Stochastic Model; Typhoon Occurrences in the North West Pacific Basin

Student: Peter Hedley
Supervisor: Thomas Kjeldsen

Department of Architecture and Civil Engineering
The University of Bath

1. Introduction

1.1. Data used in the process:

- IBTRAc's Typhoon dataset which can be found at:
https://www.ncdc.noaa.gov/ibtracs/index.php?name=ibtracs-data
- ENSO dataset containing the ENSO signal for given months of a year obtained from:
http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml

1.2. Initial Investigations into the data

Plotted – locations of typhoon formation for each year in a Stat_2d density plot (ggplot2). Combined plots into video – find within Videos directory (using ffmpeg), Figure-1.

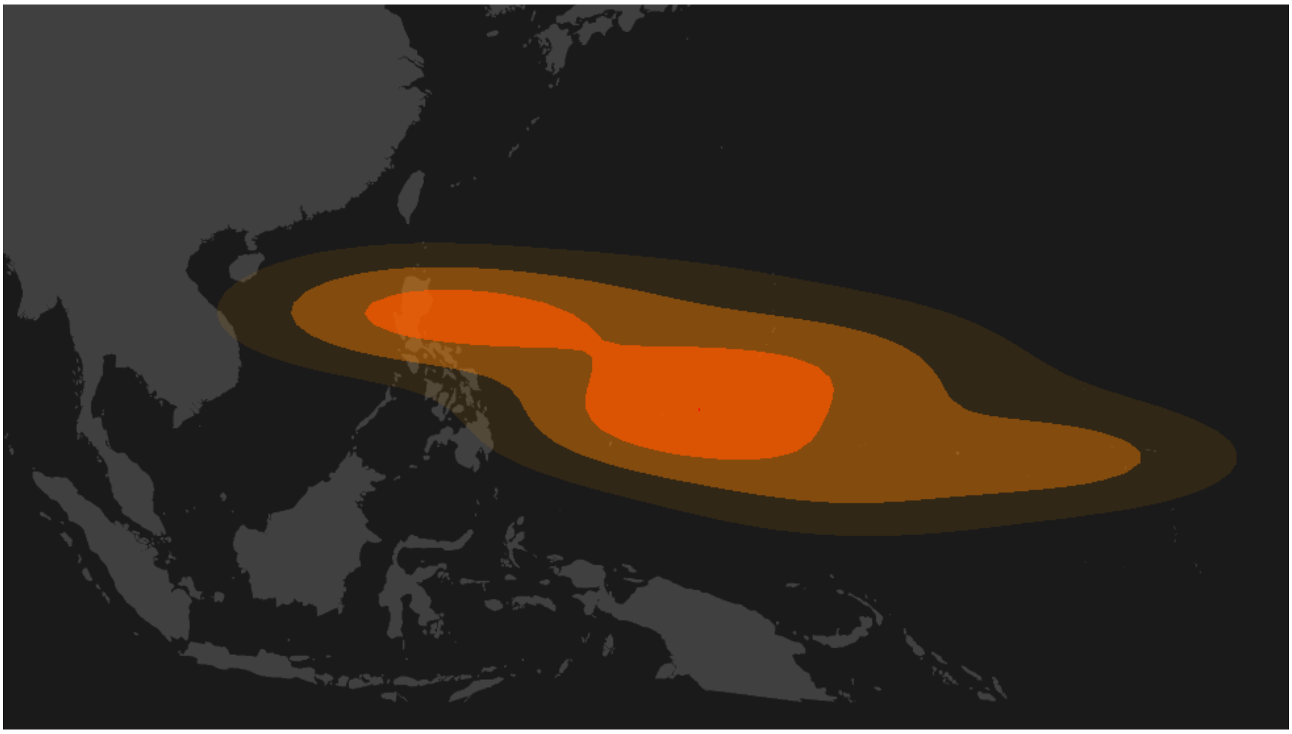


Figure-1 | Typhoon Birth Density (1959)

Note that some typhoons are generated outside of the picture at longitudes > -180. This incorporates some issues with travelling from -180 to +180 that are accounted for in the plots and program. The number of typhoon events vary year on year and are assumed to be approximated by a Poisson model as described later in this document. When the typhoons birth locations are plotted with respect to the month that they occured there is an easily visible difference between the months. As the sea warms in the latter months of the summer typhoons are able to be produced at greater latitudes and hence the density plot moves Northwards, see Figure-2.

