

Master Thesis

im Rahmen des
Universitätslehrganges „Geographical Information Science & Systems“
(UNIGIS MSc) am Zentrum für GeoInformatik (Z_GIS)
der Paris Lodron-Universität Salzburg

zum Thema

„Crime Mapping in Hamburg“ Eine Suche nach alternativen Darstellungsformen der Straftatenverteilung

vorgelegt von

Dipl. Geogr. Gabriele Dall
u1036, UNIGIS MSc Jahrgang 2003
Matrikelnummer: u0222650

Zur Erlangung des Grades
„Master of Science (Geographical Information Science & Systems) –
MSc(GIS)“

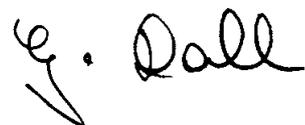
Gutachter:
Ao. Univ. Prof. Dr. Josef Strobl

Hamburg, 24.04.2005

Erklärung

Ich versichere, diese Master-Thesis ohne fremde Hilfe und ohne Verwendung anderer als der aufgeführten Quellen angefertigt zu haben. Ich habe diese Arbeit weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Alle Ausführungen der Arbeit, die wörtlich oder sinngemäß übernommen wurden, sind entsprechend gekennzeichnet.

Hamburg, den 24. 04. 2005

A handwritten signature in black ink, appearing to read 'G. Dall'. The signature is written in a cursive style with a large initial 'G' and a long horizontal stroke.

Vorwort

Crime-Mapping ist ein interessantes Thema und die Einsatzmöglichkeiten dieser – im Prinzip schon relativ alten Technik – haben sich mit dem Aufkommen von Geographischen Informationssystemen (GIS) und computergestützter Analysemethoden stark erhöht. Da die Polizei in Hamburg seit dem Jahr 2001 auch eine Crime-Mapping Anwendung (als Intranet-Anwendung auf der Basis von ArcIMS der Fa. ESRI) betreibt, diese jedoch noch nicht mit räumlichen Analysemethoden ausgestattet ist, lag der Gedanke nahe, im Rahmen meiner Master-Thesis nach geeigneten Analysemöglichkeiten zu suchen.

Die für diese Untersuchung eingesetzten Daten wurden mir freundlicherweise von der Hamburger Polizei zur Verfügung gestellt. Es handelt sich um die Diebstahl- und Einbruchdaten des Jahres 2003. Für die Erlaubnis zur Nutzung dieser Daten möchte ich mich bei Herrn Kondoch bedanken und auch für sein Interesse am Voranschreiten dieser Arbeit. Herrn Korzeniowsky danke ich für die Kopie der Datensätze, die er für mich aus dem System herausgeholt hat.

Da die Straftaten sich auf die Hamburger Straßenabschnittsdaten beziehen, welche beim Statistischen Amt für Hamburg und Schleswig-Holstein gepflegt werden, danke ich Frau Dr. Mausfeld für die Erlaubnis zur Nutzung dieser Daten und dem Datenpflegeteam des Hamburger Einsatzleitsystems „HELP“ für die Erlaubnis, die Straßenabschnittsdaten aus diesem System zu kopieren.

Abschließend danke ich meinem Arbeitgeber T-Systems GEI GmbH in Hamburg für die Finanzierung dieses Studiums und meinem Kollegen Helmut Steffens für seine Diskussionsbereitschaft und für seine Hilfe bei der Durchsicht dieser Arbeit.

Auch dem UNIGIS-Team um Herrn Prof. Dr. Josef Strobl möchte ich für die gute Organisation des UNIGIS-Studiums und die hervorragende Betreuung innerhalb der letzten 2,25 Jahre danken.

Inhaltsverzeichnis

Vorwort.....	iii
Zusammenfassung.....	ix
Summary.....	xi
Motivation und Ziele.....	1
Teil 1: Kriminalitätsanalysen und der Einsatz von GIS.....	4
1 Einsatzfelder geografischer Informationssysteme im Umfeld der Polizei.....	4
2 Crime Mapping – was ist das?.....	5
2.1 Straftatenanalyse durch „Crime Mapping“.....	6
2.2 Historische Entwicklung des „Crime Mapping“.....	7
2.3 Crime Mapping und GIS.....	9
3 Die Randbedingungen des Crime-Mappings.....	12
3.1 Datenerfassung.....	12
3.2 Datenauswertung.....	14
3.3 Neue Herausforderungen für Crime-Mapping-Systeme.....	15
4 Datenanalyse und Ergebnisdarstellung im Crime-Mapping.....	17
4.1 Straftaten als diskrete Punkteverteilung.....	18
4.2 Die Darstellung der Straftaten in einer Fläche.....	20
4.3 Die Straftatendichte als kontinuierliche Oberfläche.....	24
4.4 Weitere Methoden aus dem Umfeld der Point-Pattern-Analyse.....	26
5 Bisher eingesetzte Programme und Methoden	36
5.1 STAC – Spatial and Temporal Analysis of Crime.....	37
5.2 CrimeStat.....	38
5.3 GAM – Geographic Analysis Maschine(s).....	39
Teil II: Konkreter Fall – Hamburg.....	40
1 GISPOL – die Crime Mapping Anwendung der Hamburger Polizei	40
2 Die Erstellung eines Lagebildes mit GISPOL.....	41
3 Die Beschaffenheit der Daten.....	42
3.1 Welche Daten werden untersucht?.....	42
3.2 Repräsentativität des Datenmaterials.....	44
4 Alternativen zur „Steckkarte“	47

4.1 Nearest-Neighbour-Analyse zur ersten Untersuchung der Datensätze.....	47
4.2 Flächenhafte Darstellung der Straftatenverteilung.....	49
5 Lokale Autokorrelation und Clusterbildung.....	71
5.1 Diebstähle 2003.....	73
5.2 Einbrüche 2003.....	75
5.3 Methodenkritik.....	79
6 Raumzeitliche Verteilung der Straftaten	79
6.1 Qualität der Zeitangaben.....	80
6.2 Die raumzeitliche Verteilung der Einbrüche und der Diebstähle.....	80
6.3 Suche nach raumzeitlichen Clustern mit LISA.....	92
7 Einsatz von Methoden des Spatial Data Mining	95
7.1 Spatial Data Mining – Aufdecken räumlicher Cluster.....	96
7.2 Entdecken raum-zeitlicher Cluster durch „Spatio-Temporal“ Data Mining....	98
8 Methodendiskussion.....	99
Literaturverzeichnis.....	102
Weitere Web-Adressen.....	109

Abkürzungsverzeichnis

CSR	Complete Spatial Randomness (räumliche Zufallsverteilung) theoretische Verteilung, gegen die die gegebenen Punktverteilung getestet wird (zufällig verteilt, gleichmäßig oder räumlich konzentriert)
GIS / GI-Systeme	Geographische Informationssysteme
KDE	Kernel-Density-Estimation spezielles Glättungsverfahren zur Abschätzung der räumlichen Verteilung eines Ereignisses, gehört zum Methodeninventar der PPA
LISA	Local Indicators of Spatial Association
NNA	Nearest Neighbour Analyse
PKS	Polizeiliche Kriminalstatistik (offizielle jährliche Statistik)
PPA	Point-Pattern Analyse Methoden zur explorativen statistischen Analyse diskreter Punktverteilungen

Abbildungsverzeichnis

Abbildung 1.....	4
Abbildung 2.....	19
Abbildung 3.....	22
Abbildung 4.....	24
Abbildung 5.....	25
Abbildung 6.....	26
Abbildung 7.....	29
Abbildung 8.....	31
Abbildung 9.....	38
Abbildung 10.....	40
Abbildung 11.....	47
Abbildung 12.....	50
Abbildung 13.....	50
Abbildung 14.....	53
Abbildung 15.....	54
Abbildung 16.....	56
Abbildung 17.....	57
Abbildung 18.....	58
Abbildung 19.....	59
Abbildung 20.....	61
Abbildung 21.....	64
Abbildung 22.....	65
Abbildung 23.....	67
Abbildung 24.....	76
Abbildung 25.....	77
Abbildung 26.....	82
Abbildung 27.....	83
Abbildung 28.....	85
Abbildung 29.....	86
Abbildung 30.....	89
Abbildung 31.....	90
Abbildung 32.....	94
Abbildung 33.....	97

Tabellenverzeichnis

Tabelle 1.....	43
Tabelle 2.....	44
Tabelle 3.....	48
Tabelle 4.....	69
Tabelle 5.....	72

Zusammenfassung

Der Einsatz von Geographischen Informationssystemen (GIS) ist aus dem Umfeld der Polizeiarbeit – und insbesondere auch aus dem Themenbereich der Straftatenanalyse nicht mehr wegzudenken. Dies ist dem ersten Teil dieser Arbeit zu entnehmen, in dem die Entwicklung der Crime-Mapping Anwendungen und Methoden beschrieben wird. Im Verlauf der vergangenen 10 Jahre wurden große Anstrengungen vorgenommen, mit Methoden der explorativen Datenanalyse die räumliche Verteilung von Straftaten aufzubereiten und Schwerpunkte zu erkennen. Diese Anwendungen wurden bisher allgemein als Expertensystem an einzelnen GIS-Arbeitsplätzen eingeführt.

Die Hamburger Polizei ist mit der Anwendung GISPOL einen anderen Weg gegangen. Die Straftaten werden dort unmittelbar in ein Intranet-GIS eingestellt und den Anwendern zeitnah zur Verfügung gestellt. Die Straftaten werden als Punkte abgebildet. Die Punkte repräsentieren jedoch nicht die tatsächlichen Tatorte, sondern stellen einen Straßenabschnitt dar, der dem Tatort am nächsten liegt. Bisher sind keine räumlichen Analysen (abgesehen von den Standardfunktionen der Intranet-GIS-Anwendung) vorhanden.

Das Ziel dieser Arbeit besteht nun darin, Methoden zu untersuchen, mit deren Hilfe die räumliche Darstellung der Straftaten verändert und das Erkennen von Schwerpunkten erleichtert werden soll. Über diese Suche nach geeigneten Methoden berichtet der zweite Teil dieser Arbeit.

Neben den bekannten Verfahren der Quadrat-Methode und der Kernel-Density-Estimation (KDE) wurde auch eine bisher unbekannte Methode für die Straftatenanalyse – die Potentialanalyse – untersucht. Auch die von Anselin (1995) und Getis & Ord (1992 und 1995) vorgestellten Methoden zur Untersuchung lokaler Beziehungen in räumlichen Datensätzen (LISA-Statistiken) wurden eingesetzt, da sie in die neue ArcGIS-Generation (ArcGIS 9.x) Eingang gefunden haben.

Die Daten wurden sowohl nach räumlichen als auch nach raum-zeitlichen Kriterien untersucht. Insbesondere die Darstellung der raumzeitlichen Verteilung der untersuchten Straftaten (Einbrüche und Diebstähle des Jahres 2003) zeigten sich erhebliche Unterschiede im Hinblick auf die dargestellte raum-zeitliche Verteilung aber auch im Hinblick auf die Ergebnisse, die mit den unterschiedlichen Methoden erzielt wurden.

Im Kapitel 7 des II. Teils wurde dann das „Spatial Data Mining“ als zukunftsweisende Methode zur Untersuchung räumlicher Daten vorgestellt. Insbesondere die bisher entwickelten Verfahren zum Aufdecken räumlicher Cluster lassen einen großen Nutzen für die Hot-Spot Analyse im Crime-Mapping erwarten – auch wenn in dieser Arbeit keines der bisher in der Literatur vorgestellten Methoden zur Untersuchung der Hamburger Daten eingesetzt wurden. Es wird jedoch erwartet, dass sich diese Methoden unmittelbar in die Datenbank integrieren lassen und somit auf die Datenbank-Infrastruktur zur Verwaltung räumlicher Daten zurückgreifen können. Dies sollte zu einem Performance-Gewinn beitragen.

Diese Arbeit schließt mit einer Bewertung der verschiedenen Methoden und ihrer Eignung in Bezug auf die Integration räumlicher Analysefunktionen in GISPOL. Es lässt sich nicht eine einzige Methoden als die für alle Zwecke geeignete Methode herausfinden. Die Wahl der geeigneten Methoden hängt ab von dem Ziel, das es zu erreichen gilt, bzw. von der Ergebnisdarstellung, die man den Anwendern zumuten möchte.

Summary

Geographic Information Systems (GIS) have taken over serious parts in police management systems – among which crime mapping software plays an important role. The first part of this thesis therefore deals with a short description of the development of Crime Mapping Systems and the methods they use. During the last ten years great efforts have been taken to analyse the spatial distribution of certain crime and find out the high-crime areas, the so called „hotspots“, by using explorative data analysis. In most cases crime mapping and crime analysis software were established as expert systems on only a few desktop workplaces.

However, Hamburg Police went a different way when establishing the GISPOL-Application in the year 2001. Crime incidents are published via Intranet-GIS and local area network, so that every Head Officer is capable to get information about the current dissemination of crime in his district. Crime data are georeferenced as point data, but these points do not symbolize the exact crime location. They rather represent some special points on the Hamburg street network, which are closest to the documented crime location. Spatial analysis functions (except the Intranet-GIS standard functions like buffer, identify, etc.) have not been implemented, so far.

This master thesis aims at finding some spatial analysis functions that will be useful to analyse and display the spatial dissemination of crime data and to identify crime hotspots. The second part of this work, therefore describes these efforts in finding applicable methods. Among the well-known methods of Quadrat-Count (chapter II, 4.2.1.1) and Kernel-Density-Estimation (chapter II, 4.2.2), other techniques like Potential Analysis (chapter II, 4.2.3) or LISA (local indicator of spatial association, chapter II, 5) are tried out.

Spatial as well as spatio-temporal analyses were performed and documented. Especially analysing the spatial-temporal dissemination of crime data (2003 theft and burglary data) showed large differences according to the differing methods used. Chapter 7 (in the second part), finally described „spatial data mining“ as forward-looking method in finding clusters or patterns in the distribution of spatial data in general and crime data in particular. New methods in cluster detection, that could be implemented directly on the database layer, will be expected to become helpful in future hot-spot detection methods.

Finally, this thesis comes to the conclusion that it is not possible to decide on one single function alone. For spatial and spatio-temporal analysis different methods are suggested.

Motivation und Ziele

Die Aufgabe der Polizei ist es, Recht und Ordnung im öffentlichen Raum zu gewährleisten. Um dieser Aufgabe gerecht zu werden und die begrenzten Ressourcen effektiv einzusetzen, bedarf es der Kenntnis über das aktuelle Geschehen in diesem Raum. Also besteht auch der Bedarf, die räumliche Verteilung von Straftaten zu verfolgen. Dies geschieht mit dem Ziel, Verteilungsmuster zu erkennen, um daraus taktische und strategische Gegenmaßnahmen abzuleiten.

In der Mitte der 90er Jahre des letzten Jahrhunderts haben Geographische Informationssysteme (kurz: GI-Systeme oder GIS) in das Arbeitsumfeld der Polizei Einzug gehalten. Damit stehen Werkzeuge bereit, die es der Polizei ermöglichen, die räumliche Verteilung von Straftaten abzubilden, zu verfolgen und bei auffallenden Veränderungen darauf zu reagieren. Das Lokalisieren von Tatorten in einer Karte und die damit einhergehenden Möglichkeiten zur Visualisierung wird als „Crime Mapping“ bezeichnet. Insbesondere in den USA und im weiteren angelsächsischen Sprachraum hat sich in den letzten Jahren ein eigener Forschungszweig zu diesem Thema entwickelt. In dessen Umfeld haben sich bereits diverse Anwendungen etabliert und es werden fortlaufend neue Methoden erprobt.

Seit dem Jahr 2000 gibt es auch bei der Hamburger Polizei eine Crime-Mapping-Anwendung. Diese ist als Intranet-Anwendung konzipiert. Mit Hilfe eines Web-GIS werden die Daten zur Verfügung gestellt. Sie können von den Anwendern in den Revieren abgerufen werden. Suchmechanismen ermöglichen das Abfragen von Sachdaten zu jedem Datensatz. Auch die zeitlichen Verteilung der Straftaten ist abrufbar und lässt sich in Form von Balkendiagrammen darstellen. Eine raum-zeitliche Betrachtung der Straftatenverteilung ist bisher nicht realisiert.

Eine Besonderheit der Hamburger Anwendung besteht in der Verortung der Straftaten. Während in vielen amerikanischen Systemen die Adressen mit Hilfe eines Geocoders auf einem Straßengraphen interpoliert werden (vgl. Vann & Garson (2003), S. 27) und in den britischen Anwendungen die eingegebenen Adressen auf ein regelmäßig verteiltes Postleitzahlengitter bezogen werden (s. Gatrell et al. (1996), S. 256), sind die Hamburger Adressangaben auf Strassenabschnitte aggregiert. Diese Strassenabschnitte

sind unregelmäßig über das Stadtgebiet verteilt. Ein Straßenabschnitt wird als Punkt repräsentiert. Dieser Punkt vertritt nun alle Adressen, die beispielsweise in einem Abschnitt zwischen zwei Kreuzungen liegen (vgl. Statistisches Amt für Hamburg und Schleswig-Holstein (2004)). Diese Verortung der Daten führt dazu, dass mehrere Straftaten an einem einzigen Straßenabschnittspunkt zusammenfallen können, von denen kein einziger dort tatsächlich stattgefunden haben muss.

Da jede Straftat als einzelner Punkt repräsentiert wird und sich u. U. mehrere Punkte überlagern, ist es einem Beobachter unmöglich zu erkennen, ob an einem Punkt nur eine einzige oder mehrere Straftaten festgestellt wurden. Sichtbar ist in allen Fällen immer nur ein Punkt (der oberste). Sollte darüber hinaus an jedem Strassenabschnitt (mindestens) ein Ereignis verzeichnet worden sein, so stellt das resultierende Muster weniger die Verteilung der Straftaten, sondern die Verteilung der Strassenabschnitte dar. Spätestens hier gerät die ausschließlich visuell orientierte Interpretation des Punktmusters an ihre Grenzen.

Das Ziel der vorliegenden Arbeit ist es, den zum gegenwärtigen Zeitpunkt in GISPOL vorherrschenden Informationsverlust aufzulösen und eine andere Darstellungsform zu finden, mit deren Hilfe die räumliche Verteilung der erfassten Delikte deutlicher hervortritt. Zu diesem Zweck werden verschiedene Methoden zur Darstellung räumlicher sowie raumzeitlicher Verteilungen untersucht.

Ein weiteres Ziel der vorliegenden Arbeit besteht darin, Methoden zu erforschen, die geeignet sind, Bereiche hoher Kriminalitätsraten zu identifizieren. Die Ergebnisse der Analysen sollen (karto-)graphisch aufbereitet und in einem GIS dargestellt werden. Eine mögliche Darstellungsform könnte z. B. in der Erzeugung von „Wärmebildern“ bestehen. Der Begriff „Wärmebild“ beschreibt eine Darstellungsform durch die räumliche Konzentrationen von Ereignissen hervorgehoben werden. Der Darstellung gehen Berechnungen (räumliche Analysen) voraus. Zu diesem Zweck werden verschiedene Methoden der explorativen Datenanalyse eingesetzt. Unter dem Begriff der „explorativen Datenanalyse“ werden Methoden zusammengefasst, mit deren Hilfe Datensätze nach Strukturen und Besonderheiten durchsucht werden (vgl. Ester & Sander (2000), S. 30). Das Ergebnis dieser Berechnungen stellt eine sog. „statistische

Oberfläche“ dar, welche die Verteilung der untersuchten Ereignisse über das gesamte Untersuchungsgebiet darstellt.

Da es sich bei der Anwendung der Hamburger Polizei um eine Web-Applikation handelt, und die in dieser Arbeit untersuchten Verfahren sich – im Falle einer Übernahme – in diese Anwendung integrieren müssen, ist die Eignung der Methoden auch an den Anforderungen an eine Web-Anwendung zu messen. Dies bedeutet, dass die ausgewählten Methoden schnell zu einem darstellbaren Ergebnis kommen müssen und dass die Ergebnisse leicht darstellbar und ebenso leicht interpretierbar sein müssen. Weiterhin sollen so wenige Parameter wie möglich durch die Anwender definiert werden, denn dies erschwert die Vergleichbarkeit der Ergebnisse.

Teil 1: Kriminalitätsanalysen und der Einsatz von GIS

1 Einsatzfelder geografischer Informationssysteme im Umfeld der Polizei

Ein großer Teil der täglichen Polizeiarbeit ist geprägt von räumlichen Informationen. Ist z. B. ein Verkehrsunfall geschehen, sind verschiedene räumliche Frage zu beantworten: Wo ist das passiert? Wie kommen die Einsatzkräfte schnell dort hin? Gehen mögliche Gefahren von diesem Ereignis aus? Wer ist dann davon betroffen? Es ist deshalb nur verständlich, dass mit der zunehmenden Technisierung der Polizeiarbeit auch die GIS-Technologie Einzug in die polizeiliche Tätigkeit gehalten hat.



Abbildung 1: Law-Enforcement and GIS Enterprise (Quelle: ESRI 2005)

Aus der Abbildung 1 geht hervor, in welchen Teilbereichen der polizeilichen Arbeit geographische Informationen benötigt und durch den Einsatz eines GIS bereitgestellt werden. Eines dieser Tätigkeitsfelder, nämlich die Analyse von Straftaten und Einsätzen, wird im Rahmen der vorliegenden Arbeit näher betrachtet. Dieses Themengebiet ist in der Abbildung 1 durch eine rote Ellipse hervorgehoben.

2 Crime Mapping – was ist das?

Verbrechen treten dann auf, wenn vier Aspekte zusammentreffen (vgl. Brantingham & Brantingham (1981), S. 7):

- das Gesetz – die gesetzlichen Regeln legen den Straftatenkatalog fest, auf den sich die polizeilichen Aktivitäten beziehen
- ein oder mehrere Täter
- ein Ziel, bzw. ein Opfer
- ein Ort, welcher eine „...discrete location in **time** and **space** at which the other 3 dimensions intersect and a criminal event occurs“ (Brantingham & Brantingham (1981) S. 7) darstellt.

Diese Beschreibung weist dem Tatort und dem Tatzeitpunkt ein hohes Erklärungspotential zu, da hier alle übrigen Einflussfaktoren zusammentreffen und diese beiden mehr oder weniger exakt ermittelt werden können. Für das Zusammentreffen von Täter und Opfer zu einem gewissen Zeitpunkt, an einem gewissen Ort kann es Gründe geben. Ein Teil dieser Gründe könnte gerade in dem Tatort und/oder dem Tatzeitpunkt selber begründet sein. Die Untersuchung dieser Einheiten und deren Eigenschaften sowie das Herstellen von Beziehungen zu ähnlichen Tatorten/Tatzeitpunkten könnte Hypothesen entstehen lassen, die die Straftatenverteilung erklären. Ebenso lassen sich aus der Betrachtung von Tatorten/Tatzeiten Taktiken und Strategien zur Kriminalitätsbekämpfung entwickeln.

Vor diesem Hintergrund stellt die statistische Auswertung der erfassten Straftaten und Einsatzdaten (auch unter räumlichen und zeitlichen Gesichtspunkten) ein geeignetes Mittel dar, mit dem die Polizei sich einen Überblick über das aktuelle Aufkommen an Einsätzen und Straftaten in Ihrem Verantwortungsbereich verschaffen kann. Zusätzlich zum Erkenntnisgewinn über mögliche Trends in der Verbrechenslandschaft, dienen solche Statistiken (z. B. die Kriminalitätsstatistik) auch der Erfolgskontrolle. So werden z. B. in der Kriminalitätsstatistik die Straftaten sowohl auf administrative räumliche Einheiten als auch auf die organisatorischen Einheiten der Polizei bezogen. Auf beiden Ebenen wird auch über die jährlichen Veränderungen in den Häufigkeiten der verschiedenen Deliktgruppen berichtet (vgl. Polizei Hamburg (2003), S. 14ff.).

Die räumliche Erfassung und Analyse von Straftaten mit Hilfe kartographischer Methoden wird im englischen Sprachgebrauch als „Crime Mapping“ bezeichnet. Der Begriff „Crime Mapping“ setzt sich zusammen aus den Worten Crime – Verbrechen und „Mapping“ - Darstellung räumlicher Informationen mit Hilfe von Karten (thematische Kartographie). Ganz allgemein lässt sich „Crime Mapping“ als ein Mittel zur Untersuchung (Darstellung und Analyse) der räumlichen Verteilung von Straftaten beschreiben.

