Article:
Gezer, F orcid.org/0000-0002-8287-2763, Aykroyd, RG orcid.org/0000-0003-3700-0816 and Barber, S orcid.org/0000-0002-7611-7219 (2021) Statistical properties of Poisson-Voronoi tessellation cells in bounded regions. Journal of Statistical Computation and Simulation, 91 (5). pp. 915-933. ISSN 0094-9655

https://doi.org/10.1080/00949655.2020.1836184

Statistical properties of Poisson-Voronoi tessellation cells in bounded regions

Fatih Gezer[a], Robert G. Aykroyd[a], Stuart Barber[a]

[a]School of Mathematics, University of Leeds, Leeds, United Kingdom

ARTICLE HISTORY
Compiled October 16, 2020

Word Count: 6700

ABSTRACT
Many spatial statistics methods require neighbourhood structures such as the one determined by a Voronoi tessellation, so understanding statistical properties of Voronoi cells is crucial. While distributions of cell properties when data locations follow an unbounded homogeneous Poisson process have been studied, little attention has been given to how these properties change when a boundary is imposed. This is important when geographical data are gathered within a restricted study area, such as a national boundary or a coastline.

We study the effects of imposing a boundary on the cell properties of a Poisson Voronoi tessellation. The area, perimeter and number of edges of individual cells with and without boundary conditions are investigated by simulation. Distributions of these properties differ substantially when boundaries are imposed, and these differences are affected by proximity to the boundary. We also investigate how changes in such properties when boundaries are imposed vary over two-dimensional space.

KEYWORDS
Voronoi tessellation; Poisson point process; Bounded regions; Spatial statistics; Generalized gamma distribution

1. Introduction

Voronoi tessellation is a standard subdivision method which partitions a space into smaller disjoint regions called Voronoi cells based on point locations. Poisson point processes are useful models to portray the behaviour of spatial data points [1]. The spatial tiling produced by the Voronoi tessellation induced by points following a homogeneous Poisson process is referred to as a Poisson-Voronoi tessellation (PVT). Ripley [2] discusses the importance of space subdivision methods to investigate spatial splines, and gives examples of different spatial point patterns for both simulated and real data to relate the subject to the estimation of distributions of the locations within a region using the Voronoi tessellation. The Voronoi tessellation has been applied in different

CONTACT Fatih Gezer. Email: mmfge@leeds.ac.uk

sciences such as in seismology where Schoenberg [3] investigated the distribution of cell areas in a Voronoi tessellation based on the locations of earthquakes in Southern California; astronomy [4–6] to discover how galaxies are distributed in space; to investigate the conditions of the habitat of animals when they are establishing territories [7]; in agriculture for maximal weed suppression in plant crops [8] and to study atomic crystals [9], liquids [10], glasses [11], and wireless networks [12,13]. An application of constrained Voronoi tessellation is used in micro-structure modeling [14] where a new space subdivision method is introduced using inverse Monte Carlo based on conditions such as moving the randomly placed points until their geometric features obey a particular distribution.

Numerous authors have investigated the geometrical and statistical properties, such as the cell area, perimeter, numbers of edges or vertices, and vertex angles, of Voronoi cells in one, two, and three-dimensional spaces. Preliminary studies [15,16] investigated the mean of cell area, edge length, and number of faces for an aggregate of crystals which refers to points in the space, and Møller [17] emphasized the algebraic calculation of the first two moments. Koufos and Dettmann [13] have shown that the features of Poisson Voronoi cells do not change by conditioning on the location of a point in an infinite plane.

Even though the distributions of the properties of Voronoi cells has been investigated empirically, no exact distribution has yet been found. However, various authors have found that useful approximations can be made using the gamma distribution with appropriately chosen parameters. Kiang [18] proposed an appropriate fit for the length of Voronoi line segments in one dimension, the area of cells in two dimensions and the volume of polyhedrons in three dimensions using Monte Carlo methods. In addition, Weaire et al. [19] suggested analytic derivation of the properties and tested its performance through larger computer simulations. Other similar approaches to estimate the distribution of such properties of Voronoi cells for homogeneous Poisson point process are also done in [20–23] with increasing precision as the computational power increased over time. A generalized gamma distribution introduced by Stacy [24] was used in [21,23] and found to have a good fit. A recent paper investigated the area distribution of bounded Poisson Voronoi cells [13].

Our aim is to investigate how the properties of Voronoi cells change due to the imposition of different boundaries, which has received little attention in previous work. Our study is based on the Voronoi tessellation of a homogeneous Poisson point pattern, as discussed in Section 2. A simulation to investigate the boundary effect on the Poisson Voronoi cells is described in Section 2. Results are presented and discussed and comparisons are made to previous work in Sections 3 and 4 respectively. In Section 5, the change in behaviour of cells due to different point intensities is investigated and conclusions are made in Section 6

2. Voronoi tessellation and imposed boundaries

Consider a finite set of n recorded points $\boldsymbol{X} = \{\boldsymbol{x}_i \in \Omega; i = 1, 2, ..., n\}$ uniformly distributed within some finite region $\Omega \subset \mathbb{R}^2$ and let the points $\boldsymbol{X}$ follow a homogeneous Poisson process so that $n \sim Po(\rho|\Omega|)$ with intensity $\rho > 0$, where $|\Omega|$ denotes the area of the region Ω. The standard Voronoi tessellation partitions the region into a collection of disjoint subregions $\boldsymbol{V} = \{V_i; i = 1, \ldots, n\}$ called Voronoi cells. These cells have the

properties that $V_i \cap V_j = \varnothing$ for $i \neq j$ and $V_1 \cup \cdots \cup V_n = \Omega$. Each cell V_i is defined to be that segment of Ω which is closer to the corresponding point $\boldsymbol{x}_i$ than any other point in $\boldsymbol{X}$, hence

$$V_i = \{\boldsymbol{x} \in \Omega \big| \|\boldsymbol{x} - \boldsymbol{x}_i\| \leq \|\boldsymbol{x} - \boldsymbol{x}_j\| \quad \text{for} \quad j = 1, 2, ..., i-1, i+1, ..., n\}$$

where $\|.\|$ denotes Euclidean distance. Since the points $\boldsymbol{X}$ follow a homogeneous Poisson process, the resulting tessellation $\boldsymbol{V}$ is referred to as a *Poisson Voronoi tessellation* (PVT) and each V_i is called the *Poisson Voronoi cell* corresponding to location $\boldsymbol{x}_i$. This PVT allows key cell summaries to be defined, for example area, perimeter and number of cell edges.

The left panel of Figure 1 shows an example of the PVT with $\rho = 200$ within the unit square, $\Omega = [0, 1]^2$, where the circles (open or filled) are the locations of the recorded points. The solid rectangle is the boundary of the region within which the uniform random points are generated, and dashed lines indicate the convex hull of the recorded points. The locations shown with filled circles are the observed points located on the convex hull whereas the open circles are strictly interior to the convex hull. The small polygons, shown with solid lines, define the cells of the PVT with each cell containing exactly one point. The cell edges are perpendicular bisectors adjacent data points and the intersection of two edges defines a vertex. If all pairs of points sharing an edge are joined, then the resulting collection of lines is known as a Dirichlet triangulation.

In some applications there may be a pre-defined region within which points are recorded, such as a national boundary, or the region may be an artefact of the data collection process, such as quadrat counting. In other cases, however, the true location of the region may not be recorded and instead may need to be defined as part of the data analysis. In the analysis, the bounding region can be a suitable rectangle, the convex hull of the points, or some other specified region that plays a key role in this study. The right panel of Figure 1 is a close-up of a section of a PVT showing the range of different possible effects on PVT cells caused by imposing a boundary on a set of points. Faint grey points and cells are excluded from the data by the imposition of the vertical boundary. Open circles indicate points whose cells are unaffected by the boundary. Squares indicate data points whose original cells are intersected by the boundary. Filled circles indicate points whose original cells are not intersected by the boundary but are still changed by the imposition of the boundary due to the removal of points outside the boundary. Note that cells affected by the boundary may shrink or grow depending on the configuration of points outside the boundary.

Let N denote the number of edges, P the perimeter and A the area of a Voronoi cell. Meijering [15] showed

$$E(N) = 6, \quad E(P) = \frac{4}{\sqrt{\rho}}, \quad E(A) = \frac{1}{\rho} \tag{1}$$

for the two-dimensional case. Since the cell perimeter and area depend on ρ, it can be useful to standardize these by dividing by their expectation. We use actual, or non-standardized, values to visualize and show the summary statistics of these properties when we wish to consider the effects of changing the intensity. Standardization with respect to ρ is applied for parameter fitting purposes in later sections to yield parameter

estimates comparable across intensities, within different boundary scenarios, and with previous work.

