International Journal of Scientific Research in Mathematical and Statistical Sciences **Review Paper**
Vol.6, Issue.2, pp.80-85, April (2019) **E-ISSN:** 2348-4519
DOI: https://doi.org/10.26438/ijsrmss/v6i2.8085

A Review on the Methods of Detecting Space Time Clusters

Jaisankar R[1*], Ranjani M[2], Kausalya N[3]

Dept. of Statistics, Bharathiar University, Coimbatore-641046, Tamil Nadu, India

Corresponding Author: r_jaisankar@rediffmail.com

Available online at: www.isroset.org

Received: 20/Mar/2019, Accepted: 21/Apr/2019, Online: 30/Apr/2019

***Abstract*-**In disease, cluster is an unusually high incidence of a particular disease or disorder occurring in close proximity in terms of both time and geography. Detecting the clusters of disease cases resulting from emerging disease outbreaks is an important part of spatial epidemiology because it is used to identifying hotspots and also environmental factors related with disease and thus guide for the investigation of etiology of diseases. These disease clusters can be as purely spatial or temporal or both spatial and temporal. The present work is aims on reviewing the methods for detecting the space time clusters.

Keywords: Spatial clusters, Cluster detection, Space-time, Window shapes

I. INTRODUCTION

Spatial clusters or hotspots of disease and health-related have brought significant attention in the recent past to public health researchers and policymakers. A disease cluster is a geographical construct that can be identified, visualized and explored using GIS and spatial analysis methods based on the information on unusual number of cases within a population, place, and time period. Disease clusters can provide evidences to disease etiology and disease behaviors, local environmental or social characteristics that stimulate increased risk. Based on these one may develop selectively targeting active interventions to hotspot areas. In the past two decades, several methods for detecting disease clusters have developed that uses more robust statistical formulation and spatial search processes for accurately detecting disease hotspots. This article selectively reviews recent developments in detecting space time clusters.

Methods for detecting spatial clusters comprised of a geographical search method used to identify local concentrations of disease cases to be tested for clustering that scan the study area at regularly spaced intervals. (Cromley, E [1] and Shi [2]).

Also, in spatial searching the size and configuration of the search window/filter within which the clusters are detected is to be determined. In some methods, the window forms are like circle or ellipse etc., and some other methods employ a predefined spatial weights matrix to identify the zone in which clusters may be detected.

Cluster detection techniques also involve methodologies based on statistical or other criteria for determining if the local concentration of disease is unusual or significantly higher than expected. These testing procedures include spatial and GIS characteristics of local population and environment that might influence occurrence of clusters. Factors like uneven population distribution, transportation network and topography etc., are also important for this geographical search.

II. SPACE-TIME CLUSTER DETECTION METHODS

Earlier, local disease cluster detection methods utilize the tools like the geographical analysis machine (GAN) (Openshaw, Charlton, and Craft, [3]), the Besag and Newell method (Besag and Newell [4]) Disease Mapping and Analysis Program (DMAP) (Rushton and Loloins, [5]), the spatial scan statistic (Kulldorff, [6]) etc., which are dealt with detecting spatial disease clusters. Later, the time factor has been added and the methods for detecting disease clusters over space and time were developed.

III. METHODS USING DIFFERENT SCAN WINDOW SHAPES

For scanning in a space time cube for detecting significant spatio-temporal clusters, An extension of spatial scan statistics was made by Kulldorff [7] by adding the time factor which uses cylindrical window. The domain cylindrical Z is defined as,

$$Z = \{(x, y, t)/(x - x_c)^2 + ((y - y_c)^2) \leq r^2, t_s \leq t \leq t_e\}$$

where, (x_c, y_c) is the geographical centre of the domain, r is the radius of the domain, and $[t_s, t_e]$is the domain period.

The radius varies continuously in size and the base is centered around one of several possible centroids located throughout the study region. The study period as a whole, as well as the height reflects any possible time interval of less than or equal to half the total study period. For each possible time interval the window is then moved in space and time for each possible geographic location and size it visits.

