Some mixture models for joint analysis of wind speed and wind direction

Debarghya Jana[1] and Arnab Hazra[2]

Iowa State University[1]

Indian Institute of Technology Kanpur[2]

IISA conference, Nebraska
June 15, 2025

Outline

Table of Contents

Motivation & Introduction

- Our Primary motivation of this work came from the shocking incident of cyclone Amphan in 2020 which was the strongest Bay of Bengal storm in 21 years, made the North Indian Ocean cyclone season the costliest ever, causing massive destruction and at least 19 deaths in Kolkata *(Economic Times, 2020)*.

- In fact, the North Indian Ocean annually generates about 5.5 tropical cyclones, mostly in the Bay of Bengal and contributing roughly 7% to the global total.

Motivation & Introduction

- By jointly modelling wind speed and direction in this context using appropriate bivariate or circular-linear distributions, we can effectively capture the underlying dependency structure, anisotropic behavior, and rotational patterns present in cyclone systems.

- Furthermore, joint models of wind speed and wind dirction for cyclones are essential when simulating wind fields over the spatial domains.

Motivation & Introduction

- Now, there are few statistical models where joint modelling of wind speed and wind direction has been done in **parametric** and **nonparametric** set up but there was almost no **semiparametric** approach for joint modelling of wind speed and wind direction.

- So, we proposed two mixture models — **Isotropic Gaussian Mixture (IGM)** and **Anisotropic Gaussian Mixture (AGM)** models for jointly modeling wind speed and direction.

Table of Contents

Data and Exploratory analysis

- Data has been collected from International Best Track Archive for Climate Stewardship (IBTrACS), hosted by the National Climatic Data Center (NCDC) at www.ncdc.noaa.gov/oa/ibtracs/.

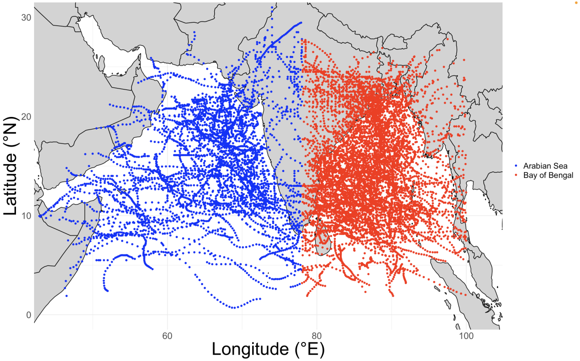

Figure: Cyclones in the North Indian Ocean basin between 1980–2022.

Data and Exploratory analysis

- Initial exploratory data analysis reveal that the Arabian Sea experiences roughly one-fourth the cyclone frequency of the Bay of Bengal, with a notable increase in cyclone occurrences in the NIO basin since 1990.

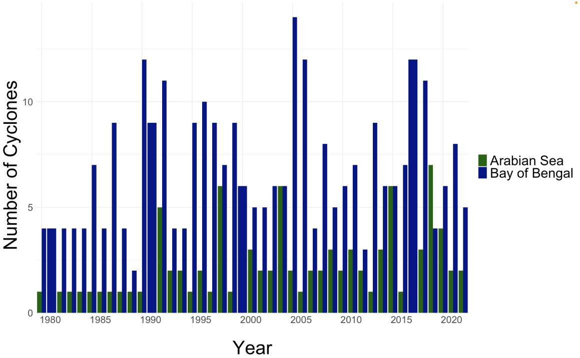

Figure: Cyclones in the North Indian Ocean basin between 1980–2022.

Table of Contents

Methodology

- We proposed two semiparametric models, Isotropic Gaussian mixture (AGM) and Anisotropic Gaussian mixture(IGM) model which were developed by using the parametric models Isotropic Gaussian model *[McWilliams and Sprevak, 1980]* and Anisotropic Gaussian model *[Weber, 1991]* respectively.
- **Isotropic Gaussian model (IG)** *[McWilliams and Sprevak, 1980]*: The joint pdf of speed (V') and angle (Θ'') can be expressed as,

$$f_{V',\Theta''}(v',\theta') = \frac{\exp(2v')}{2\pi\sigma^2_{y'}} \exp\left[-\frac{\mu^2_{y'}}{2\sigma^2_{y'}}\right] \exp\left[-\frac{\exp(2v') - 2\mu_{y'}\exp(v')\cos(\theta''+\pi)}{2\sigma^2_{y'}}\right] \exp(v') \quad (1)$$

for $v' \in \mathbb{R}$ and $-\pi < \theta'' < \pi$ and μ'_y is the mean and σ'_y is the variance of the longitudinal component.

Methodology

- **Isotropic Gaussian Mixture(IGM) model:** We proposed the joint pdf of wind speed (V') and wind direction (Θ'') as

$$f_{V',\Theta''}(v',\theta') = \sum_{j=1}^{C} \phi_j f^{(j)}_{V',\Theta''}\left(v',\theta'' \mid \mu^{(j)}_{y'}, (\sigma^{(j)}_{y'})^2\right), v' \in \mathbb{R}, -\pi < \theta'' < \pi \tag{2}$$

Here, $\phi_{i=1,2,\ldots,C}$ is the mixture weight and $f^{(j)}_{V',\Theta''}\left(v',\theta'' \mid \mu^{(j)}_{y'}, \sigma^{(j)}_{y'}\right)$ denotes the joint pdf of wind speed and wind direction in Isotropic Gaussian(IG) model respectively for the mixture $\{j\}_{\{1,2,\ldots,C\}}$.