3. Distributions of characteristic properties

3.1. General

We initially study by simulation cell properties of PVT on the infinite plane to establish a benchmark, and then impose unit square and convex hull boundaries to investigate how this affects the properties of cells compared to absence of a boundary. Estimated distributions of the properties will be compared for different boundary conditions and previous work through standardized values obtained from (1). Standardized cell area and perimeter are defined as $A^* = A \times \rho$ and $P^* = \sqrt{\rho}/4 \times P$, therefore, $E(A^*) = E(P^*) = 1$ for the standardized forms. The number of edges $E(N)$ is not standardized since it is independent of ρ.

Initially, we set the Poisson intensity of the points to $\rho = 200$. For each PVT realization, n spatially uniform random points in $[0,1]^2$ are generated with $n \sim Po(\rho)$ and their Voronoi tessellation computed. We sample from the distribution of PVT cells in the infinite plane using a process proposed by Tanemura [23]. One cell (say $\boldsymbol{x}_j$) is selected uniformly at random and moved to the centre of the region, so $\boldsymbol{x}_j$ is translated by a vector $\boldsymbol{\nu} = (0.5, 0.5)' - \boldsymbol{x}_j$. The positions of the other points are moved by the same translation using periodic boundary conditions to retain those which move outside $[0,1]^2$. Finally, properties of the selected cell are calculated. This procedure is repeated for $r = 10^6$ realizations, sampling one cell from each to generate an independent sample of Voronoi cells.

3.2. Cell area

Figure 2 shows the distribution of cell areas in the (a) infinite plane, (b) unit square and (c) convex hull cases — in the latter two, the histogram for the infinite plane is also shown, in grey, for comparison. Corresponding summary statistics are shown in Table 1. Results for the infinite plane use all 10^6 cell areas. However, our focus is on changes induced by imposing a boundary, so for the unit square and convex hull cases only cells whose area changes are included, to avoid domination by the majority of cells for which the imposed boundary has no effect. In our simulations, 28% of the cells were affected by the imposed boundaries.

In Figure 2, all three distributions are clearly positively skewed and unimodal, with a range from 0 to about 0.02 and mode slightly below the analytically calculated expectation of 0.05 for the infinite plane. Using the sample distribution for the infinite plane as a base-line, the peak in the distribution with the unit square boundary is lower and the spread greater, especially towards the right. With the convex hull boundary the distribution increases more rapidly, with a lower peak and with slower decay than in the infinite plane case.

Table 1 shows that the mean cell area for the infinite plane case equals the theoretical value, but is larger for the unit square and smaller for the convex hull among those cells affected by the boundary. The variability is similar in all cases and while all cases are positively skewed there is an increase in skewness when boundaries are imposed.

For the (excess) kurtosis, all are leptokurtic with the most extreme being the unit square case and the infinite plane case being closest to normal distribution kurtosis.

To further investigate how the mean cell area is affected by the imposition of boundaries, consider Figure 3 which shows the spatial variation of mean cell area as image plots (left) and as line plots along transects (right). Results are shown for both unit square (top row) and convex hull (bottom row) boundaries. Due to the vertical and horizontal symmetry, we have 'folded' the results into the bottom-left quadrant of the unit square to make patterns more clearly visible and reduce random sampling variation. Lighter intensities in the image plot indicate smaller cell area; recall that the expected mean area for the infinite plane is $1/\rho = 0.005$. The line plots present data following three transects through the image plot. The diagonal transect is from the origin to $(0.5, 0.5)'$; the centre transect includes data both horizontally from $(0, 0.5)'$ to $(0.5, 0.5)'$ and vertically from $(0.5, 0)'$ to $(0.5, 0.5)'$; the edge transect combines data along both axes starting from the origin. In each case the x-axis gives distance from the origin; the diagonal transect is longer than the others and the excess part of the diagonal transect is omitted since it shows no meaninful variation.

In the unit square case, the smallest values are in the corner of the region with low values along each edge. Slightly in from the edge the mean areas are higher and then return to the overall expected mean area. These patterns are more clearly visible in the line plots (right) where the diagonal starts at a very low value, peaking before returning to the average. Hence, it is clearly demonstrated that areas close to the boundary of the unit square, on average, start small then increase above the infinite plane mean when moving inwards but then return to the overall expected level.

Next, consider the bottom row of Figure 3, which shows the spatial variation in cell area when imposing the convex hull boundary. Again, the smallest mean areas are in the corner of the region, but this time the largest areas are only slightly above the overall expected mean value. There is a clear pattern of small areas near the boundary and a gradual increase moving inwards until the overall mean value is reached.

3.3. Cell perimeter

The empirical distributions of cell perimeter length for the infinite plane, unit square and convex hull cases are shown in Figure 4. All have a more symmetric appearance than the corresponding plots for cell area. Comparing the unit square case, panel (b), with the infinite plane case, the peak is lower and slightly further to the right with a slightly broader spread. In contrast, for the convex hull case, panel (c), the peak is slightly to the left, but is still lower, and again the variability is greater. From the numerical summaries in Table 1, the unit square case has the largest mean and the convex hull the smallest. The standard deviations for the infinite plane case is the smallest, with the other two having similar but much larger values. The unit square case is slightly more skew but has the only positive (excess) kurtosis. However, all skewness values are close to zero indicating symmetry and all kurtosis values indicate a distribution close to the kurtosis of the normal distribution. Hence, cell perimeter values might be approximated by a normal distribution.

Figure 5 show the spatial variation of the cell perimeter length for the unit square and convex hull cases. Recall that the expected value for the infinite plane case would be constant at $4/\sqrt{\rho} \approx 0.283$. For the unit square, the smallest perimeter occurs in the

corner with the highest value close by on the diagonal. All cells along the boundaries have shorter perimeter length with a band of higher perimeter cells slightly in from the boundary. In the interior, the perimeters follow the value of the cells in the infinite plane. For the convex hull case, again the smallest perimeter is for cells in the corner with small values along the boundary also. The peak along the diagonal from the corner is only slightly higher than surrounding cells and is only noticeable from the line plot.

3.4. Number of cell edges

Figure 6 shows the distributions for the number of edges of a cell for the infinite plane, unit square and convex hull cases with corresponding summaries in Table 1. The minimum number of edges is $N = 3$, corresponding to a triangle, and there are very few cells with more than 9 edges – the expected number of edges for the infinite plane is $E(N) = 6$. All three cases show a slightly positive skew with a lower peak for the infinite plane case which has a slightly greater spread. For the infinite plane case the mode is at 6, but it is at 5 when a boundary is imposed. These observations are supported by the summary statistics in Table 1, with the infinite plane case having higher mean and greater standard deviation. In contrast, it has lower skewness and lower (excess) kurtosis. Overall, there is little difference between the unit square and convex hull cases.

The number of cell edges is shown as a surface plot and transects for the unit square boundary and for the convex hull cases in Figures 7. Recalling that $E(N) = 6$ over the infinite plane, cells affected by an imposed boundary are likely to have fewer edges with the smallest numbers for cells at the boundary and especially in the corner. There is no place that the cells are having more than 6 edges on average. Diagonal and central slices from the surface plot show similar patterns. Even though there are many cells with more than 6 edges, they do not concentrate at particular locations but are spread over the region uniformly at random. The number of cell edges is slightly larger in the convex hull case than for the unit square for cells close to the boundary, possibly because of having irregular shapes of bounding regions in the convex hull compared to the unit square where the bounding is a straight line. This reduces the number of cell edges intersecting the boundary.

The cell area, perimeter and number of cell edges are seriously affected by the introduction of boundaries. These variations are high around the corner but milder near the sides of the region.

4. Parametric modelling

We now consider fitting parametric distributions, first to the properties of interest and then to the changes to these properties which occur when a finite boundary is imposed.

4.1. Distributions of cell area, perimeter and number of edges

In this section, our aim is to approximate the distributions of the cell properties by parametric distributions for infinite plane, unit square, and convex hull boundaries.

Then, the approximations will be compared across different boundary cases and the previous work. The standardized properties explained in Section 3.1 are used when estimating the parameters of distributions.

Summary statistics are given in Table 2 for the properties of Voronoi cells in three different cases. As in section 3, only the cells affected by the boundary are included in the calculations for the unit square and convex hull cases. Mean cell area $\bar{A}_I^*$ and perimeter $\bar{P}_I^*$ calculated for cells in the infinite plane are very close to the analytically calculated expectations, however, the mean cell area and perimeter for the unit square case are calculated as $\bar{A}_U^* = 1.137$ and $\bar{P}_U^* = 1.097$ and for the convex hull case $\bar{A}_C^* = 0.886$ and $\bar{P}_C^* = 0.951$. On average, cells bounded by a unit square boundary will have larger area and perimeter than the cells in the infinite plane, and convex hull bounded cells will have the smallest area and perimeter.