By jointly covering the entire study region, an infinite number of overlapping cylinders of different size and shape are acquired. The pre-selection of the cluster alarm is accounted for in terms of both cluster location and cluster size with each cylinder reflecting a possible cluster.

Based on the null hypothesis that the number of cases is assumed to be Poisson dispensed with constant risk over space and time and the likelihood is studied for each cylinder. The cylinder with the maximum probability, and with more than its expected number of cases, is given the most likely cluster.

Conditioning on the total number of cases observed, confounding variables can be adjusted for by calculating the expected number of cases in each area and time period through indirect standardization. The trend can be adjusted for by multiplying the confounder-adjusted expected number of cases for each census area and time period by the overall or national rate during that particular year, or by any proportion of that rate considering there is a temporal trend.

The method proposed by Kulldorff [6] used of cylindrical space-time windows for the detecting clusters can limit the fit to the disease being analyzed and the cylindrical shape cannot model growth or shrinkage of a cluster over time nor model its movement over the time. Hence the square pyramid shape was proposed by Iyenger [8] to overcome these drawbacks. The axis of the pyramid does not need to be orthogonal to the two spatial axes allowing the cluster to model movement of the disease. The pyramid is with square cross sections representing the included geographical area at each time in an interval. The pyramid cluster can be truncated (need not include the apex) and is allowed to either grow or shrink from the start to the end of the time interval.

Shiode [9] developed a network-based search window that takes the form of a sub-network, or a collection of line segments, whose total length remains the same but its form changes and flexibly follows the structure of the network as it sweeps along. While a circular search window is used for sweeping across a study area defined by the 2D Euclidean plane, a network search window moves along a network and captures incidents that are found within the extent of the network search window in a single instance.

Shiode S and Shiode N [10] proposed a street-level space-time hotspot detection method to analyze crime incidents recorded at the street-address level and provide description of the micro-level variation of crime incidents over space and time. It expanse the notion of search-window techniques widely used in crime science by developing a method that can account for the spatial temporal distribution of crime incidents measured in network distance.

The network-based search window covers part of the street network in such a way that its total length remains the same, while its space-time equivalent is formed by extruding it in the vertical direction. Since they have formed along the street network, both search windows are confined by the layout of the network and change their shape accordingly as they move along the network.

In 2008, Takahashi [11] proposed a flexible space-time scan statistic which imposes a three dimensional prismatic window with an arbitrarily shaped base *Z*. For any given region *i*, the set of arbitrarily shaped bases consisting of *k* connected regions $(1 \le k \le K)$ including *i* are created. To avoid detecting a cluster of unlikely peculiar shape, the connected regions are restricted as the subset of the *K*-nearest neighbors to the region *i*, where *K = 1* implies the region *i* itself. Let $Z_{ik(j)}, j = 1, \dots, j_{ik}$denote the *j-th* window which is a set of *k* regions connected starting from the region *i*, where j_{ik} is the number of j satisfying $Z_{ik(j)} \subseteq Z_{ik}$ for $k = 1, \dots, K$. Then, all the windows to be scanned are the prisms whose base is included in the set

$$z_2 = \{z_{ik(j)} | 1 \le i \le m, 1 \le k \le K, 1 \le j \le j_{ik}\}$$

with height in the set $y = \{[t_p - t + 1, t_p] | 1 \le t \le T\}$. That is, for any given region *i*, the cylindrical scan statistic consider *K* concentric circles for the base, whereas the flexible scan statistic consider *K* concentric circles plus all the sets of connected regions including the single region *i*, whose centroids are located within the *K*-th largest concentric circle.

Define *L(W)* as the likelihood under the alternative hypothesis that there is a cluster in the space-time window $W(\in w)$, where $w = (z_2 \times y)$ and L_0 the likelihood under the null hypothesis. Then, conditioning on the observed total number of cases, *N*, in the definition of the space-time scan statistic *S* is the maximum likelihood ratio overall possible windows *W*,

$$S = \frac{\max_{W \in w}\{L(W)\}}{L_0} = \max_{W \in w}\left\{\frac{L(W)}{L_0}\right\}$$

The window for which the likelihood ratio is maximized identifies the most likely cluster.

