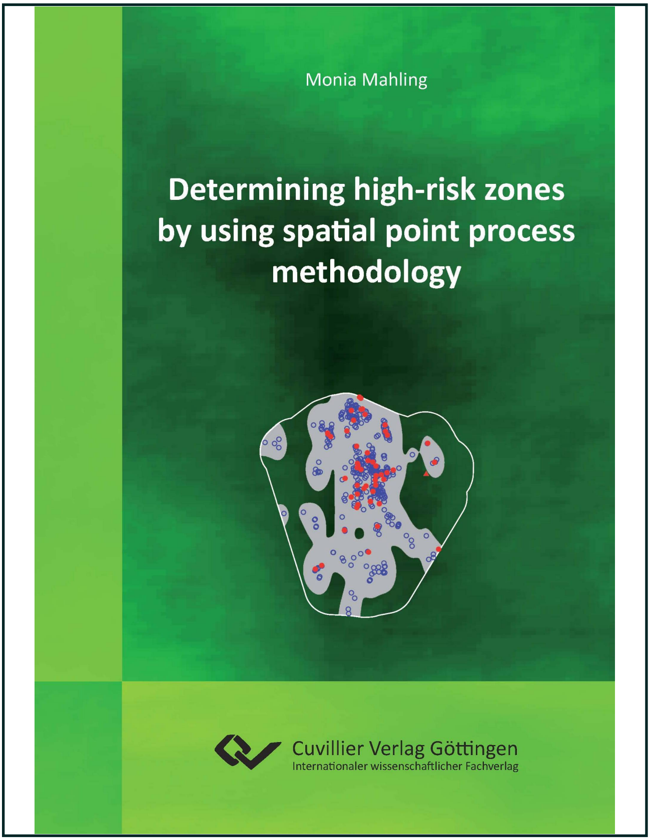
Monia Mahling (Autor)

Determining high-risk zones by using spatial point process methodology



https://cuvillier.de/de/shop/publications/6437

1. Background and motivation

High-risk zones are relevant in all situations which can be characterized as follows: Some kind of event is observed in random locations, but not all locations are known and one would like to find a region where the unobserved events will be discovered with a high probability. Such situations can arise in a variety of applications, such as epidemiology (e.g. if not all cases of a certain type of infectious disease are reported to the authority in charge) or ecology (e.g. for locations of rare plants which are difficult to detect). The focus in this thesis is on one specific application, unexploded World War II bombs.

Even more than 65 years after the end of the Second World War, unexploded bombs still represent a serious problem in Germany. Their clearance usually requires the evacuation of houses and the closing of roads and railway lines. As they are often accidentally found during construction work, these actions have to be taken quickly and often at times which are especially inconvenient. Even worse, unintended detonations have resulted in severe accidents in several cases.

To avoid accidents and render evacuations more foreseeable, it is desirable to search high-risk areas for unexploded bombs before any construction work starts. Depending on several characteristics of the subsoil, the search alone–without clearance or possible reconstruction–costs between 0.20 and 20 € per m^2, so it must be restricted to carefully selected areas.

During and after the Second World War, the Allies took aerial pictures of regions they had bombed. An example is given in Figure 1.1. Nowadays, experts analyse these aerial pictures and derive the locations of bomb craters. In some cases, smaller structures in the aerial pictures, which are difficult to detect, indicate that a bomb may have thuded in this place, but did not detonate. However, this does not necessarily mean that an unexploded bomb is located in this position, as such findings from the aerial pictures are rather vague and it is often impossible to retrace where unexploded ordnance was removed during and in the first years after World War II.