Our main tool for parametric modelling of PVT properties is the three-parameter gamma distribution due to [24], which has density function

$$f(x|a,b,c) = \frac{ab^{c/a}}{\Gamma(c/a)} x^{(c-1)} e^{-bx^a}, \qquad 0 < x < \infty \qquad a,b,c > 0, \tag{2}$$

where b is the shape parameter, c is the rate parameter, and a is the location parameter. Note that (2) reduces to the probability density function of the standard two-parameter gamma distribution when $a = 1$. In addition to the gamma distribution results presented here, we have considered the Weibull and log-normal distributions but they did not fit the data as well as the gamma distribution. Maximum likelihood estimates of parameters fitting the two and three-parameter gamma distributions to the standardized properties for each boundary scenario are listed in Table 3 in the top and bottom panels respectively. Corresponding density lines for fitted three-parameter gamma distributions are shown in in Figure 8 as solid, dashed and dotted lines for the different boundary cases.

Plotted symbols signify the mid points of the top of the histogram bins of properties for infinite plane (•), unit square (■) , and convex hull (▲) boundary cases as indicated in the figure legend. The excellent fit of the density lines to the observed properties demonstrates that the three-parameter gamma distribution shows a good fit to the statistical properties of bounded and unbounded Poisson Voronoi cells.

The parameter estimates presented in Table 3 verify the disparities of the behaviours of cells in different boundary cases. Parameter estimates for the infinite plane case agree with those reported by Hinde and Miles [21] and Tanemura [23] for the three-parameter gamma distribution, and with Weaire [19], Kumar and Kurtz [22] and Koufos and Dettmann [13] for the two-parameter gamma distribution.

Koufos and Dettmann [13] investigated the the area distribution of bulk *'interior'*, edge and corner cells rather than the mixture of these cells. Their results regarding the bulk cells can be directly compared to the interior cells in Table 3. However, further investigation is necessary to make our current results directly comparable to those in [13] by separating the edge and corner cells. The results are provided in the supplemental material. Our findings for edge and corner cells show substantial differences to those in [13], and this indicates the importance of the cell classification method used.

The separation of cell types is not always straightforward. Koufos and Dettmann [13] generate their locations in a different way to our approach. They initially fix a point at a particular location at or close to the edges or corner of the quadrant, then generate the remaining points and investigate the properties of the initially generated point. Alternatively, one could identify edge cells as those which have a vertex located on the boundary, and corner cells to have a corner point as one of its vertices. However, this is not straightforward in all boundary scenarios such as the convex hull boundary. It is also possible to identify cell types by partitioning the finite plane into sub-regions. Since this study is primarily intended to investigate the imposition of different boundaries, the the mixture of cell types is considered.

Lastly, fitted density lines for cell area and perimeter conditional on the number of cell edges are shown in Figures 9 and 10 for all boundary cases. Since the fitted gamma distributions show little difference for cells having more than 7 edges, observations for $N = 8, \ldots, 15$ are aggregated and denoted as $N > 7$. Infinite plane, unit square, and convex hull cases are signified with (——), (- - - -), and (·····) respectively. It is obvious that cells with different numbers of edges lead to different distributions of cell area and perimeter. The fitted distributions for different boundary cases never coincide as the number of cell edges change. Consequently, the mean of the fitted gamma distributions are different for different values of N. The mean μ is smallest when $N = 3$ and it increases for larger N. We note that the cell perimeter distributions are more similar for the two bounded cases than for the infinite plane case, while the cell area distributions do not show this pattern.

4.2. Distributions of changes to cell area and perimeter

We define the area reduction and ratio of an unbounded cell when it is bounded to be

$$\begin{aligned} \text{area reduction} &= (\text{area with no boundary - area with boundary}) \\ \text{area ratio} &= (\text{area with no boundary / area with boundary}), \end{aligned}$$

with equivalent definitions for cell perimeter and number of edges. Histograms and surface plots showing the reductions in and ratios of area, perimeter and number of edges when the unit square boundary is imposed on the unbounded cells are shown in Figure 11. Cells which are not affected by the boundaries are not included here and, due to a small number of extreme values, the histograms are clipped from both tails to improve the visualization. Histograms, especially of area reduction and ratio, show a high peak where the change is very small. Surface plots indicate that cells very close to the unit square boundary had shrinkage in their area, perimeter and number of edges; the cells close to the corner shrank the most. Looking at the right panel, highest ratios in the properties are observed at the corners of the surface where the cells shrank the most.

Reductions and ratios for convex hull bounded cells are visualized in Figure 12, showing significant shrinkage in the area, perimeter and number of edges near the boundary and smallest at the corner. Not all of the convex hull bounded cells shrunk. There is an exception for some cells that had expansion in their area and perimeter. These cells are located parallel but not very close to the edges and the corner where slightly darker colour is seen. The extreme reduction in area and perimeter of convex hull bounded cells causes high area and perimeter ratios.

5. Influence of different point intensities

In this section, Poisson Voronoi tessellation for different intensities is explored. All simulation results discussed so far are based on PVT with intensity $\rho = 200$, however, it is important to discover whether the patterns of cells vary as ρ changes. Hence, we conducted further simulations with $\rho = \{30, 50, 100, 300\}$ for $r = 10^6$ realizations and compared the results to those obtained with $\rho = 200$. We compare the standardized area, standardized perimeter, and number of edges.

Line plots are shown in Figure 13 following a transect across the centre of the data. Different ρ values are indicated by grey scale with darker lines representing larger ρ.

The occurrence of different patterns across the region is clear and matches the patterns seen earlier, with the variation from the expected values for the infinite plane case greatest on the cells closer to the boundary of the region. These patterns are more pronounced for smaller values of ρ. There is little deviation from the infinite plane case close to the centre of the region, corresponding to cells that are not affected by the boundary, except when $\rho = 30$. The initial intuition for $\rho = 30$ is that the number of the points generated for such low intensity are small enough that a randomly sampled cell could still be affected by the boundary even though it is located relatively close to $[0.5, 0.5]$.

To investigate the effect of having fewer points on the patterns observed in the cell properties, we conducted another simulation to estimate the probability of a randomly sampled cell being affected by the boundary when the number of points is $n \in \{10, 15, 20, 30, 50, 100, 200, 300\}$. In this simulation, the number of points n is fixed in each realization instead of being Poisson distributed to avoid the variability on the number of points generated for each specified intensity especially when the ρ is very small. For each n, the simulation was repeated for $r = 10^5$ realizations.

The proportion of simulated cells being affected by the boundary is plotted against the number of points in Figure 14. When $n = 10$, almost every randomly sampled cell is affected by the boundary and the proportion decreases as the number of points increases. The proportion of cells affected by the convex hull is slightly higher than the unit square case but the proportions converge at higher n. The different pattern observed in Figure 13 for $\rho = 30$ is also highlighted here for lower intensities in general.

6. Conclusion

We have investigated the changes in the properties of Poisson Voronoi cells in two dimensional space as boundaries are imposed on data observed in a finite region. Using the unit square and the convex hull of points as boundaries, cell area, perimeter, and the number of cell edges can be estimated using the three-parameter gamma distribution. The difference in the fitted distributions raises the importance of taking into account these boundary effects in the analysis of spatial data which usually comes with its own boundary case.

The classification of cells as interior, edge and corner as in [13] is not completely obvious for the unit square boundary which has simple properties, and is even more complicated for irregular boundary cases. Although we have briefly discussed the challenges of classification of cell types, we leave detailed consideration of the properties

of different cell types as potential future work.

Cells affected by the two different boundaries we have considered do not have identical properties. Hence, in practice, further study considering other specifically determined boundaries or real boundaries such as state borders or coastlines may need to be performed for any given data set.

Results from different intensities lead to different properties of Voronoi cells, hence using intensities representative of any given data set would be necessary. We have only considered the homogeneous Poisson point process, and further work may investigate how our results would change for other point patterns. For example, Schoenberg [3] studied the Voronoi tessellation of the locations of earthquakes in Southern California and found that the Pareto distribution to fit the cell area and perimeter well.

