Dotting the Difference – Strategies in finding the best suitable Dot Value

Annette Hey*

* University of Rostock

Abstract. Dot mapping is a cartographic representation method to visualize discrete absolute values. It is used to reveal spatial distribution patterns. Each dot in the map represents a defined value (dot value). All dots in an area add up to the depicted value. A typical application of dot mapping is a population map. Because the manual design of a dot map is very elaborate the automation of dot mapping is covered by many research projects. For example in Hey (2012) a method, which allows the automatic design of a dot map based on some predefined parameters, is presented. This method forms the basis of the tests described in this paper.

This article deals with the question: Is it possible to infer from certain characteristics of the data that there might be problems in finding a suitable dot value? The focus will be set on the heterogeneity of the data that shall be mapped. A population data set of North Rhine-Westphalia is used as example data. North Rhine-Westphalia is a German state with large city regions and rural areas as well. The tests shall find an answer to the question: Is it possible to draw conclusions concerning the reachable dot value from the relation between the data and the area used for placing the dots in the map?

Keywords: Dot mapping, Automation

1. Introduction

Designing a dot map needs the definition of dot size and dot value. These depend on the chosen scale of the map and the area in the map that is available for placing the dots. Dot placement according to the algorithm described in Hey (2012) is based upon a system of logarithmic spirals and concentric circles around the reference point of the data. Dots are placed in a centered manner around these reference points forming so-called ‘dot clusters’. The position of each dot is shifted by a pseudo-random offset to conceal the regular structure of the dot cluster. Dots are not allowed to be tangent to each other. The basic method offers the possibility to integrate distribution areas and exclusion areas. Distribution areas are those areas where the depicted objects may occur (e.g. settlement areas for population data). In contrast, exclusion areas define the space, where no dots are placed (e.g. large water bodies, place for map labels). For a visualization as accurate as possible the dot value has to be as small as possible. One way to achieve this is to be able to assign the distribution areas bijectively to the reference points. Then each dot cluster can cover the complete distribution area without being in danger of mixing up with adjacent dot clusters. If this condition is not met a circle around the reference point is used to place the dots. The radius of this circle is defined by the restriction, that the circle around the nearest neighboring reference point is justly touched. In the tests described here a dataset with bijective assignment is used.

2. Previous works on automated Dot Mapping

The time when computer science and automation methods found their way into cartography (approximately in the 1970s) solutions for different cartographic representation methods have been sought. Dot mapping often was subject to these research projects (see e.g. Hofmann 1972, Klamt 1972, Aschenbrenner 1989, Lavin 1986 und Kimerling 2009), which were focusing on different aspects. While Aschenbrenner concentrates on a strict regular dot placement, Kimerling works with randomly placed dots. Also there has been done research in the field of working with existing dot maps. For example de Berg (2004) deals with the question of how generalized dot maps can be derived from existing dot maps with random dot placement. In Hey (2012) a method is presented that allows the design of a dot map even without having the knowledge of an expert cartographer. This enables cartographic laymen to design good maps (in a cartographic sense), while the design process for expert cartographers is simplified and sped up. Because this method is based on visualization parameters and the analysis of data characteristics it is used to find out if and how specific data properties affect the reachable dot value and with this the quantitative precision of the dot map.

3. Objective

Dot mapping is used to depict spatial distribution patterns of discrete objects. If these distributions are relatively homogeneous there will be no problems in finding dot value and dot size suitable for the data and the planned map scale. But if the objects are distributed very unsteady there will be difficulties, because areas with a low object density are strongly neglected compared to areas of high object density. The dot value (and with that the quantitative accuracy of the map) is always oriented towards areas with an unfavorable relation between data value and distribution area. This affects areas with a very high data value as well as data values with a very small distribution area and areas with an unfavorable combination of these two factors. The dot value is determined in a way that allows enough dots within the distribution area to represent the data value. This often results in quite high dot values, which lead to only a few dots in areas of low density, even if there is space for a lot more. The higher the dot value is the larger the quantitative error of the map will be. This error cannot be prevented due to the fact that every dot cluster represents a multiple of the dot value and therefore the real data value only appears in a generalized manner. This rounding is inevitable when using simple dot mapping (only one dot value). Because of this, sometimes a combination of dot mapping and graduated symbols (graded or gliding symbol scale) is used. Another possibility is the 'Kleingeldmethode', where the data value is represented by a sum of a preferably little number of value unit symbols representing different values. It can be compared to paying with as little a number of coins as possible. This method is a very complex version of dot mapping. Dot mapping and the 'Kleingeldmethode' are similar in depicting the data as a sum of map symbols representing certain data values. Anyway, they do differ in the way of depicting the spatial distribution of the data. While dot mapping can give a very detailed impression the 'Kleingeldmethode' works with schematically placed map symbols. An easy method to address the problem of very heterogeneous requirements for a suitable dot value is an inset map for the area of high density with a larger scale and the same dot value and dot size like the overview map, to allow comparisons.