IV. METHODS ALLOWING CLUSTERS OF VARIOUS SHAPES

Rogerson and Yamada [12] have developed a methodology for monitoring spatial pattern of diseases for the comparison of disease surveillance in which the univariate and multivariate cumulative sum methods were used.

Let X_t represents the observed number of cases of disease in a given region during time period t. Assuming that X_t is approximately normally distributed, with mean equal to the expected number of cases E_t, and known variance σ^2, the one-sided cumulative sum at time *t* is

$$S_t = \max(S_{t-1} + z_t - k, 0)$$
$$S_0 = 0$$

where z is the standardized value of X. The parameter *k* is usually chosen to be equal to 0.5, the value of *k* minimizes the time to detect an increase of 2k standard deviations in the magnitude of the mean. The cumulative sum accumulates values of *z* that are in excess in *k*. when the cumulative sum (or cusum) exceeds a threshold, *h*, an alarm is sounded, indicating that the mean has increased.

The univariate approach consists of simultaneously and independently monitoring the disease rate in each region; the multivariate approach is used for accounting co-variation between the regions, and thus overcame the limitation of univariate approaches, that is their lack of ability to account for the spatial auto correlation of regional data. The multivariate methods are also having the limitation of difficulty in accurately specifying the multiregional covariance structure.

Multivariate monitoring is based upon the cumulative difference between the observed and expected number of cases:

$$S_t = \sum_{j=t-n_{t+1}} (O_j - E_j)$$

where O_j and E_j are vectors of observed and expected counts at time *j*. If there are *p* regions then there would be *p* elements in each vector, corresponding to entries for each of the *p* regions. The vector of observed counts is assumed to be approximately multivariate normal. The quantity n_t represents the number of time periods since the cumulative sum was last reset to zero.

The quantity monitored is $MC1_t = max\{0, \|S_t\| - kn_t\}$

$$n_t = n_{t-1} + 1, MC1_{t-1} > 0$$
$$= 1, \quad otherwise$$

where the norm of S is a scalar representing the multivariate distance of cumulated differences from the target $\|S_t\| = \sqrt{S_t' \Sigma^{-1} S_t}$, where, Σ is the variance-covariance matrix associated with the *p* regions. The value of the parameter *k* is chosen to be equal to one of the multivariate distance from the target vector to the hypothesized vector.

It is established that when the degree of spatial autocorrelation is low, the univariate method is generally better at detecting changes in rates that occur in a small number of regions; the multivariate is better when changes occurs in a large number of regions.

Jacquez et al [13] defined a spatially and temporally local case-control cluster statistic, called Q-Statistic for detecting case control clustering of residential histories given by,

$$Q_{i,k,t} = c_i \sum_{j=1}^{N} \eta_{i,j,k,t}\, c_j$$

$$\eta_{i,j,k,t} = \begin{cases} 1 & if\ and\ only\ if\ j\ is\ a\ k\ nearest \\ & neighbour\ of\ i\ at\ ime\ t \\ 0 & otherwise \end{cases}$$

This is the count, at time *t*, of the number of *k* nearest neighbors of case *i* that are cases, and not control. When *i* is a control

$$Q_{i,k,t} = 0$$

For the i^{th} residential history, the sum, overall T+1 time points of the local spatial cluster statistic is $Q_{i,k,t}$. Then then local clusters of residential histories are identified by the statistic,

$$Q_{i,k,} = \sum_{t=0}^{T} Q_{i,k,t}$$

This is the number of cases that are k-nearest neighbors of the i^{th} residential history (a case), summed overall T+1 time points. It will be large when cases tend to cluster around the i^{th} case through time. The statistics will be evaluated for each of the cases to identify those cases with low *p*-values.