2.1 Straftatenanalyse durch „Crime Mapping“

Die einfachste Form des Crime-Mapping stellt die Steckkarte dar: In einem Stadtplan wird jeder Tatort durch das Einstecken einer Nadel markiert. Mit Hilfe dieser Steckkarten wird versucht, die räumliche Verteilung von Straftaten nachzuvollziehen, Schwerpunkte zu erkennen und ggfs. auch die Bewegungsmuster eines Täters im Raum zu durchleuchten. Weiterhin werden Gegenmaßnahmen abgeleitet und eine Erfolgskontrolle dieser Maßnahmen vorgenommen. Steckkarten werden schon seit langem zur Unterstützung der Polizeiarbeit eingesetzt. So lässt sich der Einsatz von Karten in New York bis in das Jahr 1900 zurückverfolgen (vgl. Harries (1999), S. 1). Mit dem Vordringen von GIS in die Polizeidienststellen wird die analoge Form der Steck-Wandkarte (engl. „Pin Map“) mehr und mehr durch eine elektronische Variante ersetzt. Diese hat z. B. den Vorteil, dass alle jemals erfassten Daten archiviert und bei Bedarf erneut abgefragt werden können (vgl. Harries (1999), S. 94).

Neben den taktischen Anwendungsfeldern, deren Aufgabe in der unmittelbaren Verbrechensbekämpfung besteht, sind auch die strategisch ausgerichteten kriminalistischen Arbeitsfelder an Methoden des Crime-Mapping interessiert. Viele kriminalgeographische Fragestellungen können erst dann befriedigend beantwortet werden, wenn die Verteilung der Straftaten zu anderen raumbezogenen Daten in Beziehung gesetzt wird, z. B. zu Daten über die Sozialstruktur in einem Stadtteil oder über die Verkehrsinfrastruktur in der Stadt. Aus den Erkenntnissen, die auf diese Weise gewonnen werden, lassen sich Handlungsempfehlungen ableiten. Beide (Erkenntnisse und Hand-

lungsempfehlungen) können anschließend in die gesellschaftspolitische Diskussion eingebracht werden (vgl. Ratcliffe (2004b), S. 76).

2.2 Historische Entwicklung des „Crime Mapping“

Philips (1972) beschreibt in seinem Aufsatz (vgl. Philips (1972), S. 86), dass räumliche Analysen im Zusammenhang mit dem Crime Mapping bereits seit 140 Jahren betrieben werden. Die ersten Untersuchungen fanden zwischen 1830 und 1880 in Frankreich statt. Zu dieser Zeit begann die systematische Erfassung soziodemographischer Daten durch die staatlichen Organe, zu denen u. a. auch die Dokumentation von Straftaten gehörte. Diese wurden in einem Zensus auf regionaler Ebene erfasst und anschließend kartographisch dargestellt und ausgewertet.

Die Ergebnisse dieser Untersuchungen konnten zum ersten Mal Zusammenhänge zwischen der Verbrechensdichte und anderen soziodemographischen Variablen auf regionaler Ebene nachweisen. Guerry (1833) und Quetlet (1840), beide zitiert in Brantingham & Brantingham (1981, S. 8f), stellten z. B. signifikante regionale Unterschiede zwischen Eigentums- und Gewaltverbrechen fest. Während im industrialisierten Norden Frankreichs eine hohe Dichte an Eigentumsverbrechen zu beobachten war, dominierten die Gewaltverbrechen in den ländlich strukturierten Departements im Süden Frankreichs.

Bereits in diesen ersten Untersuchungen wurden nicht nur die räumlichen, sondern auch die zeitlichen Phänomene der Straftatenverteilung untersucht. Guerry (zit. in Philips 1972, S. 87) stellte dabei zu seinem Erstaunen fest, dass in jedem Jahr nahezu dieselbe Anzahl an Straftaten in denselben Regionen erhoben wurde. Spätere Studien in England konnten diese Ergebnisse bestätigen (s. Brantingham & Brantingham (1981), S. 10). Die englischen Studien zeigten darüber hinaus, dass der Grad der räumlichen Aggregation (Bezirk, Ort, Ortsteil) einen erheblichen Einfluss auf die Verteilung der Verbrechenshäufigkeiten ausüben. (Brantingham & Brantingham (1981), S. 12). Diese Untersuchungen waren die ersten, die sich mit dem Problem des Maßstabsbezugs von räumlichen Untersuchungen auseinandersetzten.

Die Chicagoer Soziologenschule

In den 20er und 30er Jahren des letzten Jahrhunderts etablierte sich in Chicago die sog. „Ökologische Schule“. Die Anhänger dieses soziologischen bzw. sozialgeographischen Forschungszweiges untersuchten die Entwicklung der Städte in den USA vor dem Hintergrund der zunehmenden Industrialisierung und der Zuwanderung neuer Bevölkerungsschichten. Um die Entwicklungsprozesse in den amerikanischen Städten der Ostküste zu kategorisieren, bedienten sie sich eines sozialdarwinistischen Gedankenmodells. Einige Vertreter benutzten Kriminalitätsdaten, um sozio-demographische Unterschiede zwischen einzelnen Vierteln aufdecken und erklären zu können. Als das bedeutendste Werk dieses Zweiges nennt Philips (1972, S. 88) „Juvenile Delinquency in Urban Areas“ von Shaw und McKay, das 1942 erschien und 1967 erneut aufgelegt wurde. Die Autoren untersuchten hier die Wohnadressen jugendlicher Straftäter, und stellten Häufungen für bestimmte Stadtviertel fest.

Computergestütztes Crime Mapping

Die ersten Computer-Karten zur Darstellung der Verbrechenshäufigkeiten in einer Stadt wurden in den 60er Jahren des letzten Jahrhunderts in St. Louis erstellt (vgl. Harries, (1999), S. 4). Doch erst seit Mitte der 90er Jahre ist eine flächenhafte Ausbreitung in der Nutzung computergestützter Crime-Mapping-Technologien in der polizeilichen Tätigkeit (in den USA) feststellbar (vgl. Vann & Garson (2003), S. 120ff.). Dies liegt vor allem daran, dass erst zu diesem Zeitpunkt die Technologie soweit entwickelt war, dass ansprechende Karten am PC erzeugt werden konnten. Darüber hinaus waren zu diesem Zeitpunkt auch die Anschaffungskosten für Hard- und Software soweit gesunken, dass diese Technologien nicht mehr ausschließlich als Expertensysteme eingesetzt, sondern einem größeren Nutzerkreis zugänglich gemacht werden konnten (Harries, (1999), S. 6).

Seit dem Ende der 90er Jahre bzw. dem Beginn des 21. Jahrhunderts setzt auch die Polizei in Deutschland vermehrt auf den Einsatz von Crime-Mapping-Anwendungen. Jedoch ist die Nutzung (bisher) auf einige Großstädte (München, Hamburg, Köln) begrenzt (vgl. Schiefer (2002), S. 16f., Pandler (2002), S. 22f.).

2.3 Crime Mapping und GIS

Geographische Informationssysteme (GIS) stellen Software-Pakete dar, die zur Darstellung räumlicher Daten und zur Analyse von räumlichen Beziehungen zwischen diesen Daten entwickelt wurden. Darstellung und Analyse räumlicher Daten sind darauf gerichtet, räumliche Entscheidungen herbeizuführen (vgl. Longley et al. (2001), S. 2).

Auf die räumliche Dimension von Straftaten wurde bereits in den beiden vorhergehenden Kapitel hingewiesen. Es lag also nahe, die Darstellung der räumlichen Verteilung von Straftaten in einem GIS vorzunehmen. Dies geschah zuerst durch die Überleitung der klassischen Steckkarte (Pin-Map) in eine elektronische Variante (vgl. Harries (1999), S. 2). Sehr schnell wuchs die Erkenntnis, dass ein GIS zusätzliche nutzbringende Funktionen für weiterreichende Untersuchungen bereithält. Diese Untersuchungen konnten mit den Standard-Methoden eines GIS – also die Buffer- und die Verschneidungsfunktionen – durchgeführt werden (vgl. Harries (1999), S. 108, Francis et al (2002), S. 166). Der Einsatz dieser Funktionen liefert z. B. Erkenntnisse im Hinblick auf Fragen nach möglichen räumlichen Beziehungen zwischen Straftaten und anderen räumlichen Ereignissen. Auf diese Weise lässt sich z. B. feststellen, wie viele Diebstahl- oder Drogendelikte in der Nähe von Schulen aufgetreten sind.

Ein Beispiel für den erfolgreichen Einsatz von Crime Mapping mit der Hilfe eines GIS im Rahmen der polizeilichen Ermittlungsarbeiten zeigt sich bei der Bekämpfung von Serienstraftaten. So konnte die Polizei in München durch die Nutzung ihres Systems „GLADIS“ einen Serienstraftäter fassen, indem die bisherigen Tatorte untersucht und anhand der Ergebnisse potentielle weitere Ziele abgeleitet werden konnten. Dies brachte den Erfolg, dass der Täter gefasst werden konnte, als er tatsächlich an einem dieser Ziele zuschlug (vgl. Okon & Doerffel (2003), S. 44f).

2.3.1 Räumliche Analysen zur Untersuchung von Straftaten

Um Straftaten in einem GIS darstellen zu können, müssen die Daten geocodiert werden. Die geschieht durch Umwandlung von Adressangaben in Punktkoordinaten. Sobald die Straftaten als geocodierte Punktobjekte vorliegen, lassen sie sich durch den Einsatz räumlich statistischer Methoden weiter untersuchen. Dies setzt jedoch die Verfügbarkeit

geeigneter räumlich-statistischer Analysemethoden voraus. Diese waren lange Zeit nicht in den kommerziellen GIS-Anwendungen verfügbar (vgl. Bailey (1994), S. 13ff.). Es wurden deshalb spezialisierte Anwendungen zur räumlichen Analyse von Straftaten geschrieben. Zur Darstellung der Analyseergebnisse war dann jedoch ein GIS erforderlich. Dies führte dazu, dass verschiedene Versuche zur Integration von räumlichen Analysemethoden für das Crime-Mapping und GIS-Anwendungen unternommen wurden. Insbesondere die GIS-Anbieter oder deren Partnerfirmen boten fachliche Erweiterungen zu ihrer GIS-Software an, mit deren Hilfe spezielle Fragen des Crime-Mapping bearbeitet werden konnten. Dies hatte den Vorteil, dass die Analysefunktionen aus dem GIS heraus aufgerufen und die Ergebnisse unmittelbar im GIS angezeigt wurden (vgl. The Omega Group (2002), Tetrad (2005)). Die zweite Alternative besteht darin, eigenständige Programme zur Durchführung räumlicher Analysen zu entwickeln (vgl. Levine (1999), Block (1995)). Diese Programme müssen so gebaut sein, dass sie Daten von allen gängigen GIS-Formate einlesen und ausspielen können, denn die Visualisierung der Analyseergebnisse muss in einem beliebigen GIS möglich sein. Die Ausweitung des Crime-Mappings um den Einsatz der räumlichen Analysemethoden führte schließlich zu den Einsatzgebieten der „Hot-Spot-Analyse“ (vgl. Harries (1999), S. 112ff., Chainey, et al. (2002), Vann & Garson (2003), S. 32ff.) oder dem „Geographic Profiling“ von Serientätern (vgl. Harries (1999) S. 152, ECRI (2001), Catalano (2000)).

Insbesondere in den USA ist bereits eine große Anzahl an Polizeidistrikten mit Crime-Mapping-Anwendungen ausgestattet. Das [National Institute of Justice \(NIJ\)](#) unterstützt die Weiterentwicklung von Crime-Mapping Methoden und deren Einsatz in den Polizeidirektionen in den gesamten USA. Zu diesem Zweck hat es das Programm [„maps“ \(mapping analysis for public safety\)](#) eingerichtet. Im Rahmen dieses Programms wird z. B. eine jährliche Konferenz zum Thema „Crime Mapping“ organisiert.

Hot-Spot-Analyse

Unter dem Begriff „Hot-Spot-Analyse“ werden Methoden zusammengefasst, mit deren Hilfe räumliche (und/oder zeitliche) Konzentrationen von Straftaten erkannt und dargestellt werden. Hotspots treten dort auf, wo die untersuchten Straftaten räumlich

konzentriert vorliegen. Da jedoch bereits die Ausgangsdaten (Bevölkerung, Parkplätze) und somit die Gelegenheiten räumlich ungleichmäßig verteilt und somit in bestimmten Regionen konzentriert vorliegen, kann nicht automatisch jede Konzentration von Straftaten als Hotspot identifiziert werden (vgl. Harries (1999), S. 112). Ratcliffe & McCullagh (2001, S. 67) sprechen deshalb von der Suche nach Hotspots als der Suche nach den stärksten Konzentrationen von Straftaten.

Auch in einer zeitlichen Dimension wird von Hotspots gesprochen. Damit sind dann i. d. R. die Zeitpunkte/Zeiträume gemeint, an denen überdurchschnittlich viele Straftaten zu verzeichnen sind (vgl. Vann & Garson (2003), S. 34).

Geographic Profiling

Als „Geographic Profiling“ wird ein Ansatz beschrieben, mit dessen Hilfe der räumliche Aktionsradius eines Täters ermittelt werden soll, um z. B. neue potentielle Tatorte zu bestimmen und weitere Verbrechen abzuwenden. In den angelsächsischen Staaten wird diese Form des Crime Mapping bereits bei der Suche nach Serientätern von Sexualverbrechen eingesetzt. Nachdem eine Reihe von Straftaten einem Serienstraftäter zugeordnet werden konnten, versucht man, durch die Lokalisierung der Tatorte den möglichen Wohnort des Täters einzugrenzen (vgl. Catalano (2000), S. 6f.). Anschließend können die Ermittlungsarbeiten gezielt an den verbleibenden Orten fortgesetzt werden, z. B. durch die Anordnung eines Gentests für alle Männer einer bestimmten Altersgruppe in diesem Zielgebiet, wenn es sich bei der untersuchten Straftat um ein Sexualdelikt handelt.

2.3.2 Ausweitung der Technologie in Intra- und Internet

Ein weiterer Meilenstein in der Nutzung des Crime-Mapping wurde durch die Verfügbarkeit der Internet-Technologie und des Internet-GIS gesetzt. Waren bisher ausschließlich Desktop-Anwendungen eingesetzt, die aufgrund ihrer verhältnismäßig hohen Kosten und des erforderlichen Erfahrungsschatzes als Expertensystem galten und deshalb spezialisierten Anwendern vorbehalten blieben, bietet nun das Internet-GIS die Möglichkeit, einen größeren Nutzerkreis unmittelbar auf die Daten zugreifen zu lassen.

Dies kann sowohl intern (innerhalb der Polizei) als auch nach außen hin wirken. Neben den Chancen, alle Beteiligten gleichermaßen an den Daten und den Ergebnissen teilhaben zu lassen, stellt die Einbeziehung von Nicht-Experten höherer Anforderungen an die Darstellung der Daten und die Aufbereitung der Thematik. Diese Aufbereitung gilt umso stärker, je weiter die Daten den ursprünglichen Bezugsrahmen verlassen, denn nun ist nicht mehr zu erwarten, dass das geeignete Fachwissen zur Interpretation der Daten vorhanden ist (vgl. Ratcliffe (2000), S. 6).

In den USA werden die räumlichen Verteilungen der Straftaten bereits im Internet veröffentlicht (Arizona Criminal Justice Commission (2002), S. 15). In Deutschland wird von dieser Möglichkeit bisher kein Gebrauch gemacht.

3 Die Randbedingungen des Crime-Mappings

In der Einleitung wurden bereits die Ziele erwähnt, die für die Einführung von Crime-Mapping-Anwendungen bei der Polizei sprechen. In diesem Kapitel werden die technischen, ablauf- und datenorientierten Anforderungen an ein Crime-Mapping-System angesprochen, die erfüllt sein müssen, bevor Crime-Mapping-Anwendungen die in sie gesteckten Erwartungen erfüllen können.

3.1 Datenerfassung

Die Datenerfassung ist Voraussetzung für die weiterführende Nutzung der Daten im Rahmen der Datenanalyse. Durch Datenerfassungssysteme gelangen die Daten in die Auswertesysteme. Fehler, die in diesem Anfangsstadium gemacht werden, wirken sich in allen anschließenden Bearbeitungsschritten aus. Hier ist deshalb ein exaktes Arbeiten durch die aufnehmenden Personen gefordert.

3.1.1 Besonderheiten bei der Erfassung räumlicher Daten

Straftaten werden an einem Tatort begangen. In der Regel werden die Angaben zu diesem Tatort bei der Aufnahme der Straftaten erfasst. Diese Ortsangabe kann als Adresse oder als andere textförmige Beschreibung des Tatortes existieren. Um den

Tatort in einem GIS darstellen zu können, muss die textförmige Ortsangabe in Koordinaten umgewandelt werden. Aus der Adressangabe wird somit ein Punkt im (2-dimensionalen) Raum.

In vielen Crime-Mapping-Anwendungen wird zu diesem Zweck ein Geo-Coder eingesetzt. Die Georeferenzierung der Adressangabe erfolgt i. A. durch Interpolation auf einem Straßengraphen. Die Adresskoordinate entspricht in diesem Fall nicht dem Grundstück, auf dem die Straftat begangen wurde, sondern einem Punkt auf dem Straßengraphen, der zur Adressfindung herangezogen wurde (vgl. Vann & Garson (2003), S. 27f).

Mit Hilfe dieser Methode können jedoch ausschließlich diejenigen Ortsangaben umgewandelt werden, die zu einer existierenden Adresse (in einer geschlossenen Ortschaft) passen. Ist keine Adresse vorhanden oder wurde die Adresse nicht eindeutig erfasst, können die Daten nicht geocodiert werden. Dies führt dazu, dass nur ein Teil der Straftaten im GIS verfügbar gemacht werden kann (vgl. Henry & Bryan (2000), S. 17). Der verbleibende Teil der Daten muss entweder manuell nacherfasst werden, oder die Daten bleiben unberücksichtigt. Daher wird in den meisten Crime-Mapping-Anwendungen die Forderung nach einer raumbezogenen Auswertung aller Straftaten nicht erfüllt werden können. Ratcliffe (2000, S. 6) weist in diesem Zusammenhang – insbesondere vor dem Hintergrund einer möglichen Veröffentlichung von Crime Maps – darauf hin, dass die Ungenauigkeiten bei der Erfassung der Ortsangaben im GIS die Qualität des Crime-Mapping negativ beeinflussen und – schlimmstenfalls – sogar die Ergebnisse verfälschen.

Vor diesem Hintergrund ist die Frage nach einer Höchstgrenze in Bezug auf den gerade noch zu tolerierenden Datenverlust dringend zu klären. Während Vann & Garson (2003), S. 28) von einer sog. „minimum acceptable Hit-Rate“ – also einer mindestens notwendigen Anzahl richtig georeferenzierter Adressdaten – von 60% sprechen, setzt Ratcliffe (2004a, S. 67) als Ergebnis einer statistischen Untersuchung zu diesem Problem, einen Wert von 85% exakt georeferenzierter Daten als Mindestanforderung für den erfolgversprechenden Einsatz geostatistischer Analysemethoden voraus.

3.1.2 Die Erfassung des Tatzeitpunktes

Ebenso wichtig wie die Betrachtung des Tatortes ist die möglichst genaue Erfassung des Tatzeitpunktes für die nachfolgenden Untersuchungen. Die Angabe des Tatzeitpunktes ist jedoch mit größeren Unsicherheiten verbunden als die Angabe des Tatortes, da sie – abhängig von der Art der Straftat – wenige Minuten bis mehrere Wochen vor der Anzeigenaufnahme geschehen sein können. Während der Zeitpunkt, an dem die Polizei Kenntnis von dem Vorfall erhält, relativ nah am eigentlichen Tatzeitpunkt liegt, ist bei Eigentumsdelikten eher davon auszugehen, dass der Zeitpunkt der Anzeige und der eigentliche Tatzeitpunkt nicht identisch sind, sondern einigen Stunden bis zu einigen Wochen auseinander liegen können (Ratcliffe (2002), S. 25).

Erfasst die Polizei in diesen Fällen ausschließlich den Zeitpunkt, an dem sie Kenntnis von dem Vorfall erhielt, dann wird eine spätere Analyse der zeitlichen Verteilung dieser Straftaten einseitig überlagert vom Anzeigeverhalten der Geschädigten (vgl. Ratcliffe (2002), S. 31). Viele Polizeibehörden sind sich dieses Problems bewusst und erfassen beim Aufnehmen der Anzeige zwei Zeitpunkte: den Zeitpunkt, an dem der Geschädigte das Objekt zuletzt unversehrt verlassen hat und den Zeitpunkt, an dem das Objekt geschädigt oder nicht mehr vorgefunden wurde. Aus diesen beiden Zeitpunkten ließe sich nun mit Hilfe statistischer Methoden ein wahrscheinlicher Tatzeitpunkt ermitteln (vgl. Ratcliffe (2002), S. 25).

3.2 Datenauswertung

Die Möglichkeiten der Datenauswertung in einer Crime-Mapping-Anwendungen hängen von den Daten ab, die im System verfügbar sind und von den Funktionen, die die Anwendung bereithält. Beide sollen nachfolgend betrachtet werden.

3.2.1 Verfügbare Daten in Crime-Mapping-Anwendungen

Die erfassten Straftaten lassen sich auf vielfältige Arten untersuchen. Um jedoch die Untersuchungsergebnisse sinnvoll interpretieren und die zu können, ist es erforderlich, die Ergebnisse zu anderen räumlichen Daten in Beziehung zu setzen. Welche Daten dies im einzelnen sein sollen, hängt dabei ab von den zu prüfenden Hypothesen und der zu

untersuchenden Straftat. Als Beispiel für solche Daten seien Objekte erwähnt, die als Ziel oder als Quelle von Straftaten bekannt sind. Hierbei könnte es sich z. B. um Schulen, Bahnhöfe, Parkplätze, wichtige Verkehrswege und weitere handeln. Darüber hinaus können weitere soziodemographische Daten, die die z. B. die Sozialstruktur von Stadtteilen charakterisieren, in das System eingebunden werden. Durch das Vorhandensein solcher Daten könnte z. B. ein raumbezogenes Risikopotential für bestimmte Stadtteile und bestimmte Straftaten abgeleitet werden. (vgl. Vann & Garson (2003), S. 29).

3.2.2 Die analytischen Möglichkeiten der Anwendungen und die Kenntnisse der Bearbeiter

Als zweite limitierende Größe im Hinblick auf die Analysemöglichkeiten einer Crime-Mapping-Anwendung ist die methodische Ausstattung der Systeme zu nennen. Je anspruchsvoller die Aussagen werden, die mit Hilfe dieser Anwendung erzielt werden sollen, desto differenzierter muss das Angebot der bereitgestellten Methoden sein. Diese können in die Crime-Mapping-Anwendung integriert sein (vgl. Williamson et al. (1999)) oder parallel zu diesem System eingesetzt werden (vgl. Levine (1999)). Im zweiten Fall, muss jedoch gewährleistet sein, dass die Ergebnisse der Berechnungen anschließend in einem GIS dargestellt werden können, sofern sich daraus neue geographische Bezüge ergeben.

Die methodisch-analytischen Möglichkeiten des Systems sollten den Anwendungsfeldern und den statistisch-analytischen Fähigkeiten der Anwender angepasst sein (vgl. Ratcliffe (2004b)).

3.3 Neue Herausforderungen für Crime-Mapping-Systeme

Eine Schwierigkeit des Crime-Mapping besteht darin, dass durch die digitale Datenerfassung immer neue und immer mehr Daten zur Verfügung stehen. Die Datenflut und die damit einhergehende Menge an Informationen (z. B. die Darstellung aller Tatorte als einzelne Punkte in einer Karte) führt dazu, dass der Betrachter sich in der Informationsflut verliert und die wesentlichen Aussagen der Darstellung nicht mehr

erkennen kann (vgl. Vann & Garson (2003), S. 82). Vor diesem Hintergrund sind neue Verfahren gefordert, die aus den Datenmengen Muster herauslesen und diese präsentieren können.

3.3.1 Die gezielte Suche nach raumzeitlichen Mustern

Ein GIS ist dazu geeignet, räumliche Ereignisse im 2- (2,5-) dimensionalen Raum darzustellen. Das Auftreten von Straftaten ist jedoch nicht nur in räumlichen, sondern auch zeitlichen Dimensionen zu untersuchen. Der Zusammenhang zwischen Raum und Zeit als Ursachefaktoren für die Entstehung von Straftaten lässt sich bisher in einem GIS schwer abbilden.

Da jedoch erst durch die gemeinsame Betrachtung der räumlichen und der zeitlichen Verteilung das Muster der Straftaten hinreichend durchleuchtet werden kann, ist es erforderlich, den zeitlichen Faktor in der Darstellung der Daten im GIS stärker als bisher zu berücksichtigen. Als eine Möglichkeit, die raumzeitliche Variabilität in der Verteilung von Straftaten darzustellen, schlagen Openshaw et al. ((1994), S. 131f.) die Verwendung von Animationen zur Darstellung von Zeitreihen vor. Um die zeitlichen Variabilitäten in der Straftatenverteilung adäquat beurteilen zu können, sollten jedoch die im Kapitel 3.1.2 genannten Besonderheiten in Bezug auf die Erfassung von Tatzeitpunkten berücksichtigt werden.