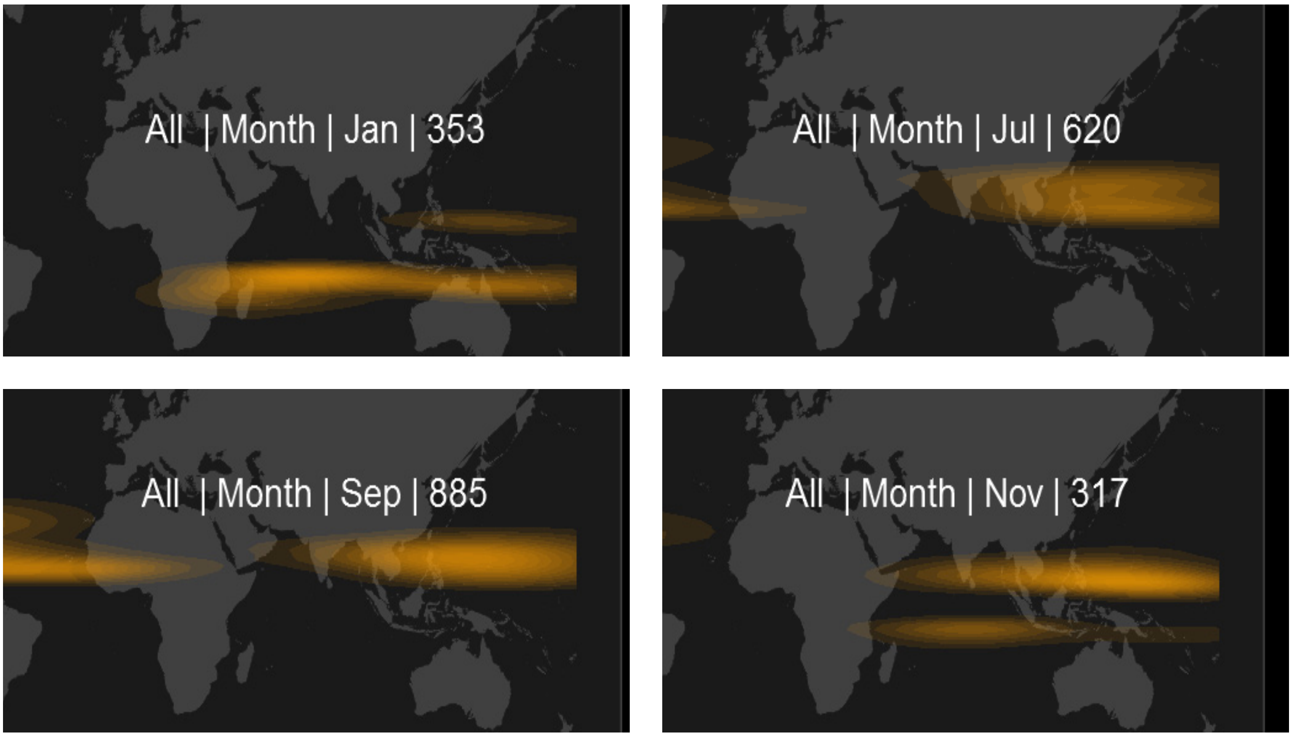


Figure-2 | Typhoon Birth Density per Month Statistical Density Plot. Numbers Represent a Count of Typhoons Recorded in that Month from Historical Data.

The genesis location varies from the southern hemisphere to the northern with the varying sea temperature. No typhoons are conceived near the equator as the correolis effect among other variables determines the typhoon genesis. Typhoons that are generated often travel according to the predominant wind (mainly). A variation in typhoon paths can be seen in Figure-3. This is because once above a certain Lattitude the typhoon experiences West to East direction winds that cause a curved profile in which the typhoon begins to travel back East into the Pacific. The other predominant type of typhoon does not reach this lattitude and so travels in a straighter path generally coming into contact with the nearest land - China.

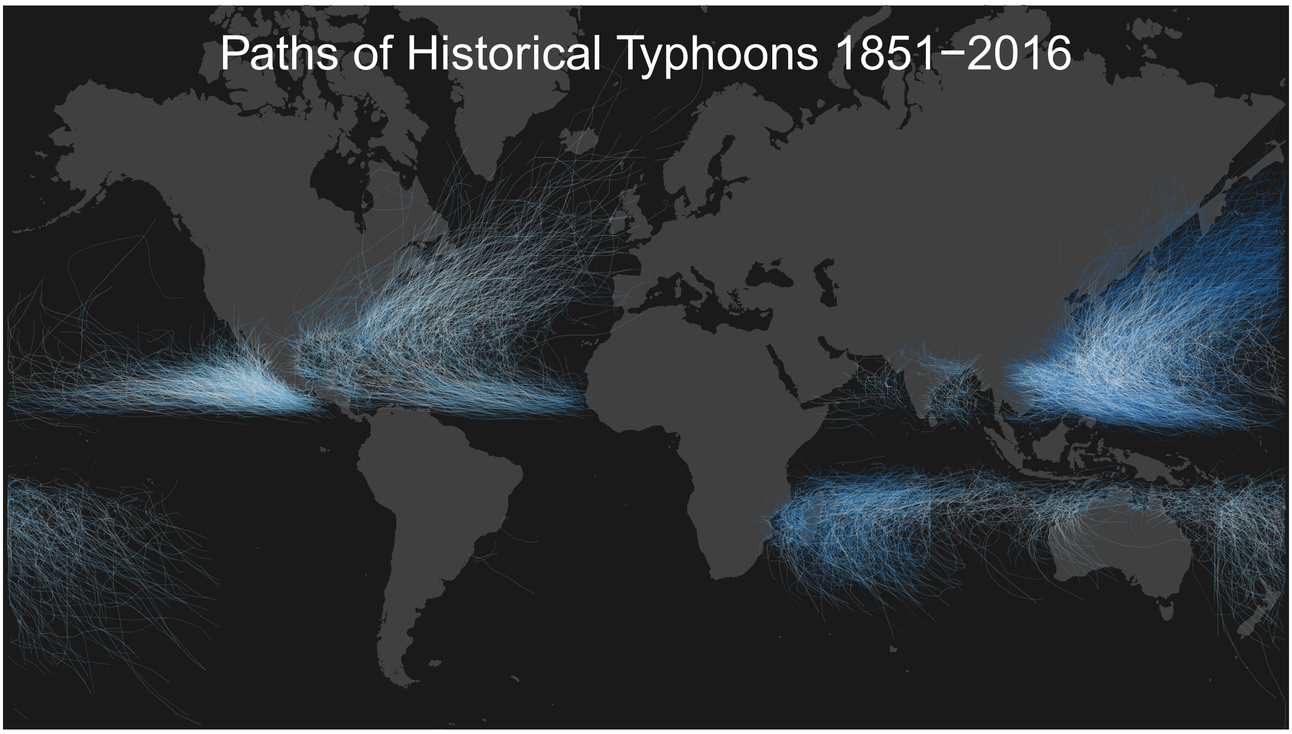
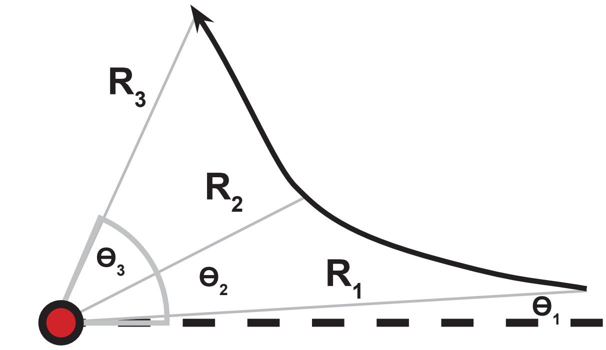