In case control studies, the methods for analyzing space-time variation in risk usually ignore residential mobility. An improved approach for analyzing case control data for mobile individuals is suggested by Jacquez et al [14]. This is done by using time-dependent nearest neighbor relationships, the global, local and focused clustering of residential histories are quantified. The Q statistics suggested by Jacquez et al [13] for case control clustering of residential histories was taken and it is extended to account for risk factors and covariates and duration weights. The duration weighted version of statistic is

$$Q^{E}_{F,k,\omega_t} = \omega_t \sum_{j=1}^{N} \eta_{F,j,k,t} c_j e_{j,t}$$

This is the count of the numbers of cases with active exposure traces that are *k*-nearest neighbors of the focus at duration timeω_t.

$$e_{j,t} = \begin{cases} 1 & if\ and\ only\ if\ time\ t\ is\ within\ the \\ & exposure\ trace\ for\ individual\ i \\ 0 & otherwise \end{cases}$$

This statistic may be tested for its significance evaluated by constructing exposure traces for the controls

and by then repeatedly allocating case-control identifiers across the *N* lifelines that are *k*-nearest neighbors of the focus in for the construction of the reference distribution for $Q^{E}_{F,k,t}$.

Cook, Gold, Li [15] proposed new spatial cluster detection method for repeated outcomes using cumulative geographic residuals. Spatial methods developed so far focused on cross-sectional outcomes that are binary or continuous and this method is proposed for longitudinal outcomes. The main advantage of this method is that information of study participant's relocation can readily be incorporated.

Another space-time variant which uses kernel density estimation and scan statistics was proposed by Tomoki Nakaya and Keiji Yano [16], giving the three dimensional disease mapping called space-time cubes. They have explored the possibility of three-dimensional mapping of crime events in space time cube with the aid of space-time variants of kernel density estimation and scan statistics for an effective interpretation of spatio-temporal patterns. The methodology proposed enables simultaneous visualization of the geographical extent and duration of clusters, by which stable ad transient space time clusters can be intuitively differentiated. The temporal inter-cluster associations showing that transient clusters alternatively appeared in a pair of hotspot regions are also revealed through this method. This method is applied to crime epidemiology from a data set of space-time events. The space time kernel density estimate proposed by Brunsdon et al. [17] is used, which is given by,

$$\hat{f}(x,y,t) = \frac{1}{nb_s^2 b_t}\sum_i K_s\left(\frac{x-x_i}{b_s},\frac{y-y_i}{b_s}\right)K_t\left(\frac{t-t_i}{b_i}\right)$$

Where, $\hat{f}(x,y,t)$ is theestimator of location density, (x,y,t), n is the number of events, and b_s is the spatial bandwidths and b_t is the temporal bandwidths.

Toshiro Tango et al [18] proposed a new space-time scan statistic which compares the observed number of cases with the unconditional expected number of cases, takes a time-to-time variation of Poisson mean into account, and implements an outbreak model to capture localized emerging disease outbreaks more timely and correctly. It is stated that Kulldorff's and Takahashi et al.'s space-time scan statistics not suitable for syndromic surveillance, because (i) it uses conditional expected number of cases for detecting emerging disease outbreaks which is not relevant, (ii) it considers the hot-spot model which is less powerful in detecting outbreaks, as the temporal pattern of emerging disease outbreaks, usually has a gradual or steep increase in the number of cases under study in the initial stage and (iii) without taking in to account of Poisson mean or temporal overdispersion, the false alarm rate increases in the field of statistical process control.

Consider the situations where an entire study area is divided into *m* regions (for example, counties, zip-codes, enumeration districts) with each region periodically reporting the number of cases n_{it} (for region *i* at time *t*) of a disease or syndrome under study. The outbreaks that are present in the following *T* time intervals or temporal windows is given by

$$I_u = [t_P - u + 1, t_P], u = 1, \dots, T$$

where *T* is a pre-specified maximum temporal length of the cluster or outbreak. It is assumed that under the null hypothesis of no outbreaks, the number of cases N_{it} in region *i* (*i*=1,…,*m*) at some surveillance time *t* follows an independent negative binomial distribution by taking the possibility of non eligible time-to-time variation of Poisson mean or temporal over dispersion, into account. The negative binomial distribution is given by,

