# Using Point Pattern Analysis Techniques to Describe Spatial Arrangement of Date Palm Trees in Iraq

[1]Bashar S. Al-Dabbagh, [2]Laith Alzubaidi, [2]Reem Ibrahim Hasan and [3]Haider A. Alwzwazy
[1]Higher Polytechnic School, Autonomous University of Madrid, Madrid, Spain
[2]Department of Smart Cities, University of Information Technology and Communications,
Baghdad, Iraq
[3]Department of H&c-it, Missan Oil Company, Missan, Iraq

**Abstract:** Date Palm Trees (DPT) are the most growing plants in Iraq and the second source of Iraqi income. Distribution of date palm trees shows a lot about the pattern of DPT over the years which helps local government to know what area needs to grow more DPT. The spatial distributions of DPT locations in Iraq are explored by different methods of spatial point pattern analysis. Many methods for the analysis of point pattern data have been applied in a wide range of scientific fields. Describing the pattern of data can show a lot of details about the data special pattern. Many point pattern techniques have been implemented to describe the pattern however not all of them provide a clear description of the data in the space. Some of them miss some points which lead to an inaccurate pattern. In this study, Different clustering techniques (G, F, K spatial point pattern analysis techniques) are implemented to describe the spatial arrangement of DPT in Iraq. We also present a review of some spatial point pattern analysis techniques. Our results show effective performance in term of special pattern. Based on results of spatial point pattern analysis techniques it can be specified that the trees are evenly spaced or clustered in groups.

**Key words:** Clustering, spatial point pattern, K-function, palm trees, performance, spatial point

## INTRODUCTION

Point pattern analysis is the study of spatial distribution of locations in space. It has been applied to describe the locations of disease, crimes and plants of different species (O'Sullivan and Unwin, 2014; Potter *et al.*, 2010). The spatial point pattern has three main types which are complete spatial randomness, regularity and clustering (Loosmore and Ford, 2006; Li and Zhang, 2007), illustrated in Fig. 1.

Point patterns can be described either by first order effects that analyze number of events per unit area (point density) or second order effects that describe interaction between points (point separation) (Gabriel, 2014; Ojiambo and Holmes, 2011). In our study, we explore use of second order distance-based methods to describe distribution pattern of points. Specifically the measures F, G and K-function described here have been implemented to describe the patterns of Date Palm Trees (DPT) in Iraq. Different techniques and devices can involve in order to have whole system along with point patterns (Ammar *et al.*, 2018; Farhan *et al.*, 2018). Iraq is the native land of the date palm and it was the domestication center of this crop. Furthermore, for a
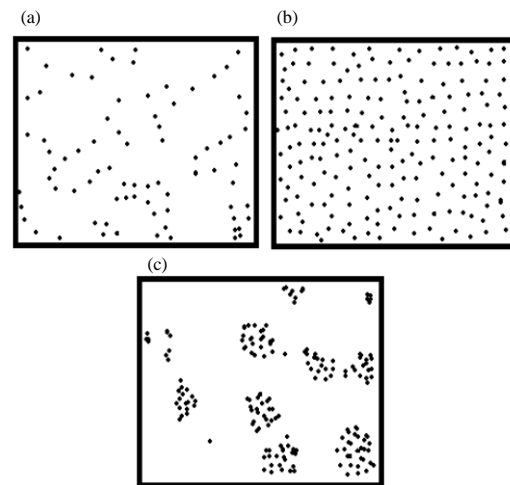


Fig. 1: Spatial point pattern types: a) Complete Spatial Randomness (CSR); b) Regularity and c) Clustering

few years, Iraq was the largest producer of dates in the world. Different factors negatively have affected both the production and natural genetic diversity of the crop (Khierallah *et al.*, 2015). In this research, we have

---

**Corresponding Author:** Bashar S. Al-Dabbagh, Autonomous University of Madrid, Madrid, Spain

developed MATLAB programs to compute these measures (G, F, K spatial point pattern analysis techniques). We have applied G, F and K-functions to describe the pattern. We found out that G and F omit a lot of points while K shows a clear description of trees in space (evenly or clustered). We present an explanation of spatial point pattern analysis techniques.