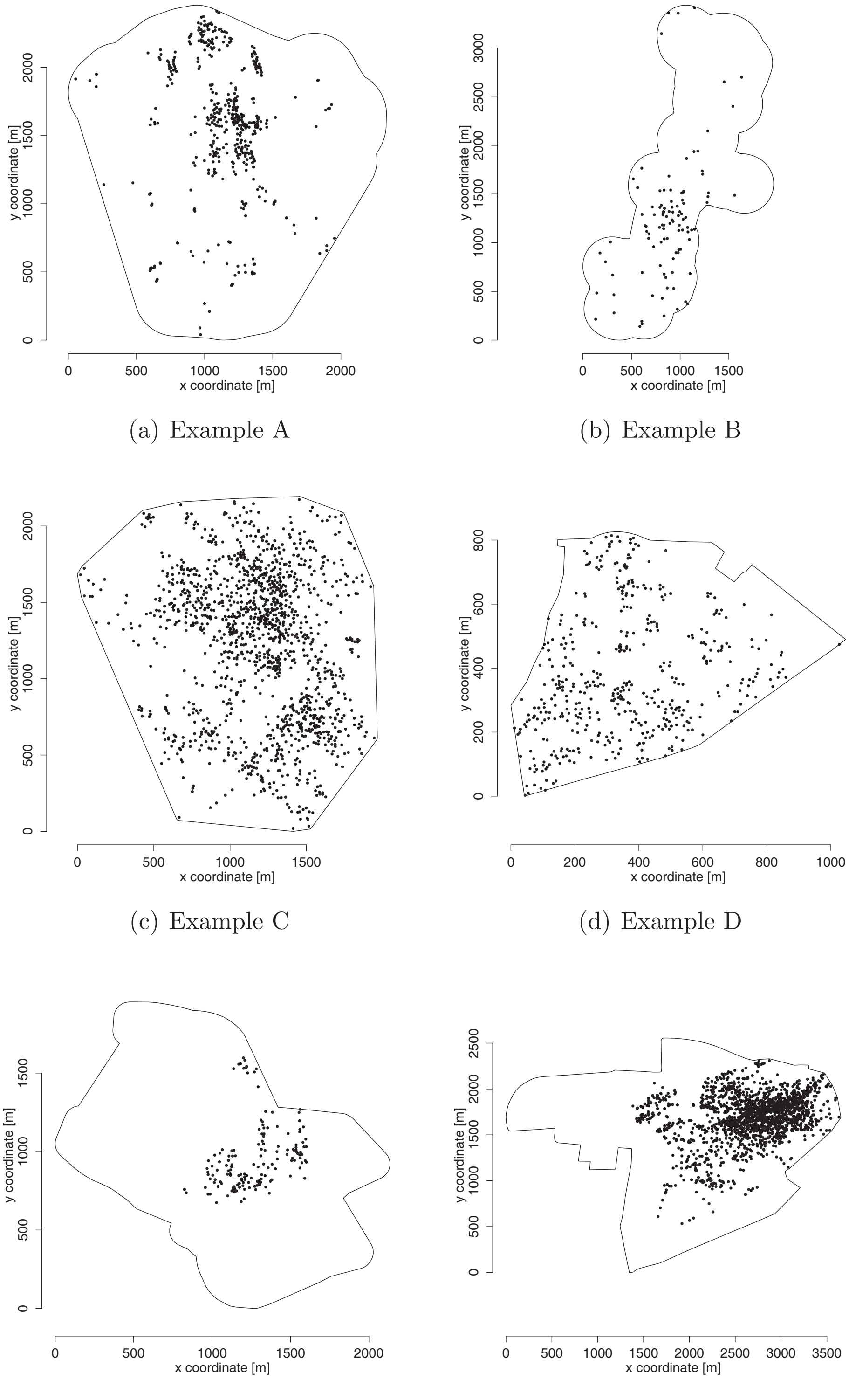
If suitable aerial pictures of the area in question exist, the locations of bomb craters can be used to determine high-risk zones for unexploded bombs. Note that this is usually not possible for properties situated in cities because the bomb craters are mostly covered by ruins of houses and therefore cannot be discerned in the pictures. The high-risk zones determined on the basis of the bomb craters are an important step in deciding where to search for unexploded bombs. Additionally, other aspects such as historical data from archives are also considered before the final decision is taken.

Up to now, high-risk zones have been defined in a way where only very little information from the data is used, namely the coordinates of every single observation, but no characteristics of the pattern in general.

Figure 1.1.: Example of an aerial picture showing bomb craters (source: National Archive - Aerial Photo, Sortie 34-3658, Date 24.03.1945 (WW II Europe)).

Therefore, *Oberfinanzdirektion Niedersachsen* (OFD), which supports the removal of unexploded ordnance in federal properties in Germany, and *Mull und Partner Ingenieurgesellschaft*, who are experts in the analysis of aerial pictures, searched for more sophisticated approaches. In a cooperation project with the Statistical Consulting Unit (*Statistisches Beratungslabor*), Department of Statistics, Ludwig-Maximilians-Universität München, the task of evaluating existing methods and developing a novel approach was addressed. A further aim was to perform a risk assessment of the investigated areas, e.g. by estimating the probability that there are unexploded bombs outside the high-risk zone. The final goal of the cooperation was to develop a procedure which can be applied automatically by users who are not experts in statistics. The cooperation project started in May 2009 and was funded by *Oberfinanzdirektion Niedersachsen*.

