A Statistical Analysis of North Indian Ocean Cyclones

Debarghya Jana   Arnab Hazra
Department Of Mathematics and Statistics, IIT Kanpur

Introduction

- A tropical cyclone is a large-scale weather system characterized by a low-pressure center and rotating winds. These cyclones can bring destructive winds and heavy rainfall and pose a critical threat to coastal regions by forming over warm ocean waters near the equator.
- We focus on tropical cyclones generated in the North Indian Ocean Basin (NIO, Bay of Bengal and Arabian Sea sub-basins), which typically form during the pre-monsoon and post-monsoon seasons and mainly impact the countries India, Bangladesh, Myanmar, and Sri Lanka.
- We have explored two non-parametric models (NP-KD and NP-JW) described in [2] to model wind speed and wind direction jointly. Further, we fit a parametric model to model them independently as well.
- We also analyze the spatial point patterns of the cyclogenesis locations.

Data and exploratory analysis

- The track data of tropical cyclones in the North Indian Ocean region were acquired from the International Best Track Archive for Climate Stewardship (IBTrACS), which is hosted by the National Climatic Data Center (NCDC, `www.ncdc.noaa.gov/oa/ibtracs/`).
- The Arabian Sea experiences a lower frequency of cyclones compared to the Bay of Bengal and ratio is approximately 1:4.

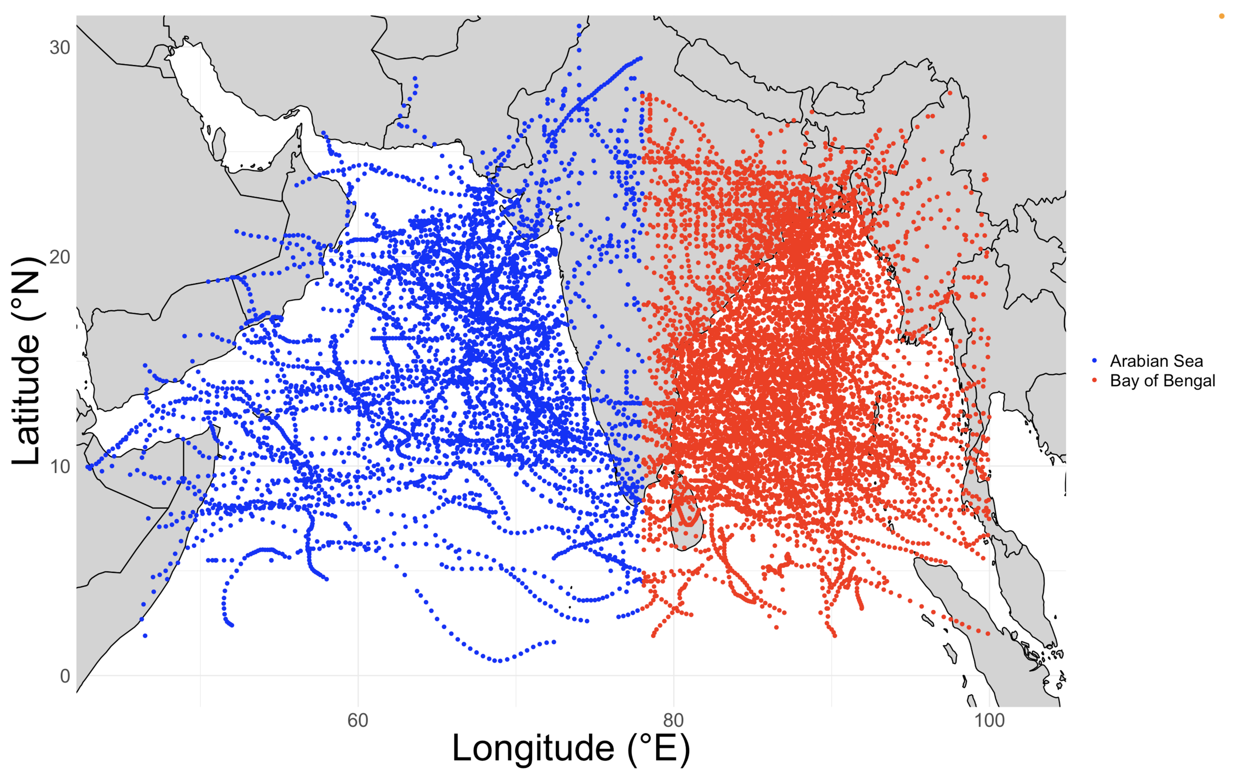


Figure 1. Cyclones in the North Indian Ocean basin between 1980–2022.

