VNU Journal of Science, Mathematics - Physics 26 (2010) 193-200

Simulation of tropical cyclone tracks in the offshore of Haiphong, Vietnam

T.M.Cuong[1,*], V.H.Sam[1], P.D.Tung[1], L.H Lan[2]

[1]*Department of Mathematics, Hanoi University of Science, VNU, 334 Nguyen Trai, Thanh Xuan Hanoi, Vietnam*

[2]*Department of Basic Sciences, University of Transport and Communication, Hanoi, Vietnam*

Received 10 October 2010

Abstract. In this paper, we simulate a large number of synthetic cyclone tracks for the offshore of Haiphong, Vietnam based on the 114 historical storms observed from 1951 to 2007 in this area. With these synthetic tracks, the assessment of damage risks can be improved.

Keywords and phrases: Stochastic model, Inhomogeneous Poisson point process, Monte-Carlo simulation, tropical cyclone.

1. Introduction

Catastrophes caused by tropical cyclones are a threat to many aspects of human lives. One needs to assess as precisely as possible the risk and extent of losses in areas affected by tropical cyclones. Since reliable data on cyclone track is only available for a relatively short of time, it is not sufficient to make a risk assessment based solely on historical storm track. Therefore, there arises a need to propose a stochastic model for the computerized generation of a large number of synthetic cyclone track. This will provide a larger dataset than previously available for the assessment of risks in areas affected by tropical cyclones. In this paper, based on the idea in [1], we propose a similar stochastic model for simulating tropical cyclone tracks in the offshore of Haiphong, Vietnam.

The original available data consists of the tracks of all tropical cyclones recorded during the period 1951-2007 in the north Gulf of Tonkin. Figure 1 shows the tracks of all 118 storms considered.

Each track is given as a polygonal trajectory connecting between two adjacent points of measurement. Besides the date and time of measurement, the storm's current position (longitude and latitude) and its current maximum wind speeds are given for each point. The measurements within each individual storm are taken at regular intervals of 6 hours, so the storm's translational speed can be easily calculated. All observations fall into an observation window that is delimited by the equator in the south, 30^oN in the North, 80^oE in the West, and 180^oE in the East. However, in this paper we are only interested in storms in the offshore of Haiphong so the observation window W determined by 19^oN in the South, 22^oN in the North, 105^oE in the West, and 111^oE in the East. For each original track, all observations outside the observation window W are elimilated. The tracks of all 114 tropical cyclones in the observation window W are shown in Figure 2 below. The simulation is only based on the data of these storms. In Figure 1, the area bounded by the small rectangular is the observation window W. In [1], Rumpf and others consider storms in a large area (the western North Pacifica)

* Corresponding authors: E-mail: cuongtm@vnu.edu.vn.