Figure-3 | Typhoon Global Plot, Typhoon Paths

1.3. Polar Coordinates Investigation

Many previous studies have purely considered the genesis and paths of typhoons from a cartesian viewpoint. Thus an additional study into Polar coordinates was conducted. This involves taking a point 0 Lattitude, 85 Longitude and plotting the polar path of each typhoon, see Figure-4.

Since the path is in the dataset as Longitudinal and Latitudinal coordinates a cartesian conversion is needed before converting to polar coordinates. This requires correct distance estimation and uses the Vincy-Ellipsoid formula to convert distances correctly before conversion to Polar coordinates.



(85,0) **Figure-4 | Typhoon Polar Conversion**

The investigation into Polar coordinates allowed the production of a graph that took 4 cases of each type of typhoon, namely the curving and straighter travelling. These 4 examples of each were taken from the original data via visual inspection and then plotted using polar coordinates, Figure-5.

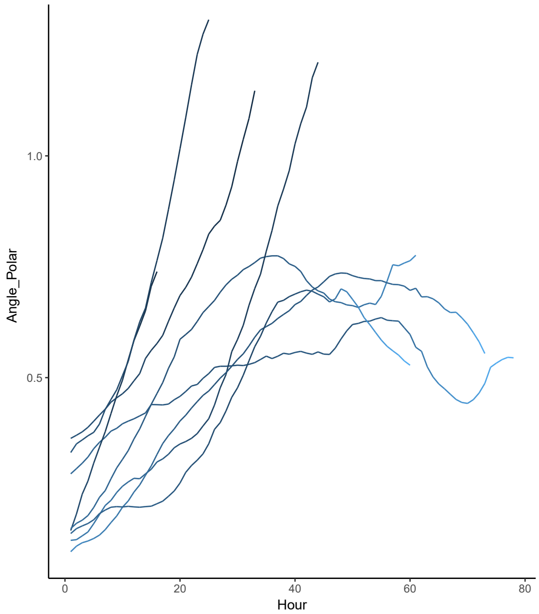


Figure-5 | Polar Plot of Typhoon Paths Polar Angle vs Time Since Genesis.

The curved paths are representative of the curved typhoons in the normal cartesian plot. These typhoons maintain a relatively constant angle once they have curved back East. This is due to their increased distance from the origin point (85,0) and the change in path that is experienced.

Although an interesting investigation it was later decided to continue exploring the data through the use of other methods. However the code for conversion to polar coordinates and the production of some related graphs is still present. This can be explored in the future.

1.4. Effects of ENSO signal

Using the ENSO signal data with a signal of +0.5 considered as an El Nino year and <-0.5 as a La Nina year the number of typhoons per calendar month were plotted. Note that since some months in the historical data have been with an El Nino signal more often than with a La Nina signal it was important to adjust the results to account for this, see Figure-6.

Although there are an increased number of Nino events during the summer compared to Nina it does not look statistically significant to prove this is the case. The study contradicts the study conducted by: Patricola and Saravanan (2013). Although a different methodology for selecting EL Nino years is used and theAtlantic meridional mode (AMM) is also incorporated in the study.

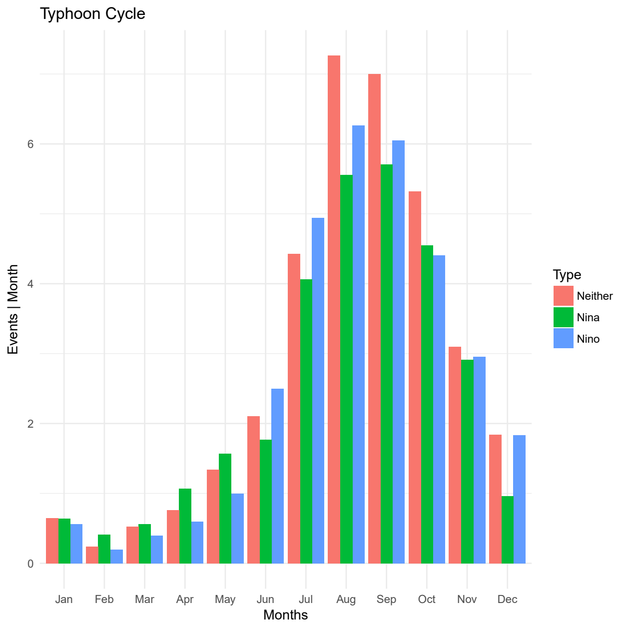


Figure-6 | Typhoon ENSO Monthly Plot Events

However, although the number of events is not proven to vary with the ENSO signal in this study, a plot of ENSO signal vs Genesis location displays a difference in where the typhoons occur. An EL Nino year is characterised by a warmer sea temperature and so generally in EL Nino years the typhoons are generated further to the East and also generally further North.