- It is evident that there has been a notable increase in the occurrence of cyclones generated in the NIO basin since 1990.

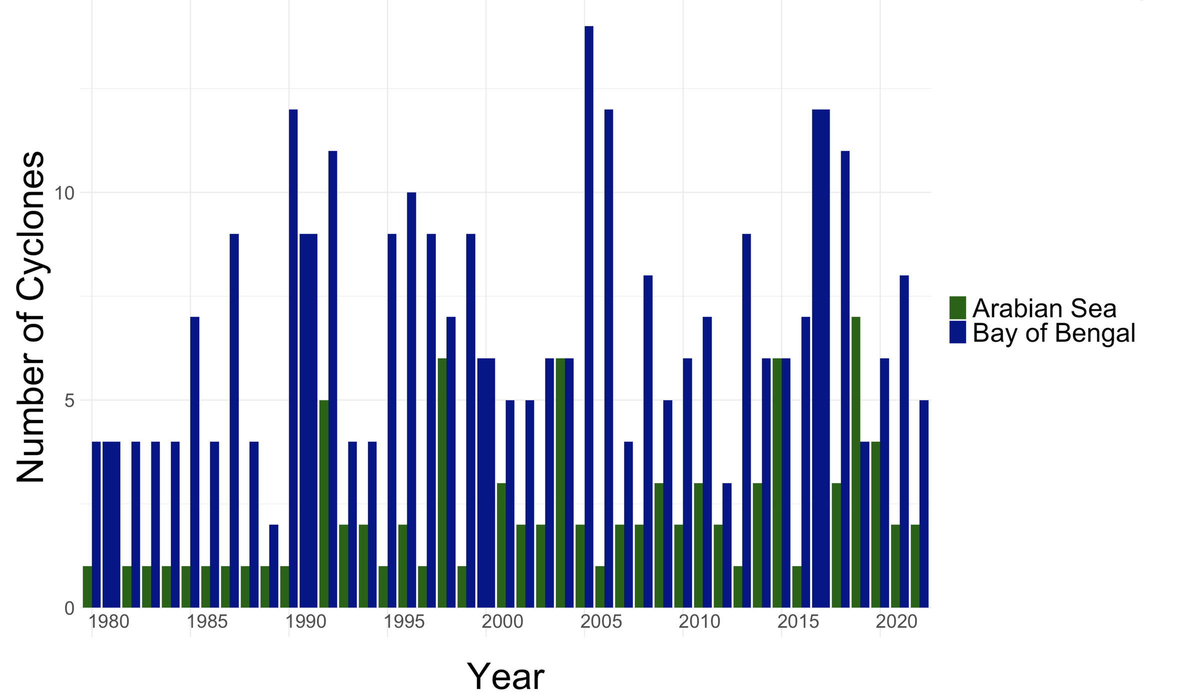


Figure 2. Annual number of Cyclones in NIO Basin between 1980–2022.

Wind speed and wind direction modeling

Our main focus is on exploring the joint probability density function (PDF) of wind speed and wind direction. In this regard, we use two non-parametric models (NP-KD and NP-JW) and one simple parametric model ignoring the dependence between the variables.

- **Non-Parametric Kernel Density (NP-KD) Model:**
Let $(v_1, \theta_1), (v_2, \theta_2), \ldots, (v_n, \theta_n)$ be a random sample drawn from an unknown bivariate population. The joint PDF of wind speed and wind direction is given by

$$f_{V,\Theta}(v, \theta) = \sum_{i=1}^{n} K_{V,\Theta}(v, \theta),$$

where the bivariate kernel $K_{V,\Theta}(v, \theta)$ can be expressed as the product of univariate kernels as $K_{V,\Theta}(v, \theta) = K_V(v) \cdot K_\Theta(\theta)$, where

$$K_V(v) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(v-\mu_v)^2}{2\sigma^2}\right),\ K_\Theta(\theta) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos\left(\theta - \mu_\theta\right)\right).$$