Acknowledgements

The authors gratefully acknowledge the helpful comments from reviewers which have led to improvements in this paper. The first author is grateful for the financial support from the Republic of Turkey, Ministry of National Education.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Kingman JFC. Poisson Processes. Oxford (UK): Clarendon Press; 1993.
[2] Ripley BD. Spatial statistics. Hoboken (NJ): John Wiley & Sons; 2005.
[3] Schoenberg FP, Barr C, Seo J. The distribution of Voronoi cells generated by Southern California earthquake epicenters. Environmetrics. 2009; 20:159–171.
[4] Icke V, Weygaert R. Fragmenting the universe. Astronomy and Astrophysics. 1987; 184:16–32.
[5] Yoshioka S, Ikeuchi S. The large-scale structure of the universe and the division of space. The Astrophysical Journal. 1989; 341:16–25.
[6] Ramella M, Boschin W, Fadda D, et al. Finding galaxy clusters using Voronoi tessellations. Astronomy & Astrophysics. 2001; 368:776–786.
[7] Tanemura M, Hasegawa M. Geometrical models of territory I. Models for synchronous and asynchronous settlement of territories. Journal of Theoretical Biology. 1980 Feb 7;82(3):477–96.
[8] Fischer RA, Miles RE. The role of spatial pattern in the competition between crop plants and weeds. A theoretical analysis. Mathematical Biosciences. 1973 Dec 1;18(3-4):335–50.
[9] Mackay AL. Stereological characteristics of atomic arrangements in crystals. Journal of Microscopy. 1972 Apr;95(2):217–27.
[10] Finney JL. Random packings and the structure of simple liquids. I. The geometry of random close packing. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences. 1970 Nov 10;319(1539):479–93.
[11] Luchnikov VA, Medvedev NN, Naberukhin YI, et al. Voronoi-Delaunay analysis of normal modes in a simple model glass. Physical Review B. 2000 Aug 1;62(5):3181.
[12] Baccelli F, Błaszczyszyn B. On a coverage process ranging from the Boolean model to

the Poisson-Voronoi tessellation with applications to wireless communications. Advances in Applied Probability. 2001 Jun 1;293–323.
[13] Koufos K, Dettmann CP. Distribution of cell area in bounded Poisson Voronoi tessellations with application to secure local connectivity. Journal of Statistical Physics. 2019 Sep 1;176(5):1296–315.
[14] Xu T, Li M. Topological and statistical properties of a constrained Voronoi tessellation. Philosophical Magazine. 2009 Feb 1;89(4):349–74.
[15] Meijering JL. Interface area, edge length, and number of vertices in crystal aggregates with random nucleation. Philips Res. Rep. 1953;8:270–90.
[16] Gilbert EN. Random subdivisions of space into crystals. The Annals of Mathematical Statistics. 1962 Sep 1;33(3):958-72.
[17] Møller J. Lectures on random Voronoi tessellations. New York (NY): Springer Science & Business Media; 2012 Dec 6.
[18] Kiang T. Random fragmentation in two and three dimensions. Zeitschrift fur Astrophysik. 1966;64:433.
[19] Weaire D, Kermode JP, Wejchert J. On the distribution of cell areas in a Voronoi network. Philosophical Magazine B. 1986 May 1;53(5):L101–5.
[20] Crain IK. The Monte-Carlo generation of random polygons. Computers & Geosciences. 1978 Jan 1;4(2):131–41.
[21] Hinde AL, Miles RE. Monte Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process. Journal of Statistical Computation and Simulation. 1980 Apr 1;10(3–4):205–23.
[22] Kumar S, Kurtz SK. Properties of a two-dimensional Poisson-Voronoi tesselation: a Monte-Carlo study. Materials Characterization. 1993 Jul 1;31(1):55–68.
[23] Tanemura M. Statistical distributions of Poisson Voronoi cells in two and three dimensions. Forma. 2003 Nov;18(4):221–47.
[24] Stacy EW. A generalization of the gamma distribution. The Annals of Mathematical Statistics. 1962;33(3):1187–92.

Table 1. Mean, standard deviation, skewness and kurtosis of area, perimeter and number of edges of Poisson Voronoi cells in infinite plane, and for unit square and convex hull boundaries. In bounded cases, only cells affected by the boundaries are included.

	Case	Mean	SD	Skewness	Kurtosis
Area A	Infinite plane	0.0050	0.0027	1.023	1.525
	Unit square	0.0057	0.0033	1.164	2.247
	Convex hull	0.0045	0.0030	1.171	1.900
Perim. P	Infinite plane	0.2835	0.0689	0.190	-0.025
	Unit square	0.3104	0.0832	0.240	0.106
	Convex hull	0.2703	0.0870	0.130	-0.150
Edges N	Infinite plane	5.996	1.334	0.432	0.204
	Unit square	5.364	1.203	0.497	0.278
	Convex hull	5.418	1.185	0.516	0.331

Table 2. Mean, standard deviation, skewness and kurtosis of standardized area, perimeter and number of edges of Poisson Voronoi cells in infinite plane, and for unit square and convex hull boundaries.

	Case	Mean	SD	Skewness	Kurtosis
Area A	Infinite plane	1.004	0.531	1.022	1.525
	Unit square	1.137	0.654	1.164	2.247
	Convex hull	0.886	0.599	1.186	1.957
Perim. P	Infinite plane	1.002	0.244	0.190	-0.025
	Unit square	1.097	0.294	0.240	0.106
	Convex hull	0.951	0.307	0.146	-0.141
Edges N	Infinite plane	6.000	1.334	0.432	0.204
	Unit square	5.364	1.203	0.497	0.278
	Convex hull	5.432	1.185	0.509	0.321

Table 3. Parameter estimates with 95% confidence intervals for the two- and three-parameter gamma distributions fitted to the area, perimeter and number of edges of Poisson Voronoi cells in infinite plane, and with unit square and convex hull boundaries. Top and bottom panels show parameter estimates of two and three-parameter gamma distributions respectively.

	Case	$\hat{a}$	$\hat{b}$	$\hat{c}$
Area	Infinite plane	-	3.510 (3.500-3.520)	3.526 (3.516-3.535)
	Unit square	-	2.626 (2.612-2.640)	2.986 (2.971-3.000)
	Convex hull	-	2.271 (2.259-2.284)	2.012 (2.003-2.022)
	Kiang [18]	-	4.0	4.0
	Weaire [19]	-	3.63	3.63
	Kumar & Kurtz [22]	-	3.7176	3.7174
	Koufos & Dettmann [13]	-	3.5691	3.5691
Perim.	Infinite plane	-	16.038 (15.993-16.082)	16.074 (16.030-16.118)
	Unit square	-	11.903 (11.840-11.965)	13.062 (12.995-13.129)
	Convex hull	-	8.892 (8.846-8.938)	8.464 (8.422-8.507)
Edges	Infinite plane	-	3.370 (3.361-3.380)	20.220 (20.164-21.276)
	Unit square	-	3.740 (3.721-3.760)	20.064 (19.960-20.167)
	Convex hull	-	3.911 (3.891-3.931)	21.245 (21.137-21.353)
Area	Infinite plane	1.080 (1.073-1.106)	3.015 (2.961-3.069)	3.311 (3.285-3.336)
	Unit square	1.089 (0.969-0.987)	2.195 (2.123-2.267)	2.792 (2.755-2.828)
	Convex hull	1.336 (1.320-1.352)	1.336 (1.304-1.367)	1.667 (1.652-1.681)
	Tanemura [23]	1.0795	3.0322	3.3112
	Hinde & Miles [21]	1.0787	3.0328	3.3095
	Koufos & Dettmann [13]	1.097	2.919	3.253
Perim.	Infinite plane	2.334 (2.313-2.356)	2.963 (2.906-3.019)	7.593 (7.533-7.654)
	Unit square	2.278 (2.242-2.314)	2.062 (1.987-2.137)	6.412 (6.326-6.498)
	Convex hull	2.917 (2.882-2.953)	1.156 (1.126-1.186)	3.779 (3.746-3.812)
	Tanemura [23]	2.33609	2.97006	7.58060
	Hinde & Miles [21]	2.3389	2.9563	7.5579
Edges	Infinite plane	0.931 (0.920-0.941)	4.406(4.222-4.590)	21.706 (21.469-21.943)
	Unit square	0.775 (0.768-0.783)	9.088 (8.807-9.368)	25.804 (25.543-26.064)
	Convex hull	0.813 (0.805-0.820)	8.089 (7.843-8.335)	25.928 (25.675-26.181)
	Tanemura [23]	0.96853	3.80078	20.86016
	Hinde & Miles [21]	1.0186	3.130	19.784

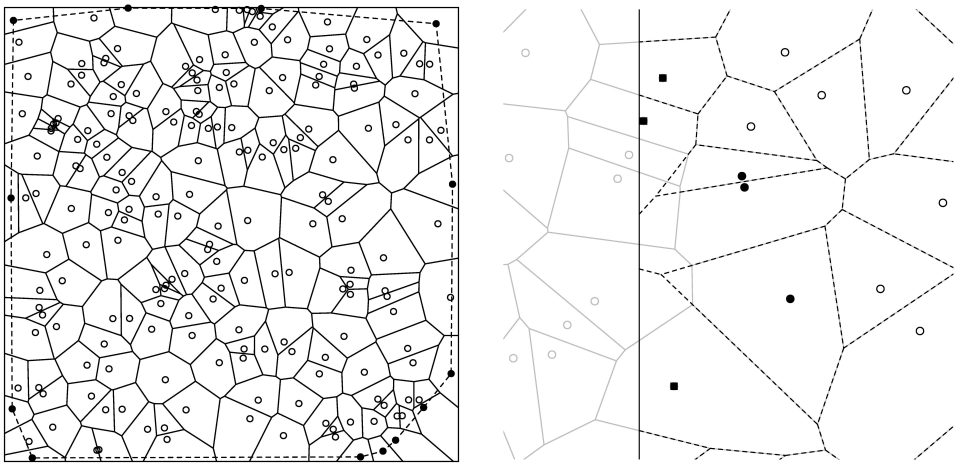

Figure 1.