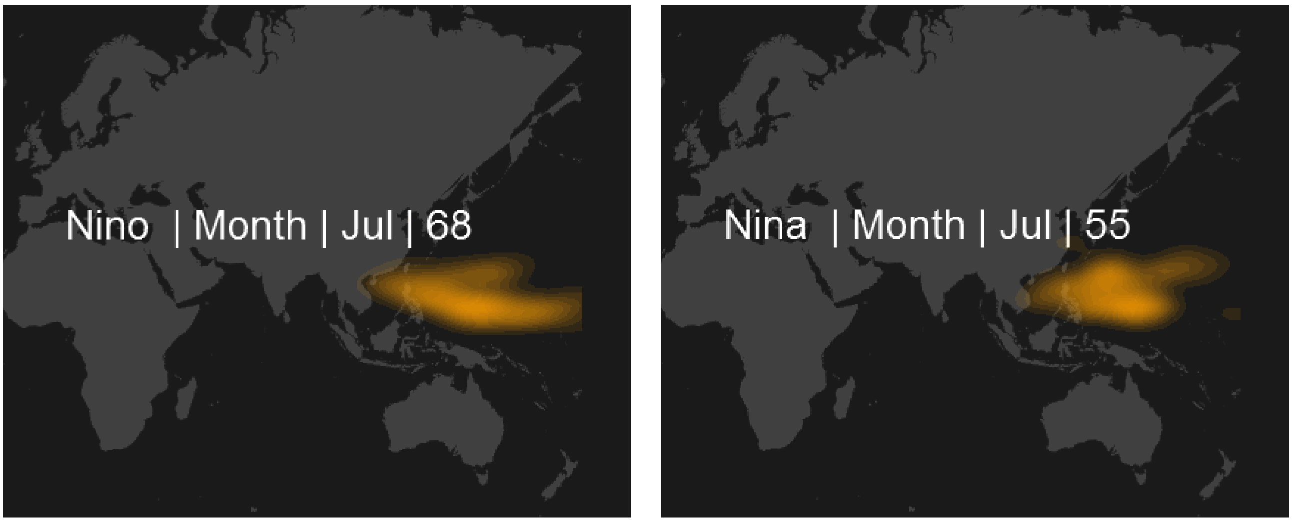


Figure-7 | Typhoon ENSO Plot. Comparison of Genesis Location for Different ENSO signals in the Month of July.

This data, if deemed to be statistically significant could be utilised in the typhoon genesis model at a later date. At the present time the model described in the latter section of this report does not incorporate the ENSO or AMM cycle.

2. Stochastic Typhoon Model

2.1. Poisson Model of Yearly Occurances

The Poisson parameters are based on the number of typhoon events per year. So the sum of events for each historical year has been modelled as a Poisson distribution to then be used later for random typhoon generation for each randomly generated year.

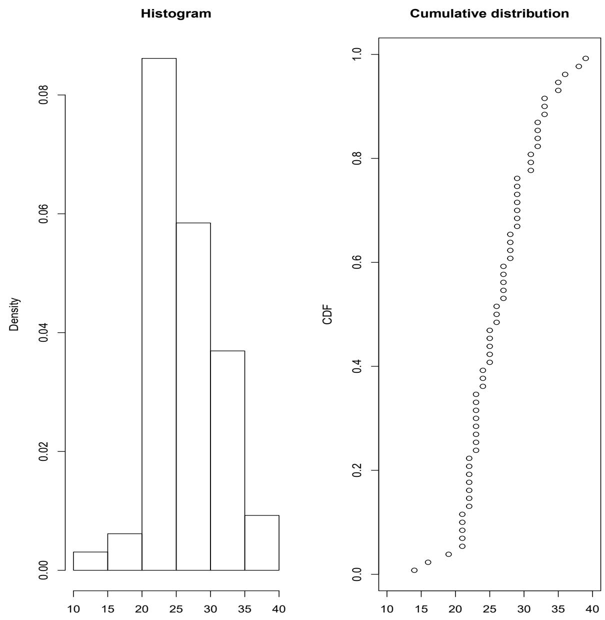


Figure-8 | Historical Tyhpoon Events per Year

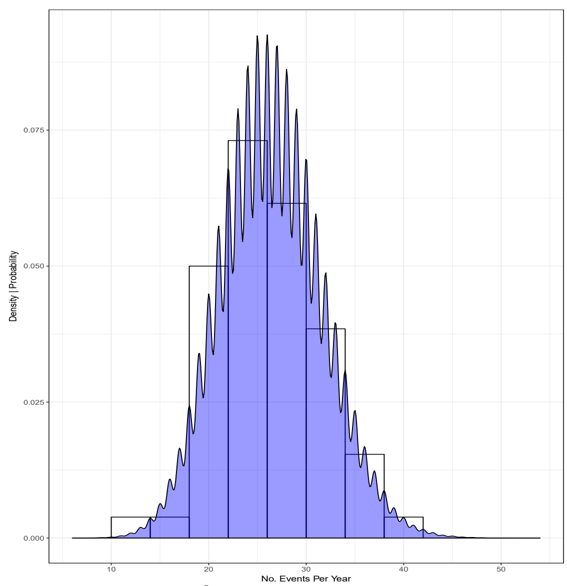


Figure-9 | Poisson Approximation of Events per Year

The Poisson distribution is discrete and takes one parameter λ. It is used in this case to predict the number of typhoon events in a given year. The formula for a Poisson distribution is given below in Equation - 1.

$$P(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \text{for x } = 0,1,2,\ldots \quad \textbf{(1)}$$

The parameter λ is estimated using the Fitdistrplus R package. This in turn used the Maximum Likelihood Estimation (MLE) function to determine the λ that produces the best fit to the data. Once the Poisson model is created it can be used to stochastically generate events for a number of artificial years. Hence this can be used in the model to provide a basis before modelling the generation location of these typhoons in an artificial year.