3.3.2 Berücksichtigung der Dunkelfeld-Problematik

Ein weiteres Problem, das sich für die Auswertung von Straftaten stellt, besteht in der Tatsache, dass die Polizeistatistiken, die für die Auswertung der Daten herangezogen werden können, nur den Teil der Straftaten abbilden, von denen die Polizei Kenntnis erhalten hat. Die Summe derjenigen Straftaten, die sich im Verborgenen ereignen, weil sie nicht angezeigt werden und auch nicht durch polizeiliche Ermittlungsarbeiten entdeckt werden, wird als Dunkelfeld (engl. „dark figure“) bezeichnet (vgl. Ratcliffe (2000), S. 6, Schwind (2003), S. 31). Dieses Dunkelfeld ist abhängig von der Art der Straftat mehr oder weniger stark ausgeprägt.

Viele kriminalistische Untersuchungen befassen sich mit diesem Phänomen (vgl. Schwind (2003), S. 31 ff). Es spricht einiges dafür, dass auch das Dunkelfeld einer bestimmten räumlichen Verteilung unterliegt, wobei sich jedoch nicht sagen lässt, ob die räumliche Verteilung des Dunkelfeldes der räumlichen Verteilung der Hellfelddaten (das sind diejenigen Daten, die in die polizeiliche Statistik eingehen) entspricht oder vollkommen anders aussieht.

Würde die räumliche Verteilung der Dunkelfelddaten sich vollkommen anders darstellen als die Daten des Hellfeldes, dann könnte eine Suche nach Schwerpunkten immer nur diejenigen Bereiche erfassen, in denen ein ausgeprägtes Anzeigeverhalten – aber nicht unbedingt ein hohes Straftatenaufkommen – zu beobachten ist. Also sind die Ergebnisse jeder (explorativen) Straftatenanalyse immer vor dem Hintergrund der Hellfeld/Dunkelfeldproblematik zu betrachten. Untersuchungen zum Anzeigeverhalten verschiedener Bevölkerungsschichten und bezogen auf verschiedene Straftaten sind heranzuziehen, um das Dunkelfeld in die Ergebnisse der Hellfeld-Analysen einfließen zu lassen.

4 Datenanalyse und Ergebnisdarstellung im Crime-Mapping

Ausgehend von der bereits mehrfach erwähnten Steckkarte werden statistische Analysemethoden auf die Straftatendarstellung angewendet. Viele der zum Einsatz kommenden Methoden entstammen dem Anwendungsfeld der „Point-Pattern-Analyse“. Die Methoden der „Point-Pattern-Analyse“ (PPA) wurden von Pflanzenökologien entwickelt. In der Zwischenzeit sind diese Methoden jedoch in viele anderer Forschungsfelder übernommen und weiterentwickelt worden (vgl. Bailey & Gatrell (1995), S. 75).

Alle Methoden der PPA setzen voraus, dass eine diskrete Punkteverteilung vorliegt. Von einer diskreten Punktverteilung spricht man immer dann, wenn das untersuchte Phänomen nur an bestimmten Punkten im Raum festgestellt werden kann. Nach Bailey & Gatrell (1995, S. 76) stellt ein solches Punktmuster ein einfaches Beispiel

räumlicher Daten dar, da die Auswertung dieser Daten sich (in den meisten Fällen) auf die Koordinaten der Punkte beschränken.

Das Ziel der Point-Pattern-Analyse besteht in der Beantwortung der Frage, ob das vorliegende Punktmuster einer zufälligen Verteilung, einer Konzentration oder einer gleichmäßigen Verteilung der Punkte im Raum folgt. Die nachfolgenden Kapitel beschreiben verschiedene Verfahren, mit denen sich die Punkteverteilung visualisieren lässt.

4.1 Straftaten als diskrete Punkteverteilung

Jede Straftat wird durch einen Punkt in einer Karte vermerkt. Das Ergebnis dieser Darstellung stellt die sog. „Pin-Map“ dar. Die Kategorie der Straftat kann durch die Vergabe von Symbolen verdeutlicht werden. Dieses Verfahren ist simpel. Der Betrachter einer Pin-Map muss jedoch mit folgenden Schwierigkeiten zurecht kommen.

1. Je länger der betrachtete Zeitraum und je größer das zu untersuchende Gebiet, desto schwieriger wird es, ein Verteilungsmuster zu erkennen, da die Vielzahl der Ereignisse die Übersichtlichkeit der Anwendung beeinträchtigt.
2. Jeder Punkt wird einzeln dargestellt. Fallen mehrere Punkte an einer Koordinate zusammen, so werden diese übereinander geschichtet. Beim Betrachten der Karte kann in diesem Fall niemand erkennen, ob an einem Ort erst eine oder bereits mehrere Straftaten aufgetreten sind – Hotspots können also nicht erkannt werden.

Die unter Punkt 1 genannte Schwierigkeit kann dadurch umgangen werden, dass nur Teilausschnitte eines Raumes und ein definierter Zeitraum betrachtet werden. Die Verantwortlichen in einer Polizeidienststelle werden sich in der Regel für ihren Verantwortungsbereich interessieren und ggfs. in die angrenzenden Reviere hineinschauen. Auch der zeitliche Rahmen ist in den meisten Fällen auf die aktuelle Woche, bzw. den aktuellen Monat einzugrenzen.

Für das unter Punkt 2 genannte Problem lässt sich eine Lösung finden, indem die Größe der dargestellten Punktsymbole proportional zur Anzahl der dort anzutreffenden

Straftaten zunimmt. Ein Vergleich der beiden Darstellungsformen ist der nachfolgenden Abbildung 2 zu entnehmen.

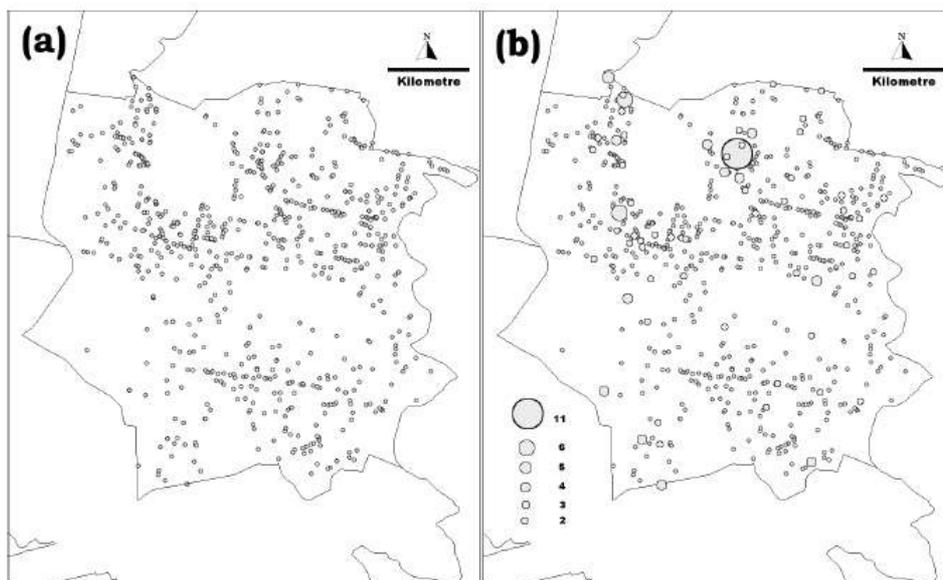


Abbildung 2: Untersuchungsgebiet mit 738 Tatorten unterschiedlich dargestellt (a) als einfache Punktdarstellung, (b) klassifiziert nach Häufigkeiten.
(Quelle: Ratcliffe (2000), S. 2)

Weitere Schwierigkeiten bei der Betrachtung einer Pin-Map zeigen sich gerade vor dem Hintergrund der Darstellungen in der Abbildung 2 in Bezug auf die Interpretation dieser Punktdarstellung. Die großen Punktsymbole zeigen einerseits örtliche Konzentrationen von Straftaten auf und weisen auf solche Orte hin, an denen das Problem wiederholter Straftaten auftaucht. Durch ihre Größe dominieren sie die Darstellung und lenken das Auge des Betrachters auf diese Bereiche in der Karte. Dies kann dazu führen, dass andere Bereiche übersehen werden, in denen viele einzelne Straftaten auf engem Raum vorkommen. Weiterhin besteht die Gefahr, dass die großen Punktsymbole andere Punkte, die weitere Tatorte kennzeichnen, verdecken könnten. Diese Gründe führen zu der Überlegung, die Straftatereignisse in einer Fläche darzustellen (vgl. Ratcliffe (2000, S. 2)).

Obwohl die punktbezogene Darstellung der Straftaten schwer zu interpretieren ist, bildet die Georeferenzierung der Straftaten als Punktobjekte die Voraussetzung für alle im weiteren Verlauf dieses Kapitels vorzustellenden Verfahren.

4.2 Die Darstellung der Straftaten in einer Fläche

Die Darstellung der Straftaten in einer Fläche kann auf unterschiedliche Arten gelingen. Die Punkte werden auf diskrete Flächen bezogen, oder es werden Methoden eingesetzt, die aus den Punktdaten eine kontinuierliche Oberfläche generieren. Nachfolgend werden beide Methoden erläutert. Dabei wird zuerst auf die Aggregation der Daten in diskrete Flächen eingegangen. Im anschließenden Kapitel werden einige Methoden vorgestellt, die zur Erzeugung kontinuierlicher Oberflächen eingesetzt werden. Als Ergebnis dieser Datenaggregation wird die Ereignisdichte, also die Anzahl der Ereignisse pro Flächeneinheit, beschrieben.

4.2.1 Aggregation der Punktdaten zur Darstellung in diskreten Flächen

Zur Aggregation von Punktdaten sind (diskrete) Bezugsflächen erforderlich. Die Straftaten, die sich innerhalb der jeweiligen Flächen befinden, werden ausgezählt und als Attribut der Flächen gespeichert. In der Regel werden die Straftaten anschließend in eine relative Beziehung entweder zur Flächengröße oder zur Bevölkerungszahl (wenn diese für die besagten Flächen vorliegt) gebracht. Nun erst ist ein Vergleich der Flächenwerte möglich, da sie eine gemeinsame Basis erhalten haben.

4.2.1.1 Aggregation auf vorhandene räumliche Einheiten

Die räumliche Darstellung soziodemographischer Daten geschieht fast ausschließlich durch Aggregation auf bestehende administrative Einheiten. Eine Aggregation der Straftaten sollte jedoch die polizeilich relevanten Flächen (Reviere, Direktionen, etc.) vorziehen (vgl. Hirschfeld, et al. (2001), S. 332f.).

Nachdem die Daten auf die Flächen aggregiert wurden, erfolgt die Darstellung der Straftatendichte in Form einer Choropletenkarte. Jede Teilfläche wird mit einem Wert (z. B. die Straftatendichte) belegt. Anschließend wird eine Klassifizierung vorgenommen. Durch die Klassifizierung wird erreicht, dass nicht jeder Einzelwert dargestellt werden muss und das Gesamtbild leichter zu interpretieren sein wird, denn die Darstellung in der Karte wird übersichtlicher, wenn wenige Wertegruppen zu berücksichtigen sind. Zur Abgrenzung der Wertegruppen werden Farbskalen eingesetzt, die einerseits

die Unterschiede zwischen den Klassen hervorheben und andererseits die Interpretation der Aussage unterstützen (vgl. Hake et al. (2002), S. 473).

Beispiele für Choropletenkarten werden im zweiten Teil dieser Arbeit vorgestellt, so dass ein Verweis auf die nachfolgenden Kapitel (Teil II, Kapitel 4.2.1) an dieser Stelle ausreichen soll.

4.2.1.2 Aggregation mit Hilfe der Quadratmethode

Die Quadratmethoden stellt ein häufig verwendetes Mittel der Point-Pattern-Analyse zur Auswertung der räumlichen Verteilung von Punkten dar. Sie arbeitet folgendermaßen. Über das Untersuchungsgebiet wird ein Raster definierter Größe gelegt. Obwohl der Name dieses Ansatzes – Quadratmethode – suggeriert, dass ein quadratisches Raster zu verwenden ist, können auch andere geometrische Formen, die ein gleichmäßiges Raster erzeugen (z. B. Dreiecke oder Hexagone), eingesetzt werden. Die meisten Raster bestehen jedoch aus einer quadratischen Grundform. Dies ist der Tatsache geschuldet, dass sich quadratische Raster leichter erstellen lassen als die anderen Formen, obwohl diese durchaus besser geeignet sein könnten (vgl. O'Sullivan & Unwin (2003), S. 83f., Hirschfeld et al. (2001), S. 333f.). Die Anzahl der Straftaten wird für jede Rasterzellen ermittelt. Sie kann anschließend als Choropletenkarte abgebildet werden.

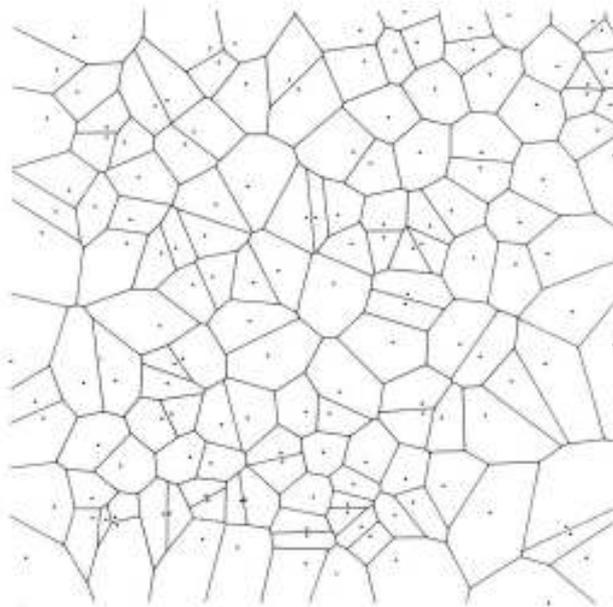
Darüber hinaus bietet die Quadratmethode jedoch auch die Möglichkeit, die räumliche Verteilung der untersuchten Ereignisse gegen die Annahme einer räumlichen Zufallsverteilung (Complete Spatial Randomness, CSR) zu testen (s. O'Sullivan & Unwin (2003) S. 98f.).

Auch Beispiele für die Quadratmethode werden im zweiten Teil dieser Arbeit vorgestellt (vgl. Teil II, Kapitel 4.2.1.1).

4.2.1.3 Aggregation mit Hilfe von Thiessen-Polygonen

Während die Quadrat-Methode die Gesamtfläche in gleichmäßige Teilflächen aufteilt, beschreibt dieses Kapitel eine weitere Möglichkeit zur Unterteilung der Gesamtfläche – nun jedoch in ungleichmäßige Teilflächen. Jedoch folgt auch hier die Erzeugung der

Teilflächen einem vorgegebenen Schema. Die Größe und die Gestalt jeder Teilfläche ist durch die Verteilung der zu untersuchenden Punktereignisse vorgegeben: In einem 2-dimensionalen euklidischen Raum lassen sich aus der räumlichen Verteilung der Beobachtungspunkte Flächen ableiten. Jede Fläche umschließt genau einen Beobachtungspunkt und beschreibt auf diese Weise das Einflussgebiet dieses Punktes, denn innerhalb dieser Fläche ist der Beobachtungspunkt der einzige nächste Nachbar bezogen auf den gesamten Datensatz. Nur die Grenzen zwischen zwei Flächen haben zwei Beobachtungspunkte als nächste Nachbarn (vgl. Okabe, et al. (1992), S. 66). Diese Flächen werden als „Thiessen-Polygone“ oder „Voronoi-Diagramme“ bezeichnet. Das Verfahren zur Erzeugung dieser Flächen wird „Dirichlet-Tessellation“ genannt (vgl. Upton & Fingleton (1985), S. 96). Die Abbildung 3 stellt typische Thiessen-Polygone vor.



*Abbildung 3: Ausschnitt aus einem planaren Thiessen -Polygon
(Quelle: Okabe, et al. (1992), S. 274)*

Allgemein wird für die Erzeugung der Thiessen-Polygone erwartet, dass jeder Flächenbezugspunkt nur einmal in dem Datensatz vorkommt. Punkte, die doppelt vorliegen, müssen deshalb vor der Erstellung der Thiessen-Polygone aus dem Datensatz entfernt werden. Nach der Erstellung der Flächen können diese Punkte dann dem Datensatz wieder zugeführt werden. Auch für die Thiessen-Polygone ließe sich dann die Punktdichte pro (Teil-) Fläche berechnen und als Choropletenkarte ausgeben. Die

Thiessen-Polygone machen darüber hinaus die räumliche Verteilung der Punkte erkennbar. Kleine Flächen, die dicht gedrängt liegen, zeigen eine räumliche Konzentration der beobachteten Ereignisse. Große Thiessen-Polygone weisen auf eine disperse Verteilung der Ereignisse hin.

Während die Thiessen-Polygone diejenigen Flächen beschreiben, die sich um eine Punkteverteilung herum anordnen, erstellt das Verfahren der Delaunay-Triangulation Dreiecksverbindungen zwischen den Mittelpunkten der Thiessen-Polygone. Die Linien, die dabei von dem Bezugspunkt eines Thiessen-Polygons ausstrahlen, schneiden die Kanten der Thiessen-Polygone genau in deren Mittelpunkt. Die Thiessen-Polygone stellen also den dualen Graphen der Delaunay-Triangulation dar (vgl. Upton & Fingleton (1985), S. 101). Beide Graphen können ineinander umgewandelt werden.

4.2.2 Das „Modifiable Areal Unit Problem“ als Konsequenz räumlicher Aggregation

Räumliche Aggregationen sehen vor, dass die Ergebnisse einer Beobachtung oder empirischen Untersuchung auf eine Fläche bezogen werden. Dieser Bezug wird dabei sehr oft über Flächen hergestellt, die nicht in einer objektiven Beziehung zu den untersuchten räumlichen Prozessen stehen. Dies ist z. B. häufig der Fall, wenn soziale Daten auf administrative Einheiten (Gemeinden oder Wahlbezirke) bezogen werden.

Die Größe, bzw. die räumliche Ausrichtung der Flächen spielt eine entscheidende Rolle in Bezug auf die Interpretation der Ergebnisse. Dies geschieht insbesondere dann, wenn Beziehungen zwischen verschiedenen Daten über diese Aggregationen hergestellt werden sollen. Dieses Phänomen wurde zuerst von Openshaw (1984), zitiert in Longley et al (2001), S. 138, beschrieben und wirkt sich insbesondere dann aus, wenn sich in vergleichenden Untersuchungen die Flächenbezüge ändern oder wenn Daten verglichen werden, die auf unterschiedlichen Flächen aggregiert wurden (vgl. Abbildung 4).

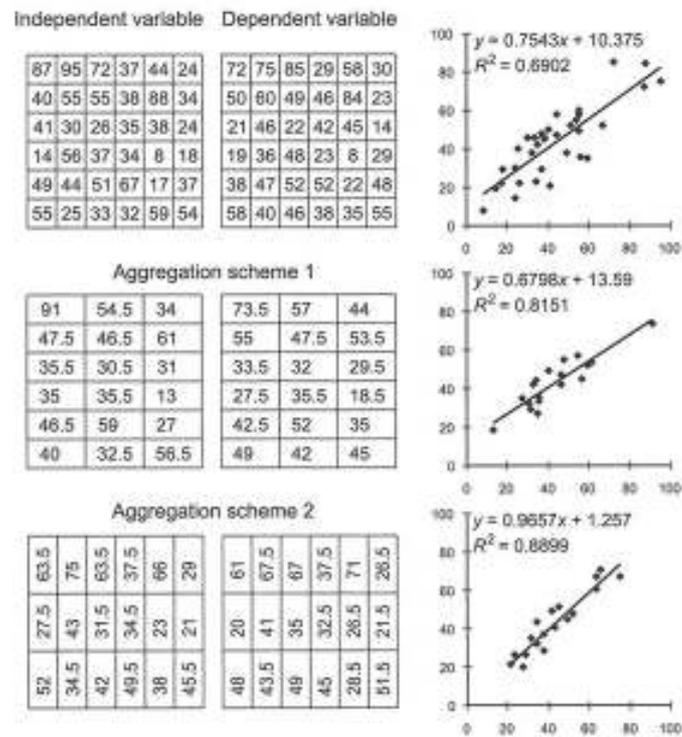


Abbildung 4: Auswirkungen des MAUP (Quelle: O'Sullivan & Unwin (2003), S. 31)

4.3 Die Straftatendichte als kontinuierliche Oberfläche

Da Straftaten – ebenso wie die Verteilung von Bäumen im Wald oder die Verteilung von Vogelnestern in Bäumen – nur an definierten Orten im Raum zu beobachten sind, bedarf es spezieller Verfahren, um die räumliche Verteilung der Straftaten als kontinuierliche flächenhafte Verteilung darzustellen. Ein geeignetes und häufig eingesetztes Verfahren ist mit der Methode der Kernel-Density Estimation (KDE) gegeben. Auch dieses Verfahren stammt aus dem Umfeld der Point-Pattern-Analyse und ist inzwischen in viele GI-Systeme integriert. Jedoch mangelt es dort vielfach an einer ausreichenden Dokumentation des Verfahrens (O'Sullivan & Unwin (2003), S. 88). Die Funktionsweise der KDE zeigt die Abbildung 5 in Form eines eindimensionalen Schnittes durch einen Kernel.

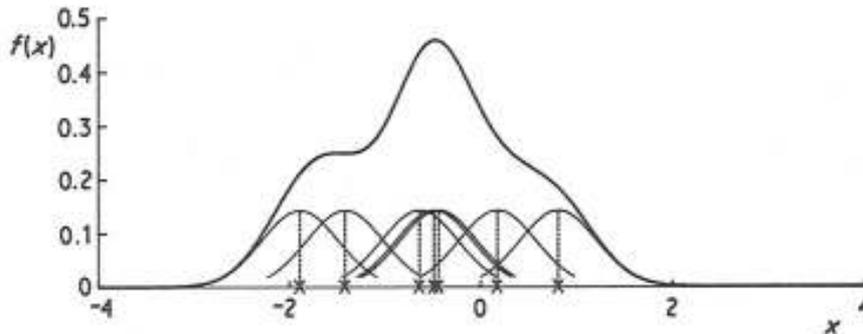


Abbildung 5: Arbeitsweise der KDE in eindimensionaler Darstellung.
(Quelle: Silverman (1986), S. 14)

Die Kernel-Density-Estimation (KDE) wirkt als Glättungsfunktion und stellt eine Weiterentwicklung der Quadrat-Methode dar. Die Idee hinter dieser Methode geht davon aus, dass die (theoretische) Ereignisdichte für jeden beliebigen Punkt im Untersuchungsgebiet aus dem vorliegenden Punktmuster abgeleitet werden kann (O'Sullivan & Unwin (2003), S. 85). Über das Untersuchungsgebiet wird ein Raster gelegt. Jeder Punkt des Rasters wird aufgesucht. Durch die Anwendung einer geeigneten KDE-Funktion wird dann die Punktdichte für jeden einzelnen Rasterpunkt bestimmt. Eine der am häufigsten verwendeten Kernel-Funktionen ist der sog. „Quartic Kernel“, dessen Formel folgendermaßen aussieht:

$$\lambda_{\tau}(s) = \sum_{h_i \leq \tau} \frac{3}{\pi \tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2 \quad (1)$$

Der Parameter h_i beschreibt die Distanz zwischen dem aktuellen Punkt s und dem Ort des Ereignisses s_i . Die Summierung der Ereignisse wird ausschließlich für die Werte von h_i durchgeführt, die sich innerhalb des Suchradius τ befinden. Der Suchradius bestimmt also die Größe des Gebietes, in welchem die Ereignisse liegen müssen, die zur Berechnung von $\lambda_{\tau}(s)$ herangezogen werden. Je größer τ gewählt wird, desto gleichmäßiger wird die geschätzte Dichteverteilung des beobachteten Ereignisses. Umgekehrt wird mit kleinerem τ die Dichteverteilung der untersuchten Ereignisse als extrem unruhige Oberfläche abgebildet. Dieses Phänomen zeigt die Abbildung 6.