For all these methods it is necessary to identify problematic areas, to be able to address the problems properly. The objective of this article is to find and evaluate criteria that can be used as problem indicators.

4. Basics

In the tests population data was used. It was combined with Corine Landcover Data[1] to allocate the population data of the administrative settlements to the actual settlement areas. The whole method is described in Hey (2006). The test area is North Rhine-Westphalia, because there are areas with a very high population density as well as rural areas. To define the

[1] Corine Landcover is an European project which provides landcover data derived from satellite images. For more information see www.corine.dfd.dlr.de.

scope of this test it shall be pointed out, that the purpose is the representation of population figures and their spatial distribution. Population density is not explicitly addressed, even though it can be developed indirectly.

As dot size a radius of 0.8 mm is chosen. This value is taken from the range of dot sizes used in dot maps proposed by Koch et al. (2002). The dots are not allowed to touch each other. There is always a gap of at least 0.2 mm between adjacent dots. This contributes to a better legibility.

In the tests the relation between data value and the area in the map (distribution area) was probed. Especially, the shape deviation of these areas compared to a coextensive circle has been considered. The shape deviation F is measured by using a perimeter-area-ratio (1).

$$(1)\ F = \frac{U^2}{A}$$

For a circle this ratio is 4π. More complex forms always have a ratio higher than that. A circle does not only have a small perimeter-area-ratio. It also offers the greatest chance to reach a small dot value, because the dot placement in the method described in Hey (2012) is circular around the reference point.

For the following considerations the formulas for calculating perimeter (2) and area (3) of a circle are used for the coextensive circle.

$$(2)\ U_0 = 2\pi r \qquad (3)\ A_0 = \pi r^2$$

Another indicator of the shape of areas is the contour index K (4). It compares the perimeter of the area U to the perimeter of the coextensive circle U_0.

$$(4)\ K = \frac{U}{U_0}$$

With an increasing radius of the circle the perimeter-area-ratio is changed (5).

$$(5)\ \frac{U}{A} = \frac{2\pi r}{\pi r^2} = \frac{2}{r} \qquad \frac{\tilde{U}}{\tilde{A}} = \frac{2\pi(r+x)}{\pi(r+x)^2} = \frac{2}{r+x} \qquad \rightarrow \qquad \frac{U}{A} > \frac{\tilde{U}}{\tilde{A}}$$

The larger the circle is the smaller the perimeter-area-ratio will be. Thus, different area sizes and different area shapes will result in changes of the perimeter-area-ratio. The impact of different shapes will be considered more closely in the tests.

Because distribution areas, which are used to place the dots, normally are not circular (e.g. settlement areas), the deviation of these areas from the circle shape shall be highlighted briefly. With a constant area the perimeter-area-ratio of a complex shape is always higher than for the circle, because the perimeter grows with an increasing complexity of the shape. Especially when buffer zones towards the areas' edges shall be kept clear to allow a strict distinction of adjacent dot clusters, the complexity of the area shape has a high influence on the obtainable dot value. Within the example data set the inner buffering lead to an area loss of up to 36 %. Because in reality distribution areas are not area-wide and due to the fact that the dots where placed around the reference points within the distribution areas no inner buffering were applied, thus ensuring a greater chance of reaching a smaller dot value (high quantitative accuracy).

Expectations in advance of the test:

- With the help of value-area-ratios it is possible to infer that there might be problems in finding a dot value suitable for all areas.
- The perimeter-area-ratio, the shape deviation and the contour index permit statements on the reachable dot value. Within a circle more dots can be placed than within any other coextensive area of an arbitrary shape.
- The combination of data-area-ratio, perimeter-area-ratio, shape deviation and contour index offers the possibility to state if there might be problems in finding a suitable dot value.