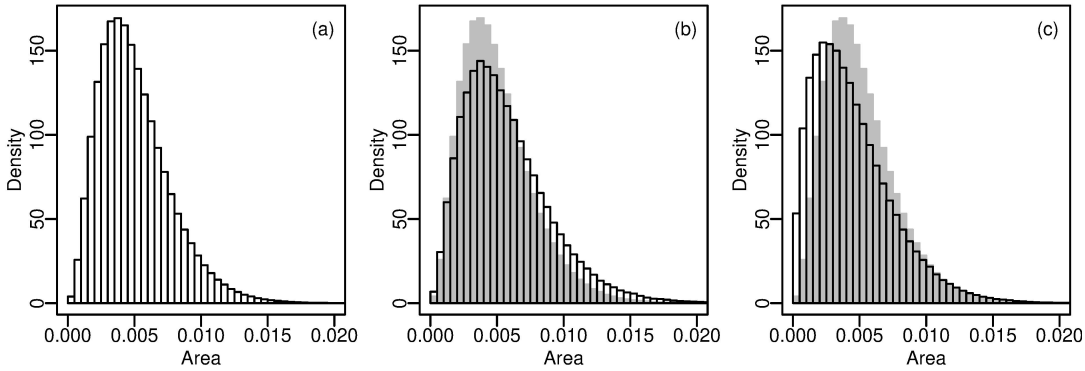

Figure 2.

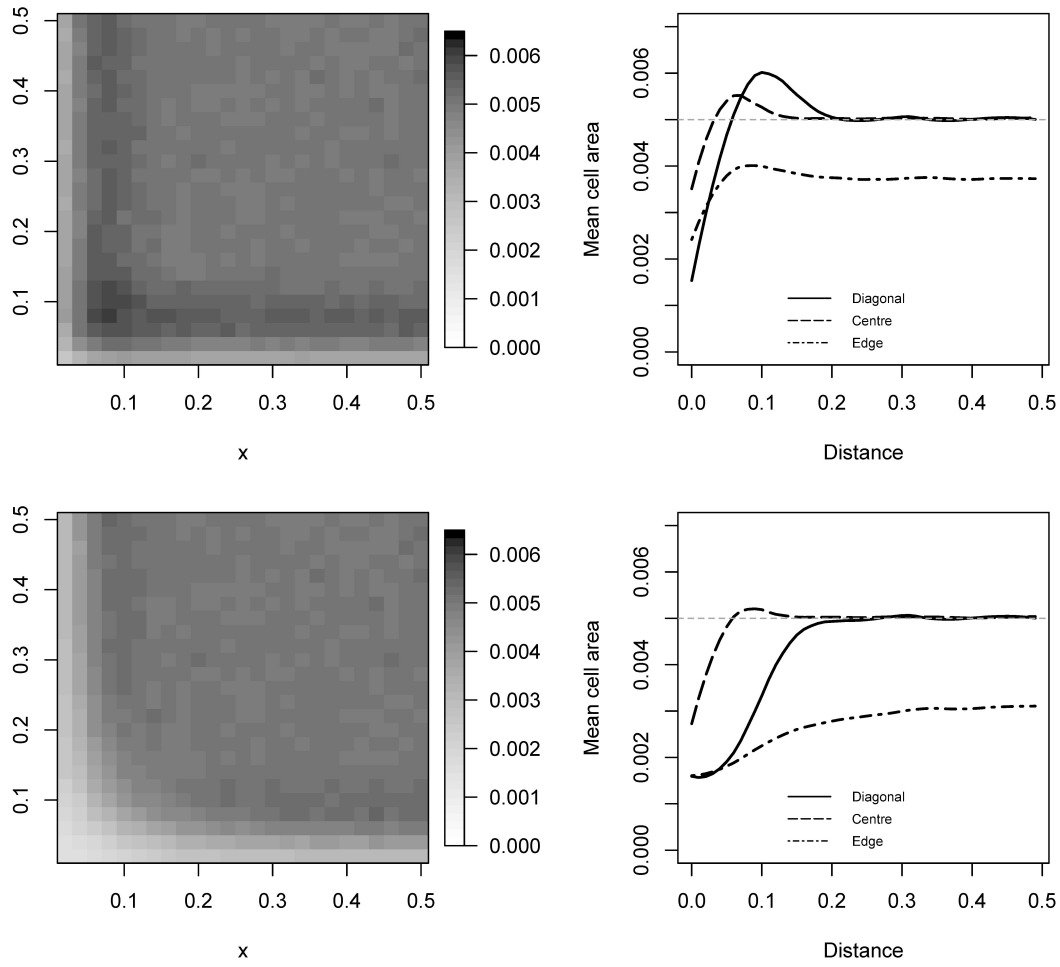

Figure 3.

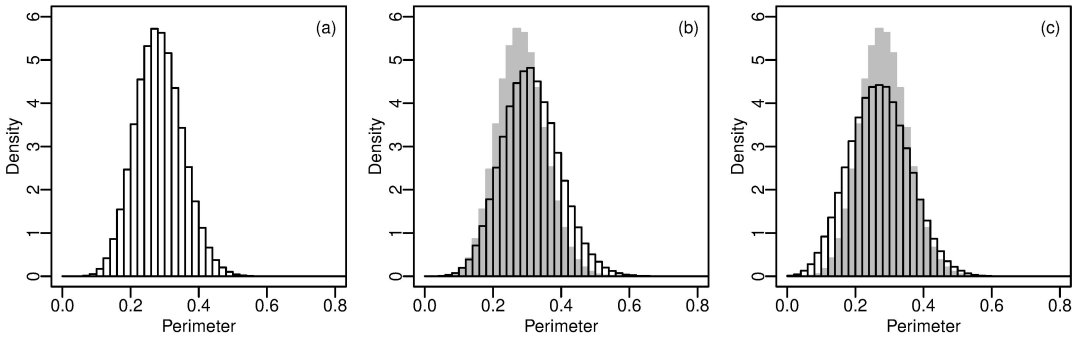

Figure 4.