Methodology

Algorithm 1 EM Algorithm for Mixture of IG model to find estimates of the unknown parameters

[1] $\boldsymbol{\theta_{(0)}} = \left(\phi_{1,(0)}, \dots, \phi_{C,(0)}, \mu_{y'}^{(1,(0))}, \dots, \mu_{y'}^{(C,(0))}, \sigma_{y'}^{(1,(0))}, \dots, \sigma_{y'}^{(C,(0))}\right)$ [Using K-means clusturing method]

[2] **E-step:** For all $c = 1, 2 \dots, C$ and all $i = 1, \dots, n$, calculate

$$\gamma_{i,c,k} = \frac{f_{V',\Theta''}^{(c)}\left(v', \theta'' \mid \mu_{y'}^{(c,k)}, (\sigma_{y'}^{(c,k)}\right)\phi_{c,k}}{\sum_{j=1}^{C} f_{V',\Theta''}^{(j)}\left(v', \theta'' \mid \mu_{y'}^{(j,k)}, (\sigma_{y'}^{(j,k)})^2\right)\phi_{j,k}},.$$

and

$$q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta_{(k)}}\right) = \sum_{i=1}^{n}\sum_{c=1}^{C} \log\left\{f_{V',\Theta''}^{(c)}\left(v', \theta'' \mid \mu_{y'}^{(c)}, (\sigma_{y'}^{(c)})^2\right)\phi_c\right\}\gamma_{i,c,k}$$

[3] **M-step:**

$$\boldsymbol{\theta_{(k+1)}} = \arg\max q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta_{(k)}}\right)$$

[4] Stop when $\|\boldsymbol{\theta_{(k+1)}} - \boldsymbol{\theta_{(k)}}\| < \epsilon$.

Methodology

Anisotropic Gaussian model (AG) *[Weber, 1991]:* The joint pdf of wind speed (V') and wind direction (Θ'') where, μ'_x is the mean of the lateral component, μ'_y is the mean and σ'_y is the variance of the longitudinal component can be expressed as

$$f_{V',\Theta''}(v',\theta'') = \frac{\exp(2v')}{2\pi\sigma_{y'}\sigma_{x'}} \exp\left[-\frac{(\exp(v)\cos(\theta'') - \mu_{y'})^2}{2\sigma_{y'}^2} - \frac{v^2\sin^2(\theta''+\pi)}{2\sigma_{x'}^2}\right], v' \in \mathbb{R}, -\pi < \theta'' < 2\pi \quad (3)$$

Methodology

Anisotropic Gaussian Mixture(AGM) model: Here, We proposed the joint pdf of wind speed (V') and wind direction (Θ'') as

$$f_{V',\Theta''}(v',\theta') = \sum_{j=1}^{2} \pi_j f^{(j)}_{V',\Theta''}\left(v',\theta'' \mid \mu^{(j)}_{x'},\sigma^{(j)}_{x'},\sigma^{(j)}_{y'}\right), v' \in \mathbb{R}, -\pi < \theta'' < \pi \tag{4}$$

Here, $\phi_{i=1,2,\ldots,C}$ is the mixture weight and $f^{(j)}_{V',\Theta''}\left(v',\theta'' \mid \mu^{(j)}_{x'},\sigma^{(j)}_{x'},\sigma^{(j)}_{y'}\right)$ denotes the joint pdf of wind speed and wind direction in Anisotropic Gaussian(AG) model in 2 respectively for the mixture $\{j\}_{\{1,2,\ldots,C\}}$.

Methodology

Algorithm 2 EM Algorithm for Mixture of AG model to find estimates of unknown parameters

[1] $\boldsymbol{\theta_{(0)}} = \left(\phi_{1,(0)}, \ldots, \phi_{C,k}, \mu_{y'}^{(1,(0))}, \ldots, \mu_{y'}^{(C,(0))}, \sigma_{x'}^{(1,(0))}, \ldots, \sigma_{x'}^{(C,(0))} \sigma_{y'}^{(1,(0))}, \ldots, \sigma_{y'}^{(C,(0))}\right)$ [Using K-means clustering method]

[2] **E-step:** For all $c = 1, 2$ and all $i = 1, \ldots, n$, calculate

$$\gamma_{i,c,k} = \frac{f_{V',\Theta''}^{(c)}\left(v', \theta'' \mid \mu_{y'}^{(c,k)}, \sigma_{x'}^{(c,k)} \sigma_{y'}^{(c,k)}\right) \phi_{c,k}}{\sum_{j=1}^{2} f_{V',\Theta''}^{(j)}\left(v', \theta'' \mid \mu_{y'}^{(j,k)}, \sigma_{x'}^{(j,k)}\right) \sigma_{y'}^{(j,k)}} \phi_{j,k}, .$$

and

$$q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta_{(k)}}\right) = \sum_{i=1}^{n} \sum_{c=1}^{C} \log \left\{ f_{V',\Theta''}^{(c)}\left(v', \theta'' \mid \mu_{y'}^{(c)}, \sigma_{x'}^{(c)} \sigma_{y'}^{(c)}\right) \phi_c \right\} \gamma_{i,c,k}$$

[3] **M-step:**

$$\boldsymbol{\theta_{(k+1)}} = \arg\max q\left(\boldsymbol{\theta} \mid \boldsymbol{\theta_{(k)}}\right)$$

[4] Stop when $\|\boldsymbol{\theta_{(k+1)}} - \boldsymbol{\theta_{(k)}}\| < \epsilon$.

Table of Contents

Result

Acknowledgement

Thank you!