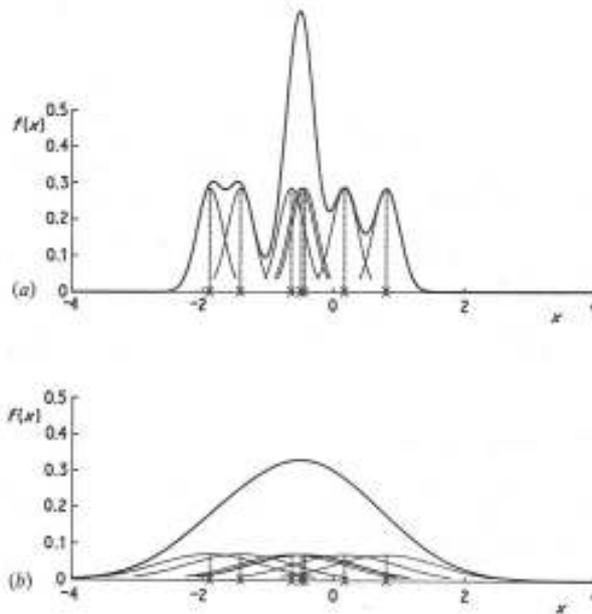


Abbildung 6: Der Einfluss des Parameters τ .
(Quelle: Silverman (1986), S. 14)

4.4 Weitere Methoden aus dem Umfeld der Point-Pattern-Analyse

4.4.1 Eigenschaften eines Punktmusters

Die bisher beschriebenen Methoden untersuchen die Verteilung der Punkte, indem sie die sog. „first-order“ Eigenschaften des Punktmusters herausstellen. Diese „first-order“ Eigenschaften des Punktmusters beschreiben die „Intensität“ ($\lambda(\mathbf{s})$) des Prozesses, der zu diesem Punktmuster führt. Unter dem Begriff „Intensität“ ist die durchschnittliche Dichte der Ereignisse pro Flächeneinheit an jedem Punkt \mathbf{s} im untersuchten Raum zu verstehen (Bailey & Gatrell (1995), S. 77). Betrachtet man die durchschnittliche Dichte der Ereignisse pro Flächeneinheit, so wird man feststellen, dass die räumliche Dichteverteilung nicht zufällig erfolgt, sondern einen Trend erkennen lässt. Dieser Trend führt entweder zu einer gleichmäßigen Verteilung oder zu einer räumlichen Konzentration der Ereignisse. Diese Tendenzen der Dichteverteilung weisen darauf hin, dass die räumliche Verteilung des untersuchten Ereignisses von anderen Einfluss-

faktoren abhängig ist, die in der aktuellen Untersuchung jedoch nicht berücksichtigt wurden. So kann z. B. die räumliche Variabilität des Straftatenaufkommens in einer Stadt auf Unterschiede in der Bevölkerungsverteilung zurückzuführen sein (O'Sullivan & Unwin (2003), S. 29).

Die Verteilung des Punktmusters kann jedoch auch durch andere Ursachen beeinflusst werden. Diese Einflüssen stellen die sog. „second-order“ Eigenschaften dar. Als „second-order“ Eigenschaften werden Beziehungen zwischen Ereignispaaren bezeichnet. So ist davon auszugehen, dass benachbarte Ereignisse sich gegenseitig beeinflussen. Diese Einflüsse sind abhängig von der Distanz (und u. U. auch der Richtung), die zwischen diesen beiden Punktereignissen liegt (Bailey & Gatrell (1995), S. 77). Vor diesem Hintergrund ist z. B. das Auftreten weiterer Straftaten in der Nähe von Orten, an denen bereits andere Straftaten erfolgreich ausgeübt wurden, als wahrscheinlich anzunehmen. Dies kann z. B. in Gegenden, die als Hot-Spot Areas bekannt sind oder dort, wo Drogen umgeschlagen werden (O'Sullivan & Unwin (2003), S. 29), beobachtet werden.

Da in der Regel beide Eigenschaften an der Ausprägung des Punktmusters beteiligt sind, werden nun einige Methoden der Point-Pattern-Analyse vorgestellt, die die „second-order-Eigenschaften“ untersuchen.

4.4.2 Methoden zur Ermittlung der Second-Order-Effekte

Der wichtigste Parameter, der in die Berechnung der Second-Order-Effekte eingeht, ist die Berechnung der Distanz zwischen den jeweiligen Punktereignissen. An dieser Stelle sollen zwei Verfahren vorgestellt werden, die bereits in verschiedenen Crime-Mapping-Anwendungen integriert sind (vgl. Levine (1999), Block, (1995)). Mit Hilfe dieser Methoden wird versucht, räumliche Cluster in den Datensätzen zu entdecken. Von einem „Cluster“ spricht man dann, wenn in einem Gebiet eine stärkere Ereigniskonzentration als im Gebietsdurchschnitt festgestellt wird.

Die nachfolgend vorgestellten Methoden geben lediglich einen globalen Eindruck über die Verteilung der Punkte. Dies tun sie, indem sie einen Wert ausgeben, der den gesamten Datensatz beschreibt. Dieser Wert lässt sich nicht unmittelbar in der

kartographischen Visualisierung nutzen. Aus diesem Grund werden die Methoden nur kurz angesprochen.

4.4.2.1 Die Nearest-Neighbour-Analyse

Im Rahmen der Nearest-Neighbour-Analyse wird eine Distanzmatrix erstellt, die die Entfernungen von jedem Punkt zu allen übrigen Punkten des Datensatzes abbildet. Anschließend wird über alle ermittelten Distanzen der Mittelwert berechnet. Diese sog. „Mean Nearest Neighbour Distance“ geht in die Berechnung eines Index-Wertes ein, der Auskunft gibt über die räumliche Verteilung der Daten in diesem Datensatz. Die Berechnung des Indexwertes erfolgt nach folgender Formel:

$$R = 2\bar{d} \sqrt{n/a} \quad (2)$$

\bar{d} (quer) - durchschnittliche Nearest-Neighbour-Distanz
 n - Anzahl der untersuchten Punktobjekte
 a - Flächengröße des untersuchten Gebietes

Der R-Wert gibt Auskunft über die räumliche Verteilung des gesamten Prozesses. Die Interpretation dieses Wertes geschieht folgendermaßen. Liegen alle Punkte in einer definierten Ebene übereinander, dann ist keine Entfernung zwischen den Punkten zu ermitteln (\bar{d} (quer) = 0) und somit wird auch $R = 0$. Ein Nearest-Neighbour-Wert um 0 deutet demnach darauf hin, dass die Punktereignisse in dem untersuchten Gebiet gehäuft vorkommen, also Cluster vorliegen. Liegt keine räumliche Beziehung zwischen den Punkten vor, dann ist ihr Muster zufällig verteilt und der R-Wert = 1. Wenn jedoch die einzelnen Punkte soweit wie (physikalisch) möglich voneinander entfernt liegen und daraus eine gleichmäßige Punktverteilung resultiert, dann konvergiert R gegen einen Wert von 2,15. Demnach beschreibt ein R-Wert von 2,15 eine gleichmäßige Punktverteilung im untersuchten Gebiet (vgl. Griffith & Amrhein (1991) S. 127).

Da jedoch ausschließlich die Distanzen zwischen den Punkten berücksichtigt werden, können unterschiedliche Punktverteilungen zu demselben räumlichen Mittelwert führen. Dieser Verlust an Informationen über die Verteilung der Punkte im Raum ist einer der Haupt-Kritikpunkte an diesem Verfahren (vgl. (O'Sullivan & Unwin (2003), S. 89).

4.4.2.2 Die K-Funktion

Die K-Funktion beschreibt den Grad der Abhängigkeit in der räumlichen Verteilung der beobachteten Ereignisse (Gatrell et al. (1996), S. 261). Im Gegensatz zur Nearest-Neighbour-Analyse werden hier die Beziehungen zwischen den beobachteten Punkten und ihrer jeweiligen Nachbarnpunkte in verschiedenen Entfernungen ermittelt. Die Definition der K-Funktion stellt sich nach Bailey & Gatrell (1995, S. 92) folgendermaßen dar.

$$\lambda K(h) = E(\#(\text{Ereignisse innerhalb der Entfernung } h \text{ zum aktuellen Ereignis})),$$

wobei # die „Menge der ...“ bedeutet, $E()$ den Operator der Normalverteilung und λ die durchschnittliche Anzahl (Intensität) der Ereignisse pro Flächeneinheit beschreibt, die über die gesamte Fläche (\mathcal{R}) des Untersuchungsgebietes als konstant angenommen wird. Wie die Berechnung der K-Funktion konkret durchgeführt wird, zeigt die nachfolgende Abbildung 7.

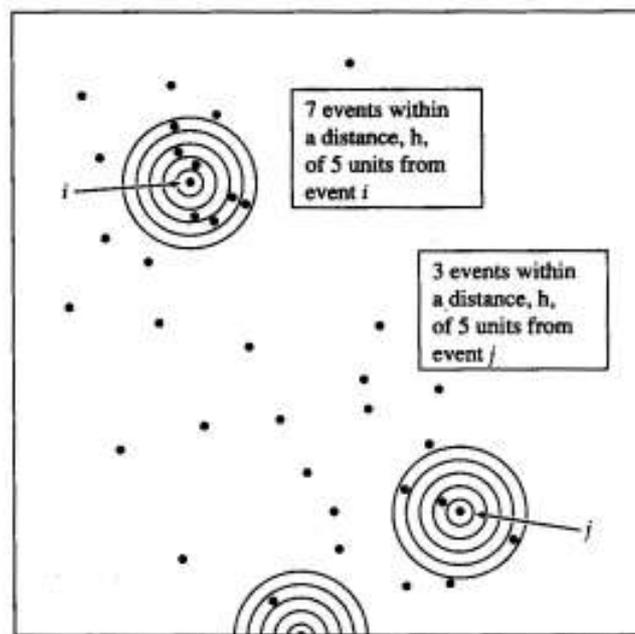


Abbildung 7: Schätzung der K-Funktion
(Quelle: Bailey & Gatrell (1995), S. 93)

4.4.3 Methoden zur Berechnung der räumlichen Autokorrelation

Räumliche und zeitliche Daten unterscheiden sich von Daten der klassischen Statistik aufgrund besonderer Eigenschaften. Diese Eigenschaften führen dazu, dass die Forderung nach der statistischen Unabhängigkeit der erhobenen Daten (und somit einer Gleichverteilung der Wahrscheinlichkeit für das Eintreten eines bestimmten Ereignisses) sowohl für raum- als auch zeitbezogenen Daten nicht gegeben ist. Werden z. B. in bestimmten Zeitintervallen Beobachtungen gemacht, so ist der zum Zeitpunkt z beobachtete Wert eines Ereignisses abhängig von dem Wert, den das Ereignis zum Zeitpunkt $z-1$ angenommen hatte. Die zum Zeitpunkt z ermittelte Ausprägung des Ereignisses beeinflusst wiederum den Wert, der zum nachfolgenden Zeitpunkt ($z + 1$) gemessen wird. Während die Abhängigkeiten auf der Zeitachse gerichtet sind (nachfolgende Ereignisse werden von den Vorgängern beeinflusst, sie können ihrerseits die Vorgänger nicht beeinflussen), so sind die räumlichen Abhängigkeiten in alle Richtungen des zwei- bzw. dreidimensionalen Raumes möglich (vgl. Longley et al. (2001), S. 99).

Das Phänomen, das diese Beeinflussung beschreibt, wird Autokorrelation genannt. Autokorrelation tritt in drei Variationen auf (vgl. O'Sullivan & Unwin (2003), S. 29; Haining (1990), S. 229). Diese werden nachfolgend aufgelistet und in der Abbildung 8 dargestellt.

1. **Positive Autokorrelation** ist der am häufigsten zu beobachtende Fall. Sie beschreibt solche Situationen, in denen die Werte benachbarter Beobachtungen ähnliche Ausprägungen aufweisen
2. **Negative Autokorrelation** ist weniger häufig, aber doch vorhanden. In diesem Fall weichen die Werte von benachbarten Beobachtungen mit hoher Wahrscheinlichkeit voneinander ab
3. **Keine Autokorrelation** tritt dann auf, wenn die Beobachtungen einer Zufallsverteilung unterliegen. Nur in diesem Fall ist die **räumliche Unabhängigkeit** der Daten gegeben und die Methoden der klassischen Statistik anwendbar

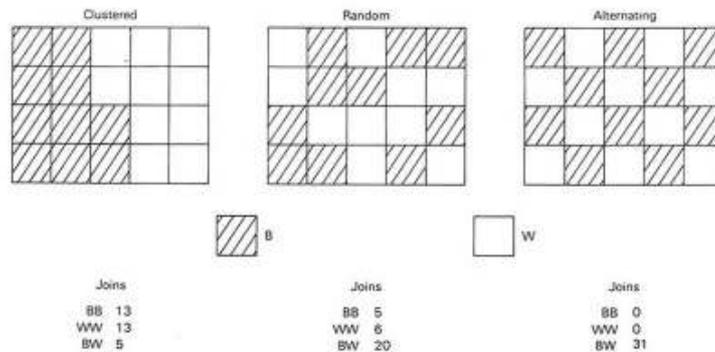


Abbildung 8: Die drei Möglichkeiten der räumlichen Verteilung und Autokorrelation.
(Quelle: Haining (1990), S. 229)

Da die Autokorrelation einen erheblichen Einfluss auf die weiteren statistischen Untersuchungen und die Interpretation der Ergebnisse hat, ist es unentbehrlich, diese vor weitergehenden Untersuchungen eines Datensatzes zu ermitteln (O'Sullivan & Unwin (2003), S. 29). Waren zuerst nur globale Methoden zur Berechnung der Autokorrelation in einem Datensatz vorhanden, so wurden später auch Methoden entwickelt, die es ermöglichen, tiefer in die räumliche Struktur des zu untersuchenden Datensatzes einzudringen und die Autokorrelation auf lokaler Ebene zu untersuchen (Anselin (1995), S. 93).

4.4.3.1 Globale Methoden

Globale Methoden zur Ermittlung der Autokorrelation werden auf den gesamten Datensatz angewendet. Die Berechnung liefert im Ergebnis einen einzigen Wert. Dieser zeigt an, ob für den gesamten Datensatz die Autokorrelation festgestellt wurde und wie stark diese ist. Die räumliche Autokorrelation kann mit verschiedenen Herangehensweisen berechnet werden. Diese werden z. B. bei Upton and Fingleton (1985, S. 151ff.) ausführlich beschrieben. Nachfolgend werden die beiden am häufigsten eingesetzten Verfahren vorgestellt.

Moran's I

Eine Methode zur Ermittlung der räumlichen Autokorrelation in einem Datensatz ist durch den Moran Index (I) gegeben. Dieser berechnet sich wie folgt (vgl. Fotheringham et al. (2000), S. 202):

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_i (z_i - \bar{z})^2} \quad (3)$$

- n - die Menge aller Elemente des zu untersuchenden Datensatzes
- i, j - die aktuell zu untersuchenden Objekte (Features)
- z_i, z_j - die Merkmalsausprägung der Objekte i und j
- \bar{z} - der Mittelwert aller Merkmalsausprägungen
- w_{ij} - das Gewicht der Nachbarschaftsbeziehung zwischen den Objekten i und j ($w_{ii} = 0, i (i \neq j)$)

Der Index (I) wird positiv, wenn (räumlich) benachbarte Datensätze ähnliche Werte aufweisen (positive Autokorrelation) und negativ, wenn diese Datensätze stark voneinander abweichende Werte erkennen lassen (negative Autokorrelation). Lässt sich keine räumliche Autokorrelation feststellen, so nimmt I einen Wert $I \cong 0$ an (vgl. Longley et al. (2001), S. S. 114).

Geary's C

Auch dieser Index beschreibt den Grad der räumlichen Autokorrelation benachbarter Werte. Er berechnet sich entsprechend der Formel (4) (vgl. Bailey & Gatrell (1995), S. 270).

$$C = \frac{n-1}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (z_i - z_j)^2}{2 \sum_i (z_i - \bar{z})^2} \quad (4)$$

Bei positiver räumlicher Autokorrelation nimmt C einen Wert zwischen 0 und 1 an und bei negativer räumlicher Autokorrelation wird $C < 1$. Ist keine Autokorrelation feststellbar, ist der Wert von $C > 1$ (vgl. O'Sullivan & Unwin (2003), S. 201).

Im Gegensatz zu Moran's I, der sensibel auf große z-Werte reagiert, ist Geary's C-Index stärker durch die absoluten Differenzen zwischen den Wertepaaren zu beeinflussen. Anhand von Untersuchungen konnten Cliff & Ord (1975, 1981 zitiert in Upton &

Fingleton (1985), S. 171) zeigen, dass Tests, die auf der Berechnung des Moran's I beruhen, sich als zuverlässiger erwiesen als andere Tests, die auf der Berechnung von Geary's C beruhen. Diese Erkenntnis mag ein Grund dafür sein, warum z. B. in dem Produkt ArcGIS 9.x der Firma ESRI nur die Berechnung des Moran's I aufgenommen wurde. Beide Indexwerte sind jedoch z. B. in der Anwendung CrimeStat (vgl. Levine 1999, S. 3) implementiert.

4.4.3.2 Lokale Methoden

Die im vorhergehenden Kapitel beschriebenen globalen Methoden zur Bestimmung der räumlichen Autokorrelation geben erste Hinweise in Bezug auf die räumlichen Beziehungen der Daten. Sie ermöglichen jedoch keine detaillierten Aussagen, wenn es z. B. um die Frage geht, ob überall im untersuchten Gebiet (Datensatz) dieselben Bedingungen herrschen. Zeigen die zurückgegebenen Werte an, dass positive Autokorrelation der Daten vorliegt, so können sie nicht beantworten, wo sie vorliegt, d.h. welche Daten davon betroffen sind. So kann z. B. in einem Datensatz, der ein großes Gebiet abdeckt, ein Teil des Gebietes durchaus positive Autokorrelation aufweisen, obwohl für den gesamten Datensatz keine Autokorrelation festgestellt werden konnte.

Nachfolgend werden zwei Methoden vorgestellt, die versuchen, diesen Aspekt der Verteilung der Autokorrelation in einem Datensatz genauer zu untersuchen. Diese Methoden untersuchen also die lokalen Beziehungen, die auf einem kleinen Raum innerhalb des gesamten Datensatzes herrschen. Diese Methoden wurden von Anselin (1995, S. 93) als LISA bezeichnet (Local Indicators of Spatial Association).

Anselin & Bao (1997, S. 42) benennen zwei Ziele, die durch den Einsatz eines LISA erfüllt werden können.

- Mit Hilfe des LISA können sowohl Hot-Spots als auch räumliche Ausreißer (z.B. hohe Werte, die von niedrigen Werte umgeben sind) identifiziert werden
- Ein LISA erlaubt darüber hinaus, die räumliche Aussagekraft eines globalen (vgl. Moran's I) Parameters zu überprüfen und dessen Reichweite in Bezug auf den gesamten Datensatz darzustellen.

Da diese Methoden für jedes einzelne Element des gesamten Datensatzes einen Wert zurückgeben, der die Beziehung dieses Elementes zu seinen Nachbarn beschreibt,

können die Ergebnisse dieser Berechnungen anschließend in einem GIS dargestellt werden (O'Sullivan & Unwin (2003), S. 204). Nachfolgend werden zwei Verfahren beschrieben, die in den 90er Jahren entwickelt wurden und auch bereits in verschiedenen Crime-Mapping-Anwendungen zum Einsatz gekommen sind (vgl. Chainey et al. (2002), Murray et al. (2001), Ratcliffe & McCullagh (1999)).

Local Moran I

Anselin (1995, S. 98) nannte das von ihm begründete Verfahren den Lokalen Moran I. Dieser lokale Index ergibt sich unter Verwendung der nachfolgenden Formel (5). Der lokale Index I wird für jedes beobachtete Ereignis des Datensatzes ermittelt. Wird anschließend über alle lokalen Indizes die Summe gebildet, ergibt sich exakt der globale Moran I.

$$I_i = z_i \sum_j w_{ij} z_j \quad (5)$$

z_i, z_j - Werteausprägung für die Ereignisse z_i und z_j (in Abweichung vom Mittelwert)
 w_{ij} - eine Gewichtungsmatrix (i.d.R. zeilenstandardisiert, wobei $w_{ii} = 0$)

Zur Interpretation der Werte ist zu sagen, dass ähnlich wie beim globalen Indikator, Moran's I, ein positiver Wert dann ermittelt wird, wenn **ähnliche** Werte (entweder niedrige oder hohe Werte) um das Ereignis herum auftreten. Negative Werte werden dann ermittelt, wenn **unähnliche** Werte um den untersuchten Punkt herum angetroffen wurden (Anselin (1995), S. 102).

G_i/G_i*

Auch mit Hilfe der G_i/G_i*-Statistiken können lokale Assoziationen zwischen Daten aufgedeckt werden, die nicht zum Vorschein kämen, wenn der Datensatz ausschließlich mit globalen Methoden betrachtet würde (vgl. Ord & Getis (1995), S. 288). Da die Berechnung von G_i/G_i* jedoch ausschließlich auf die Aufdeckung von Hotspots zielt, stellt diese Statistik keinen eigentlichen LISA (im Sinne der Definition von Anselin & Bao (1997), S. 42) dar. Die Berechnung dieser Indizes wurde erstmals von Getis & Ord (1992) vorgestellt und dann im Jahre 1995 in einer überarbeiteten Form erneut präsentiert.

Die Werte G_i und G_i^* sind zwei unterschiedliche Ausprägungen derselben Rechenregel. In der folgenden Formel ist zuerst der Indikator G_i erklärt. Anschließend wird der Unterschied für die Berechnung des Wertes G_i^* erläutert.

$$G_i(d) = \frac{\sum_j w_{ij}(d) x_j}{\sum_j x_j}, \text{ für alle } j (j \neq i) \quad (6)$$

w_{ij} - eine symmetrische Gewichtungsmatrix, die den Wert 1 erhält, wenn der Punkt j innerhalb der Entfernung d zum Punkt i liegt. Für alle übrigen Beziehungen wird der Wert 0 eingetragen. Das gilt auch für die Beziehung des Punktes i mit sich selber.

Die Berechnung des Wertes G_i^* erweitert die oben dargestellte Formel für den Fall, dass $i = j$. Dies bedeutet, dass zur Ermittlung der Summe aller x_j auch der am aktuell zu untersuchenden Ort i vorgefundene Wert x_i einbezogen wird.

Die Interpretation der Ergebnisse der G_i/G_i^* -Statistik erfolgt folgendermaßen (vgl. Anselin (1995), S. 102): Ein positiver G_i/G_i^* weist auf eine räumliche Konzentration hoher Werte hin, ein negativer G_i/G_i^* deutet darauf hin, dass um diesen Punkt herum ein Cluster mit niedrigen Werten vorliegt.

Nach Getis & Ord (1992, S. 192) erfolgt die Berechnung von G_i/G_i^* maßstabs-, jedoch nicht ortsunabhängig. Dies bedeutet, dass auf allen Aggregationsebenen dieselben Werte für G_i/G_i^* herauskommen, dass aber die Lage der jeweiligen Ereignisse einen erheblichen Einfluss auf die Berechnung des Index ausübt.

4.4.4 Die Problematik der Randeffekte

Alle Methoden, die auf einen Distanz-Parameter zurückgreifen müssen, haben ein gemeinsames Problem. Dieses Problem tritt auf, sobald sie an die Grenzen des Untersuchungsgebietes stoßen, da diese oft willkürlich festgelegt wurden und deshalb nicht mit dem natürlichen Verlauf des zu untersuchenden Prozesses übereinstimmen. Sehr oft werden die Grenzen eines Untersuchungsgebietes durch administrative Grenzen festgelegt. Es ist jedoch anzunehmen, dass sich der zu untersuchende räumliche Prozess außerhalb der gegebenen Grenzen fortsetzen wird. Bezogen auf das in der vorliegenden Arbeit zu untersuchende Thema Kriminalität ist zu erwarten, dass sich diese auch

jenseits der Grenzen einer Polizeidienststelle, eines Präsidiums oder eines Landes fortsetzen wird. Da aber die Verantwortlichkeiten an den Gebietsgrenzen aufhören, werden Daten außerhalb dieser Grenzen nicht erfasst. Dies beeinflusst die Ergebnisdarstellung an den Rändern, da nur die innerhalb der Grenzen liegenden Ereignisse in die Berechnungen eingehen.

Um diese so genannten Randeffekte auszuschließen, bzw. zu minimieren, wird häufig folgendes Vorgehen vorgeschlagen (O'Sullivan & Unwin (2003), S. 95). Entlang der Grenzen wird ein Puffer nach innen eingezogen. Diejenigen Daten, die innerhalb dieses Puffers liegen, werden nicht in die Ergebnisdarstellung integriert. Sie gehen ausschließlich in die Berechnung des Prozesses ein (z. B. der Kernel Density Estimation oder der K-Funktion). Durch dieses Vorgehen hat sich die Fläche reduziert, die zur Ergebnisdarstellung eingesetzt werden kann. Die Ergebnisse sind jedoch um eine mögliche Fehlerquelle bereinigt.