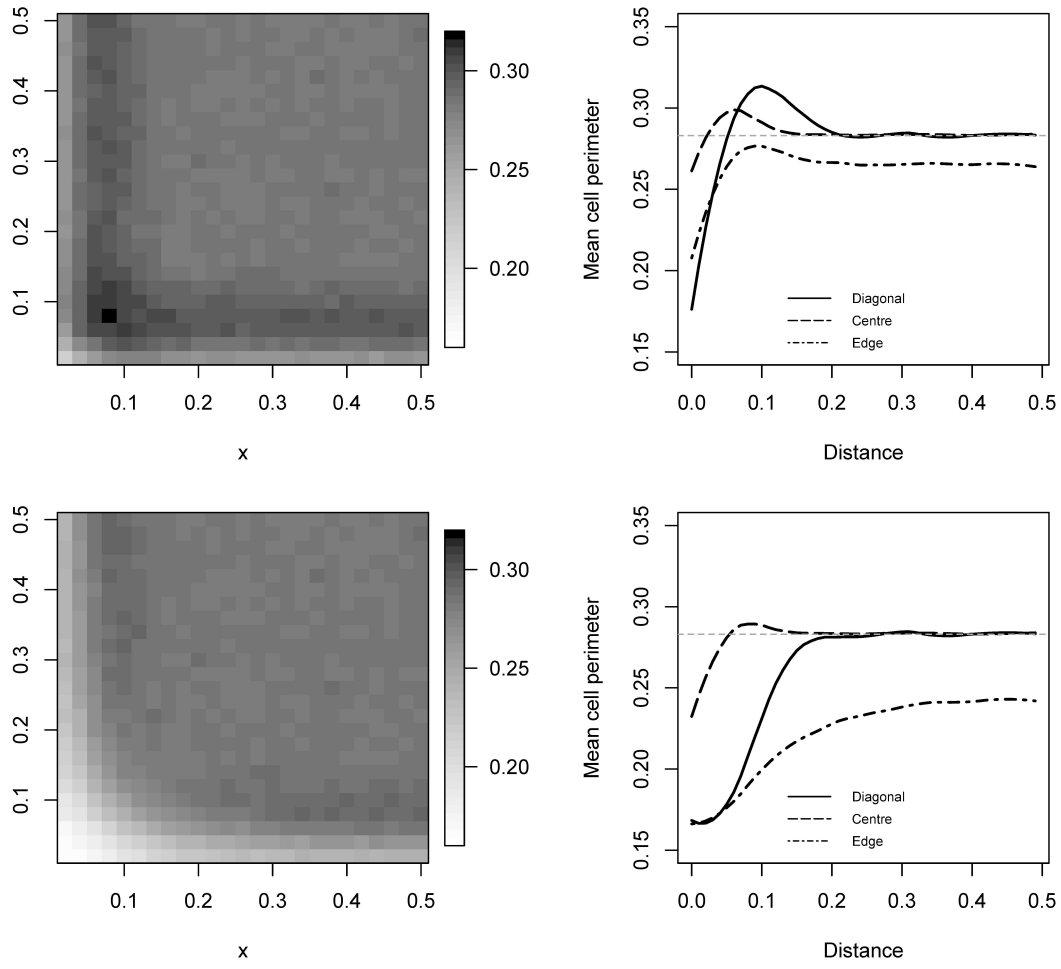

Figure 5.

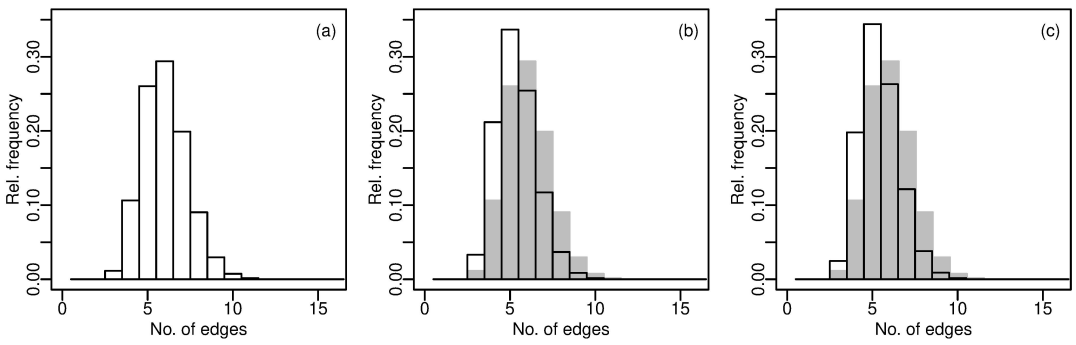

Figure 6.

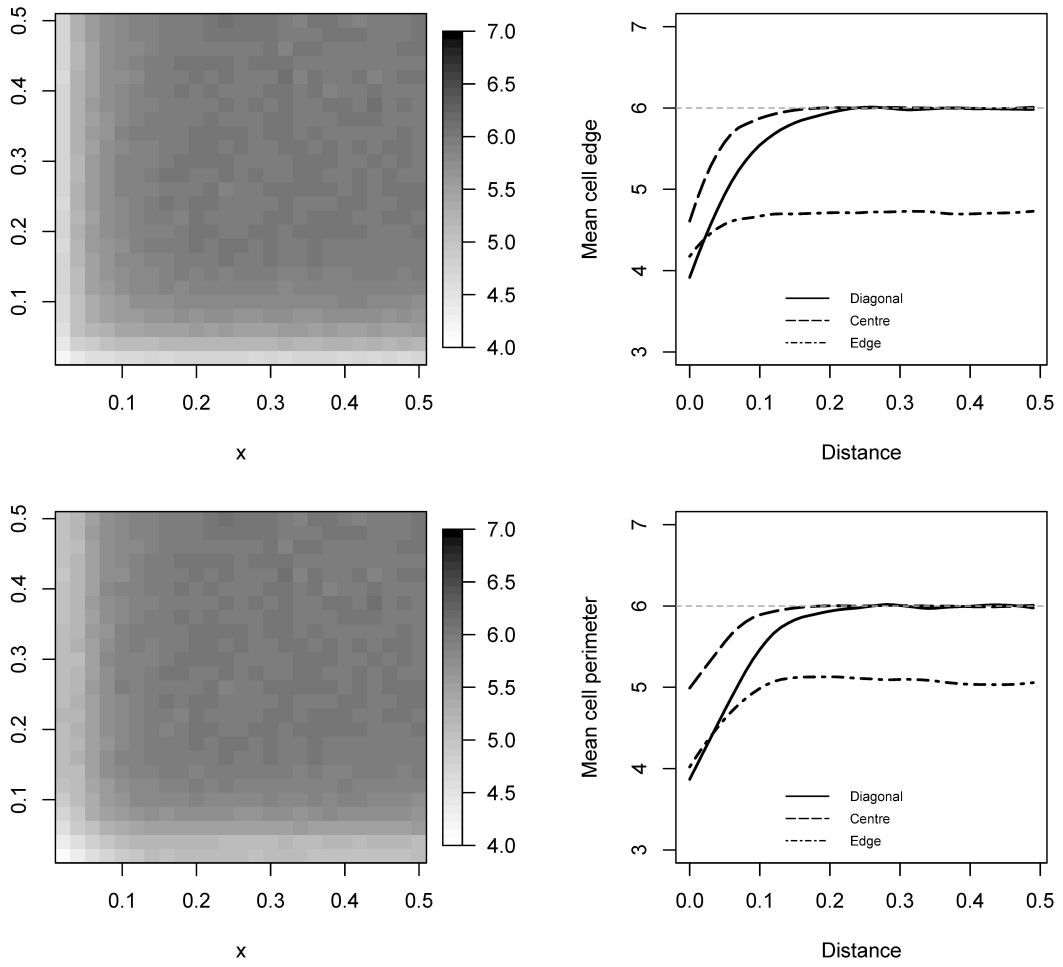

Figure 7.

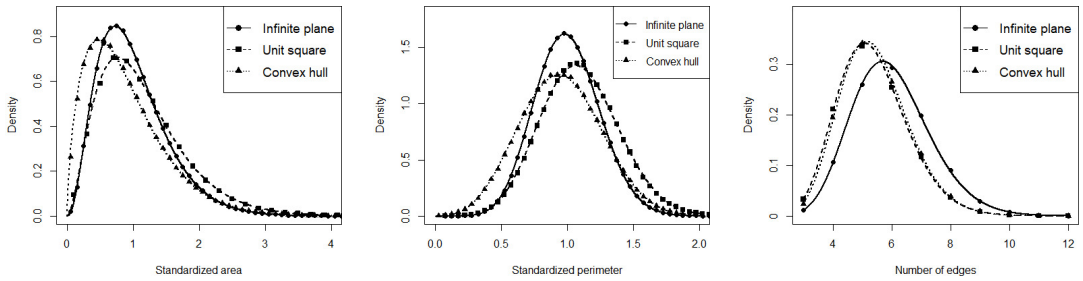

Figure 8.

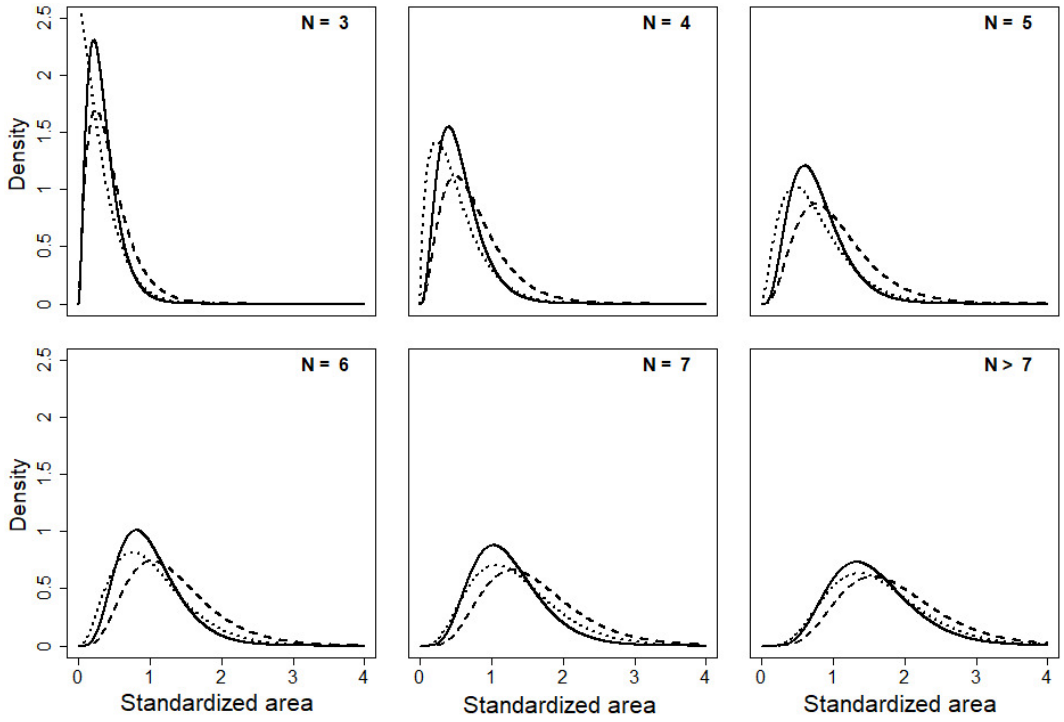

Figure 9.