$$\Pr\{N_{it} = n_{it}|\mu_{it},\phi_{it}\},$$
$$= \frac{\Gamma(\phi_{it}+n_{it})}{\Gamma(\phi_{it})n_{it}!}\left(\frac{\phi_{it}}{\phi_{it}+\mu_{it}}\right)^{\phi_{it}}\left(\frac{\mu_{it}}{\phi_{it}+\mu_{it}}\right)^{n_{it}}$$

and $E(N_{it}) = \mu_{it}$ and $\mathrm{Var}(N_{it}) = \mu_{it} + {\mu_{it}}^2/\phi_{it} = \mu_{it}\omega_{it}$
The temporal overdispersion is given by

$$\omega_{it} = 1 + \mu_{it}/\phi_{it}$$

and ϕ_{it} denote the parameter regulating overdispersion.

Dong et al. [19] designed a grid-based method to detect irregular shaped space-time clusters, stating that regularly shaped clusters such as circular, elliptic and rectangular etc. cannot work well on irregularly shaped clusters. In Grid based Scan method, a cluster is asymptotically described by a set of connected grid cells and is computed by a fast greedy region-growing algorithm with elaborating cluster merging in the process. The time complexity of GridScan is linear to the number of grids, making it scalable to very large datasets. It is demonstrated that this approach greatly outperform when compared with existing ones in terms of accuracy, efficiency, and scalability.

In GridScan, uses the density measures in prior arts, the likelihood function and for each region *Z*, the likelihood *L(Z)*, is calculated either under Bernoulli model or Poisson model.

Let n_Z denote the number of cases inside a region *Z*, n_G the global number of cases of the entire dataset, N_Z the number of population inside, *Z*, and N_G is the global number of population of the dataset. When using the Bernoulli model, *L(Z)* is defined as follows.,

$$L(Z) = \left(\frac{n_Z}{N_Z}\right)^{n_Z}\cdot\left(1-\frac{n_Z}{N_Z}\right)^{N_Z-n_Z}\cdot\left(\frac{n_G-n_Z}{N_G-N_Z}\right)^{n_G-n_Z}\cdot\left(1-\frac{n_G-n_Z}{N_G-N_Z}\right)^{(N_G-N_Z)-(n_G-n_Z)}$$

When using the Poisson model,

$$L(Z) = \left(\frac{n_Z}{E(n_Z)}\right)^{n_Z}\cdot\left(\frac{n_G-n_Z}{N_G-E(n_Z)}\right)^{n_G-n_Z}$$

where $E(n_Z) = \frac{n_G}{N_G} N_Z$ denotes the expected number of cases inside Z under the null hypothesis. Z with $\frac{n_Z}{N_Z} > \frac{n_G - n_Z}{N_G - N_Z}$ means more cases are observed than expected under the null hypothesis and Z with $\frac{n_Z}{N_Z} < \frac{n_G - n_Z}{N_G - N_Z}$ represents the underdensity. Note that this method can be extended for prospective cluster detection also.

Fanaee-T and Gama [20] have developed new methodology for disease clustering called Eigenspot method, instead of the usual method of exhaustive search over the space, tracking the changes in space time correlation structure. This new approach has more computational efficacy, and also makes no assumption about the data distribution, hotspot shape or the data quality. The primary idea is that joint combination of abnormal elements in the principal spatial and the temporal singular vectors, the location of hotspots in the spatiotemporal space can be approximated. It is claimed that the STScan methods are more complex in computation, associated with strong parametric model assumptions, which may be irrelavant in real time applications. Also their inefficient for reduction of irregular clusters and vulnerable against the noises and outliers. The new method consists of tracking the changes in the space time correlation structure using the Eigen-space technique without any parametric model assumptions or prior knowledge about the hotspots and hence suitable for detecting irregular shaped hotspots even from noisy datasets.