5 Bisher eingesetzte Programme und Methoden

Mit Hilfe eines GIS können die Straftaten und andere räumliche Ereignisse, an denen die Polizei ein Interesse hat, grafisch aufbereitet und die räumliche Verteilung dieser Ereignisse dargestellt werden. Sehr schnell zeigte sich jedoch, dass eine ausschließlich visuelle Interpretation für die Suche nach Mustern in der Straftatenverteilung nicht ausreichend war, da sie keine hinreichend genauen räumlichen Aussagen zuließen. Da die GIS-Hersteller lange nicht bereit waren, die für diese Untersuchungen erforderlichen Spezialmethoden in ihre Systeme zu integrieren, wurden alternative Formen der Mustererkennung außerhalb der GIS-Systeme gesucht. Es wurden verschiedene Anwendungen entwickelt, die die räumliche Analyse der Datensätze durchführten. Die Ergebnisse dieser Analysen wurden anschließend – sofern sie sich räumlich darstellen ließen – wieder in die GIS-Systeme importiert und dort zu den Original-Datensätzen und anderen räumlichen Daten in Beziehung gesetzt.

5.1 STAC – Spatial and Temporal Analysis of Crime

Dieses Programm wurde von der Illinois Criminal Justice Information Authority entwickelt und 1988 fertiggestellt (vgl. ICJIA (2005)). STAC war eine der ersten Anwendungen für die Kriminalitätsanalyse und das Crime-Mapping. Da es für die US-amerikanischen Polizeibehörden kostenlos im Internet zum Download erhältlich ist, hat es in den USA eine weite Verbreitung erfahren (vgl. Bowers & Hirschfeld (2001), S. 2).

Das Programm besitzt eine Vielzahl von Methoden der räumlichen Statistik, die unter anderem einige der bisher vorgestellten Methoden der Point-Pattern-Analyse beinhaltet (vgl. Block (1995), S. 2ff). Das Ziel, das mit der Entwicklung dieses Programmes verfolgt wurde, bestand darin, Methoden bereitzustellen, mit deren Hilfe die Gebiete mit der höchsten Kriminalitätsdichte identifiziert werden konnten (Hot Spots, vgl. Block (1995), S. 7).

Mit Methoden der Nearest Neighbour Analyse (vgl. Teil I, Kapitel 4.3.2.1) wird zuerst getestet, ob das zu untersuchende Gebiet eine räumliche Konzentration von Ereignissen aufweist. Ist dies der Fall, werden anschließend in mehreren Simulationsläufen die Räume mit der höchsten Ereignisdichte herausgefiltert. Die Darstellung dieser Gebiete erfolgt als Ellipse. Diese beschreibt die räumliche Standardabweichung vom Mittelwert der betreffenden Ereignisse. Wurde in einem Gebiet mehrere Cluster entdeckt, dann werden diese nach der Anzahl der auftretenden Ereignisse geordnet (vgl. Block (1995), S. 9). Das Ergebnis eines STAC-Laufes (s. Abbildung 9) wird in eine Datei geschrieben. Diese Datei kann anschließend in ein GIS geladen und dort betrachtet werden.

Im Verlauf der Zeit wurden einige gravierende Nachteile des in STAC eingesetzten Verfahrens festgestellt (vgl. Hirschfeld et al. (2001), S. 323ff.). So wird z.B. die Darstellung der identifizierten Hot-Spot-Flächen in Form einer Ellipse abgelegt, da sie eine theoretische Annahme darstellt, die i. A. nicht der tatsächlichen Ausdehnung der Hot-Spots entspricht. Darüber hinaus steht STAC im Verdacht, unter dem Modifiable Areal Unit Problem (MAUP, vgl. Teil I, Kapitel 4.2.2) zu leiden.

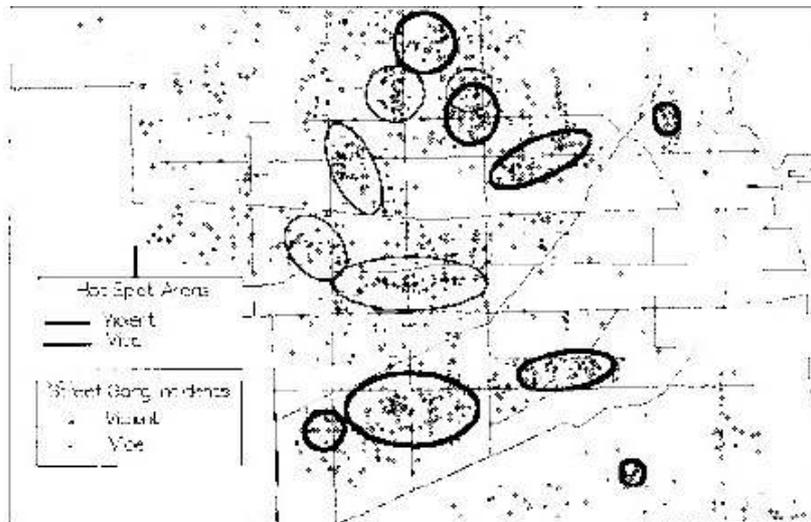


Abbildung 9: Ergebnis eines STAC-Laufes (Quelle Block (1995), S. 16)

5.2 CrimeStat

Dieses Softwarepaket wurde von Ned Levine (vgl. Levine (1999)) für das Crime Mapping Research Center (jetzt Mapping and Analysis for Public Safety) des National Institute of Justice, Washington, DC. entwickelt. Es handelt sich um eine stand-alone-Anwendung, die auf einem Windows PC betrieben wird. Auch diese Anwendung ist im Internet unter der URL <http://www.icpsr.umich.edu/NACJD/crimestat.html> kostenlos herunterzuladen. Mit CrimeStat können die verschiedensten Berechnungen der Point-Pattern-Analyse (von einfacher Berechnung des räumlichen Mittelwertes über Nearest-Neighbour-Analysen, K-Funktionen, Kernel-Density- und Cluster-Berechnungen, sowie die Ermittlung der räumlichen Autokorrelation) durchgeführt werden. Viele Ergebnisse lassen sich in unterschiedlichen GIS-Formaten abspeichern und anschließend dort einlesen (Levine (1999), S. 1).

Mit der Vielfalt der angebotenen Methoden, die CrimeStat zur Untersuchung eines Datensatzes anbietet, gehört es zweifelsfrei zu den umfangreichsten Anwendungen, die für die Untersuchung der räumlichen Verteilung von Straftaten zur Verfügung stehen.

5.3 GAM – Geographic Analysis Maschine(s)

Die Geographic Analysis Machine wurde von Openshaw et al. (1987) entwickelt. Sie stellte einen ersten Versuch zur Automatisierung der explorativen räumlichen Datenanalyse dar und wurde gebaut, weil die herkömmlichen Methoden der PPA aufgrund diverser technischer Probleme die Validität (Verlässlichkeit) der Ergebnisse nicht garantieren konnten (vgl. Openshaw et al. (1987), S. 337). Das Ziel der Analyse besteht in der Aufdeckung räumlicher Cluster. Von einem Cluster wird dann gesprochen, wenn für die zu untersuchenden Ereignisse eine räumliche Konzentration nachgewiesen wurde. Da jedoch die zu untersuchenden Ereignisse bereits von ungleichmäßig im Raum verteilten Ausgangsvariablen abhängen, wird ein Cluster erst dann akzeptiert, wenn er auch unter Einbeziehung der ungleichmäßig verteilten Risikobevölkerung („population-at-risk“) bestehen bleibt (vgl. Openshaw et al. (1987), S. 339). Dieses Vorgehen ist insbesondere im (umwelt-)medizinischen Umfeld wichtig, wenn es um die Klärung der Frage geht, ob eine bestimmte Häufung von Krankheiten z. B. Krebs auf umweltschädigende Einflüsse einer Industrieanlage (vgl. Openshaw et al. (1987), S. 355) zurückzuführen sei. Das Berechnungsverfahren ist sehr aufwändig und besteht aus den folgenden Arbeitsschritten:

1. Ein Punkt auf der Karte wird gesucht (hier: als Teil eines Rasters) und ein Kreis mit dem Radius r wird um diesen Punkt gelegt
2. Die Anzahl der vorgefundenen Ereignisse innerhalb des Kreises wird gezählt
3. Diese Zahl wird verglichen mit einem erwarteten Wert (unter der Annahme eines Zufallsprozesses und vor dem Hintergrund der „population-at-risk“)
4. Der Kreis wird beibehalten (und in einer Karte vermerkt), wenn die Zahl der festgestellten Ereignisse signifikant höher ist als der erwartete Wert
5. Schritte 1 – 4 werden für alle weiteren Punkte in der Karte durchgeführt.

Seit 1987 wurde GAM mehrfach getestet. Es wurden Ungenauigkeiten aufgedeckt und alternative Methoden für die Suche nach Clustern und die Signifikanztests vorgeschlagen (vgl. Fotheringham & Zhan (1996)). Mittlerweile ist eine Version der GAM (GAM/K) im Internet verfügbar (vgl. <http://www.ccg.leeds.ac.uk/smart/gam/gam.html>).

GAM soll auch von solchen Anwendern bedienbar sein, die keine ausgeprägten statistischen Vorkenntnisse haben (vgl. Turton & Openshaw (2001), S. 12).

Teil II: Konkreter Fall – Hamburg

1 GISPOL – die Crime Mapping Anwendung der Hamburger Polizei

Seit dem Jahr 2000 besitzt die Polizei Hamburg mit dem Projekt GISPOL eine Anwendung zur Darstellung der räumlichen Verteilung von Straftaten innerhalb des Stadtgebietes. GISPOL ist als Intranet-Anwendung konzipiert und wurde unter Verwendung der ESRI-Produkte ArcIMS und ArcSDE eingerichtet. Die Daten, die in die GISPOL-Datenbank eingehen, werden in dem Vorgangsbearbeitungssystem COMVOR erfasst und in regelmäßigen Aktualisierungsläufen an die GISPOL-Datenbank übergeben (vgl. Pandler (2002), S. 22f).

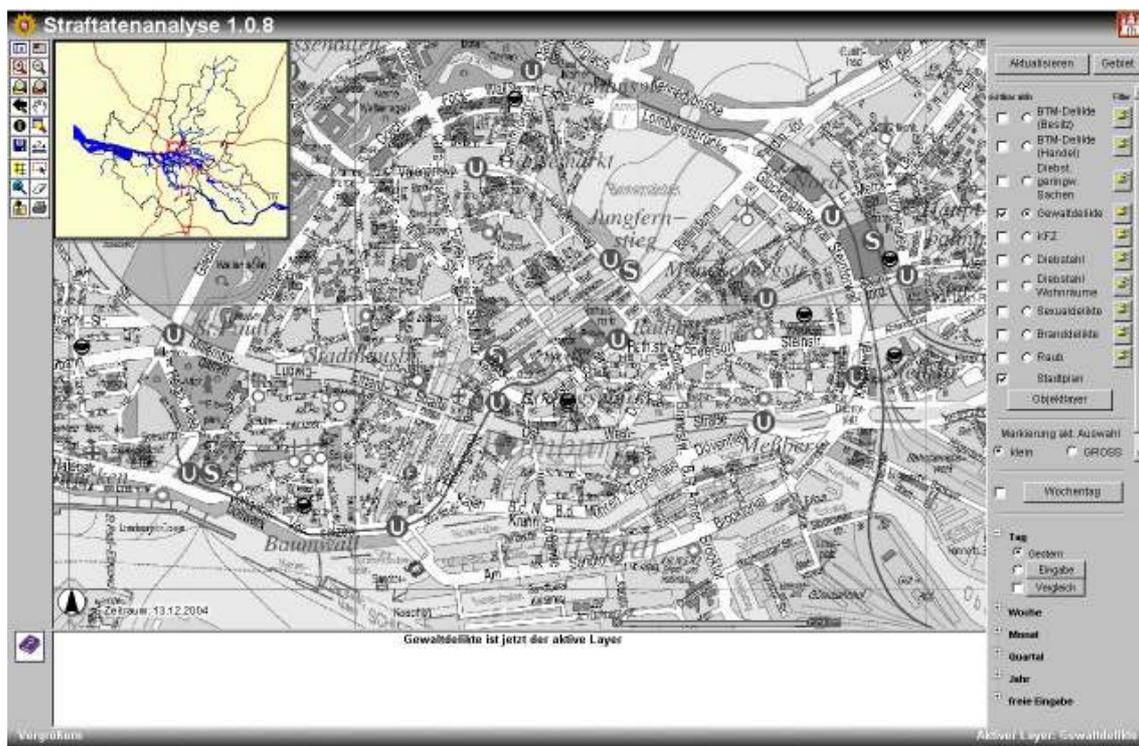


Abbildung 10: Die ArcIMS-Oberfläche der Straftatenanalyse in GISPOL

Wie aus der Abbildung 10 ersichtlich, werden die Daten – getrennt nach Deliktarten – als Layer aufbereitet und mit Hilfe des Produktes ArcIMS der Fa. ESRI im Intranet der Polizei bereitgestellt. Durch diese Möglichkeit des Datenzugriffs über das Intranet

haben die Verantwortlichen in den Polizeikommissariaten jederzeit einen Überblick über das aktuelle Geschehen (wo entstehen lokale Schwerpunkte) in ihrem Verantwortungsbereich. Dies ermöglicht es ihnen, unmittelbar und schnell auf sich verändernde Situationen innerhalb ihrer Zuständigkeitsbereiche zu reagieren.

2 Die Erstellung eines Lagebildes mit GISPOL

In der deutschen Polizeiarbeit ist der Begriff der „Lage“, bzw. des „Lagebildes“ weit verbreitet. Was ist unter einer Lage zu verstehen und wie lassen sich die neuen Technologien, wie z. B. das Crime Mapping in diesen Kontext einordnen?

Laut Brisach et al. (2001, S. 99) stellt ein Lagebild ein „reduziertes, aktuelles Abbild der gesellschaftlichen Wirklichkeit zum Zwecke ihrer Beurteilung“ dar. Im Vergleich zur Kriminalitätsstatistik, die das vergangene Kriminalitätsgeschehens abbildet, soll ein Lagebild den aktuellen Zustand der Kriminalitätsverteilung (unter Berücksichtigung der Beurteilung durch die Bevölkerung sowie der örtlichen Polizeikräfte) darstellen. (Brisach et al. (2001, S. 109).

Okon (1999) nennt in einem Vortrag auf der deutschen ESRI-Userconference folgende Ziele für die Erstellung eines Lagebildes:

- es bereitet die Grundlage für taktische und strategische Führungsentscheidungen
- es soll das frühzeitige Erkennen von Brennpunkten, Straftatenserien und neuer Deliktphänomene ermöglichen
- es soll Fahndungsinformationen bereitstellen
- es soll der Erkennung von Ermittlungsansätzen dienen
- es soll die Erarbeitung von realistischen Prognosen der Kriminalitätsentwicklung und deren Bekämpfungsstrategie ermöglichen
- es dient zur Begründung der für die Aufgabenerfüllungen notwendigen personellen und materiellen Ressourcen gegenüber politischen Entscheidungsträgern

Die Auflistung dieser Ziele macht eines deutlich: die Darstellung der räumlichen Straftatenverteilung ist ein (wichtiger) Bestandteil eines Lagebildes, sie reicht zur Erstellung und Nutzung des Lagebildes jedoch keinesfalls aus. Für die Einordnung der

Anwendung GISPOL in den Kontext der Lagebilderstellung bedeutet dies, dass diese Anwendung eine Vielzahl an Daten und Methoden bereitstellen muss, damit sich die Anwender über räumliche sowie die zeitliche Verteilung von Straftaten, der Suche nach Gemeinsamkeiten und herausragender neuer Erscheinungen informieren können. Die Darstellung der Sachdaten und eine sachorientierte Auswertung der Ereignisse sind bereits jetzt in GISPOL integriert. Räumliche Analysemethoden sind – bis auf die Standardfunktionen des ArcIMS (z. B. Buffer) noch nicht implementiert. Die Aussagekraft der Anwendung würde durch das Einbinden solcher Methoden erhöht.

3 Die Beschaffenheit der Daten

Wie im Namen bereits anklingt, wird im Rahmen des Projektes GISPOL ein GIS, bzw. GIS-Funktionalitäten eingesetzt, die zu einer räumlichen Darstellung und Interpretation des Sachverhaltes „Verteilung des Straftatenaufkommens“ beitragen. Um ein GIS (Geographisches Informationssystem) erfolgreich betreiben zu können, bedarf es zusätzlich zu der erforderlichen Software, der Hardware und den Methoden vor allem auch der Daten, die mit diesem System verwaltet und bearbeitet werden sollen. Nicht zuletzt sind die Menschen, die an diesen Systemen arbeiten und z. B. die Daten analysieren und interpretieren entscheidend für den erfolgreichen Einsatz dieser Systeme. Dieses Kapitel beschreibt die in der vorliegenden Arbeit eingesetzten Daten. Dabei werden insbesondere die Fragen der Repräsentativität, der Validität und der Reliabilität des Datenmaterials zu berücksichtigen sein.

3.1 Welche Daten werden untersucht?

Für die vorliegende Arbeit wurden drei verschiedene Punktdatensätze bereitgestellt.

- Diebstähle aus dem Jahr 2003
- Einbrüche aus dem Jahr 2003
- Straßenabschnitte aus dem ersten Halbjahr des Jahres 2004

Zusätzlich zu den Punktdaten wurden die Grenzdaten der sieben Hamburger Bezirke aus der digitalen Stadtgrundkarte hinzugezogen. Diese Daten liegen im Original als Linien-

datensätze vor. Nach der Wandlung der Linien- in Polygondaten wurden die Teilflächen zu einer einzigen Fläche zusammengefügt. Auf diese Weise war das Stadtgebiet Hamburgs als geschlossene Fläche vorhanden.

Die Punktdatensätze werden nachfolgend näher beschrieben, da sie das grundlegende Material für die durchgeführten Untersuchungen darstellen.

Diebstähle 2003

Der Datensatz „Diebstähle 2003“ enthält 61876 Einzeldaten. Zusätzlich zu den Geometriedaten waren die folgenden Attributdaten vorhanden.

Tabelle 1: Diebstähle 2003 - Sachattribute

Spaltenüberschrift	Bedeutung
LawTextInf	Beschreibung der Tat unter Nennung der Paragraphen
Deliktart	Beschreibt die „kriminologische Falleigenschaft“ (mdl. Mitteilung Kondoch (22. 03. 2005))
Bekanntsac	Boolscher Ausdruck; J = Täter bekannt, N = unbekannt
Factbegin	Tatzeit (der frühestmögliche Tatzeitpunkt, auf den Tag beschränkt)
OEPNV	Boolscher Ausdruck; J = Tatort war im ÖPNV
Versuch	Boolscher Ausdruck; J = Tat wurde nicht ausgeführt

Einbrüche 2003

Der Datensatz „Einbruch2003“ enthält 9419 Einzeldaten, von denen zwei Einträge keine Georeferenzierung aufwiesen (bzw. mit den Koordinaten X = 0 und Y = 0 versehen waren). Diese Daten wurden aus dem Datensatz entfernt, da sie die räumlichen Berechnungen verfälscht hätten. Die Attributdaten liegen in denselben Kategorien vor, wie für die Diebstähle beschrieben. Die Tatzeit wird in vielen Fällen mit Uhrzeit erfasst - leider nicht für alle Datensätze.

Straßenabschnitte 2004

Der Datensatz Straßenabschnitte enthält keine Daten, die für die Kriminalitätsanalyse zwingend erforderlich wären. Aus diesem Grund gehören sie nicht zum Datenumfang des Projektes GISPOL. Da sich jedoch die Darstellung der Straftaten auf die Straßenabschnitte bezieht, konnte im Rahmen dieser Untersuchungen nicht auf die

Straßenabschnitte verzichtet werden. Die Straßenabschnitte wurden – nach Klärung der Nutzungsrechte mit dem Statistischen Amt für Hamburg und Schleswig-Holstein – aus dem Einsatzleitsystem HELP der Hamburger Polizei beschafft. Sie werden im weiteren Verlauf als Bezugspunkte eingesetzt.

Der Datensatz enthält 26472 Einzeldaten. Die Attributdaten sind in der Tabelle 2 aufgelistet. Sie spielen jedoch für die nachfolgenden Untersuchungen keine Rolle.

Tabelle 2: Straßenabschnitte - Attribute

Spaltenüberschrift	Bedeutung
STRSCHL	Alphanumerischer Bezeichner für die Straße
OBJSCHL	Alphanumerischer Bezeichner für den Abschnitt (nicht immer gesetzt)
SAL	Länge des Straßenabschnittes (kann auch 0 sein)

Zusätzlich zu den Abschnitten auf „echten“ Straßen enthält dieser Datensatz weitere Daten, die Abschnitte auf Wegen z. B. in Kleingartenvereinen beschreiben oder auch besondere Punkte im Stadtgebiet kennzeichnen. So ist z. B. auch die Außenalster als Straßenabschnittspunkt eingetragen. Weiterhin geht die Darstellung der Straßenabschnitte im Norden, Westen und Osten über das Stadtgebiet Hamburgs hinaus. Dies ist im Süden nicht der Fall.

Die Straßenabschnitte, die bei der Polizei als Bezugspunkte eingesetzt werden, werden halbjährlich aktualisiert. Dies bedeutet, dass einige Straßenabschnitte verschwinden oder neu hinzukommen können. Der Bezug der Straftaten wird jedoch nicht der geänderten Straßenabschnittsverteilung angepasst. Da die Straßenabschnittsdaten jüngeren Datums sind als die untersuchten Straftaten, stellte sich für einige Straftaten heraus, dass der räumliche Bezug zu einem Straßenabschnitt (der bei der Erfassung dieser Straftaten vorhanden gewesen sein muss) mittlerweile verloren gegangen ist.

3.2 Repräsentativität des Datenmaterials

Die Frage nach der Repräsentativität des Datenmaterials beschäftigt sich zuerst mit der Frage, inwiefern die Hamburger Kriminalitätsdaten das tatsächliche Kriminalitätsaufkommen in Hamburg abbilden können. Anschließend liegt der Schwerpunkt der

Betrachtung auf den Besonderheiten, die mit der räumlichen Datenanalyse verbunden sind. In diesem Zusammenhang wird die Genauigkeit der Georeferenzierung besonders zu berücksichtigen sein.

3.2.1 Kriminalitätsaufkommen und Dunkelfeld

Die Daten, die in das System GISPOL eingehen, werden über das zentrale Straftatenerfassungs- und Vorgangsbearbeitungssystem COMVOR erfasst. Hier werden alle Straftaten bearbeitet, von denen die Polizei Kenntnis erhält. Kenntnis erhält die Polizei immer dann, wenn Straftaten durch die Bürger zur Anzeige gebracht werden, oder wenn die Polizei im Rahmen ihrer Tätigkeit (z. B. bei Polizeieinsätzen, Fahnungen, etc.) selber Anzeigen ausspricht.

Straftaten, die der Polizei nicht zur Kenntnis gebracht werden, sind somit auch nicht erfasst. Dieses „Dunkelfeldproblem“ haftet jeder Polizeistatistik an. In der Kriminalstatistik 2003 des LKA Hamburg ist zu diesem Problem folgendes zu lesen:

„Die Aussagekraft der Polizeilichen Kriminalstatistik erfährt dadurch eine wesentliche Einschränkung, dass der Polizei ein Teil der Straftaten verborgen bleibt. Der Umfang dieses Dunkelfeldes ist abhängig von der Art des Deliktes und u.a. vom Anzeigeverhalten der Bevölkerung und der Intensität der Verbrechensbekämpfung. Da sich die einzelnen statistikbeeinflussenden Faktoren unabhängig voneinander ändern, gibt es auch keine feststehenden Relationen zwischen statistisch erfassten und tatsächlich begangenen Straftaten in verschiedenen Zeiträumen. Die Polizeiliche Kriminalstatistik bietet demzufolge kein Abbild der Verbrechenswirklichkeit, sondern eine je nach Deliktart mehr oder weniger ausgeprägte Annäherung an die Realität.“ (vgl. Polizei Hamburg (2003), S. 5).

Auch die Untersuchung der räumlichen Straftatenverteilung kann erheblich durch das Dunkelfeld-Problem beeinträchtigt sein.

3.2.2 Georeferenzierung und Genauigkeit

Um die erfassten Straftaten in einer Karte darstellen zu können, bedarf es der Georeferenzierung. Das Prinzip des Georeferenzierens von Straftaten und die Randbedingungen, die den Erfolg der Georeferenzierung und der darauf aufbauenden räumlichen Darstellung und Analyse der Straftatendaten beeinflussen, wurde bereits im

Kapitel 3.1.1 des ersten Teils beschrieben. An dieser Stelle sei deshalb darauf verwiesen.