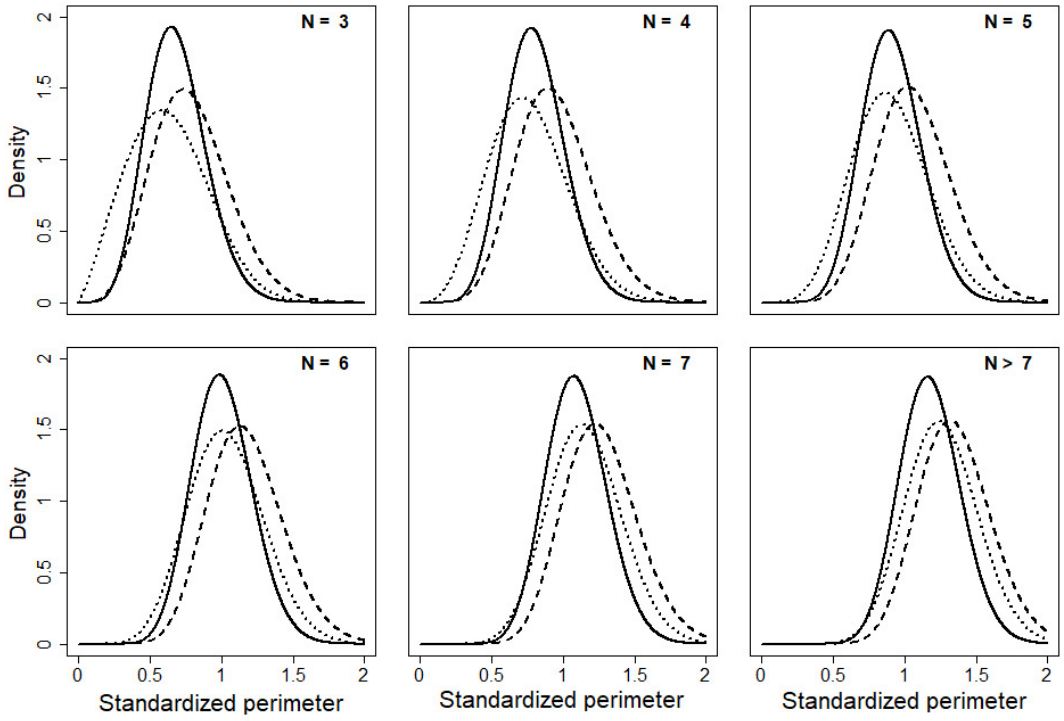

Figure 10.

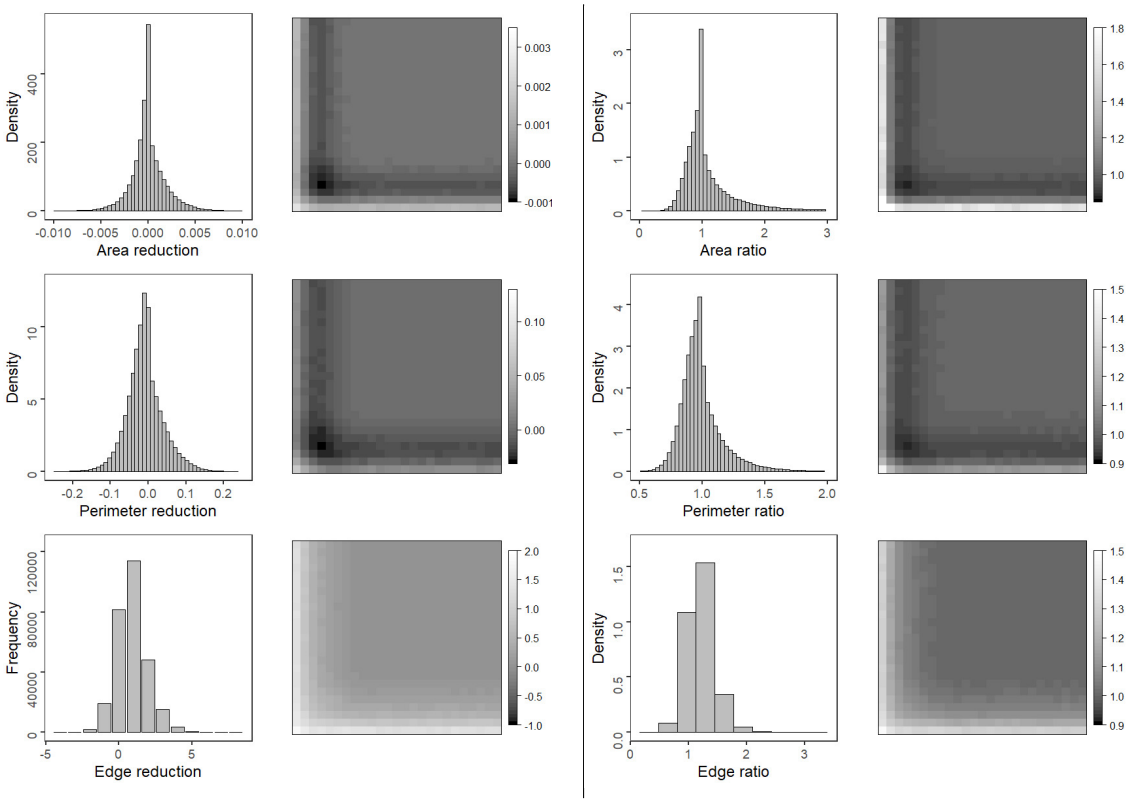

Figure 11.

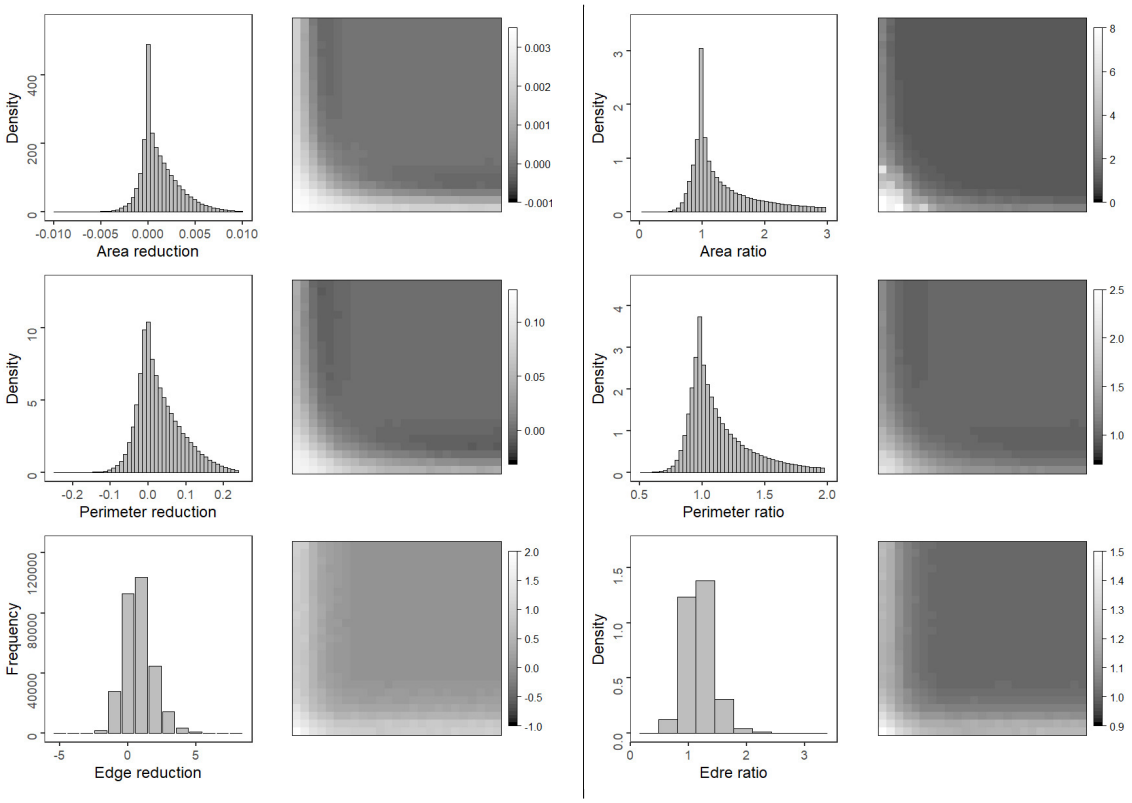

Figure 12.