5. Test procedure

5.1. Influence of the Value-Area-Ratio

With the help of the example data set the influence of the value-area-ratio onto the reachable dot value was examined. As a start the smallest and biggest ratios were determined. In the present case the ratio is fluctuating between 0,000714 and 0,007243. That means the relation from smallest to biggest ratio is approximately 1:10. If divided into classes each class of this range will allow another dot value. The purpose is to find the class limits which will cause a change of the dot value to be able to analyze the data in advance concerning the obtainable accuracy (preferably small dot value). Possibly a recommendation can be given if using a combined representation method (e.g. dots and graduated symbols) will be useful because the range of the value-area-ratio is bigger than a certain number.

To find class limits, where the dot value rises, the data set was split into 10 equal interval classes concerning the value-area-ratio. The achieved dot values are listed in *Table 1*.

Class	Data-area-ratio Class limits	Obtained dot value
1	0,000714 < 0,001367	25
2	0,001367 < 0,002020	25
3	0,002020 < 0,002673	50
4	0,002673 < 0,003326	50
5	0,003326 < 0,003979	50
6	0,003979 < 0,004632	100
7	0,004632 < 0,005285	100
8	0,005285 < 0,005938	100
9	0,005938 < 0,006590	250
10	0,006590 ≤ 0,007243	50

Table 1. Obtained dot values depending on the value-area-ratio.

Like originally presumed the dot value at first rises when the value-area-ratio grows. But in the last class a sudden and strong decrease of the dot value can be observed. For areas with a rather unfavorable ratio of data value towards area in the map again a dot value of 50 is possible. The settlements concerned by this are neither extremely large nor extremely small. They do not possess especially high or especially little numbers of inhabitants, too. They are settlements which can be found in the middle of all that. The explanation for the observed effect has to be looked for exactly in this fact. The ratio of data value towards area may be found in the extremes (very large/small area with peculiar high/little numbers of inhabitants). But this ratio may also be somewhere in between. If the ratio is not among the extremes it gives almost no information about the reachable dot value, as seen in the test. *Figure 1* shows the link between the value-area-ratio and the obtainable dot value. Although the average ratio shows a clear trend the maximum and minimum values make clear that there is no unique link between the ratio and the obtainable dot value.

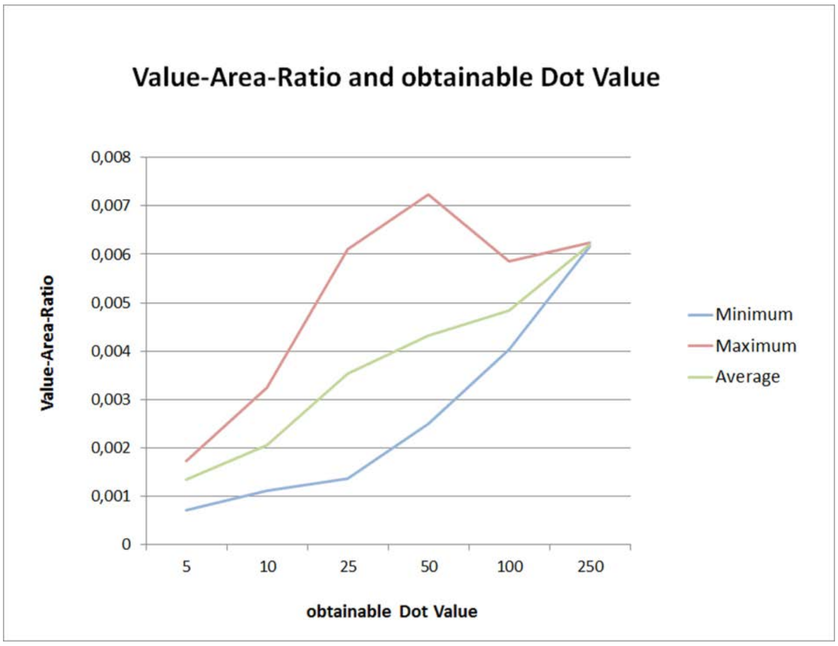

Figure 1. Relation between value-area-ratio and obtainable dot value.

According to that the value-area-ratio cannot be used as indicator for problems in finding a suitable dot value for heterogeneous data. Because of that the shape of the area, which is used to place the dots in the map has been considered. The more the shape of an area resembles a circle the more likely is a smaller dot value (considering coextensive areas). This bases on the placement method used in Hey (2012). The influence of the shape shall be investigated in the following test.