In some situations, where the disease occurrences tend to cluster very irregularly shaped areas, the conventional circular, elliptical or square shaped scanning windows to discover disease clusters are not flexible. To overcome this problem, a new algorithm is proposed by Sami Ullah [21] which uses a co-clustering strategy to prospective and retrospective space-time disease clusters with no restriction on shape and size. Instead of an in-depth search over space the proposed method detects space-time disease clusters by tracking the changes in space-time occurrence structure. In Kulldorff [6] method the likelihood ratio is calculated for each sub-region over each time point to allow for space-time cluster detection. The likelihood ratio scores can be organized in the form of a data matrix in which the rows indicate the sub-regions and the columns the time points. The Bregman Block Average Co-clustering algorithm (BBAC) is applied to the density matrix to get an optimum co-clustered matrix by the concurrent analysis of rows and columns similarity.

The contents of the rows and columns associated with the largest element of the co-cluster matrix are combined to approximate the most likely clusters, while those of the rows and columns related to the second largest element are combined to approximate secondary clusters. This can be visualized on heat maps using different colors for the different clusters.

This algorithm simultaneously partitions rows of a data matrix into clusters based on the similarities along all columns, and columns into clusters based on the similarity along all rows. It concurrently assigns rows into row-clusters and columns into column-clusters. By first initializing *m* row-clusters and *n* column-clusters, this method calculates the co-cluster matrix, where *m* and *n* are desired number of row-clusters and column-clusters respectively. The loss function taken is difference of the co-cluster matrix and the original matrix. To obtain an optimal co-clustered matrix, this difference is minimized by repeatedly allocating every row to the nearest row-cluster and every column to the nearest column-cluster until a convergence is seen.

V. CONCLUSION

Spatial cluster detection methods have developed significantly in the past decade. Improvements are made in terms of scanning window of flexible shapes and incorporating more related variables in the study. Different characteristics of spatial associations are also utilized for effective disease monitoring. However by incorporating more information about the environmental, social and biological factors, cluster detection process and also surveillance would be made more meaningful. Also, more study and process-based understandings in particular with geographic contexts will greatly enhance the potential application the methods for public health planning, policy making and intervention.

REFERENCES

[1] Cromley, E., and McLafferty: *GIS and Public Health.* 2nd edition. New York, NY: Guilford **2011**.

[2] Shi, X: *Selection of Bandwidth Type and Adjustment Side in Kernel Density Estimation over Inhomogeneous Backgrounds***.** International Journal of Geographical Information Science, **24: 643-660, 2010.**

[3] Openshow, S., M. Charlton, and A.Craft: *Searching for Leukaemia Clusters using a Geographical Analysis Machine,* Papers in Regional Science, **64: 95-106, 1988.**

[4] Besag, J., and J. Newell: *The Detection of Clusters in Rare Diseases,* Journal of the Royal Statistical Society, Series A **154(1):143-155, 1991**.

[5] Rushton, G., and P. Lolonis: *Exploratory Spatial Analysis of Birth Defect Rates in an Urban Population.* Statistics in Medicine, **15:717-726, 1996**.

[6] Kulldorff M: *A Spatial scan statistic,* Communication in Statistics-Theory and Methods 1997, **26(6):**1481-1496.

[7] Kulldorff, Martin, Athas WF, Feuer EJ, Miller BA, Key CR: *Evaluating cluster alarms: Space time scan statistic and brain cancer in Los Alamos, New Mexico.* Am J Public Health,**88(9):1377-1380, 1998.**

[8] Iyengar VS: *On detecting space-time clusters*. Proceedings of the 2004 ACM SIGKDD International non Knowledge Discovery and Data Mining, **1:587-592, 2004.**

[9] Shiode, S: *Street-level Spatial Scan Statistic and STAC for analyzing Street Crime Concentrations*. Transactions in GIS, **15 (3): 365-383, 2011.**

[10] Shino Shiode and Narushige Shiode: *Network based Space-time Search-window technique for hotspot detection of street-level crime incidents*. International Journal of Geographic Information Science, **5:866-882, 2013.**

[11] Kunihiko Takahashi, Martin Kulldorff, Toshiro Tango and Katherine Yih: *A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring*. International Journal of Health Geographics, **7:1-14, 2008**.