## MATERIALS AND METHODS

**Spatial point pattern analysis using distance-based measures:** Distance-based point pattern analysis methods rely on distances between points (often nearest-point distance) (Marcon and Puech, 2009). Let $C_i$ and $C_j$ be two points (centroids). The distance between $C_i$ and $C_j$ is computed as Eq. 1:

$$d\left(C_i, C_j\right) = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \qquad (1)$$

Once the distances between each pair of points in an space are computed, the nearest distance for each point can be found. Let $d_{min}(C_i)$ denote the distance of point $C_i$ to its nearest point. Mean nearest-point distance is then computed as:

$$\overline{d}_{min} = \frac{\sum_{i=1}^{n} d_{min(C_i)}}{n} \qquad (2)$$

$D_{min}$ summarizes all the nearest-point distances by a single mean value but omits many details about the point pattern. More detailed measurements are needed to capture the differences between point distribution patterns. The following subsections will describe some well-established point pattern measures using the simple synthetic case illustrated in Fig. 2 (O'Sullivan and Unwin, 2014) where for each point; its nearest point is marked with an arrow pointing from the original point to its nearest point. Table 1 list for each point in Fig. 2, ID of the nearest point and associated distance. Note that the nearest-point relationship is not symmetric.

**G-function**: Also known as refined nearest point, uses information in Table 1. In-stead of calculating the mean minimum distance, computes the cumulative frequency distribution of the nearest-point distances. It is described by Eq. 3:

$$G\left(d\right) = \frac{\#\left(d_{min}\left(C_i\right) < d\right)}{n} \qquad (3)$$

G(d) gives the fraction of all the nearest-point distances in the pattern that is lower than d. Figure 3 illustrated the G-function corresponding to points in
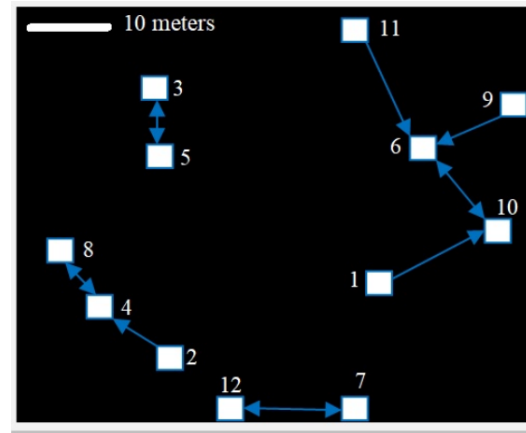


Fig. 2: Distances between points and their nearest neighbors (O'Sullivan and Unwin, 2014)
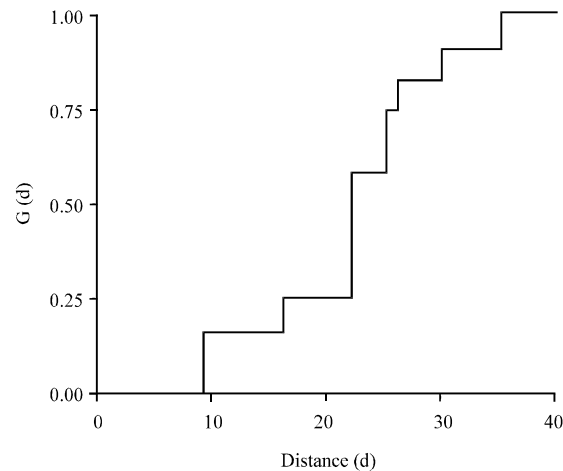


Fig. 3: G-function for the points shown in Fig. 2 (O'Sullivan and Unwin, 2014)

Table 1: Sample computations for the nearest-point distances for the point pattern shown in Fig. 2 (1)

| Points | X | Y | Nearest point | $d_{min}$ |
|---|---|---|---|---|
| 1 | 66.33 | 32.54 | 10 | 12.59 |
| 2 | 22.52 | 22.39 | 4 | 15.64 |
| 3 | 31.01 | 81.21 | 5 | 21.11 |
| 4 | 9.47 | 30.02 | 8 | 9.00 |
| 5 | 30.78 | 60.10 | 3 | 21.14 |

Fig. 2. The shape of the G-function gives information on how points are spaced in a pattern. If points are closely clustered together, then G-function boosts rapidly at short distances. If points tend to be evenly spaced, then G boosts slowly up to the range of distances at which most points are spaced and only then boosts rapidly (O'Sullivan and Unwin, 2014).