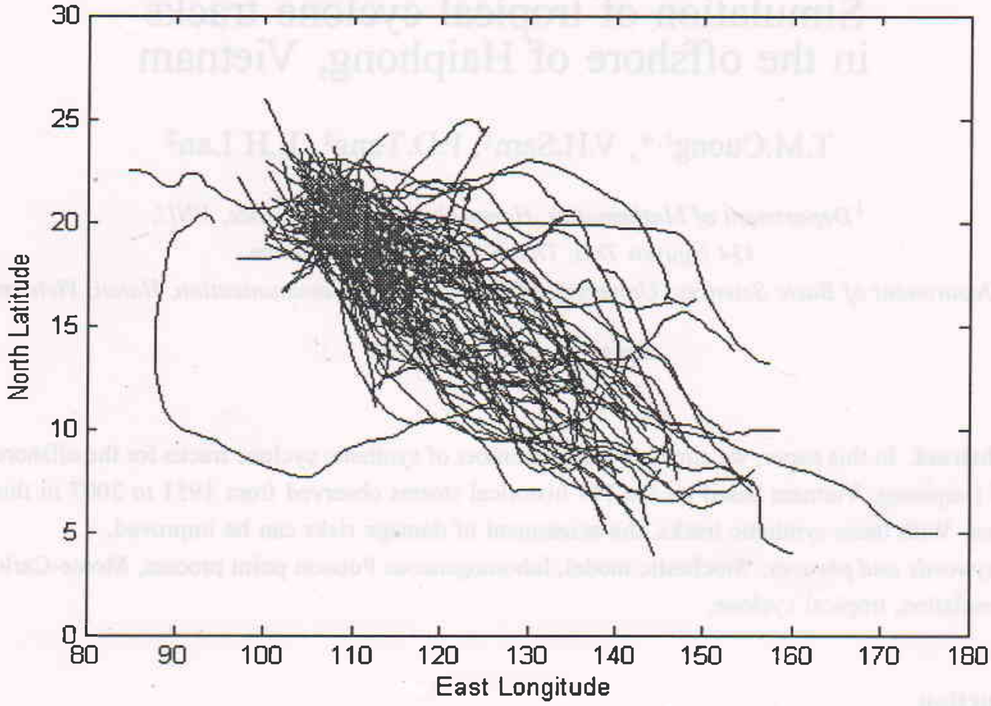

Fig. 1. Tracks of all storms in the dataset.

so there are strong inhomogeneities in the shapes of the cyclone tracks. To improve the quality of the simulation, they split storm tracks into 6 more homogeneous classes. All subsequent steps of the modelling process are done separately for each class. As can be easily seen from Figure 1 and Figure 2., the shapes of the cyclone tracks in the north Gulf of Tonkin as well as the observation window W is rather homogeneous, we do not need to split 114 storm tracks.

2. Point of genesis

For a stochastic model of the tracks of tropical cyclones, first a model for the points of cyclone genesis, i.e. the first point of the track, is needed. In this paper, because the track of a storm is restricted by the observation window W the point of genesis of a track is considered as the first observation of the storm in the observation window. Figure 3. shows the points of genesis of 114 storms in the observation window W.

The points are clearly distributed inhomogeneously within the observation window. Therefore, and inhomogeneous Poisson point process is chosen as a model. To test the hypothesis that points of genesis are located as a inhomogeneous Poisson point process, we consider the following hypothesis testing

$$\begin{cases} H_0 : \text{it is an inhomogeneous Poisson point process} \\ H_1 : \text{it is not an inhomogeneous Poisson point process} \end{cases}$$

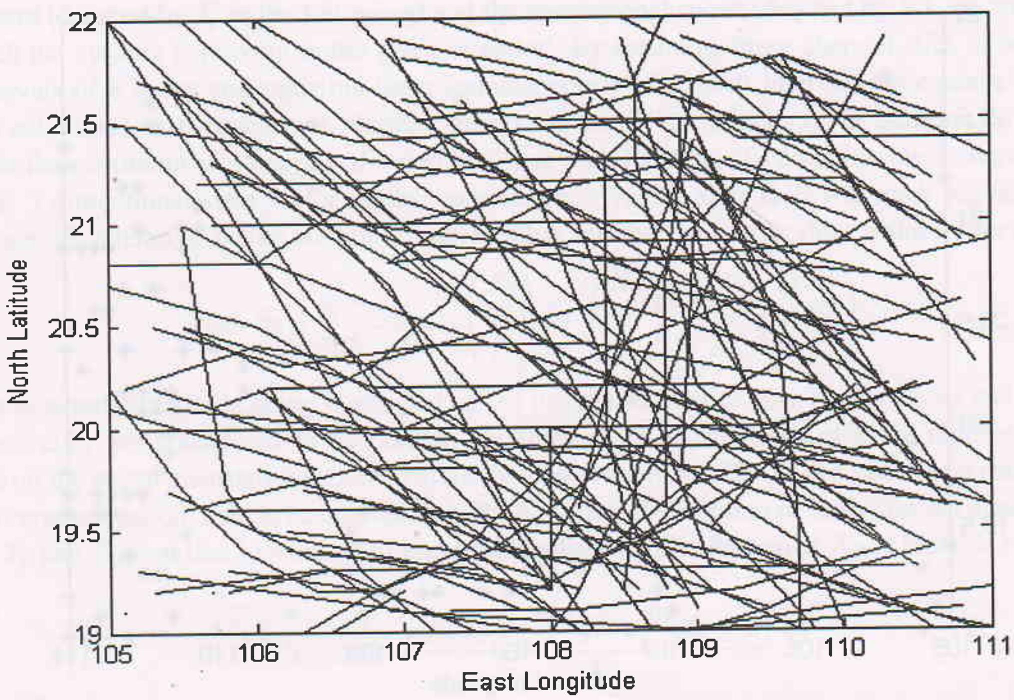

Fig. 2. Tracks of all storms in the observation window W.