2.2. Split Generated Events into Months

The number of typhoons also vary with the months. So there are a greater number in August/ September when the sea surface temperature is higher. So the number of events per year can be disaggregated between the months. This is achieved by using historical data to create profiles for each year (the proportion of events in each month across a year). These profiles were created for every historical year. Each generated year is then randomly assigned a historical profile with equal probability, see Figure-10,11.

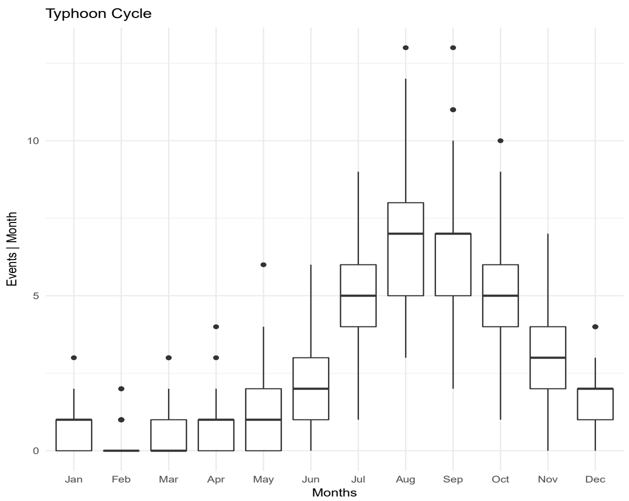


Figure-10 | Historical Typhoon Events per Month

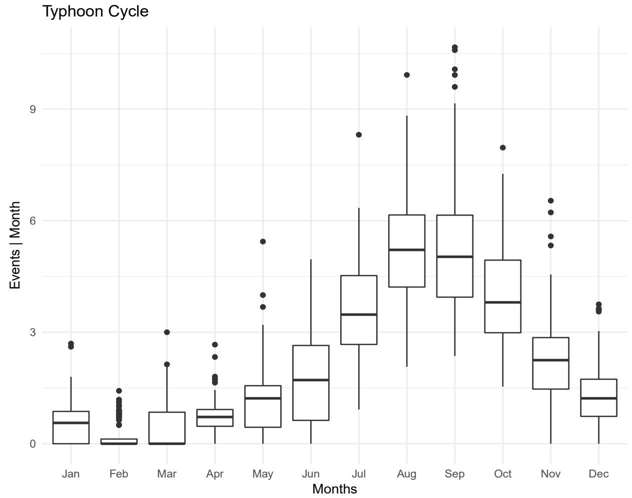


Figure-11 | Generated Typhoon Events per Month

2.3. Model Genesis Location Using Kernel Smoothing

Model the location of typhoon genesis using a Kernel density model. This places a normal distribution over each of the datapoints to smoothe the data, Figure-12.