**F-function:** It is closely similar to G-function but rather than accumulating the fraction of nearest-point distances

between points in a pattern, selects random point locations within study region, computes minimum distance from these locations to any point in the original point set. F-function is computed in Eq. 4:
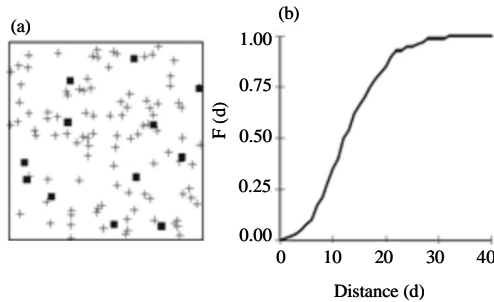
$$F(d) = \frac{\#\left(d_{min}\left(P_i, C\right) < d\right)}{m} \qquad (4)$$



Fig. 4: F-function for the points shown in Fig. 2 (O'Sullivan and Unwin, 2014)

Where:

| | | |
|---|---|---|
| $P_i$ | = | The randomly selected Point locations |
| $d_{min}(P_i, C)$ | = | The minimum distance between the randomly selected location and the points in the original point set |

The advantage of F-function is its ability to boost sample size m to get a smoother cumulative frequency curve (Fig. 4).

**K-function**: Also called as Ripley's K-function is a more descriptive point pattern measure (Cetin, 2008). G-function and F-function rely only on the nearest point distances to describe a point pattern. K-function uses all of the distances between point pairs in the pattern. K(d) corresponds to area normalized mean of number of points within distanced of another point. K-function is computed in Eq. 5. Figure 5 illustrates the K-functions for two point patterns:
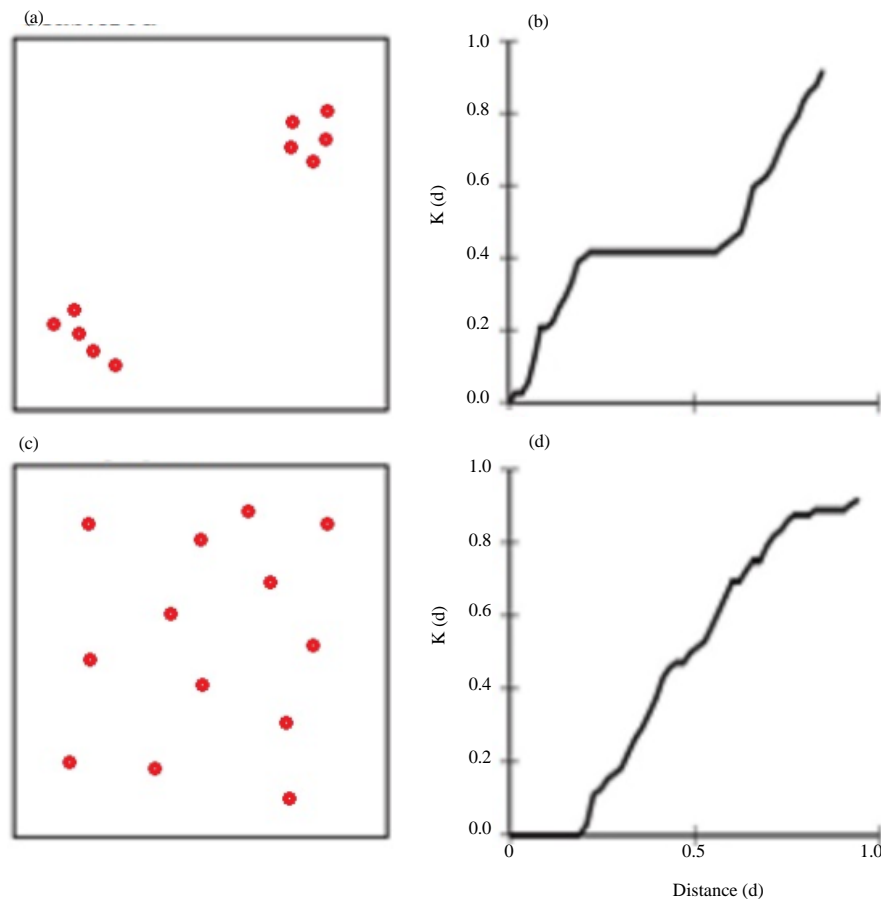


Fig. 5: K-function for evenly and clustered patterns (O'Sullivan and Unwin, 2014)

## RESULTS AND DISCUSSION

In this study, we present the following, dataset, experimental results.

**Dataset:** Random points (synthetic points) with evenly space and clustered points that present the palm trees location in Iraq.

**Experimental results:** We implemented the point pattern description functions F-, G- and K-described in MATLAB and used on random points (synthetic points) we show sample results. As can be seen particularly in sample points K-function nicely captures point distribution pattern, differentiates between regularly spaced points versus clustered point's distributions. Spatial distribution of points can reveal useful information

about pattern of points. We will use point pattern analysis and the described measures to quantitatively study spatial distributions of points.