The distribution of points of this process is determined by its intensity function $\lambda(t)$, where t is the position of the point of genesis determined by atitude and longitude. This function can be interpreted in a way that $\lambda(t)dt$ describes the infinitesimal probability of a point of the Poisson point process being located in the infinitesimally small disc with area dt centered at t (see [2]). Since there is no obvious parametric trend visible in the data, a non-parametric estimation technique was chosen. The generalised nearest neighbour estimation (see [3]) is given by

$$\hat{\lambda}(t) = r_k^{-2}(t) \sum_{i=1}^{m} K_e \left\{ r_k^{-1}(t)(t - T_i) \right\}. \tag{1}$$

The parameter $k = [\sqrt{m}]$, where m is the number of historical points of genesis of all storms in the observation window, $r_k(t)$ is the distance to the k-th nearest point of genesis from the location t, T_i is the location of the i-th historical point of cyclone genesis, and K_e is the Epanechnikov kernel defined by

$$K_e(t) = \begin{cases} \frac{2}{\pi}(1 - t^T t) & \text{if } t^T t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

A simplified interpretation of this estimator is given in the following: while the kernel K_e determines the size and the shape of the "probability mass" which is assigned to a measurement point, the bandwidth $r_k(t)$ is the radius over which this mass is spread. Note that the estimation $\hat{\lambda}(t)$ is nowhere zero at all points within the observation window, there is non-zero probability mass from exactly k historical points of genesis. This probability mass decreases with increasing distance to the k-th nearest

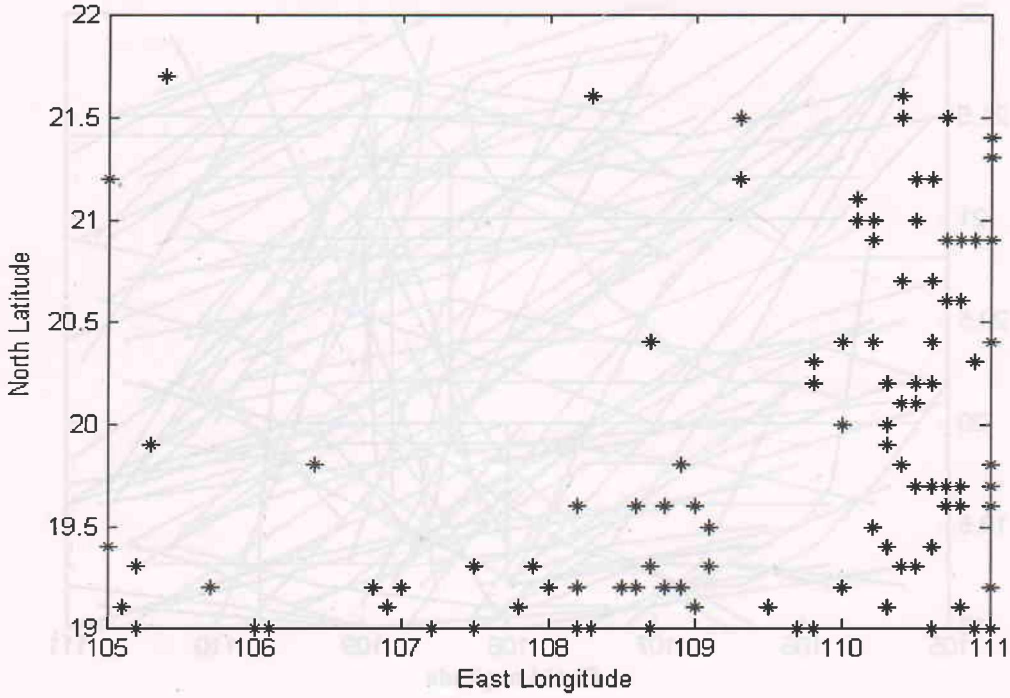

Fig. 3. Points of genesis of storms in the observation window W.

historical point of genesis, but in theory never reaches zero. This effect is intended, because it only rarely allows for the genesis of tropical cyclones within the model at locations that are far away from most historical initial points of cyclones.