5.2. Influence of the Shape of the Distribution Areas

The perimeter-area-ratio, the shape deviation and the contour index shall serve as indicators in this test. The most favorable perimeter-area-ratio is found in a circle. The inner area of a circle is surrounded by a minimal perimeter. The more complex an area is (e.g. multi-part areas, areas with inner courtyards) the larger the perimeter is compared to the size of the area (see *Figure 2*). With this the possibility to place the same number of dots within this area as within a coextensive circle dwindles.

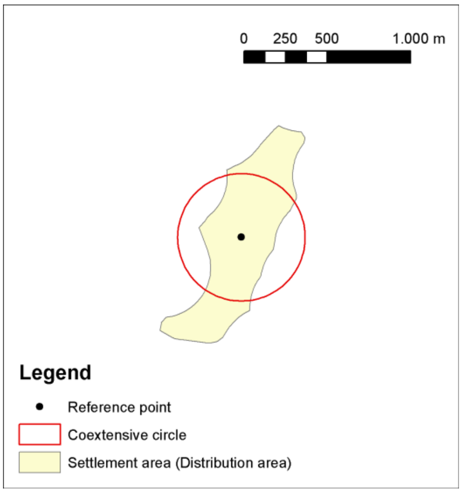

Figure 2. Real distribution area and coextensive circle.

Sometimes the smaller number of dots does not result in a changed dot value (see *Figure* 3).

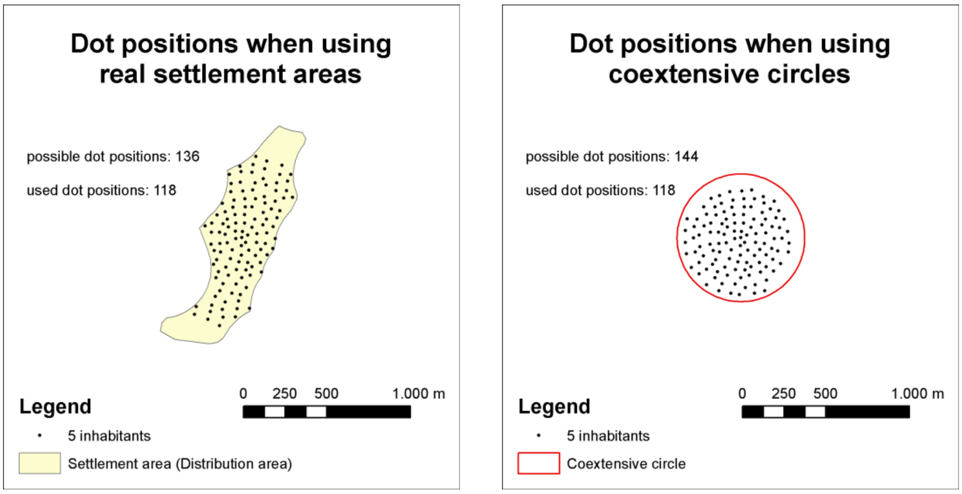

Figure 3. Differences in possible dot positions.

To find out how big the influence of the shape of distribution areas is the dot map was designed on the basis of the real distribution areas and with coextensive circles. The behavior of the dot value in dependence of the indicators was observed. The results of this test are not explicit. Figure 4 shows the relation between perimeter-area-ratio and the obtainable dot value. Although the average shows quite a clear trend, the extreme values cannot be assigned definitely.