According to what we have explained above K-function describes the input points in evenly spaced better then G, F-functions. In order to emphases on K-function we describe it with different input images. Based on the output we can know a lot about the pattern.we have tested K-function with three cases, clustered points as shown in Fig. 6, evenly space points as shown in Fig. 7, evenly space points with more points to show a smooth curve shown in Fig. 8. K-function described input samples in efficient way.

Figure 9a shows Iraqi map with synthetic points which present three groups of date palm trees location on the map. As a result Fig. 9b, demonstrates three groups of trees.
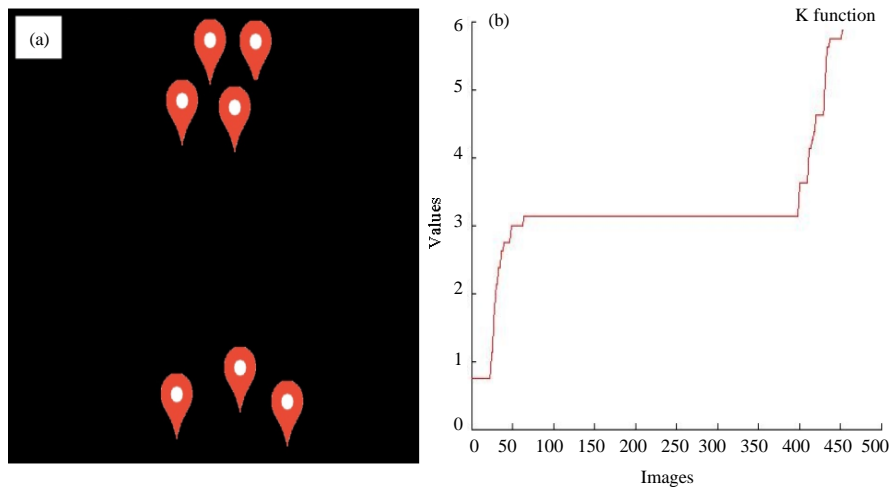


Fig. 6: K-function with clustered sample points; a) Input image and b) Output graph
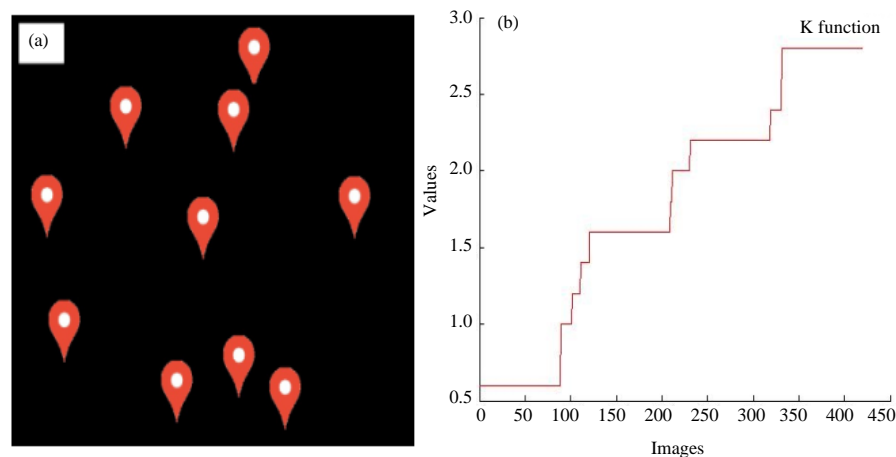


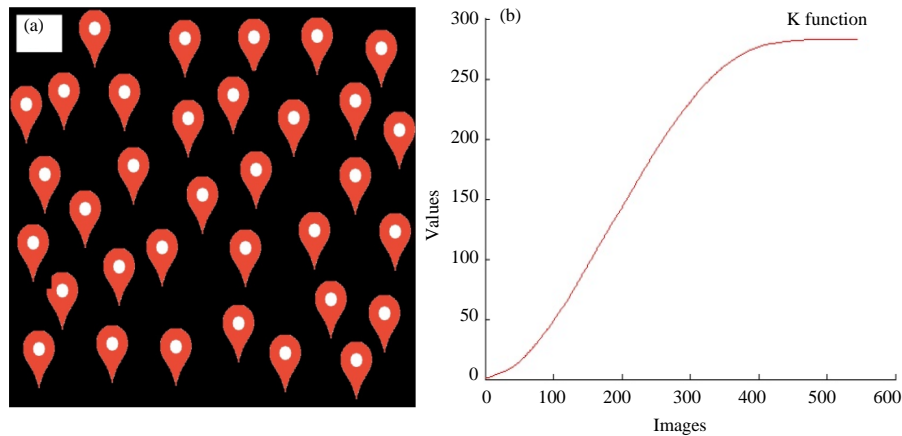Fig. 7: K-function with evenly sample points: a) Input image and b) Output graph