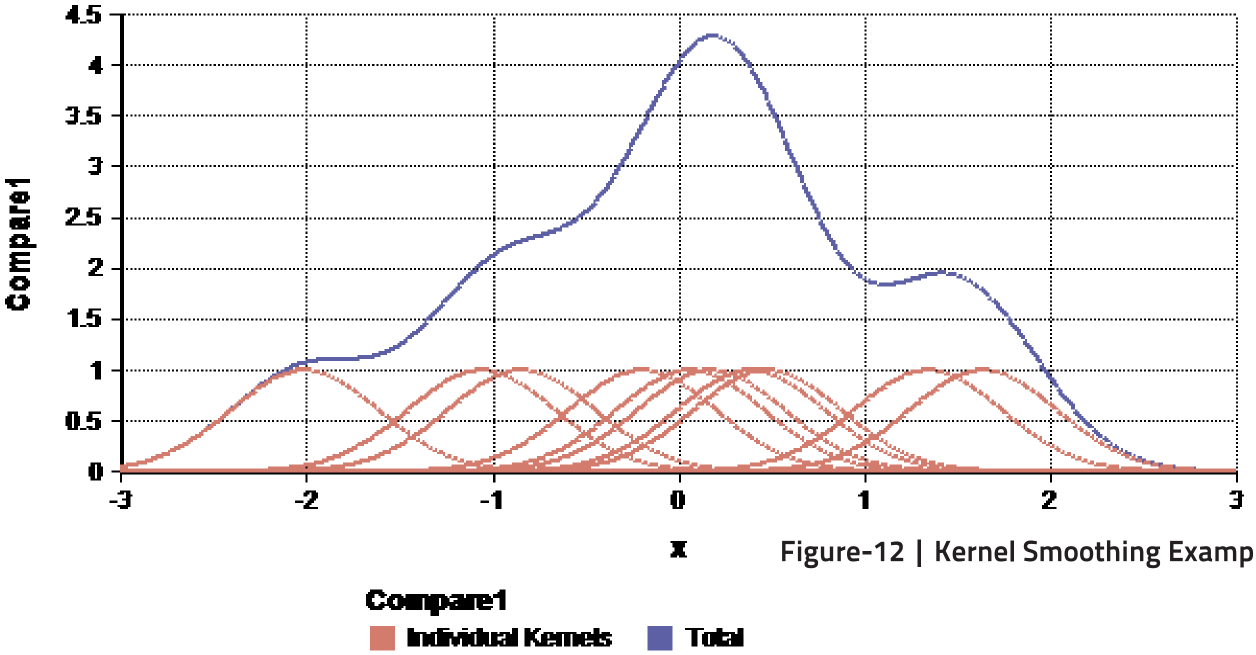


Figure-12 | Kernel Smoothing Example

This methodology can be applied in 2D to the typhoon genesis locations and produce a stochastic model. First a grid is generated that encompasses the area of historical typhoons. Then the probability of a typhoon being generated at each point can be determined from the smoothed model. The grid is plotted in Figure-13 with the historical data as a density plot behind (the contours). From this grid the probability of each grid point is also returned. This can then be used to randomly sample datapoints from the grid using the sample() R function which samples points from a distribution (in this case the points are the grid and the distribution is the kernel smoothed model).

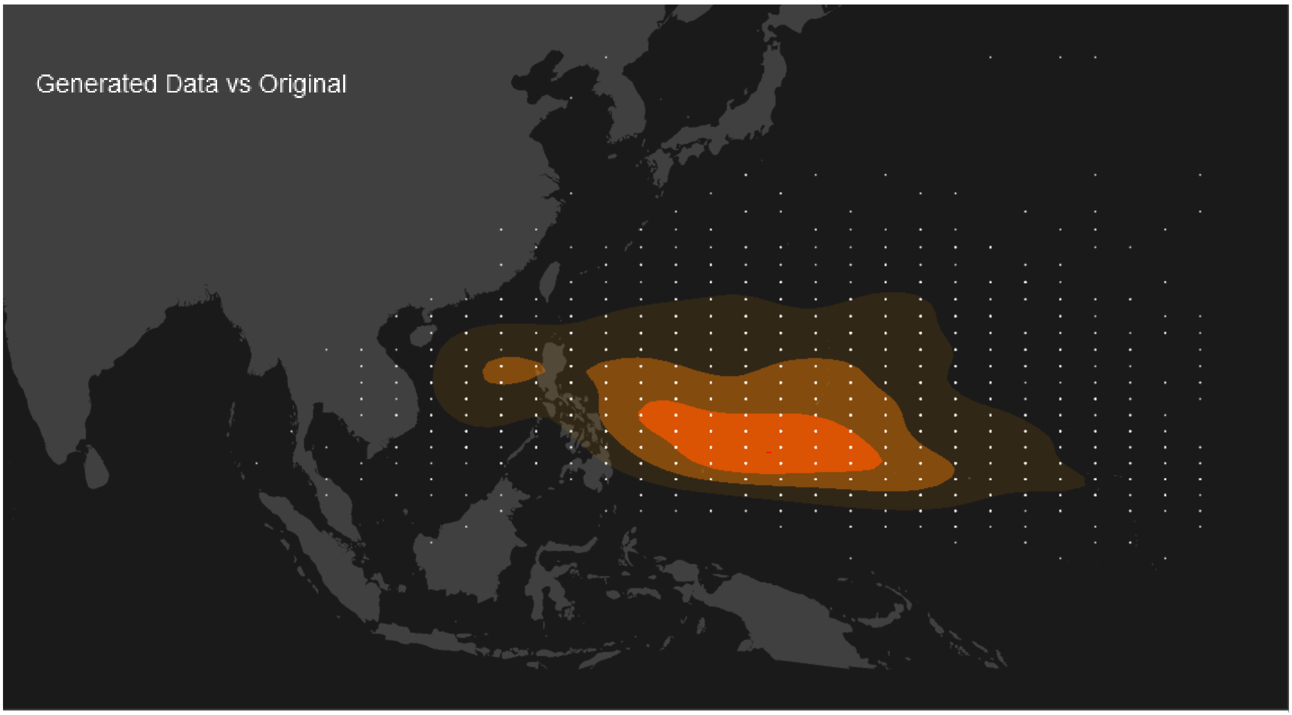


Figure-13 | Typhoon Generation Grid, The Contours are Historical Data and the White Points are the Genesis Locations of the Generated Data.

In Figure-14 the process is reversed with white points signifying the historical data and the contours are the generated model.The smoothing was conducted using the KernSmooth R package
https://cran.r-project.org/web/packages/KernSmooth/KernSmooth.pdf
The function dkde2D was used to produce a 2D smoothed kernel (Latitude and Longitude). Dpik function of the same library was used to automatically calculate the best kernel parameters – e.g. the standard dev of normal distributions used to approximate the data. The dkde2D function takes a input (xrange) and then outputs the results as a grid. So it converts a continuous distribution into discrete points. These discrete points on a grid are displayed below. The contours are created as previously from historical data and the white points are the grid outputted from dkde2D.