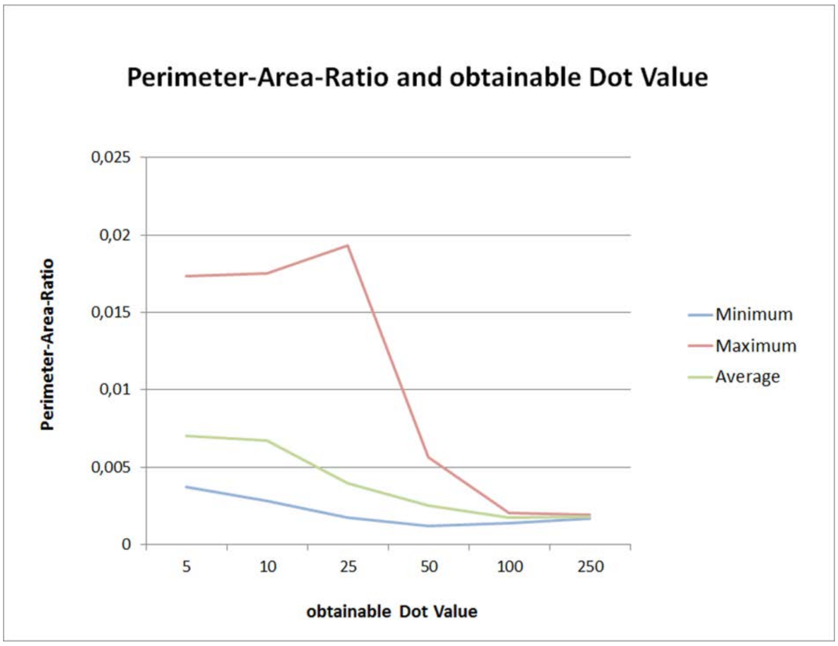

Figure 4. Relation between perimeter-area-ratio and obtainable dot value.

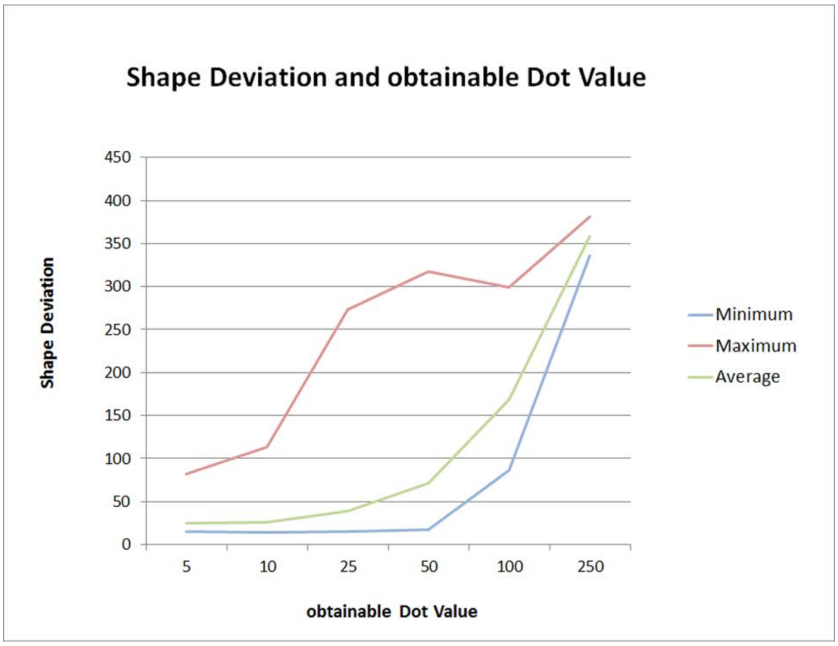

Figure 5. Relation between shape deviation and obtainable dot value.

As mentioned before the perimeter-area-ratio is influenced by shape and size of the area. To separate both impacts two shape indices were considered – shape deviation and contour index. *Figures 5* and *6* show the relation between dot value and these two indicators. In *Figure 5* the shape deviation value of about 12.6 marks the coextensive circle. In *Figure 6* the coextensive circle can be found at the contour index value 1.

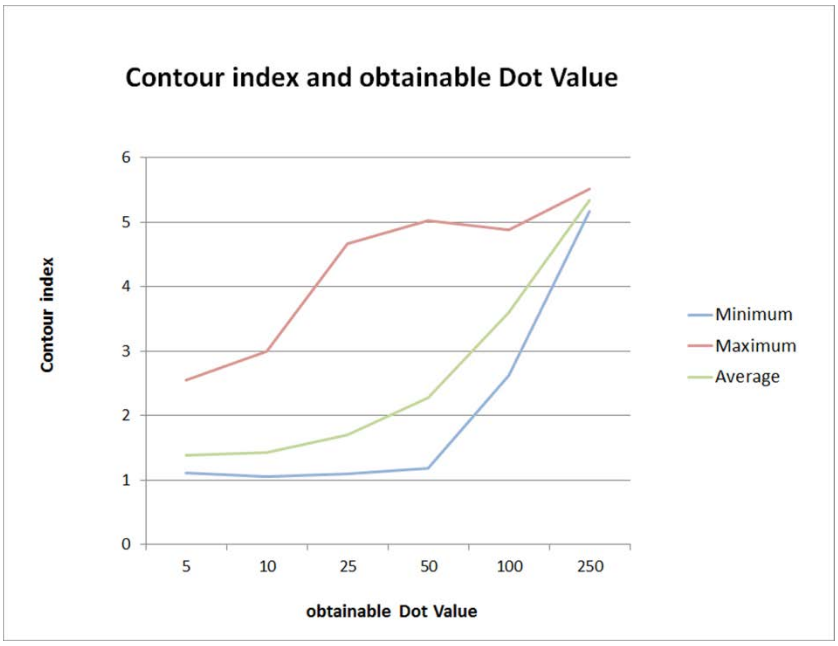

Figure 6. Relation between contour index and obtainable dot value.