To test the null hypothesis that the poits of genesis are distributed in the observation window as an inhomogeneous Poisson point process with the intensity function $\hat{\lambda}(t)$, we performed the chi-square test of goodness-of-fit using R software. In this test, the window observation is divided into tiles, and the number of data points in each tile is counted, as described in the quadratcount. In R, the quadrats are rectangulars by default, or may be regions of arbitrary shapes. In this case, we chose quadrats as rectangulars by default. Below is the result of this test:

$$\chi^2 = 5.32237, \mathrm{df} = 23, \text{p-value} = 0.3417.$$

This test shows that the historical data is in favor of the null hypothesis.

3. Cyclone tracks and wind speeds

3.1. Direction, translational speed and wind speed

Once a model for the points of cyclone genesis is available, the propagation of the tracks is the next step in the modelling process. Here our model relies on the basic assumption as the models introduced in [1] that cyclones located in similar areas of the observation window behave similarly.

An appropriate model of the tracks following the points of genesis need to include the direction of movement (denoted by X in the following) and the translational speed (denoted by Y), i.e. the velocity at which the cyclone is moving in the given direction. By assuming these charateristics to be constant for intervals of 6 hours and updating them instantaneously after each interval, the cyclone's location can be calculated in 6 hour-steps, thereby generating complete trajectory. In addition, we need to simulate the maximum wind speed (denoted by Z) at each location. To combine these characteristics, consider 3-dimentional state vector S_i that contains their values after their i-th track segment. These values are considered to be the sum of an initial value and the changes in these values after each step:

$$S_i = S_0 + \sum_{j=1}^{i} \Delta S_j = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \sum_{j=1}^{i} \begin{pmatrix} \Delta X_j \\ \Delta Y_j \\ \Delta Z_j \end{pmatrix}$$

Since a stochastic model is being developed, all of the characteristics X_0, Y_0 and Z_0 as well as ΔX_j, ΔY_j and ΔZ_j are considered to be random variables. The distributions of these random variables depend on the storm's current location t within the observation window W. To generate a realisation of S_0 at a certain location, data is resampled from the empirical distribution of the historical measurements of X_0, Y_0 and Z_0 near that location. The empirical probability distribution of X_0 at location t is defined by

$$F_{X_0}(x,t) = \frac{\#\{l : 1 \le l \le k, x_l^{(0)}(t) \le x\}}{k}, \tag{2}$$

where $x_l^{(0)}(t)$, $l = 1, 2, ..., k$ denote k historical realisations of X_0 closest to the location t. In short, the distribution of the initial direction of a track in the model is determined by all historical initial directions of storm tracks. Similar formulae are used in order to estimate the location-dependent distributions of Y_0 and Z_0, respectively

$$F_{Y_0}(y,t) = \frac{\#\{l : 1 \le l \le k, y_l^{(0)}(t) \le y\}}{k}, F_{Z_0}(z,t) = \frac{\#\{l : 1 \le l \le k, z_l^{(0)}(t) \le z\}}{k}. \tag{3}$$

In analogy to this, the probability distribution functions of a change in direction ΔX_j is given by

$$F_{\Delta X}(x,t) = \frac{\#\{l : 1 \le l \le k, \Delta x_l(t) \le x\}}{k}, \tag{4}$$

where $\Delta x_l(t)$, $l = 1, 2, ..., k$ denote k historical realisations of ΔX_j $\forall j$ closest to the location t. This means that the distribution of any change in direction ΔX_j depends on historical realisations of all changes in directions of tropical cyclones, no matter after which step of a storm they occurred. A similar formulae are used for ΔY_j and ΔZ_j, respectively

$$F_{\Delta Y}(x,t) = \frac{\#\{l : 1 \le l \le k, \Delta y_l(t) \le y\}}{k}, F_{\Delta Z}(x,t) = \frac{\#\{l : 1 \le l \le k, \Delta z(t) \le z\}}{k}. \tag{5}$$

3.2. Termination probability

Since the proposed model creates synthetic cyclone tracks in 6-hour steps, a mechanism is needed to determined whether the track should be terminated after the current step or if it should be continued. This is done stochastically by Bernoulli experiment with a success probability $p(t, Z)$ depending on the storm's current location t and wind speed Z. The termination probability is determined as maximum of two probabilities p_Z and p_t where p_Z represents the termination probability caused by wind speed