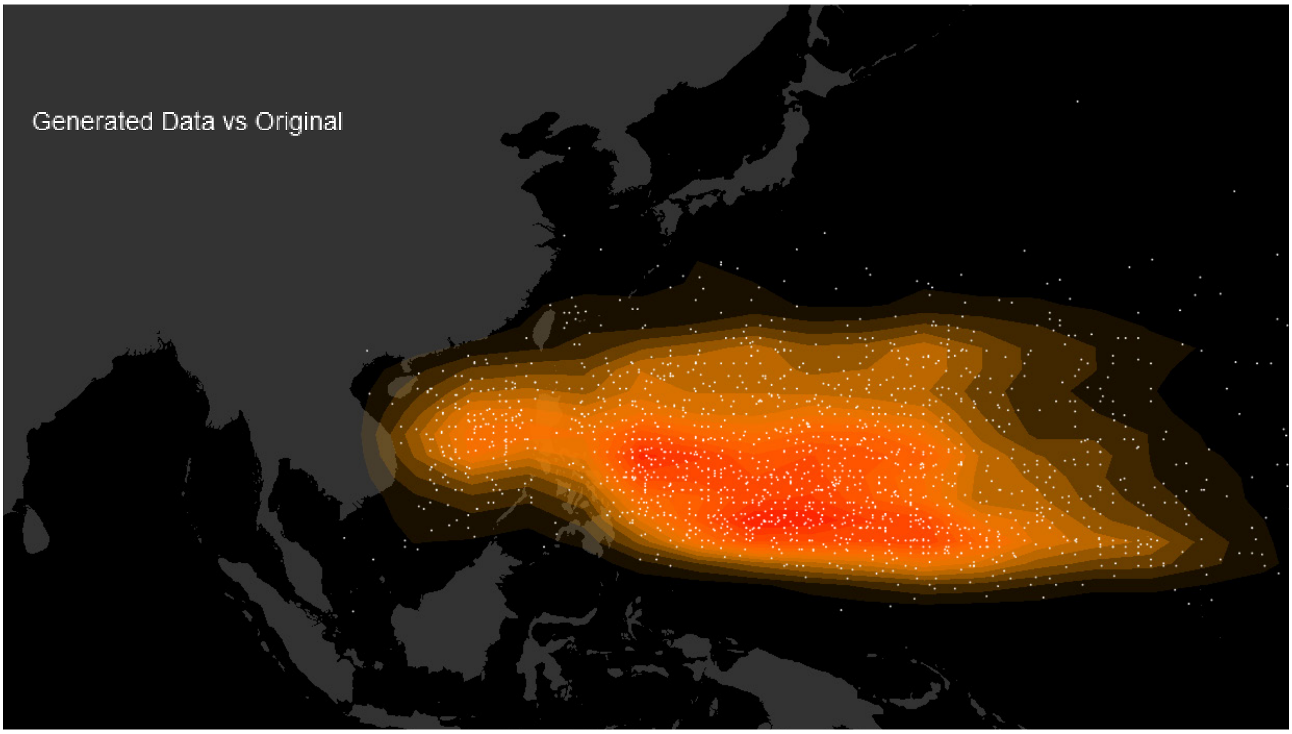


Figure-14 | Typhoon Generation Density vs Historical Data, Contours are the Generated Data, White Points Represent Historical Genesis Locations.

2.4. Regression Model Path Tracking

In order to generate the typhoon paths, an understanding of how historical typhoon paths have been affected by various conditions they experience. For example how Speed, current Longitude, Latitude etc. affect the typhoon's predicted trajectory. In order to obtain this the Pacific basin is split into 5x5 cells. Each of these cells behave differently and a typhoon would behave differently in each cell. Hence separate regression models are obtained for each cell and they determine the behaviour of a typhoon in that location, Figure-15.

