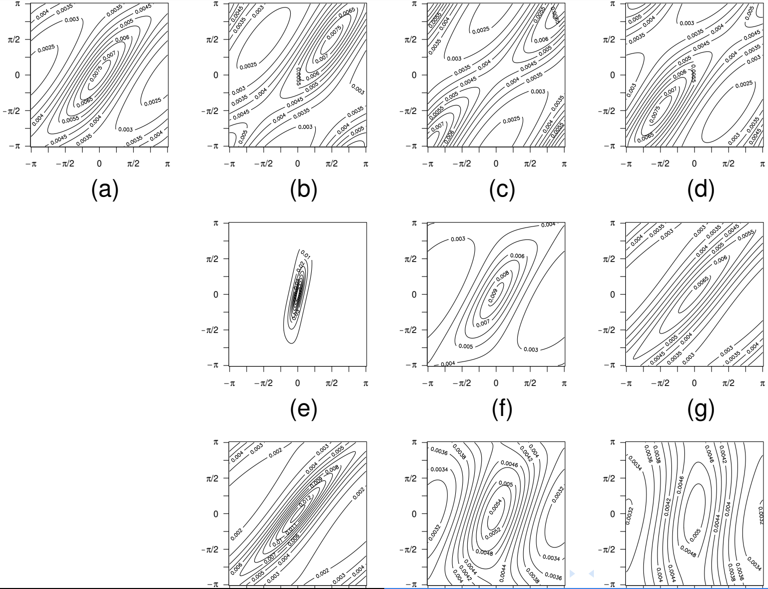
If $\rho_{jk}, \rho_{j\ell}, \rho_{k\ell} > 0$, then u_ℓ in the modes (4) and antimodes (5) is replaced by $u_\ell + \pi$.

Moment expressions and correlation

We have obtained tractable trigonometric moment expressions, which in some cases become very simple.

We have calculated different correlation coefficients which turn out to be very concise expressions (either $|\phi_{jk}|$ or ϕ_{jk}^2).

Shapes



Parameter estimation

Method of moments or maximum likelihood - needs to be done numerically.

For the algorithm, see Sophia's GitHub repository:
https://github.com/Sophia-Loizidou/ Trivariate-wrapped-Cauchy-copula.

Consistency and finite-sample performance are established by means of Monte Carlo simulations.

Plan

Unleashing the copula power

Let (U_1, U_2, U_3) be a random vector which follows the trivariate wrapped Cauchy copula (1). Assume that either

- f_j is a density on the circle $[0, 2\pi)$ $(j = 1, 2, 3)$ and F_j its distribution function ;
- f_j is a density on (a subset of) the real line $\mathbb{R}$ $(j = 1, 2, 3)$ and F_j its distribution function.

Define

$$(\Theta_1, \Theta_2, \Theta_3) = \left(F_1^{-1}\left(\frac{U_1}{2\pi}\right), F_2^{-1}\left(\frac{U_2}{2\pi}\right), F_3^{-1}\left(\frac{U_3}{2\pi}\right)\right).$$

Then it follows that $(\Theta_1, \Theta_2, \Theta_3)$ has the joint density

$$f(\theta_1, \theta_2, \theta_3) = (2\pi)^3 c_2 \Bigg(c_1 + 2 \sum_{1 \le j < k \le 3} \rho_{jk} \cos[2\pi\{F_j(\theta_j) - F_k(\theta_k)\}]\Bigg)^{-1} \prod_{1 \le \ell \le 3} f_\ell(\theta_\ell), \tag{6}$$

where each θ_j either belongs to $[0, 2\pi)$ or to (a subset of) $\mathbb{R}$.

Data analysis: 2 angles and 1 linear component

We consider a time series of 1326 observations of semi-hourly wave directions and heights, recorded in the period 15/02/2010 – 16/03/2010 by the buoy of Ancona, located in the Adriatic Sea at about 30 km from the coast.

The data points are collected far spaced enough from each other to be considered independent.

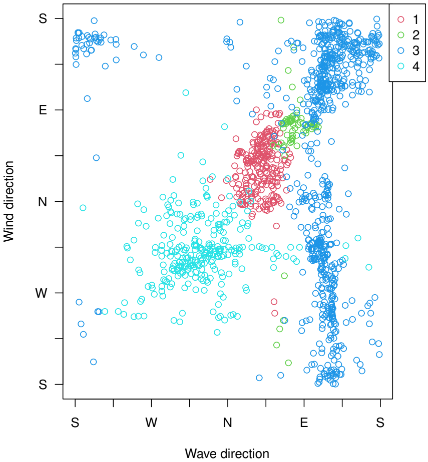
We choose as circular marginals the wrapped Cauchy distribution and the Weibull distribution for the linear marginal.

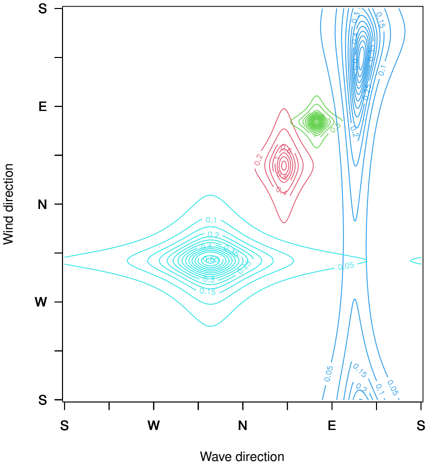
Due to the multimodality of the data, a mixture model is required, leading to densities of the form

$$f(\theta_1, \theta_2, x) = \sum_{i=1}^{K} \pi_i f_i(\theta_1, \theta_2, x), \quad \sum_{i=1}^{K} \pi_i = 1, \tag{7}$$

where K is the number of components of the mixture model, π_i is the weight of each class.

In order to estimate the parameters, we use a variant of the Expectation-Maximization algorithm.





- Bora wind: north/north-east
- Sirocco wind: south-east

Conclusion

We have proposed with the trivariate wrapped Cauchy copula a model that

- works with both angles and linear components
- is highly flexible yet tractable
- has well-known conditional and marginal distributions
- has a simple data-generating mechanism and well-identified modes
- has well-estimatable parameters
- improves on competitors from the literature
- allows tackling essential real-data problems

Extensions to higher dimensions and time series models are planned.