In this equation, σ and κ are the two bandwidth parameters associated with the tensor product kernel function and the term $I_0(\kappa)$ corresponds to the modified Bessel function of the first kind and order zero. For wind speed (v), non-parametric models typically utilize the univariate Gaussian function, and for angular data like wind direction of cyclones (θ), the von Mises kernel function is employed for fitting purposes. Thus,

$$\hat{f}_{V,\Theta}(v, \theta) = \frac{1}{(2\pi)^{3/2} I_0(\kappa)\sigma} \sum_{i=1}^{n} \exp\left(-\frac{(v-v_i)^2}{2\sigma^2}\right) \exp\left(\kappa \cos\left(\theta - \theta_i\right)\right)$$

is the estimated joint PDF of wind speed and wind direction, based on the NP-KD model as desribed in [2].

- **Non-Parametric Johnson-Wehrly (NP-JW) Model:**
The NP-JW model is an extension of the classical JW model ([3]) that combines wind speed and wind direction data into a joint PDF as described in [2]. The expression of the JW model is

$$f_{V,\Theta}(v, \theta) = 2\pi g(\zeta) f_V(v) f_\Theta(\theta) \quad 0 \le \theta < 2\pi, -\infty \le v < \infty,$$

where $f_V(v)$ and $f_\Theta(\theta)$ are the PDFs of wind speed and direction, respectively. Here ζ represents the circular variable between the wind speed and direction, defined as

$$\zeta = \begin{cases} 2\pi\left[F_V(v) - F_\Theta(\theta)\right], & F_V(v) \ge F_\Theta(\theta) \\ 2\pi\left[F_V(v) - F_\Theta(\theta)\right] + 2\pi, & F_V(v) < F_\Theta(\theta), \end{cases}$$

where $F_V(v)$ and $F_\Theta(\theta)$ are the distribution functions of wind speed and direction, respectively, and $g(\zeta)$ represents the PDF of ζ. In the NP-JW model, the wind speed and direction data are fitted using univariate Gaussian and von Mises kernel functions, respectively. By estimating the PDFs, the corresponding distribution functions for wind speed and direction are computed.

- **Independent parametric modeling:**
The joint PDF for wind speed and wind direction, incorporating the von Mises distribution for wind direction and the Gaussian distribution for wind speed given wind direction, can be expressed as

$$f\left(S, \Theta; \mu_S, \sigma_S, \mu_\Theta, \kappa\right) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos\left(\Theta - \mu_\Theta\right)\right) \frac{1}{\sigma_S\sqrt{2\pi}} \exp\left(-\frac{\left(S-\mu_S\right)^2}{2\sigma_S^2}\right),$$

where $0 \le \theta < 2\pi, v \in \mathbb{R}$. Here, S represents the wind speed, Θ represents the wind direction, μ_S represents the mean wind speed, σ_S represents the standard deviation of the wind speed, μ_Θ represents the mean wind direction, and κ represents the concentration parameter of the von Mises distribution. The parameters can be estimated using maximum likelihood estimation.

Cyclogenesis point pattern analysis

- Spatial point patterns refer to the arrangement or distribution of individual points or events in a geographical space. They consist of a set of locations of specific events or objects.
- **Kernel density estimation :** Suppose the n observed data points be denoted as $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \in \mathcal{D} \subset \mathbb{R}^2$, where $\mathcal{D}$ denotes the study domain. Then, the kernel density is the weighted average of the kernels centered at each data point, with the weights representing the contribution of each data point to the density estimation. A common choice for the kernel function is the Gaussian kernel given by:

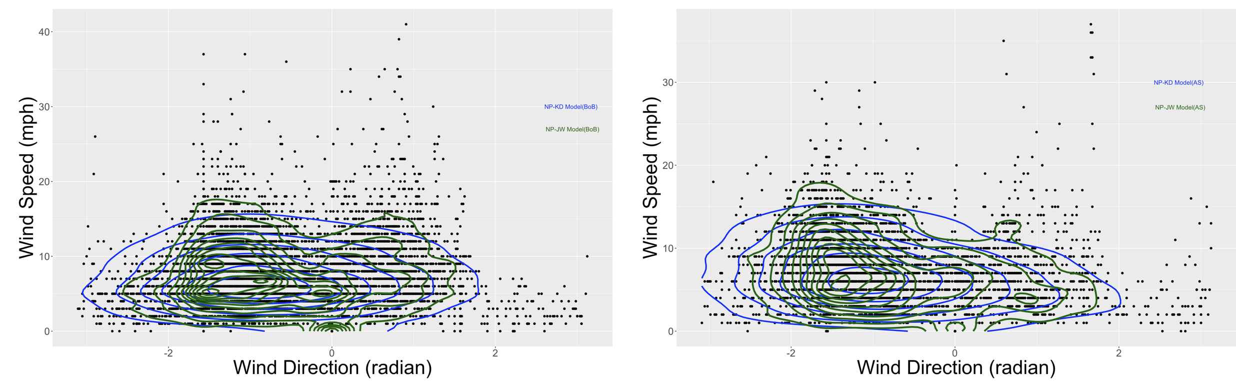
$$K(\boldsymbol{u}; h) = \frac{1}{\sqrt{2\pi h^2}} \exp\left(-\frac{\|u\|^2}{2h^2}\right),$$

where h is the bandwidth, a smoothing parameter. The estimated kernel density at a point $\boldsymbol{u}$ can be expressed as

$$\hat{f}(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n} K\left(\boldsymbol{u} - \boldsymbol{u}_i; h\right).$$

Results

Non-Parametric Models: In each of the joint density contour plot of the NP-KD and NP-JW models, only one cluster is visible. For both models, we observe that the contours closely match the points on the scatter plot. However, the NP-KD model performs slightly better than the NP-JW model in terms of fitting joint distribution. Since the NP-KD model provides the ideal bandwidth, that mode has been captured effectively, and we can only discern one cluster, which was previously discussed in [2]. We can observe a similar pattern of cyclone formation in the Bay of Bengal and Arabian Sea sub-basins.



(a) Contour plot comparison(BoB)   (b) Contour plot comparison(AS)

Figure 3. Contour plots for NP-KD and NP-JW models across the sub-basins.

Joint parametric modeling of storm speed and storm direction: The contour plot depicts the joint density of storm speed and storm direction, exhibiting an elliptical shape. Notably, the variability is more pronounced in the wind speed dimension compared to the wind direction dimension. Additionally, the data does not exhibit multiple modes.

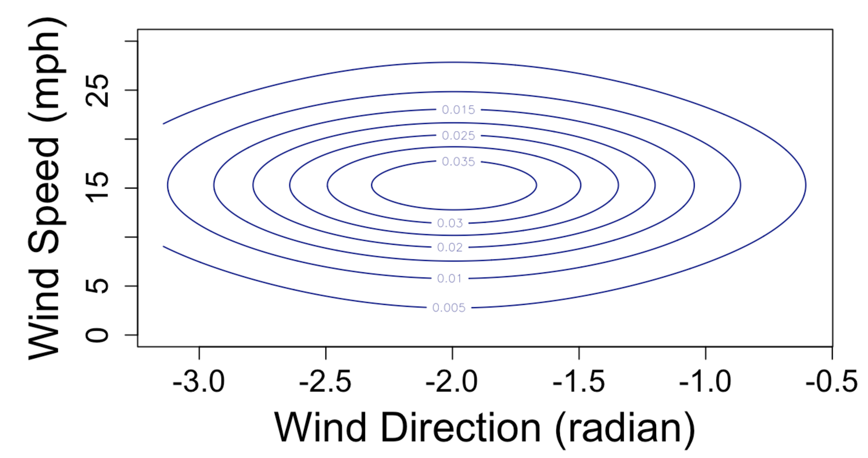
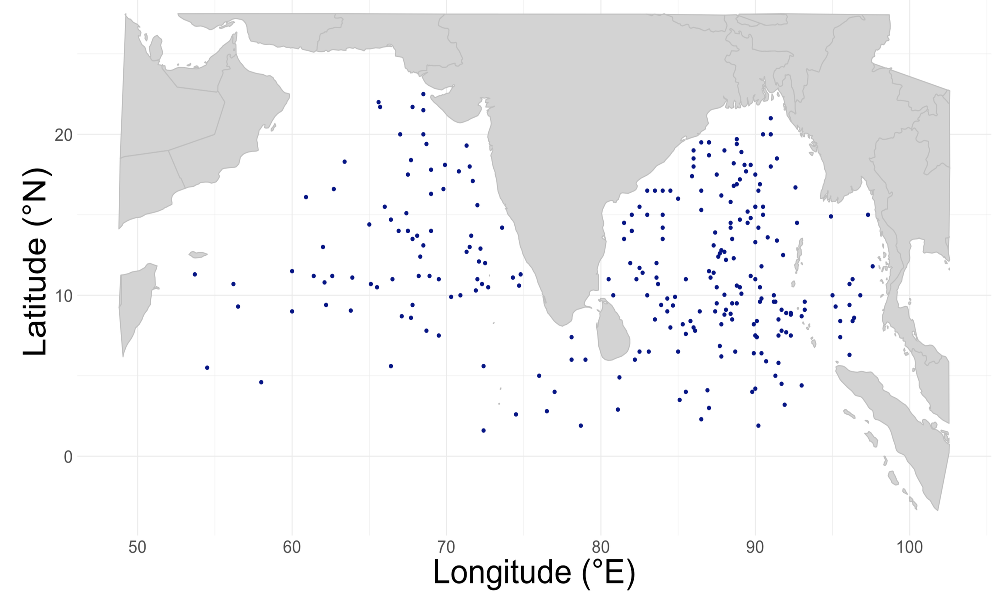