[12] Peter A. Rogerson and Ikuho Yamada: *Monitoring change in spatial patterns of disease: comparing univariate and multivariate cumulative sum approaches.* Statistics in Medicine, **23:2195-2214, 2004**.

[13] Jacquez GM, Kaufmann A, Meliker J, Goovaerts P, AuRuskin G, Nriagu J: *Global, local and focused geographic clustering for cease control data with residential histories*. Environmental Health, **4:4 2005**.

[14] Geoffrey M Jacquez, Jaymie R Meliker, Gilllian A AvRuskin Pierre Goovaerts, Andy Kaufmann, Mark L Wilson and Jerome Nriagu: *Case-Control Geographic Clustering for Residential Histories Accounting for Risk factors and Covariates*. International Journal of Health Geographics, **5:32, 2006.**

[15] Andrea J. Cook, Diane R Gold and Yi Li: *Spatial Cluster Detection for Repeatedly Measured Outcomes while Accounting for Residential History*, Biometric Journal, **51 (5): 801-818, 2009**.

[16] Tomoki Nakaya, Keiji Yano: *Visualising Crime in a Space-Time cube: An explanatory Data-analysis Approach Using Space-Time Kernel Density Estimation and Scan Statistics*. Transactions in GIS, **14(3): 223-239, 2010.**

[17] Brunsdon C, Corcoran J, and Higgs G: *Visualising Space time in crime patterns: A comparison of methods. Computers*, Environment and Urban Systems, **31: 52-75, 2007**.

[18] Toshiro Tango, Kunihiko Takahashi and Kazuaki Kohriyama: *Space-Time Scan Statistics for Detecting Emerging Outbreaks*, Biometrics, International Biometric Society, **67, 106-115, 2011**.

[19] Weishan Dong, Xin Zhang, Li Li, Changhua Sun, Lei Shi, Wei Sun: *Detecting Irregularly Shaped Significant Spatial and Spatio-Temporal Clusters*. SIAM International Conference on Data Mining, **732-743, 2012. DOI:10.1137.1.9781611972825.63.**

[20] Hadi Fanaee-T and Joae Gama: *Eigenspace method for spatiotemporal hotspot detection*, Expert Systems, **32(3):454-464, 2014. doi:10.1111/exsy.12088.**

[21] Sami Ullah, Hanita Daud, Sarat C. Dass, Habib Nawaz Khan, Alamgir Khalil: *Detecting space-time disease clusters with arbitrary shapes and sizes using a co-clustering approach*, Geospatial Health, **12(2):567, 2017**.

[22] Martin Kulldorff: *Prospective time geographical disease surveillance using a scan statistic*, Royal Statistical Society, **164(1):61-72, 2001**.

[23] G. Li, R. Haining, S. Richardson, N. Beset: *Space-Time Variability in Burglary Risk: A Bayesian Spatio-Temporal Modelling Approach*, Spatial Statistics, Elsevier, **9:180-191, 2014**.

[24] Sara Mclafferty: *Disease Cluster detection methods: recent developments and public health implications*, Annals of GIS, **21(2): 127-133, 2015**.

AUTHORS PROFILE

R. Jaisankar was awarded Doctor of Philosophy from Annamalai University in 2004. At present he is working as a Professor of Statistics in Bharathiar University. His research interests are Biostatistics, Spatial analysis in Epidemiology, Stochastic modeling and Design of Experiments. He has published more than 30 research papers in reputed national and international journals and guided many students for their research work.

M. Ranjani, who has got Master Degree in Statistics in 2017, is doing Doctor of Philosophy at Bharathiar University at present. Her research interests are Biostatistics and Spatial Analysis.

N. Kausalya has obtained Master Degree in Statistics in 2018 from PSG college of Arts and Science Coimbatore. Now she is doing Master of Philosophy at Bharathiar University. Her research interests are Biostatistics and Spatial Analysis.