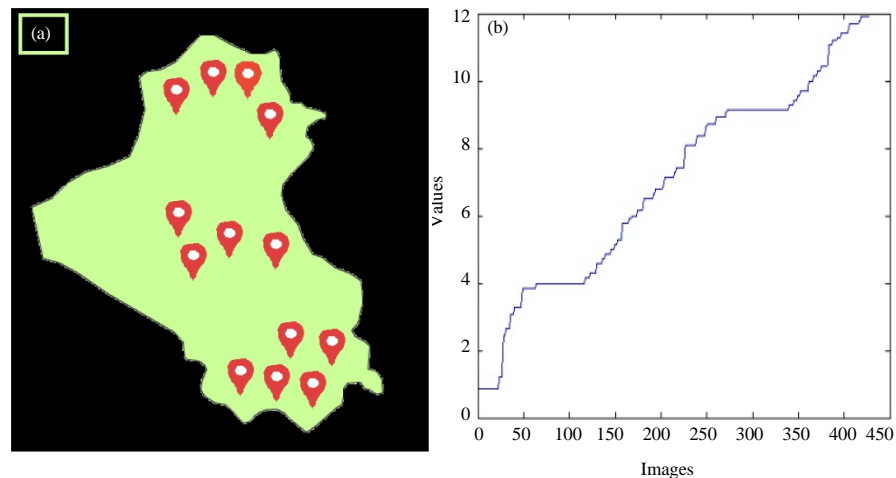Fig. 8: K-function with intense evenly sample points: a) Input image and b) Output graph



Fig. 9: K-function on sample points of Iraqi map: a) Input image map and b) Output graph

## CONCLUSION

In this study, we have developed a clearer idea of the concept of pattern and how it can be related to process. In principle, any pattern can be described using different measures. We have implemented G, F, K-functions which are spatial point pattern analysis techniques to describe spatial distribution of Date Palm Trees (DPT). We also presented a brief review of spatial point pattern analysis techniques. We applied these functions using MATLAB and we found that K-function has the best results compared to other functions.

## ACKNOWLEDGEMENTS

## REFERENCES

Ammar, A.A., S. Bashar, Al-Dabbagh and L. Alzubaidi, 2018. Review paper on wireless network-on-chip architecture. J. Eng. Appl. Sci., 13: 5142-5149.

Cetin, A., 2008. Assessment and modelling of particle clustering in cast aluminum matrix composites. Ph.D Thesis, Middle East Technical University, Ankara, Turkey.

Farhan, L., L. Alzubaidi, M. Abdulsalam, A.J. Abboud and M. Hammoudeh *et al.*, 2018. An efficient data packet scheduling scheme for Internet of Things networks. Proceedings of the 2018 1st and 3rd International Scientific Conference on Engineering Sciences (ISCES), January 10-11, 2018, IEEE, Diyala, Iraq, ISBN:978-1-5386-1499-0, pp: 1-6.

Gabriel, E., 2014. Estimating second-order characteristics of inhomogeneous spatio-temporal point processes. Method. Comput. Appl. Probab., 16: 411-431.

Khierallah, H.S.M., S.M. Bader, K.M. Ibrahim and I.J. Al-Jboory, 2015. Date Palm Status and Perspective in Iraq. Date Palm Genetic Resources and Utilization, Al-Khayri, J., J.S. Mohan and J. Dennis, (Eds.). Springer, Dordrecht, Netherlands, pp: 97-152.

Li, F. and L. Zhang, 2007. Comparison of point pattern analysis methods for classifying the spatial distributions of spruce-fir stands in the North-East USA. For., 80: 337-349.

Loosmore, N.B. and E.D. Ford, 2006. Statistical inference using the G or K point pattern spatial statistics. Ecol., 87: 1925-1931.

Marcon, E. and F. Puech, 2009. Measures of the geographic concentration of industries: Improving distance-based methods. J. Econ. Geogr., 10: 745-762.

Ojiambo, P.S. and G.J. Holmes, 2011. Spatiotemporal spread of cucurbit downy mildew in the Eastern United States. Phytopathol., 101: 451-461.

O'sullivan, D. and D. Unwin, 2014. Geographic Information Analysis. 2nd Edn., John Wiley and Sons, Hoboken, New Jersey, USA., ISBN:9781119023876, Pages: 432.

Potter, P., N. Ramankutty, E.M. Bennett and S.D. Donner, 2010. Characterizing the spatial patterns of global fertilizer application and manure production. Earth Interact., 14: 1-22.