The indicators in *Figure 5* and *6* again share the problem of the former ones. The average shows a trend but the extreme values do not. So only some very general conclusions can be drawn. Both diagrams confirm the presumption that a greater deviation from a circle leads to a smaller number of dots and with this to a higher dot value. In many cases the dot value remained unchanged. Only a few cases (7.6%) allowed a smaller dot value when using the coextensive circles.

Because none of the considered ratios show a direct connection to the obtainable dot value, all seem to be insufficient to serve as indicator. There has to be other numbers and more testing to find out if predictions considering the dot value are possible.

6. Conclusions

The purpose was to find out, to which extent the value-area-ratio, the perimeter-area-ratio, the shape deviation and the contour index allow predictions considering the dot value in dot maps. For the value-area-ratio no explicit connection could be found. If the area or the data value is extreme compared to the whole data set, then this ratio may indicate problems in finding a suitable dot value for heterogeneous data sets. As soon as the values are somewhere in the middle the explanatory power fades.

The idea, that the shape of the distribution areas compared to a circle has an impact on the obtainable dot value was derived from the expectation that the greater the deviation is the smaller the number of dots is that can be placed. Thus the obtainable dot value increases (lesser quantitative accuracy of the map). With the example data set this could only be affirmed for average values. But extreme values did not show clear trends.

The correlation between the area in the map, which is available for dot placement (distribution areas), and the obtainable dot value (number of dots) is not as obvious as expected. This may be caused by the fact that the dots are not dispersed equally across the distribution areas and that the reference points are not necessarily in the center of the areas. The pseudo-random dot shifting and the progression of the gaps between adjacent dots with a growing distance from the center of the dot cluster cause differences in the number of dots that can be placed. But these differences are only sometimes clearly enough to result in another dot value. No matter if the shape deviation was large or small a change of the dot value compared to the original distribution areas occurred quite seldom.

The found connections are still very vague and need to be complemented by other indicators to be able to make reliable predictions. Changing the dot size or the map scale might have a much larger impact. In general, it can be said, that great differences in the data set are very likely to require different dot values. If this will lead to bigger problems cannot be said.

7. Outlook

Besides the tested impact factors of value-area-ratio and shape of the distribution areas also classic elements like dot size and map scale perform a large impact on the reachable dot value. Additionally the amount of the pseudo-random shift of dot positions which is part of the used dot placement algorithm contributes to this impact.

The performed tests can serve as basis for more extensive tests. Especially considering the automation of combined representation methods, like the mentioned ‘Kleingeldmethode’, some rules may be found for the implementation.

References

Aschenbrenner J (1989) Die EDV-unterstützte Herstellung von Punktstreuungskarten auf der Basis kleinster Bezugseinheiten. Wien

De Berg M (2004) On simplifying dot maps. Computational Geometry 27: 43-62

Hey A (2006) Gestaltung und automatisierte Bearbeitung der Karte Bevölkerungsverteilung für Band 12 des Nationalatlas Bundesrepublik Deutschland (Diploma Thesis). TU Dresden

Hey A (2012) Untersuchungen zur automatisierten Visualisierung statistischer Geodaten mittels Punktmethode (Dissertation). University of Rostock

Hofmann B (1972) Entwicklung von Modellen zur Darstellung gitternetzbezogener Erscheinungen mit der Punktmethode (Diploma Thesis). TU Dresden

Kimerling A J (2009) Dotting the Dot Map, Revisited. Cartography and Geographic Information Science 36(2): 165-182

Klamt I (1972) Untersuchungen zur Punktmethode (Diploma Thesis). TU Dresden

Koch W G, Stams W (2002) Punktmethode. Lexikon der Kartographie und Geomatik, Bollmann J/Koch W G (editors), Heidelberg, Berlin: 242f.

Lavin S (1986) Mapping continuous geographical distributions using dot-density shading. American Cartographer 13(2): 140-150