Daten, die weder automatisch noch manuell georeferenziert werden, sind für die anschließenden räumlichen Untersuchungen verloren. Das Fehlen dieser Daten kann nachfolgende Darstellungen und Auswertungen negativ beeinflussen und sogar verfälschen (vgl. Ratcliffe (2004a), S. 70). In der Literatur werden einige Zahlen in Bezug auf hinnehmbaren Datenverluste im Zusammenhang mit unvollständiger Georeferenzierung genannt. Während Vann & Garson (2003, S. 28) als „Daumenregel“ von einem tolerablen Verlust von bis zu 40% der Daten sprechen, ohne diese Zahl jedoch zu belegen, schränkt Ratcliffe (2004a, S. 69) aufgrund eigener experimenteller Untersuchungen den gerade noch zu tolerierenden Datenverlust durch fehlende Georeferenzierung auf 15% ein. Die Hamburger Datenverluste, die im Zusammenhang mit erfolgloser Georeferenzierung stehen, werden auf 28% beziffert (LKA Hamburg (2003)). Es ist also zu befürchten, dass fehlende Daten aufgrund erfolgloser Georeferenzierung das Gesamtergebnis negativ beeinflussen können. Dies gilt insbesondere dann, wenn die fehlenden Daten nicht zufällig über das Untersuchungsgebiet verteilt sind, sondern sich in bestimmten Gegenden häufen. Im letzten Fall würde dieser einseitige Datenverlust die gesamte Darstellung der Straftatenverteilung verfälschen.

Im Gegensatz zu den bisher vorgestellten Georeferenzierungsmethoden (vgl. Teil I, Kapitel 3.1.1) werden die Straftaten in Hamburg auf Straßenabschnitte bezogen. Die Straßenabschnitte werden vom „Statistischen Amt für Hamburg und Schleswig-Holstein“ erhoben. Sie sind als Punktdaten auf den Mittelachsen der Straßen abgebildet und mit metrischer Genauigkeit erfasst. Jede Straße besitzt mindestens einen Straßenabschnitt. Lange Straßen, bzw. alle Straßen mit Kreuzungen oder Einmündungen erhalten pro Kreuzung/Einmündung einen weiteren Straßenabschnitt. Die Bezugnahme auf die Straßenabschnitte im Rahmen der Georeferenzierung bewirkt, dass die Straftaten nicht unmittelbar dort dargestellt werden, wo sie passiert sind, sondern auf den Straßenabschnitt verschoben werden, der dem Tatort am nächsten liegt. Im Innenstadtbereich fallen bei diesem Vorgehen viele Straftaten an einem Straßenabschnitt zusammen. Da die Daten anschließend als Einzelereignisse durch Punktsymbole dargestellt

werden, ist für den Betrachter durch den alleinigen Blick in die Karte nicht zu erkennen, ob an einem Straßenabschnitt, an dem ein Straftatereignis dargestellt wird, tatsächlich nur das eine oder weitere Ereignisse stattgefunden haben (vgl. Abbildung 11).



Abbildung 11: Verteilung der Diebstahlereignisse auf die Straßenabschnitte. Hinter den gelb markierten Symbolen verbergen sich 426 Einzeldelikte - nur wo genau?

4 Alternativen zur „Steckkarte“

Die Projektion der Punktdaten in die Fläche könnte die Unklarheiten, die in der aktuellen Darstellungsform der Steckkarte begründet sind, beheben. Zu diesem Zweck werden die nachfolgend vorgestellten Untersuchungen durchgeführt. Einige dieser Untersuchungen entsprechen den in den vorhergehenden Kapiteln vorgestellten Methoden, andere verfolgen einen neuen Ansatz.

4.1 Nearest-Neighbour-Analyse zur ersten Untersuchung der Datensätze

Um einen ersten Eindruck über die räumlichen Beziehungen der Daten zu gewinnen, wurden Sie mit Hilfe des Programms CrimeStat (vgl. Teil I, Kapitel 5.2) einer Nearest-Neighbour-Analyse (NNA) unterzogen. Die NNA gibt einen Hinweis auf die räumliche

Verteilung der Daten. Zuerst wird die durchschnittliche Distanz aller Punkte zu ihren nächsten Nachbarn ermittelt (vgl. Teil I, Kapitel 4.3.2.1).

Das Ergebnis wird anschließend mit einer erwarteten Entfernung verglichen. Die erwartete Entfernung gilt unter der Annahme, dass die Punktdaten rein zufällig verteilt sind (Complete Spatial Randomness – CSR). Für den Vergleich der tatsächlichen mit der Erwarteten mittleren Nachbarschaftsentfernung wird das Signifikanzniveau ermittelt. Die Ergebnisse der NNA für die drei Datensätze sind der nachfolgenden Tabelle 3 zu entnehmen.

Tabelle 3: Ergebnisse der Nearest-Neighbour-Analyse

Datensatz	Durchschnittliche Nearest Neighbour Distanz	NNA-Index	Signifikanzniveau einseitig/zweiseitig
Straßenabschnitt	80.12 m	0.6781	0.0001/0.0001
Einbrüche 2003	51.44 m	0.2799	0.0001/0.0001
Diebstähle 2003	7.20 m	0.0978	0.0001/0.0001

Der Nearest Neighbour Index beschreibt das Verhältnis zwischen der durchschnittlichen Nearest Neighbour Entfernung und der erwarteten Entfernung. Ein NNA-Index von 1 weist darauf hin, dass die Punktdaten zufällig verteilt sind. Wird der Index kleiner als 1, dann deutet er auf eine gehäufte Verteilung der Punktdaten hin; ein Index größer als 1 steht für eine gleichmäßige Verteilung der Punkte im Raum (vgl. Walford (1996), S. 358). Aus der Tabelle ist zu entnehmen, dass alle Datensätze eine gehäufte Verteilung aufweisen. Da auch die Straßenabschnitte eine räumliche Konzentration anzeigen, sind also die Bezugspunkte für die Diebstahl- und Einbruchdaten konzentriert. Wie wirkt sich die Verteilung der Bezugspunkte auf die räumliche Verteilung der Einbruch- und Diebstahldaten aus?

Entsprechend der eingehenden Erläuterungen zum Darstellungsproblem der Punktdaten in GISPOL erklärt sich auch die zunehmende Konzentrierung in den beiden Datensätzen. Die Einbruchs- und Diebstahlereignisse beziehen sich auf die Straßenabschnitte. Da an einigen Straßenabschnitten mehrere Straftaten übereinander liegen, ergibt sich eine Nearest-Neighbour-Distanz mit dem Wert 0 für diese Ereignisse. Dieser Wert beeinflusst die Berechnung der Nachbarschaftsdistanzen und sorgt dafür, dass die

Einbruchs- und Diebstahldaten stärkere räumliche Punkthäufungen erkennen lassen. Aus den Ergebnisse lässt sich schließen, dass mehr Diebstähle als Einbrüche an Straßenabschnitten übereinander liegen. Nun stellt sich die Frage, wo diese Häufungen vorkommen und wie die lokalen Häufungen besser dargestellt werden können.

4.2 Flächenhafte Darstellung der Straftatenverteilung

Eine Alternative zur Überlagerung von Punkten besteht in der Projektion der Punktdaten in die Fläche. Die Vorteile, die sich durch dieses Verfahren ergeben, wurden bereits im ersten Teil dieser Arbeit (vgl. Teil I, Kapitel 4.1) genannt.

Die Art der zur Verfügung stehenden Daten bietet zwei Darstellungsalternativen an: zum einen die Darstellung der Straftatenverteilung in einer Choropletenkarte mit einem Bezug auf diskrete Flächen und zum anderen die Berechnung einer statistischen Oberfläche mit Hilfe der Kernel-Density-Estimation (vgl. Teil I, Kapitel 4.2.3) oder des Potentialansatzes, der später zu erläutern sein wird.

4.2.1 Darstellung der Straftaten in diskreten Flächen

Als diskrete Flächen werden solche Flächen bezeichnet, die sich klar durch eine definierte Grenze von den Nachbarflächen abheben. Solche Flächen werden in den zwei nachfolgend beschriebenen Ansätzen zur Darstellung der Straftatenverteilung eingesetzt. Der erste Ansatz entspricht der Quadrat-Methode, wie er bei Haggett (1977, S. 276f.) beschrieben ist. Im zweiten Ansatz werden aus den Straßenabschnitten Thiessen-Polygone erzeugt und diese als Projektionsflächen zur Darstellung der Straftaten genutzt.

4.2.1.1 Die Quadrat-Methode

Über die Hamburger Stadtfläche wird ein gleichmäßiges quadratisches Raster gelegt. Für jede Rasterzelle wird die Anzahl der Straftaten ausgezählt, die innerhalb dieser Zelle liegen. Das Ergebnis wird grafisch aufbereitet und dargestellt. Die Abbildungen 12 und 13 zeigen die räumliche Verteilung der Diebstähle und Einbrüche als Ergebnis dieser Methode.

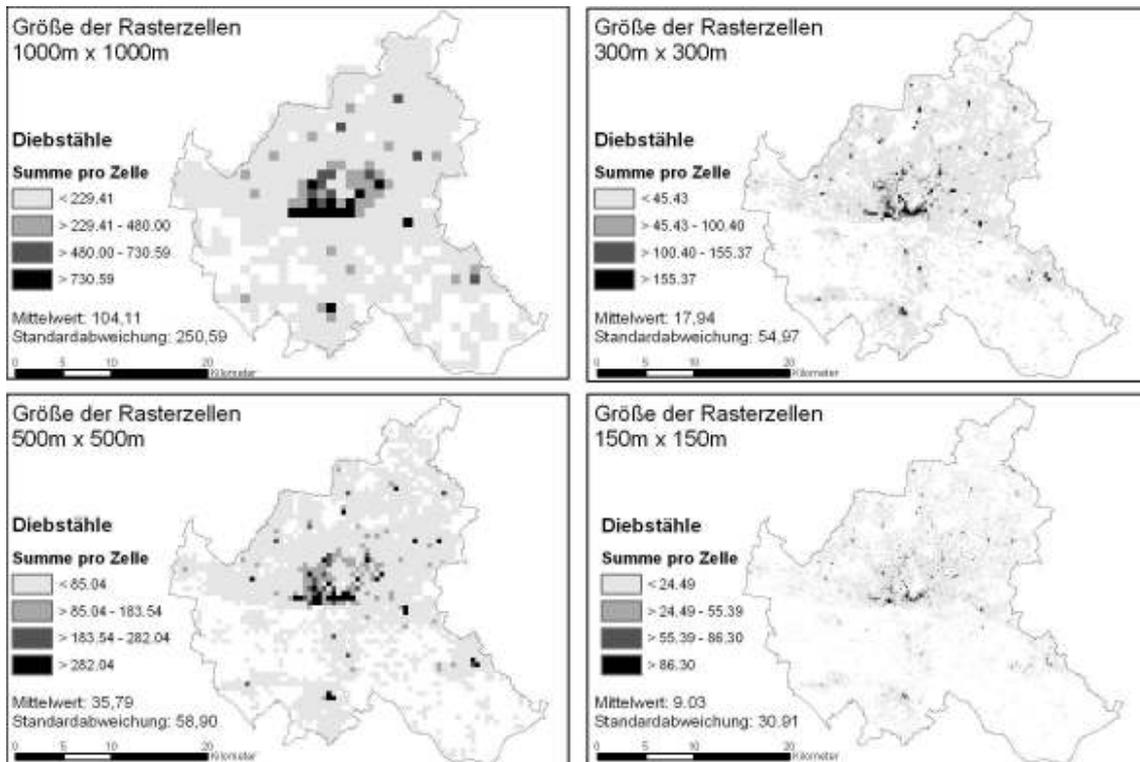


Abbildung 12: Darstellung der Diebstahlsverteilung in Abhängigkeit von der Größe der Rasterzellen

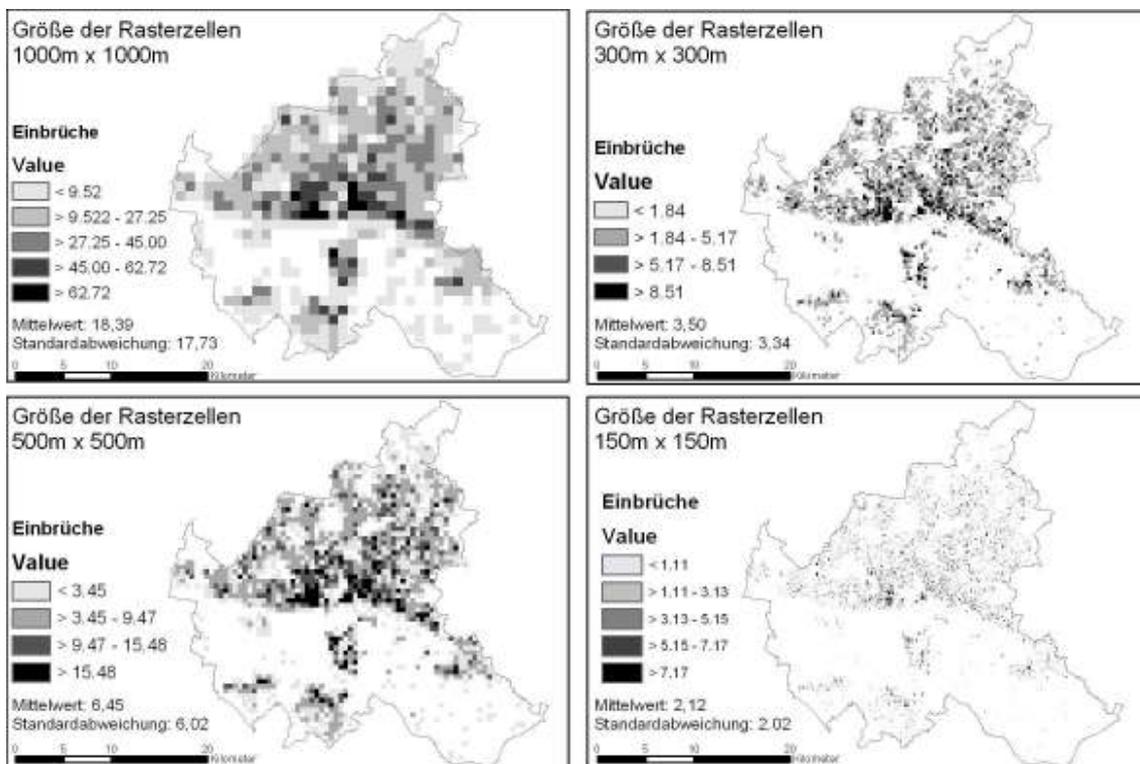


Abbildung 13: Darstellung der Einbruchverteilung in Abhängigkeit von der Größe der Rasterzellen

Mit Hilfe dieser Darstellung lässt sich die räumliche Verteilung der Straftaten nachvollziehen. Wie aus den Abbildungen 12 und 13 erkennbar ist, zeigen die Diebstähle des Jahres 2003 ein vollkommen anderes räumliches Muster als die Einbrüche desselben Jahres. Dennoch lassen sich sowohl für die Einbrüche, als auch für die Diebstähle Regionen mit einem höheren Straftatenaufkommen identifizieren. Ausschlaggebend für das dargestellte Ergebnis ist die Kantenlänge der Rasterzellen. Bei einem Vergleich dieser Kantenlängen lässt sich folgendes feststellen:

1. Die Quadrate mit der größten Kantenlänge zeigen größere Werte-Häufigkeiten. Da das Raster jedoch ziemlich grob ist, ist die Aussagekraft in Bezug auf den betroffenen Raumes relativ gering.
2. Die Quadrate mit einer mittleren Kantenlänge zeigen ebenfalls Bereiche, mit einer hohen Anzahl an Straftaten. Da diese Raster einen kleineren Raum abdecken, ist hier der räumliche Bezug besser herstellbar.
3. Im Gegensatz dazu besitzen die Rasterzellen mit der geringsten Kantenlänge (150 x 150 m) kaum noch eine Aussagekraft in Bezug auf die räumliche Verteilung der Straftaten. Das Bild, das sich aus diesem Raster ergibt, ähnelt bereits dem Bild der Punktverteilung.

Dieses Ergebnis entspricht dem, was bereits Haggett (1977, S. 277) beschrieben hat. Der Versuch, mit Hilfe der Quadratmethode eine räumliche Aussage in Bezug auf die Verteilung der Daten zu treffen, wird durch die Größe der Quadrate beeinflusst. Diese Rastergröße bestimmt das Aussehen des zu erzielenden Ergebnisses und die darauf aufbauende Interpretation.

4.2.1.2 Thiessen-Polygone als Bezugsflächen der Straftatenverteilung

Die Darstellung der Straftaten in Hamburg orientiert sich an den Straßenabschnitten. Die Straßenabschnitte werden als Punkte dargestellt. Dennoch lassen sich aus ihnen Flächen generieren.

Die Verteilung der Straßenabschnitte hat zur Folge, dass zwischen den Straßenabschnitten Bereiche existieren, die genau einen Straßenabschnitt als nächsten Nachbarn haben. Diese Bereiche stellen Flächen dar. Sie lassen sich aus der Punktverteilung ableiten und werden Thiessen-Polygone oder Voroni-Diagramme genannt. Dieses sind die häufigsten Namen, dennoch werden in der Literatur andere Bezeichnungen genutzt (vgl. Okabe et al. (1992), S. 6ff).

Die Thiessen-Polygone wurden mit Hilfe des Programmes ArcView (3D-Analyst) der Fa. ESRI aus den Straßenabschnitten ermittelt und mit der Stadtgrenze Hamburgs verschnitten. Auf diese Weise blieben nur diejenigen Straßenabschnitte und die Polygonflächen zurück, die sich innerhalb der Stadtgrenze Hamburgs befanden. Es stellte sich heraus, dass die Punkte der Grenz-Straßenabschnitte sich in den meisten Fällen genau auf der Stadtgrenze befanden. Die Verschneidung ließ deshalb nur einen Teil der Polygonflächen übrig. Hierbei handelte es sich um den Teil der Fläche, der sich innerhalb der Stadtgrenze befand. Dies führte zu einer Flächenreduzierung bei den Randpolygonen.

Wären jedoch die Thiessen-Polygone nicht mit der Stadtfläche verschnitten worden, hätte es ein anderes Problem gegeben. Den Polygonen an der Stadtgrenze wäre eine unverhältnismäßig große Fläche zugewiesen worden, da eine Begrenzung dieser Flächen nur noch durch das umschließende Rechteck des Punktdatensatzes existiert hätte. Die Randpolygone hätten sich also in diesem Fall bis zum Rand des umschließenden Rechtecks erstreckt. Auch hier existiert also eine Art „Randeffekt“ wie er für verschiedene Methoden der Point-Pattern-Analyse beschrieben wurde.

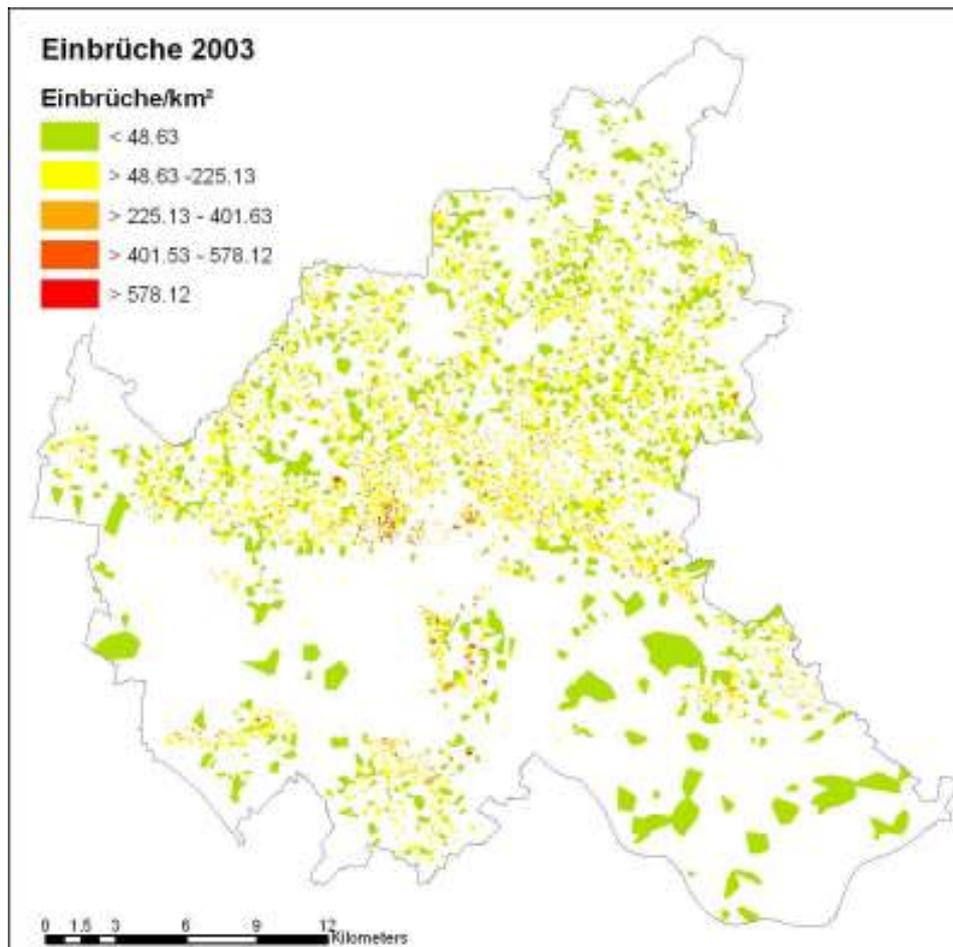


Abbildung 14: Räumliche Verteilung der Einbrüche des Jahres 2003 im Hamburger Stadtgebiet (bezogen auf Straßenabschnittsflächen, Mittelwert 136,88, Standardabweichung 176,5)

Ein Straßenabschnitt, der offensichtlich noch zum Hamburger Stadtgebiet gehören sollte, da an diesem Straßenabschnitt Straftaten festgestellt wurden, befand sich hinter der Hamburger Stadtgrenze. Dieser – und somit auch die an ihm hängenden Straftaten – ging den nachfolgenden Untersuchungen verloren. Nachdem die Thiessen-Polygone der Straßenabschnitte erstellt waren, wurden die Einbrüche, bzw. Diebstähle, die sich innerhalb der jeweiligen Polygonflächen befanden, durch einen Spatial-Join bzw. durch eine räumliche Verschneidung ausgezählt, und den Straßenabschnitten als Attribut hinzugefügt. Die räumliche Verteilung der Einbrüche und Diebstähle des Jahres 2003 zeigen die Abbildungen 14 und 15.

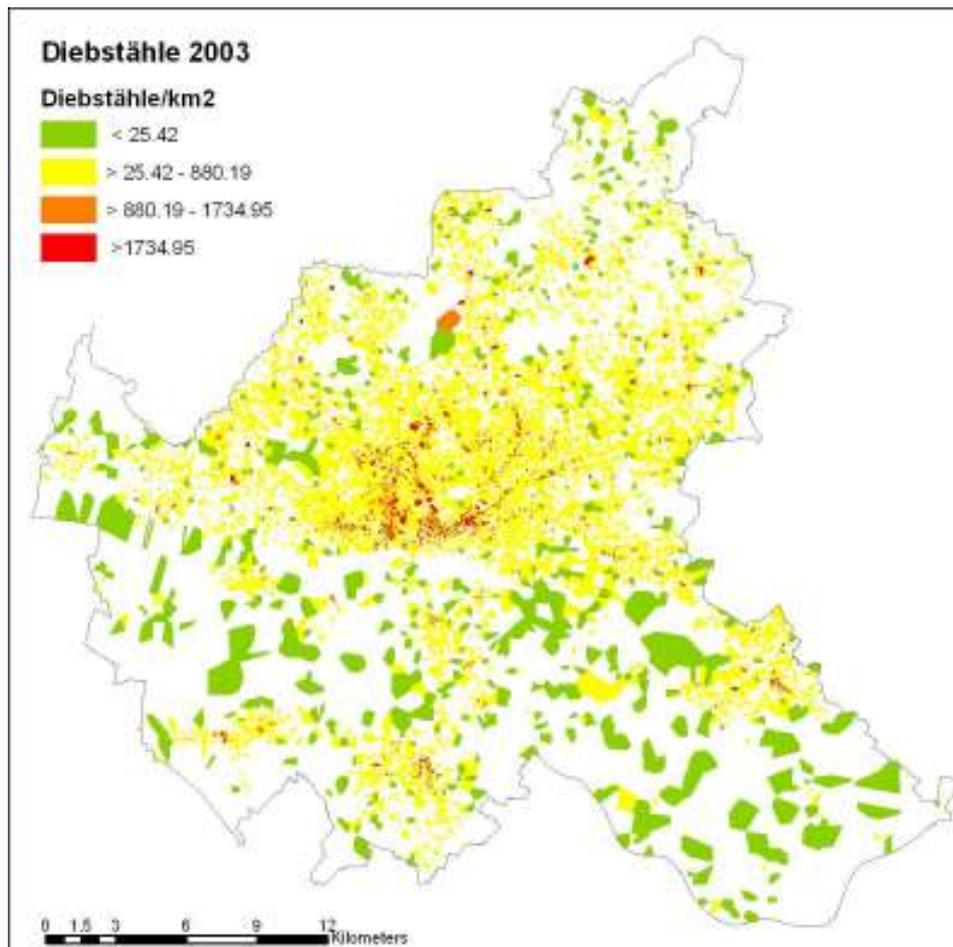


Abbildung 15: Räumliche Verteilung der Diebstähle des Jahres 2003 im Hamburger Stadtgebiet (bezogen auf Straßenabschnittsflächen, Mittelwert 452,9, Standardabweichung 854,77)

Die Größe der Thiessen-Polygone ergibt sich aus der räumliche Verteilung der Straßenabschnitte. Dort, wo viele Straßenabschnitte konzentriert vorliegen, sind die Flächen der Thiessen-Polygone kleiner als dort, wo die Straßenabschnitte weit verstreut liegen. Werden nun die Straftaten auf diese Flächen bezogen dargestellt, ist die Größe der jeweiligen Teilflächen in den Vergleich der Straftatenhäufungen einzubeziehen. Aus diesem Grund werden keine absoluten Werte, sondern die relativen Häufigkeiten der Straftaten (Einbrüche/km², Diebstähle/km²) dargestellt.