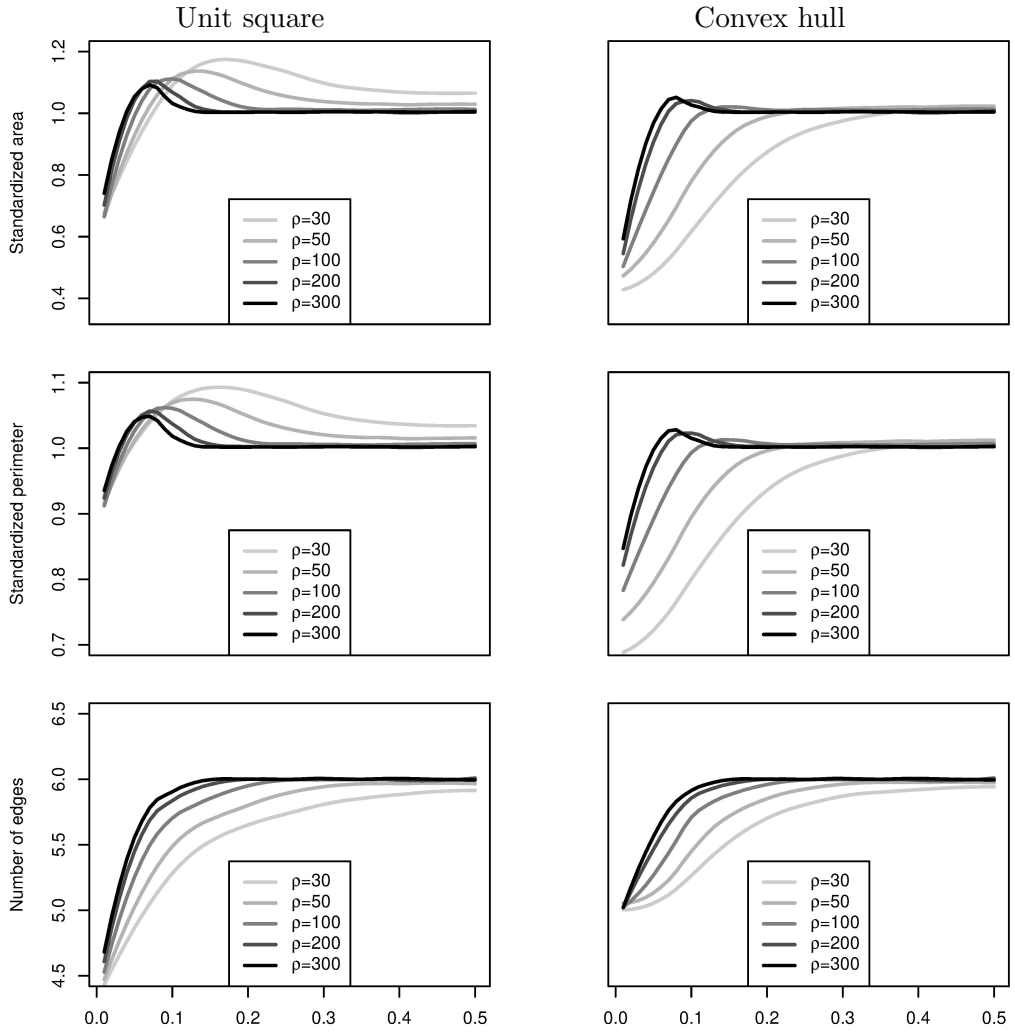

Figure 13.

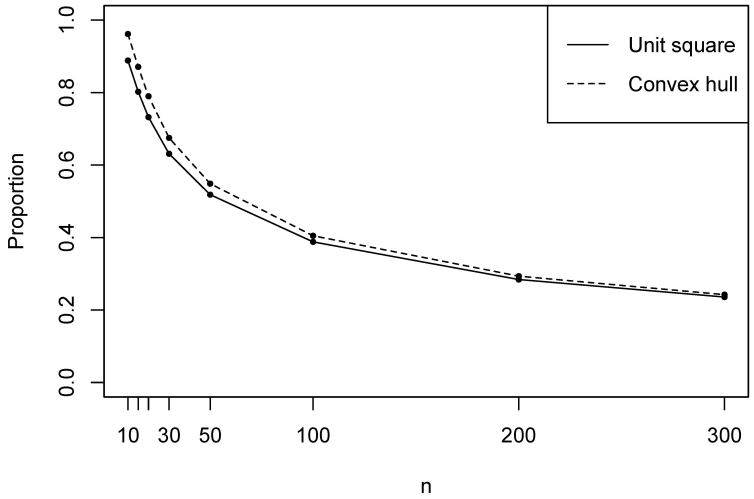

Figure 14.

(1) Left: Voronoi tessellation of homogeneous Poisson points with intensity $\rho = 200$ bounded by the unit square (black solid square) and the convex hull of points (dashed polygon). Right: Zoomed in version of a Voronoi tessellation showing the effects of imposing a boundary. Gray lines are the original tessellation lines before any boundary is used. Dashed lines are the tessellation lines after the vertical solid line is imposed as a boundary to the points on the right side. Cells with solid black square points had an intersection with the boundary, and cells with black solid circle points did not have an intersection with the boundary but have had a transformation in their shapes. Gray and black circles indicate the points of remaining cells.

(2) Empirical distribution of cell area in (a) infinite plane, (b) unit square and (c) convex hull.

(3) Surface plot of cell area in bounded region (left). Averaged cell area over the grids against the distance of the direction being followed from the origin (right). Results shown for unit square (top row) and convex hull (bottom row). The gray horizontal dashed line shows the expected cell area.

(4) Empirical distribution of cell perimeter in (a) infinite plane, (b) unit square and (c) convex hull.

(5) Surface plot of cell perimeter in bounded region (left). Averaged cell perimeter over the grids against the distance of the direction being followed from the origin (right). Results shown for unit square (top row) and convex hull (bottom row). The gray horizontal dashed line shows the expected cell perimeter.

(6) Empirical distribution of number of cell edges in (a) infinite plane, (b) unit square and (c) convex hull.

(7) Surface plot of number of cell edges in bounded region (left). Averaged number of cell edges over the grids against the distance of the direction being followed from the origin (right). Results shown for unit square (top row) and convex hull (bottom row). The gray horizontal dashed line shows the expected number of cell edges.

(8) Density functions of three-parameter gamma distributions using the parameter estimates from Table 3 for standardized cell area (left), standardized perimeter (centre), and number of edges (right). Points show the mid points of the histogram bins of simulated cell properties and different line types show infinite plane, unit square and convex hull cases.

(9) Density lines of three-parameter gamma distribution for the estimated parameters of standardized area for infinite plane (——), unit square (- - - -), and convex hull (·····) cases conditioned on the number of cell edges $N = 3, 4, ..., 7$ and $N > 7$.

(10) Density lines of three-parameter gamma distribution for the estimated parameters of standardized perimeter for infinite plane (——), unit square (- - - -), and convex hull (·····) cases conditioned on the number of cell edges $N = 3, 4, ..., 7$ and $N > 7$.

(11) Histograms and surface plots of reduction (left panel) and ratio (right panel) of the cell properties when unbounded cells are bounded with a unit square.

(12) Histograms and surface plots of reduction (left panel) and ratio (right panel) of the cell properties when unbounded cells are bounded with convex hull of points.

(13) Standardized PVT properties for intensities $\rho = 30, 50, 100, 200, 300$ for points across the centre of the region. Each ρ value is indicated by a different shade of grey showing the pattern of different intensities over the surface for standardized properties. Left panel show the results for unit square and right panel is for the convex hull bounded cells.

(14) Proportion of boundary-affected cells in 10^6 realizations for varying numbers of points $n \in \{10, 15, 20, 30, 50, 100, 200, 300\}$.