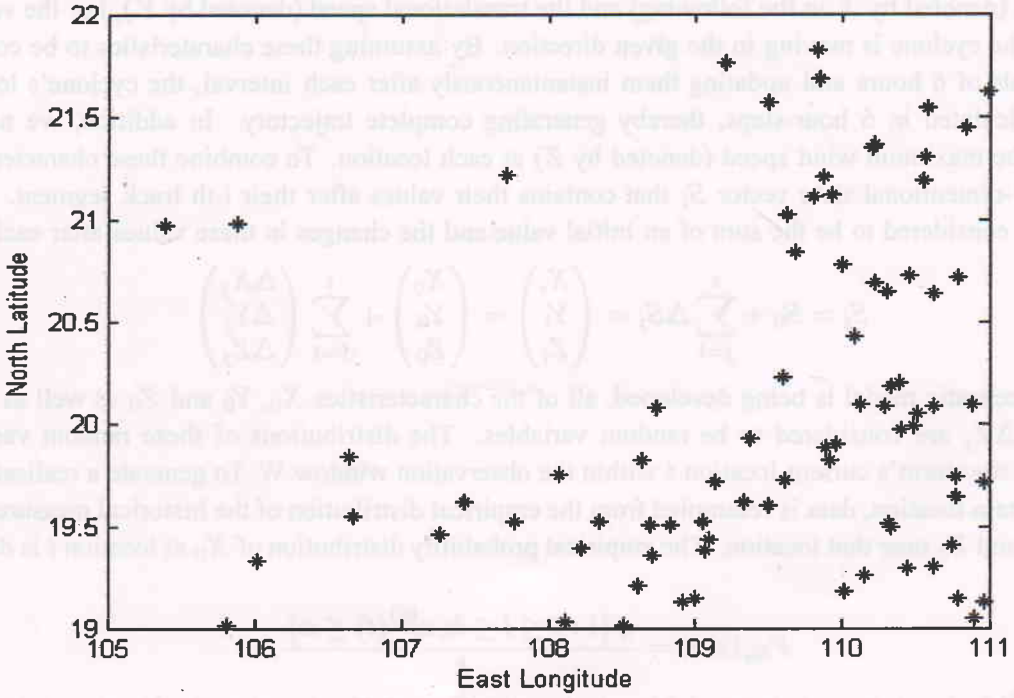

Fig. 4. Simulated points of genesis of storms in the observation window W.

and p_t represents the termination probability caused by location. Based on the historical storms in the observation window, we estimate these probability by their point estimators. Specifically, at each point of measurement, p_t is approximated by the fraction of termination points among the k nearest points of measurement of the location t and p_Z is approximated by the fraction of termination points among the k windspeed-nearest points of measurement. The termination probability used in Bernoulli trial is then taken to be

$$p(t, Z) = \max\{p_t, p_Z\}. \tag{6}$$

Noting that we terminate the synthetic cyclone track if it gets out of the observation window.

4. Semulation and Results

In this section, an algorithm for generating synthetic cyclone track is described. First, we introduce the algorithm for simulating an inhomogeneous Poisson point process with intensity function $\hat{\lambda}(t)$ defined by (1) (see [4]). It consists of 3 main steps.

Firstly, we calculate the above bound λ^* of $\hat{\lambda}(t)$ by putting $\lambda^* = \max\{\hat{\lambda}(t)\}$ for all the points t of measurement in the observation window.

Secondly, simulating a stationary Poisson point process of intensity λ^*. To do this, the number of points in the observation window W is determined by simulating a Poisson random variable with parameter λ^* (see [5]) and then determine the position of the points in W by simulating a binomial point process in W with that number of points.