Der Vorteil dieser Darstellung besteht darin, dass nicht nur die Verteilung der Straftaten, sondern auch die Beziehungen der Flächen zueinander dargestellt werden. So sind im Innenstadtbereich viele kleine Straßenabschnitte mit einer hohen Straftatendichte erkennbar. Drüber hinaus bilden die Flächen hier eine annähernd geschlossene

Oberfläche. In den Vier- und Marschlanden (im Südosten des Stadtgebietes) zeigen sich große Flächen mit einer geringen Straftatendichte. Darüber hinaus treten hier nur einzelne „Inseln“ auf, an denen Straftaten festgestellt wurden.

Problematisch an dieser Darstellung ist jedoch, dass diejenigen Gebiete, die die höchsten Dichten aufweisen – nämlich die Innenstadtbereiche – aufgrund der hohen Straßenabschnittsdichte in der Darstellung nicht deutlich genug zum Vorschein kommen. Sie werden erst erkennbar, wenn der Betrachter diesen Ausschnitt vergrößert, bzw. in den Innenstadtbereich hinein zoomt. Dann jedoch geht der Gesamteindruck verloren.

4.2.1.3 Die Ergebnisdarstellung – das Problem der „richtigen“ Klassifizierung

Die Ergebnisse, die mit Hilfe der bereits vorgestellten Methoden erzielt wurden, bzw. durch die später noch vorzustellenden Methoden, erzielt werden, sollen kartographisch aufbereitet und dargestellt werden. In der Regel werden nicht alle ermittelten Einzelwerte präsentiert, sondern es werden Ergebnisklassen gebildet und diese werden dargestellt. Da es jedoch diverse Möglichkeiten der Klassifizierung gibt, stellt sich die Frage nach der geeigneten Klassifizierung für die Darstellung der Straftatenverteilung. Die Wahl der Klassifizierungsmethode hat einen erheblichen Einfluss auf die Ergebnispräsentation und somit auch auf die Schlüsse, die aus dieser Darstellung gezogen werden.

Das Ziel der Klassifizierung sollte immer darin bestehen, die räumliche Verteilung der dargestellten Daten so weit wie möglich zu erhalten (vgl. Kraak & Ormeling (2003), S. 116). Also sollte die Werteverteilung so klassifiziert werden, dass ähnliche Werte einander entsprechen und unähnliche Werte sich möglichst stark unterscheiden. Die Klassifizierung der in dieser Arbeit erstellten Daten wurde mit Hilfe des Programms ArcView der Firma ESRI durchgeführt. Diese Programm bietet fünf verschiedene Klassifizierungsverfahren an. Die Auswirkungen der gewählten Klassifizierungsmethoden auf die Ergebnisdarstellung und auf den dadurch gewonnenen Gesamteindruck der Daten wird nachfolgend am Beispiel der Diebstahldichten der Thiessen-Polygon vorgestellt.

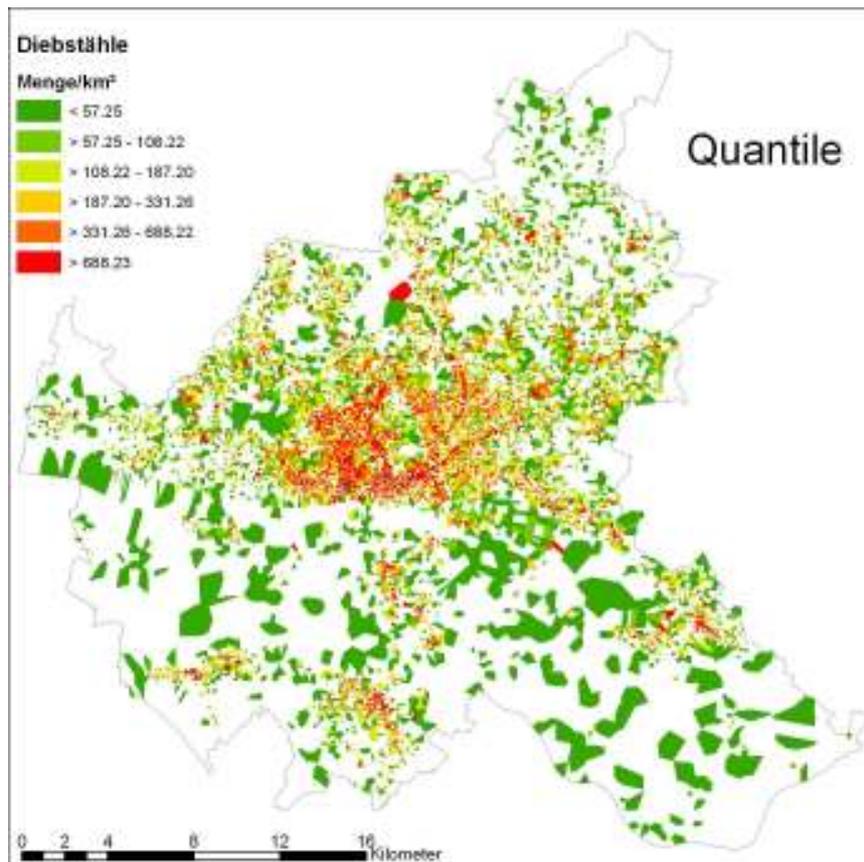


Abbildung 16: Diebstahldichten pro Straßenabschnittspolygon - Quantile

Die Abbildung 16 zeigt die Verteilung der Diebstähle in der Quantil-Klassifizierung. Es wurden 6 Klassen (Hexile) gebildet. In jeder dieser 6 Klassen befindet sich die gleiche Menge an Werten. Eine gerade Klassenzahl wurde gewählt, da auf diese Weise der Median (der Wert, der genau in der Mitte des Datensatzes liegt) erkennbar wird. Dieser Wert bildet die Klassengrenze zwischen der dritten und der vierten Klasse. Da allen Klassen dieselben Anzahl an Werten besitzen und die Daten L-förmig verteilt sind, haben die ersten 3 Klassen relativ geringe Klassenweiten, wogegen die Klassenbreite der vorletzten Klasse bereits 6 mal größer ist als die Klassenbreite der ersten Klasse. Da das Gesamtbild eine konzentrische Abnahme der Daten vom Stadtkern an die Ränder darstellt, häufen sich im Zentrum die hohen Werte. Besondere Schwerpunkte sind im Zentrum nicht erkennbar. Das ist problematisch.

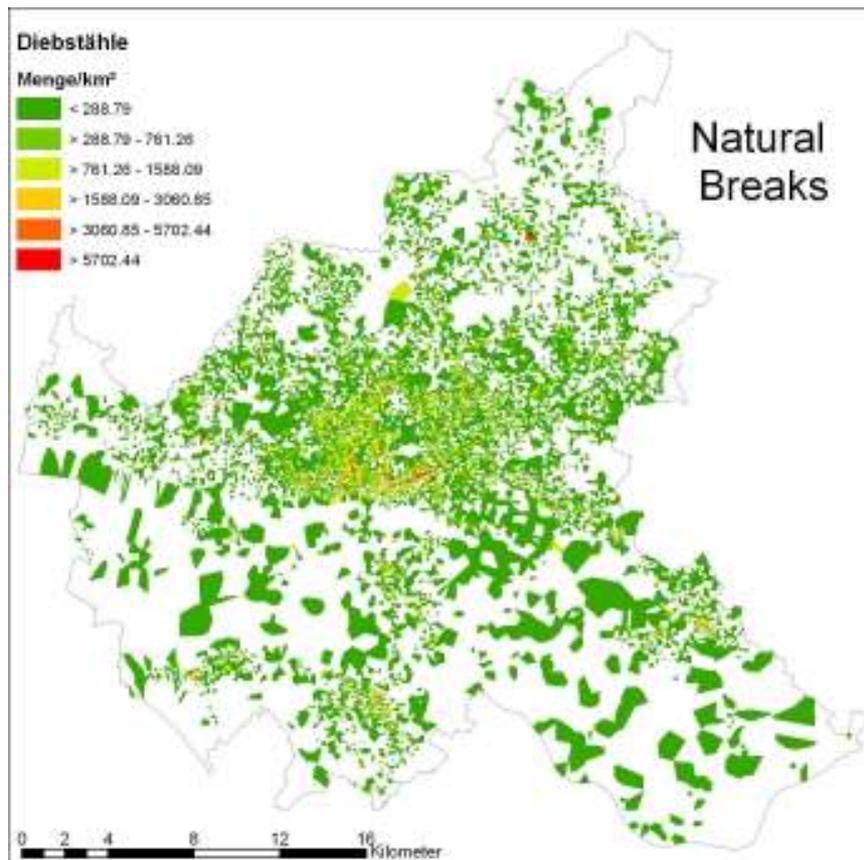


Abbildung 17: Diebstahldichten pro Straßenabschnittspolygon - Natural Breaks

Die Abbildung 17 zeigt die Klassifizierung mit Hilfe der Natural-Breaks-Methode. Wiederum wurden 6 Klassen gebildet. Dieses Klassifizierungsverfahren wählt die Klassengrenzen derart, dass die Variabilität der Werte innerhalb der Klassen minimiert, aber zwischen den Klassen maximiert wird. Auf diese Weise erhält man homogene Klassen. In Bezug auf die Darstellung der Diebstähle in Hamburg liefert diese Klassifizierungsmethode jedoch auch ein relativ homogenes Gesamtbild mit wenigen Schwerpunkten. Die Flächen mit hohen Diebstahldichten fallen kaum auf. Es stellt sich die Frage, ob ein solches Bild bei der Suche nach Schwerpunkten hilfreich sein kann, oder ob diese Darstellung die Gesamtsituation nicht zu positiv wiedergibt.

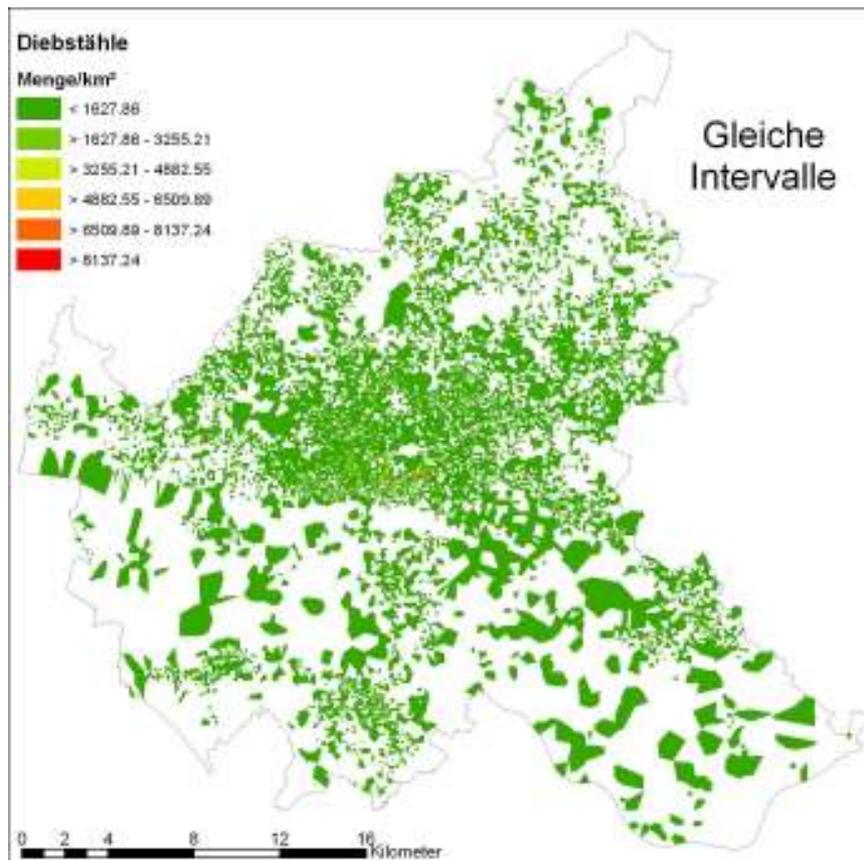


Abbildung 18: Diebstahldichten pro Straßenabschnittspolygon - gleiche Intervalle

Als dritte Klassifizierungsart sei in der Abbildung 18 die Variante zur Bildung gleicher Klassenintervalle vorgestellt. Auch hier wurden wieder 6 Klassen erzeugt. Bedingt durch die L-förmige Verteilung der Werte fallen jedoch fast alle Werte bereits in die erste Klasse. In die übrigen Klassen gelangen nur noch einzelne Werte. Auch diese Darstellung zeigt – ebenso wie das Natural-Breaks-Verfahren – ein relativ einseitiges Bild. Eine Mustererkennung ist hier nicht möglich.

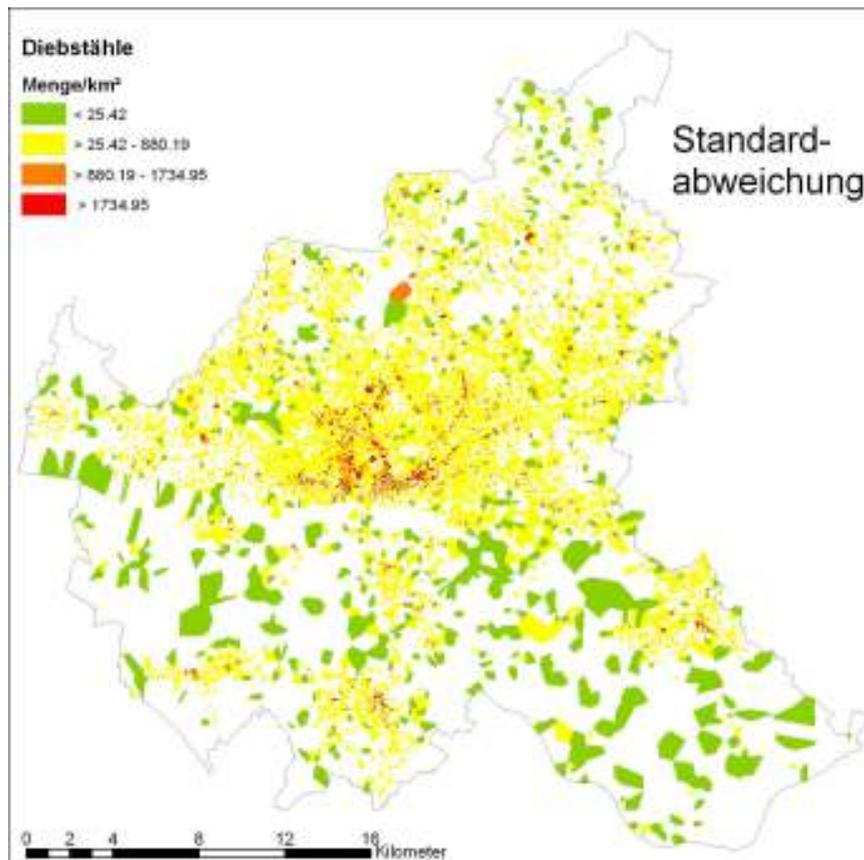


Abbildung 19: Diebstähle pro Straßenabschnittspolygon - Standardabweichung

Als letzte Variante wird mit der Abbildung 19 das Klassifizierungsverfahren der Standardabweichung behandelt. Aus allen Attributwerten werden Mittelwert und Standardabweichung berechnet. Die Standardabweichung beschreibt dann das Intervall zwischen zwei Klassengrenzen. Der Anwender kann festlegen, wie groß die Klassenintervalle sein sollen (eine Standardabweichung, $\frac{1}{2}$ Standardabweichung oder $\frac{1}{3}$ Standardabweichung). Die nachfolgenden Erläuterungen gehen von einem Klassenintervall mit einer Standardabweichung aus. Die Klassengrenzen werden vom Mittelwert ausgehend nach links und nach rechts definiert. Der Mittelwert bildet das Zentrum der ersten Klasse, weshalb die Grenzen dieser Klasse anschließend um $\frac{1}{2}$ Standardabweichung nach links und um $\frac{1}{2}$ Standardabweichung nach rechts vom Mittelwert entfernt gesetzt werden. Die weiteren Klassen schließen sich an. Diese Klassifizierung führt dazu, dass der größte Teil der Daten in den Bereich um den Mittelwert herum, bzw. links vom Mittelwert eingruppiert wird. Die den Leser interessierenden Abweichungen, die rechts vom Mittelwert liegen, sind noch in hinrei-

chender Zahl vorhanden. Aus diesem Grund eignet sich diese Klassifizierung gut zur Darstellung derjenigen Gebiete, die unter einem übermäßig hohen Aufkommen an Straftaten leiden.

Harries (1999, S. 54) beantwortet in seiner Arbeit die Frage nach der im Crime-Mapping einzusetzenden Klassifizierungsmethode folgendermaßen: „Generally, natural breaks or equal intervals will be the best methods for creating area-type maps in crime analysis“. Weiter unten zitiert er jedoch McEachren, der 1994 folgendes zur Nutzung der Standardabweichung gesagt hat „... Standard deviation classes may be helpful in some situations where the distribution is not extremely skewed“ (vgl. Harries (1999), S. 54). Für die Darstellung der vorliegenden Daten wird in den Fällen, in denen nicht explizit darauf hingewiesen wird, das Verfahren der Standardabweichungen eingesetzt, da es sich als das Verfahren erwiesen hat, mit dem die vorliegenden Daten graphisch am besten erläutert werden können.

4.2.2 Straftatenanalyse durch die Kernel-Density-Estimation (KDE)

Das Verfahren der Kernel-Density-Estimation wurde bereits im ersten Teil dieser Arbeit vorgestellt (vgl. Teil I, Kapitel 4.2.3). Diese Methode stellt ein Verfahren zur Glättung von Extremereignissen dar. Sie wird sowohl zur Untersuchung räumlicher sowie zeitlicher Verteilungen eingesetzt. Dieses Glättungsverfahren lässt vorhandene Trends hervortreten.

Das Verfahren der Kernel-Density-Estimation ist in die Spatial-Analyst Extension des Programms ArcView der Firma ESRI integriert und lässt sich auch mit Hilfe des Programmes CrimeStat berechnen. Das Ergebnis der durchgeführten Berechnung liefert eine statistische Oberfläche in Form eines quadratischen Rasters, die die geschätzte relative Straftatenhäufigkeit pro Flächeneinheit wiedergibt.

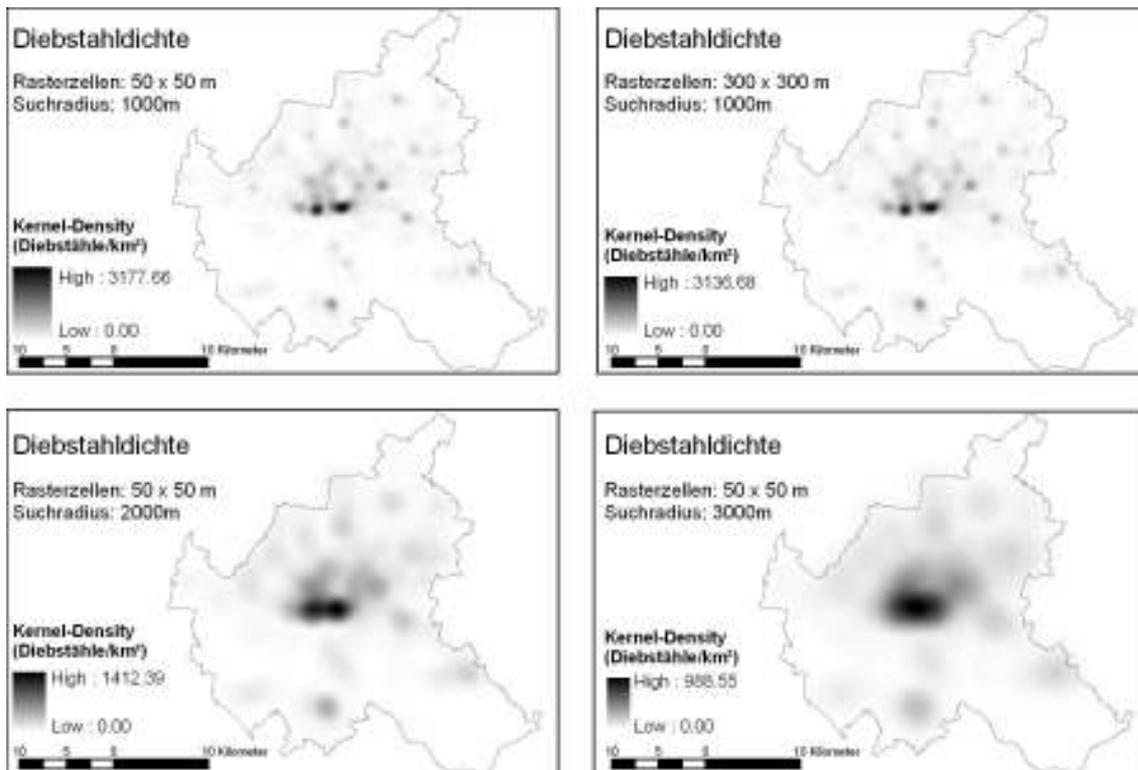


Abbildung 20: Einflussgrößen, die auf die Ergebnisse der Kernel-Density Estimation wirken. (Dargestellt für die Diebstähle des Jahres 2003)

Bevor die Kernel-Berechnung durchgeführt werden kann, definiert der Benutzer die Größe der zugrunde liegenden Rasterzellen, sowie einen Distanzparameter, der den Suchradius des Kreises vorgibt. Dieser Kreis definiert den Bereich, innerhalb dessen alle angetroffenen Ereignisse in die Werteermittlung eingehen. Dabei nimmt der Einfluss eines Ereignisses auf das Ergebnis der Kernelberechnung mit zunehmender Distanz vom Ausgangspunkt der Berechnung ab. Die Größe des durch den Anwender zu bestimmenden Suchradius hat einen erheblichen Einfluss auf das Endergebnis: je größer der Suchradius, desto höher die geschätzten Dichten, desto stärker nähern sich benachbarte Minimal- und Maximalwerte an und desto geringer ist die Chance, Extremereignisse zu erkennen. Je kleiner der Suchradius wird, desto stärker treten die Extremwerte hervor und desto schwieriger wird es, Gemeinsamkeiten in den Daten zu erkennen. Die Abbildung 20 stellt verschiedene Ergebnisse der Kernel-Density-Estimation für die Diebstahldaten des Jahres 2003 dar. Es zeigt sich, dass mit zunehmendem Suchradius die Schwerpunkte sich ausweiten, bzw. ineinander verschmel-

zen. Gleichzeitig reduziert sich die berechnete Straftatendichte pro Rasterzelle. Auch die Größe der Rasterzellen übt einen Einfluss auf die Berechnung der Straftatendichten aus.

4.2.3 Darstellung der Straftaten mit Hilfe des Potentialansatzes

Die Potentialanalyse wurde bisher überwiegend in kleinmaßstäbigen Untersuchungen zur Berechnung des Bevölkerungspotentials (z. B. für die USA) eingesetzt. In der Literatur wurde dieser Ansatz bisher nicht als Mittel zur Untersuchung der räumlichen Straftatenverteilung vorgestellt. Deshalb ist es interessant zu erfahren, ob und inwieweit sich dieses Verfahren für die explorative Analyse der Straftatenverteilung eignet.