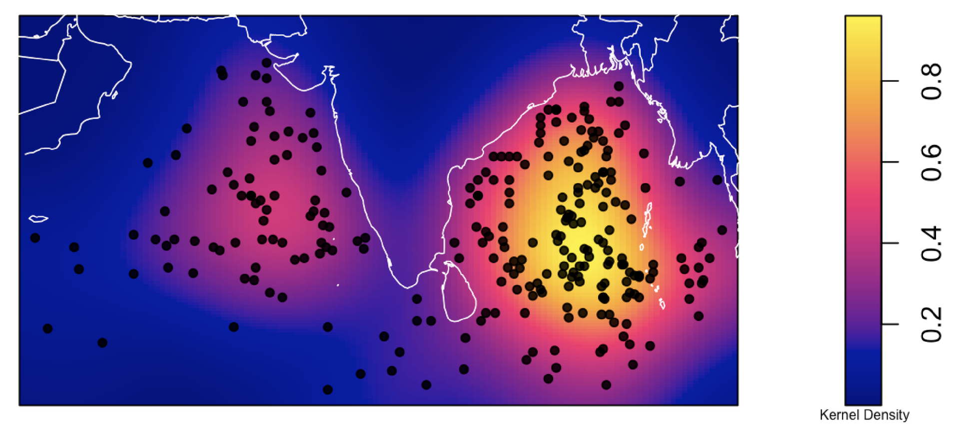
Figure 4. Contour plot of Joint Density of Storm Speed and Storm Direction

Results (contd.)

Spatial Point Pattern Analysis: We observe a clustered point pattern that has mainly two clusters surrounding the regions of Arabian Sea and Bay Of Bengal sub-basins. The kernel density within the Bay of Bengal Basin is higher than that in the Arabian Sea on an average.



(a) Originating locations of cyclones in NIO Basin



(b) Heat map of Kernel density

Figure 5. Spatial Point pattern Analysis

Conclusions and future work

- We observe a similar pattern of joint density of wind speed and wind direction in the Bay of Bengal and Arabian Sea sub-basins.
- There is a significant likelihood of cyclone development in Tamilnadu's coastline region and its surrounding regions in the Bay of Bengal basin. Similar to this, Kerala's coastline region and its neighbouring coastal areas have a higher cyclone tendency.
- We need to model wind directions at different positions of a cyclonic track jointly using a multivariate von Mises distribution.
- Modeling the cyclogenesis point patterns using a Poisson point process would be a future endeavor.

References

[1] Ju Feng and Wen Zhong Shen.
Modelling wind for wind farm layout optimization using joint distribution of wind speed and wind direction.
Energies, 8(4):3075–3092, 2015.

[2] Qinkai Han, Zhuolin Hao, Tao Hu, and Fulei Chu.
Non-parametric models for joint probabilistic distributions of wind speed and direction data.
Renewable Energy, 126:1032–1042, 2018.

[3] Richard A Johnson and Thomas E Wehrly.
Some angular-linear distributions and related regression models.
Journal of the American Statistical Association, 73(363):602–606, 1978.

[4] B McWilliams, MM Newmann, and D Sprevak.
The probability distribution of wind velocity and direction.
Wind Engineering, pages 269–273, 1979.