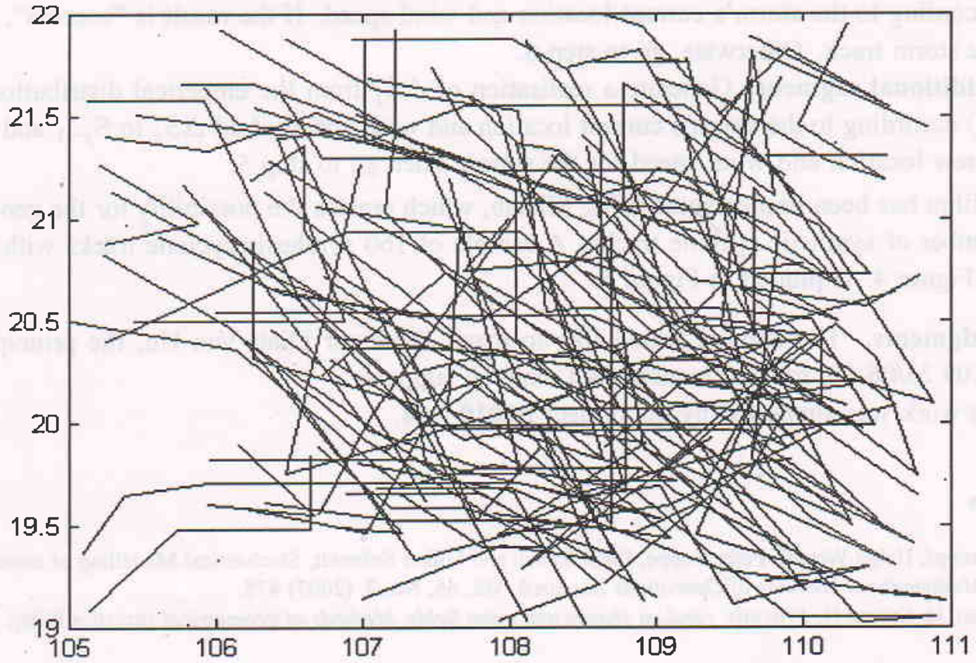

Fig. 5. Synthetic tracks of storms in the observation window W.

Finally, the resulting point pattern is thinned by deleting each point t independently of the others with probability $[1 - \hat{\lambda}(t)/\lambda^*]$. If the points of the stationary Poisson point process pattern are $\{t_1, t_2, ...\}$ then this thinning can be performed with the aid of an independent sequence $(U_1, U_2, ...)$ of random numbers uniformly distributed over $[0, 1]$. The point t_k is deleted if $U_k > \hat{\lambda}(t_k)/\lambda^*$.

Figure 4. shows 160 simulated points of genesis in the observation window.

To create a complete set of synthetic storm tracks from the model described above, the procedure is as follows.

1. **Initialisation:** Find all needed estimators and probabilities including the intensity function $\hat{\lambda}(t)$ in (1), emperical distribution functions defined by (2)-(5) and termination probability in (6) and go to step 2.
2. **Points of genesis.** Generate a realisation of the inhomogeneous Poisson point process with the density function $\hat{\lambda}(t)$ in (1) by the above algorithm and go to step 3.
3. **Choose a point.** From the point process realisation generated in step 1, pick one point that does not yet have a corresponding cyclone track and go to step 4. If there are no such points left, terminate the algorithm.
4. **Initial segment.** Generate a realisation of S_0 from the emperical distribution functions (2) and (3) according to the location t of the cyclone's starting point from step 3. With this, find the storm's new location after its first segment and go to step 5.

5. **Termination probability.** Perform a Bernoulli trial with the success probability given by (6) according to the storm's current location and wind speed. If the result is "success", terminate the storm track. Otherwise, go to step 6.
6. **Additional segment.** Generate a realisation of ΔS_j from the emperical distribution (4) and (5) according to the storm's current location and wind speed. Add ΔS_j to S_{j-1} and from this a new location and wind speed for the storm. Then go to step 5.

This algorithm has been implemented using Matlab, which creates the possibility for the generation of a large number of synthetic cyclone tracks. A sample of 160 synthesis cyclone tracks with points of genesis in Figure 4. is plotted in Figure 5.

Acknowledgments. The authors would like to thank Professor Dinh Van Uu, the principal of the project KC09.23/06-10, for his financial and material support.

This work was supported by the project **B2010 - 04**.

References

[1] Jonas Rumpf, Helga Weindl, Peter Hoppe, Ernst Rauch and Volker Schmidt, Stochastical Modelling of tropical cyclone tracks, *Mathematical Methods of Operations Research*, Vol. 66, No. 3, (2007) 475.

[2] D. Stoyan, H. Stoyan H, *Fractals, random shapes and point fields. Methods of geometrical statistics.* Wiley, Chichester, (1994).

[3] B.W. Silverman, *Density estimation for statistics and data analysis*, Chapman& Hall, NewYork, (1986).

[4] Dietrich Stoyan, Wilfrid S. Kendall and Joseph Mecke, *Stochastic geometry and its application*, Wiley Publisher 2nd edition (1996).

[5] Sheldon M. Ross, *Introduction to probabilistic models*, Academic Press 7th edition, (2000).