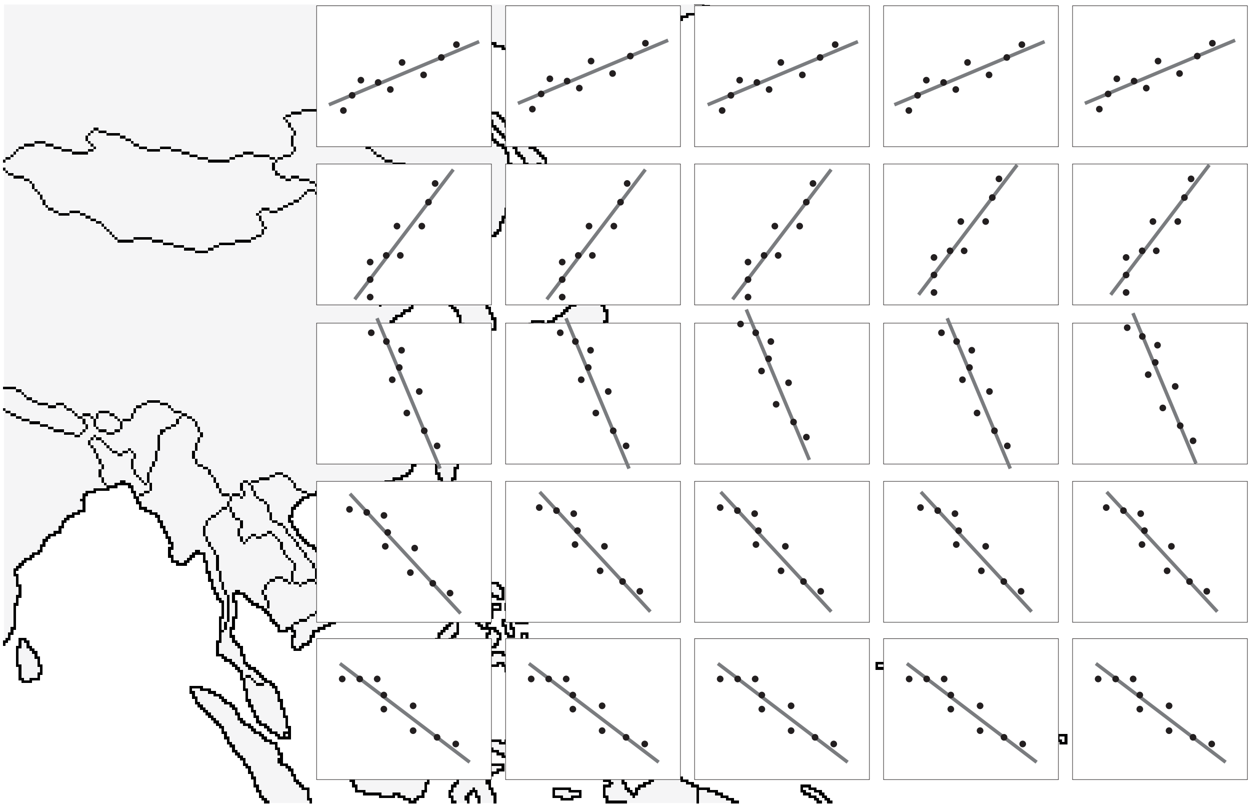


Figure-15 | Typhoon Regression Grid . Lines are Estimated from Historical Data which are the points. (Path Direction Example)

Figure 15 is purely a representative example of how the regression models work. This is a path direction example which would be a combination of the change in Longitude and change in Latitude models. Each cell contains 4 regression models that help to predict the state of the typhoon in 6 hours time (t+6).

$$\Delta Longitude\ (t, t+6) = \gamma_1 \Delta Longitude_{(t-6,t)} + \gamma_2 \Delta Longitude_{(t-12,t-6)} + \gamma_3 Longitude_{(t)} + \gamma_4 Latitude_{(t)} \quad (2)$$

$$\Delta Latitude\ (t, t+6) = \alpha_1 \Delta Latitude_{(t-6,t)} + \alpha_2 \Delta Latitude_{(t-12,t-6)} + \alpha_3 Latitude_{(t)} + \alpha_4 Longitude_{(t)} \quad (3)$$

$$Intensity\ (t+6) = \mu_1 Intensity_{(t)} + \mu_2 Intensity_{(t-6)} + \mu_3 Latitude_{(t)} + \mu_4 Longitude_{(t)} \quad (4)$$

$$Wind\ Speed\ (t+6) = \beta_1 Wind\_Speed_{(t)} + \beta_2 Wind\_Speed_{(t-6)} + \beta_3 Latitude_{(t)} + \beta_4 Longitude_{(t)} \quad (5)$$

Equations 2-4 are the regression models used where (t,t+6) is the difference between the values at time t compared to t+6. These models, once fitted to historical data can be used to progressively predict the typhoons position at t+6 recursively until the entire path has been generated.

2.5. Generate Initial Typhoon Data

In order to predict the typhoon path via the regression models in section 2.4 some initial data is needed to predict from. These intial values are predicted from a Weibull distribution (Equation - 6) that is fitted to the historical data using maximum likelihood estimation. This is the same process as before which was used to fit the Poisson model to the historical frequency.

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^{k}} \quad \textbf{(6)}$$

The PDF of the Weibull distribution is displayed in Equation-6. This takes two parameters, k which alters the shape and λ which alters the scale of the distribution. Once this is fitted to the historical data for each variable in the regression models, it can be sampled from to provide the starting data for the typhoon. Then this starting data is fed into the regression models until a complete typhoon path is generated.

2.6. Final Result

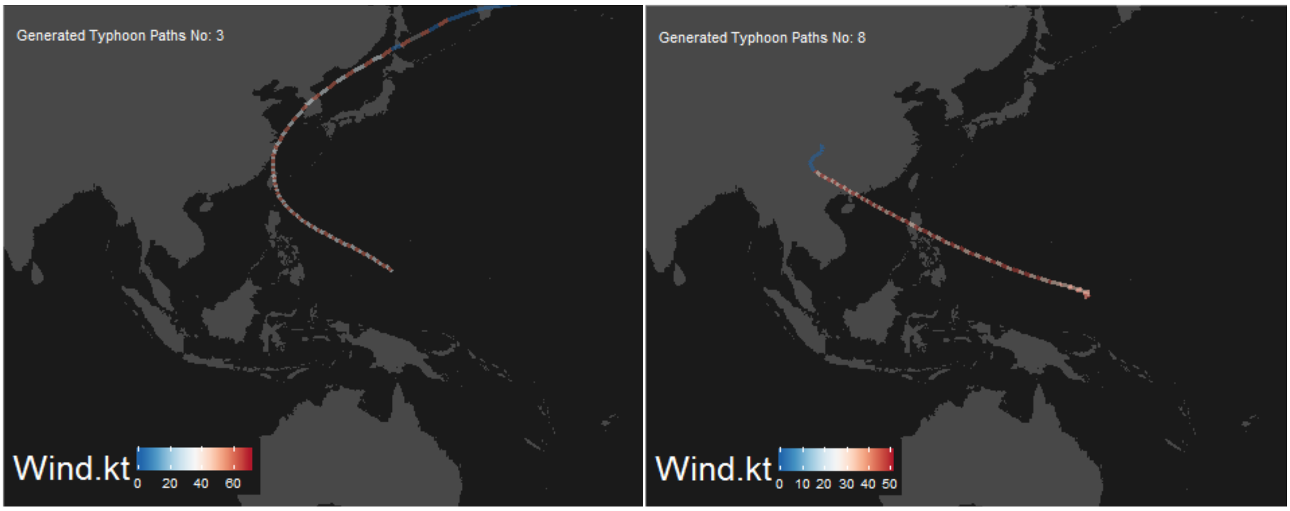


Figure-16 | Final Generated Typhoon Path Examples.