Examples of the data are presented in Figure 1.2. The georeferenced locations of the bomb craters have been provided by *Mull und Partner*. As all information is derived from the aerial pictures, no data are available about the locations of unexploded bombs that have been found for the specific areas of interest. Example A comprises 443 observations of bomb craters in an area of approximately 400 *ha*. Example B consists of 104 observations in an area of approximately 350 *ha*. The bomb craters are mainly located in the southern part of the property. The 1369 observations of Example C are scattered over large parts of the property with an area of 334 *ha*. Example D consists of 451 observations on 52 *ha*. They seem to be more dense in the south of the property. The 152 bomb craters of Example E are concentrated on a rather small part of the property, which has an area of 239 *ha*. Example F comprises 1706 observations on 504 *ha*. Most of them are located in the north-east of the property.



(a) Example A

(b) Example B

(c) Example C

(d) Example D

(e) Example E

(f) Example F

Figure 1.2.: Properties to be cleared: The solid lines represent the border of the areas for which data are available, the points illustrate the locations of bomb craters.

The properties to be cleared typically had military importance during war (i.e. barracks, airfields, military training areas) or were important from a strategic or economic point of view (i.e. rivers, floodgates, roads and industrial plants). Since detailed information on specific premise type and location could facilitate the identification of the property and hence have an economic impact, no further information other than relative coordinates on the specific property was provided. In general, it is not justified to assume several targets in a property. In some cases, like for Example B, the target of the attack even seems to be situated outside the property to be cleared.

The probability of non-explosion of every bomb can vary depending for example on the subsoil and the year of the attack. A well-established reference value is 0.1. In some cases, it is possible to estimate the probability of non-explosion for a given property from historical records. It is usually assumed to be constant on the whole property, as there is little data from which more specific assumptions could be derived.

The novel approach for constructing high-risk zones consists in interpreting the observed pattern of bomb craters as a realisation of a spatial point process. This point of view is widespread in the analysis of the locations of lightning strikes (Schabenberger and Gotway, 2005) or earthquakes (Vere-Jones, 1970; Choi and Hall, 1999). McDonald and Small (2006, 2009) used spatial point process methodology for analysing patterns of unexploded ordnance at former air force bombing ranges. Neyman and Scott (1972) mention patterns resulting from bombing as an example for a special type of spatial point process (see Section 2.5 for Neyman-Scott processes). These point process models had been used to optimise the formation bombing strategy for clearing land mines from the landing beaches in Normandy during World War II.

Interpreting the observed bomb crater patterns as realisations of spatial point processes provides a rich methodology to analyse the patterns and develop a construction method for high-risk zones: Various point process characteristics can be considered to investigate the properties of the observed patterns. In particular, the intensity function serves as basis for high-risk zones.

This thesis is organised as follows: Chapter 2 procures selected notation and properties of spatial point processes. The properties of the six real-data examples are investigated in an exploratory analysis via functional summary characteristics in Chapter 3 in order to find an appropriate model for the data. In Chapter 4, three methods for constructing high-risk zones are presented: The traditional method, the quantile-based method, which is not entirely new, but based on the considerations of *Mull und Partner*, and–as a completely novel approach–the intensity-based method. The behaviour of the construction methods is investigated in Chapter 5, which also contains a model check and a comparison of the theoretical properties of intensity-based and quantile-based high-risk zones. The chapter finishes with a recommendation for the intensity-based method. In Chapter 6, the risk associated with an intensity-based high-risk zone is assessed. As this risk does not reliably equal the parameter which is intended to specify it, a correction method is proposed. The consequences of spatial clustering are investigated in Chapter 7. A sensitivity analysis is performed and different types of models which account for clustering are fitted to the data. The R package `highriskzone` comprising an implementation of the main methods of this thesis is introduced in Chapter 8. Finally, a summary of the most important results and a review of open research questions is given in Chapter 9.

All analyses were performed by using the statistical software R (R Development Core Team, 2012). In particular, the R package `spatstat` (Baddeley and Turner, 2005, 2006) for the analysis of spatial point patterns was employed.

Parts of Chapters 3 to 7 have been published in an article in the Journal of the Royal Statistical Society (Series C) (Mahling et al., 2013). This article contains contributions by Michael Höhle and Helmut Küchenhoff. Most of the ideas are my own. I performed all analyses and wrote the article. Helmut Küchenhoff and Michael Höhle commented on the manuscript.

The R package `highriskzone` which is introduced in Chapter 8 was created by Heidi Seibold on the basis of my implementation of the methods for constructing and evaluating high-risk zones. The package is the major part of Heidi Seibold's bachelor thesis (Seibold, 2012), which was supervised by Helmut Küchenhoff and me jointly.