4.2.3.1 Das Verfahren der Potentialberechnung

Nach Pooler (1987, S. 269) stellt die Potentialanalyse „... a conceptually clear and straightforward index of relative location...“ dar. Schumacher (1989, S. 130) betrachtet das Ergebnis der Potentialanalyse als ein „...Potentialfeld ..., welches von einem bestimmten Punkt ausgeht und dessen Anziehungskraft in der Umgebung charakterisiert.“ In Bezug auf das Crime Mapping soll die Potentialanalyse also die für die jeweiligen Straftaten attraktivsten Orte (gemessen an der beobachteten Verteilung dieser Straftaten) herausfiltern. Die Potentialanalyse lässt sich maßstabsunabhängig einsetzen, da das Potential „... at any spatial scale...“ (vgl. Pooler (1987), S. 270) berechnet werden kann.

Die theoretischen Grundlagen dieses Ansatzes entstammen der physikalischen Gravitationslehre (vgl. Kraak & Ormeling (2003) S. 113f.). Das Potential wird nach der folgenden Formel berechnet (vgl. Schumacher, (1987), S. 161).

$$P_i = \sum_{j=1}^n \frac{m_j}{d_{ij}^b}, d_{ii} > 0 \quad (7)$$

P_i	-	Potential des Ortes i
m_j	-	Merkmalswert des Ortes j
d_{ij}	-	Distanz zwischen den Orten i und j
b	-	Exponent der Distanz
n	-	Anzahl der betrachteten Orte

Einen Sonderfall stellt in dieser Gleichung die Berechnung des so genannten „Eigenpotentials“ der jeweiligen Orte i dar (gemeint ist hier der Term m_i / d_{ii}^b). Da

dieses Eigenpotential in die Berechnung des Potentialwertes eingeht, ist folgendes zu beachten: Die Distanz d_{ii} ist in diesem Fall Null. Blicke es bei diesem Wert, wäre die Gleichung nicht lösbar. Aus diesem Grund wird für d_{ii} ein anderer Wert gewählt, der ≥ 1 ist. Sehr oft wird $d_{ii} = 1$ gesetzt. Auf diese Weise entspricht das Eigenpotential des Ortes dem Merkmalswert m_i des Ortes i . Diese Behandlung des Eigenpotentials führt jedoch zu einer Überbewertung des Eigenpotentials im Verhältnis zu den übrigen Werten (vgl. Pooler (1987), S. 274). Aus diesem Grund werden in der Literatur auch andere Ansätze zur Berechnung des Eigenpotentials beschrieben (Pooler (1987), S. 277ff., Kraak & Ormeling (2003), S. 115).

Die Quantifizierung des Exponenten b stellt eine weitere Herausforderung in der Anwendung dieser Methode dar. In klassischen Modellen der Potentialanalyse wird $b = 1$ gesetzt. Dies kann dazu führen, dass ein dominierendes Zentrum zu stark in seine Nachbarschaft ausstrahlt und auf diese Weise die Variabilität in seiner unmittelbaren Umgebung überdeckt. Mit zunehmendem b nimmt der Einfluss der Entfernung auf die Potentialberechnung zu und lokale Variabilitäten sind leichter zu identifizieren. Die Schwierigkeit besteht darin, einen geeigneten Wert für b zu finden.

Im Rahmen der vorliegenden Untersuchung wurde getestet, ob diese Methode einen Beitrag zur explorativen Untersuchung der Straftatenverteilung in Hamburg leisten kann. Es wurden zwei Ansätze ausprobiert. In beiden Ansätzen wird der Wert des Exponenten $b = 1,5$ gesetzt.

4.2.3.2 Punktbezogene Potentialberechnung auf der Basis der Straßenabschnitte

Dieses Verfahren stellt die Potentialberechnung nach dem Punktprinzip (vgl. Schumacher (1989), S. 131) dar. Für jeden Straßenabschnittspunkt wurde das Straftatenpotential berechnet. Die Darstellung der Ergebnisse geschieht durch eine Übertragung der Potentialwerte in die Thiessen-Polygone, die sich aus den Straßenabschnitten ableiten. Auf diese Weise lässt sich die Darstellung des Straftatenpotentials exakt auf das Stadtgebiet von Hamburg beschränken.

Die Potentialberechnung kann nur unter Einbeziehung des Eigenpotentials erfolgen. Dieses ergibt sich für diejenigen Straßenabschnitte, an denen Straftaten festgestellt

wurden. Der Distanzwert „Straßenabschnittskordinate – Koordinate des Ereignisses“ ist in diesem Fall gleich Null. Wird nun die Distanz $d_{ii} = 1$ gesetzt, dann wird das Eigenpotential vieler Straßenabschnitte überschätzt: Es findet keine Glättung statt, sondern der Absolutwert der Straftaten an dieser Stelle setzt sich durch und das Potential dieser Fläche ist im Vergleich zu ihren Nachbarflächen zu dominant.

Um dieses zu verhindern, wird die Berechnung des Eigenpotentials angelehnt an das bei Kraak & Ormeling ((2003), S. 114) beschriebene Verfahren. Wenn die Koordinate des Straßenabschnittes mit der Koordiante des aktuellen Ereignisses zusammenfällt, dann wird der Distanzwert folgendermaßen ermittelt.

1. Das zu diesem Straßenabschnitt gehörende Thiessen-Polygon wird gesucht.
2. Die Fläche des Polygons wird ermittelt und einer Kreisfläche gleichgesetzt.
3. Aus dieser Fläche wird der Radius des Kreises ermittelt.
4. Der Radius wird nun als Distanzfaktor d_{ii} in den Term m_i/d_{ii}^b eingesetzt.

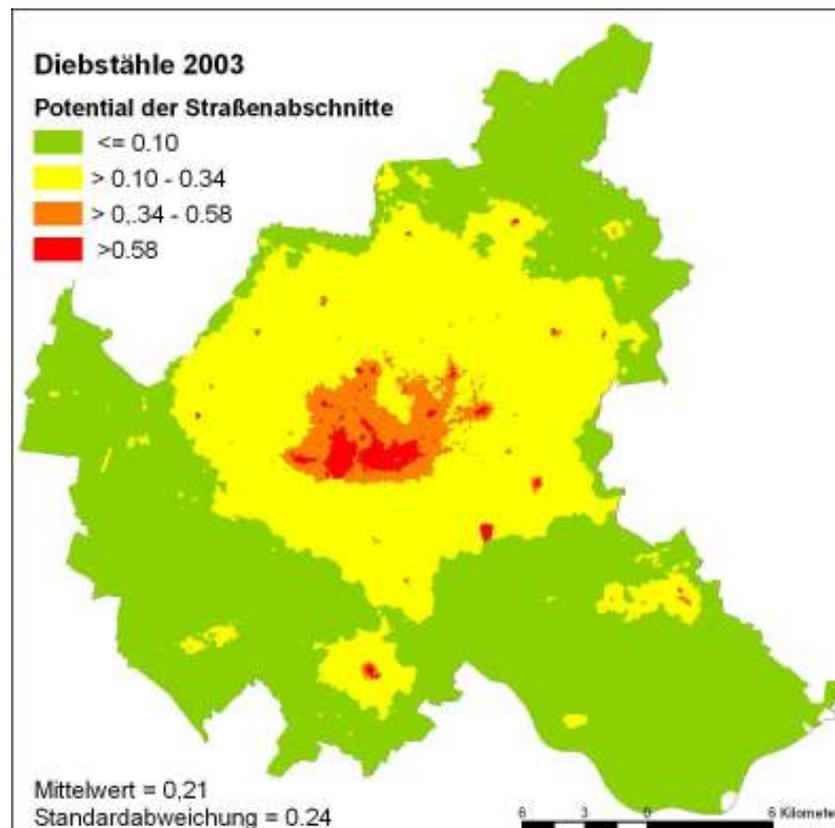


Abbildung 21: Das Diebstahlpotential bezogen auf die Straßenabschnittsflächen

Das Ergebnis dieser Berechnung liefert das Potential für jeden Straßenabschnittspunkt. Zur Darstellung der Potentiale wird nun auf die Thiessen-Polygone zurückgegriffen. Die Potentiale jedes Punktes werden in die Fläche des zu diesem Punkt gehörenden Thiessen-Polygons übernommen. Diese werden anschließend in Form einer Choropletenkarte dargestellt. Die Klassifikation wird wiederum mit Hilfe der Standardabweichungen durchgeführt, da auch hier eine L-förmige Verteilung der Daten vorliegt. Die Ergebnisse dieser Berechnungen zeigen die Abbildungen 21 (Diebstahlpotential) und 22 (Einbruchpotential).

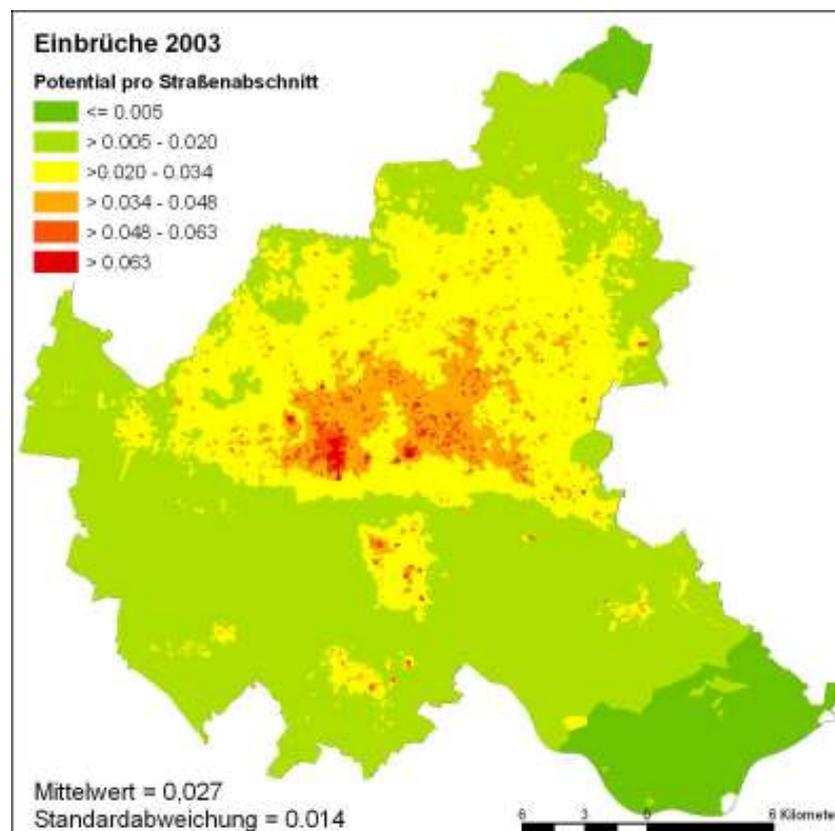


Abbildung 22: Das Einbruchpotential bezogen auf die Straßenabschnittsflächen

4.2.3.3 Die Berechnung der Potentiale für Rasterflächen

Im Gegensatz zur punktbezogenen Potentialberechnung wird hier ein (quadratisches) Punktraster über das Untersuchungsgebiet gelegt. Als Bezugspunkte der Potentialberechnung werden nun nicht mehr die Beobachtungspunkte als solche, sondern die Rasterpunkte genutzt. Diese Form der Potentialberechnung wurde in Pooler (1987,

S. 280) und in Schuhmacher (1989, S. 130) als Alternative zum Punktansatz vorgestellt.

Dieses Verfahren hat gegenüber dem Punktansatz zwei große Vorteile:

1. Die Schwierigkeiten bei der Berechnung des Eigenpotentials fallen weg, weil die Wahrscheinlichkeit, dass zwei Punkte exakt dieselben Koordinaten besitzen, relativ gering ist (vgl. Pooler (1987), S. 280).
2. Das Raster stellt bereits eine kontinuierliche Oberfläche dar, so dass die Ergebnisse der Potentialberechnung unmittelbar im Anschluss z. B. als „Potentialgebirge“ (vgl. Schumacher (1989), S. 131) dargestellt werden können. Eine nachfolgende Interpolation, wie bei Leykauf (1983), S. 163ff für den Punktansatz vorgeschlagen, ist also nicht mehr erforderlich.

Die Berechnung des Potentials mit Bezugnahme auf eine Rasterfläche geschah unter Einsatz des Programms ArcMap der Fa. ESRI, welches um Methoden zur Durchführung der Potentialberechnung für Rasterflächen erweitert wurde. Das Raster wurde über das gesamte Untersuchungsgebiet ausgebreitet – was bedeutet, dass die Ausdehnung des Rasterlayers der Ausdehnung des zu untersuchenden Shapefiles entspricht.

Der Einsatz eines ESRI-Grids zur Berechnung des Potentials führte zu einer anderen Anordnung der Bezugspunkte, als bei (Pooler (1987), S. 280) beschrieben. Dort sollte das Potential für jeden Kreuzungspunkt von zwei Rasterflächen berechnet werden. Da die ESRI-Grids, die zur Erstellung der Rasterflächen genutzt wurden, eine Rasterzelle jedoch durch den Zellenmittelpunkt repräsentieren, wurde die Potentialberechnung auf diesen Punkt bezogen: Das Potential wird für den Mittelpunkt der Rasterzelle berechnet und dann für die gesamte Rasterzelle dargestellt.

Eine große Herausforderung für diesen Ansatz stellt jedoch die Bestimmung der Rasterzellengröße dar. Je größer jede einzelne Rasterzelle wird, desto weniger Rasterzellen besitzt das gesamte Raster. Da jede Rasterzelle mit jedem Straßenabschnitt, an dem eine Straftat geschehen ist, in Beziehung gesetzt werden muss, beeinflusst die Anzahl der vorhandenen Rasterzellen die Dauer der Potentialberechnung. Diese wächst linear mit der Zunahme der Rasterzellen. Da die Größe einer Rasterzelle jedoch durch die Angabe ihrer Kantenlänge bestimmt wird, wächst die Zahl der Rasterzellen um das vierfache an, wenn die Kantenlängen um die Hälfte reduziert werden.

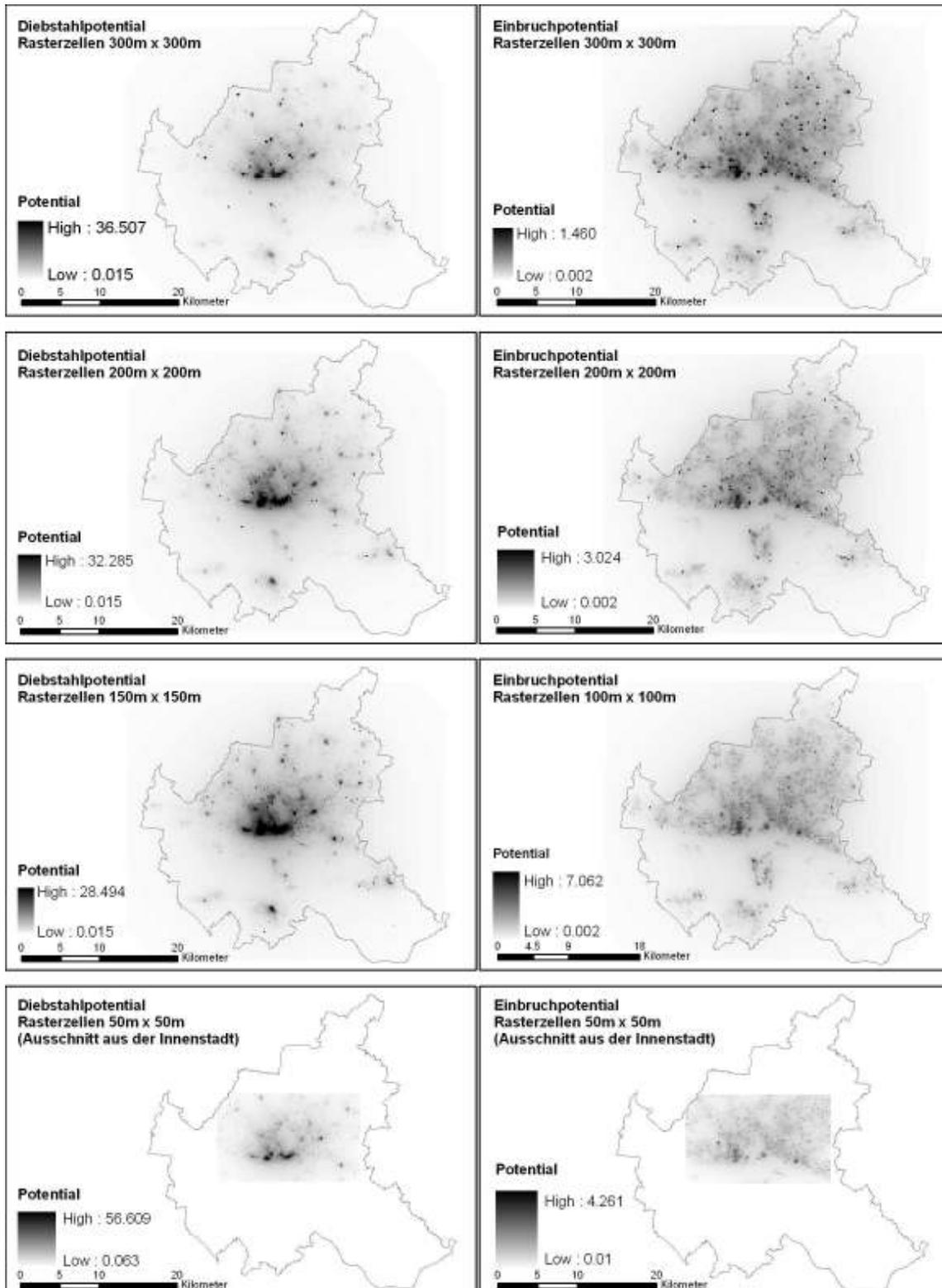


Abbildung 23: Ergebnisvergleich der Potentialberechnung unter Verwendung verschiedener Rasterzellengrößen für Diebstähle und Einbrüche. Darstellung (Standardabweichung, fließende Klassenübergänge)

Sobald das Raster angelegt ist, beginnt die Berechnung des Potentials. Diese geschieht nach der in Kapitel 4.2.3.1 dargestellten Formel (7). Berechnet werden nun die Entfernungen der jeweiligen Ereignisse zu den jeweiligen Rasterpunkten. Da regelmäßig verteilte Rasterzellen nun in Beziehung gesetzt werden zu unregelmäßig verteilten Ereignissen, sind die Entfernungen der jeweiligen Punkte zueinander ungleichmäßig verteilt. So kann es vorkommen, dass für einige Rasterzellen Extremwerte ermittelt werden, weil die jeweiligen Punkte zufällig recht nah beieinander liegen. Diese Verteilung der Extremwerte in der Karte erschwert die Darstellung und die Interpretation des Ergebnisses.

Um zu testen, ob der Anspruch der Maßstabsunabhängigkeit (vgl. Pooler (1987), S. 130) tatsächlich gegeben ist, wird die Potentialberechnung mit Rastern unterschiedlicher Zellengrößen durchgeführt. Es wurden Rasterzellen mit Kantenlängen von 300, 200, 100/150 und 50 Metern eingesetzt. Für das Raster mit den kleinsten Rasterzellen (50m x 50m) konnte die Berechnung nicht mehr für das gesamte Stadtgebiet durchgeführt werden, da die Anzahl der Zellen so groß geworden war, dass die Rechenleistung des verfügbaren Rechners (2,4 GHz, 512 MB RAM) überfordert war. Die Potentialberechnung für dieses Raster wurde deshalb nur im Innenstadtbereich durchgeführt (vgl. Schumacher (1987), S. 130).

Die Potentialberechnung für die verschiedenen Raster ergaben zwar unterschiedliche Werte für die einzelnen Rasterzellen, die Klassifizierung der jeweiligen Werte mit Hilfe der Standardabweichung und die anschließende Darstellung der Klassen zeigte jedoch in allen Fällen eine ähnliche räumliche Verteilung der Hotspots (vgl. Abbildung 23). Dies weist im Vergleich zu den Ergebnissen der Quadrat-Methode (s. Teil II, Kapitel 4.1.1.1.) auf eine maßstabsunabhängige Berechnung des Potentials in der Fläche hin.

4.2.4 Vergleichende Bewertung der vorgestellten Methoden

In den drei vorhergehenden Kapiteln wurden verschiedene Verfahren vor- und gegenübergestellt, mit deren Hilfe es gelingen kann, die punktbezogene Darstellung der Straftaten durch eine flächenbezogene Darstellung abzulösen.

Sollte diese Flächendarstellung in die ArcIMS-Anwendung integriert werden, dann ist davon auszugehen, dass Anwender damit arbeiten, die bisher keine oder nur geringe Erfahrungen mit der die Interpretation räumlicher Daten haben. Da es sich bei GISPOL um eine Intranet-Anwendung handelt, muss weiterhin darauf geachtet werden, dass die Berechnungen entweder wenig Zeit beanspruchen oder die Ergebnisse nicht unmittelbar nach Aufruf dieser Funktionen durch den Anwender geliefert werden müssen. Entweder sind also die Berechnungen bereits durchgeführt worden und die Ergebnisse darstellbar oder sie werden nach „Beantragung“ durch den Anwender ausgeführt und ihm zu einem späteren Zeitpunkt zur Anzeige übermittelt.

In der nachfolgende Tabelle 4 werden Bewertungskriterien für die Beurteilung der vorgestellten Methoden aufgestellt und die Methoden an diesen Kriterien gemessen.

Tabelle 4: Vergleich der verschiedenen Methoden zur Darstellung der Straftaten bezogen auf die Fläche gemessen an einigen Parametern (- gering, + lange, ++ sehr lange)

	Maßstabs-unabhängig?	Wie viele Parameter durch Benutzer zu definieren?	Benötigte Rechenzeit
Quadratmethode	nein	1	+
Thiessen-Polygone	ja	0	+
KDE	nein	2	-
Potential / Raster	ja	1	+
Potential / Thiessen	ja	0	++

Ein Vergleich der verschiedenen Methoden anhand der in der Tabelle 4 vorgegebenen Kriterien lässt die Quadratmethode als Alternative zur Darstellung der Daten in der Fläche ungeeignet erscheinen. Sie erweist sich als maßstabsabhängig und diese Maßstabsabhängigkeit hängt ausgerechnet am einzugebenden Parameter. Die Kantenlänge der Rasterzellen ist durch den Benutzer vorzugeben und diese beeinflusst das Ergebnis.

Besser eignen sich die Thiessen-Polygone, die sich aus den Straßenabschnitten ergeben. Diese Polygone stellen die Dichte der pro Straßenabschnitt ermittelten Straftaten im Verhältnis zu ihrer Flächengröße dar. Die Berechnung des Spatial Join dauert mit ArcView 8.2 recht lange (abhängig von der Anzahl der Straftaten). Sie könnte jedoch in einer Datenbank unmittelbar im Anschluss an die Datenübernahme durchgeführt werden und wäre dann schneller abgeschlossen.

Die Kernel-Density-Estimation liefert relativ eindeutige, d.h. leicht interpretierbare Ergebnisse und bewies unter den vorgestellten Methoden (in ArcView) die kürzesten Rechenzeit. Dennoch wären auch diese Rechenzeiten zu langsam für die Forderung nach einer Ergebnisdarstellung unmittelbar im Anschluss an den Funktionsaufruf. Die Tatsache jedoch, dass diese Methoden mindestens 2 Eingabeparameter erwartet, von denen das Ergebnis abhängt, lässt sie eher für solche Arbeitsplätze geeignet erscheinen, an denen schnell mehrere Ergebnisse nacheinander abgerufen und verglichen werden können (Desktop-GIS). Ein Einsatz der KDE im Intranet der Polizei wäre nur zur Präsentation von Ergebnissen ratsam, die an anderen Rechnern erstellt und aufbereitet wurden.

Die Potentialberechnung liefert interessante Ergebnisse, die den Einsatz dieses Verfahrens zur flächenhaften Darstellung der Straftaten in Hamburg als geeignet erscheinen lassen. Da die Potentialberechnung, die auf der Basis der Straßenabschnitte durchgeführt wird, ohne Parameter auskommt und in der Darstellung auch keine Artefakte – bedingt durch die Verteilung der Bezugspunkte – zu erwarten sein werden, ist dieser Ansatz auch dem Rasteransatz vorzuziehen. Leider benötigt dieser Ansatz jedoch die längsten Rechenzeiten unter allen untersuchten Verfahren. Dies macht den unmittelbaren Einsatz in der Intranetanwendung nicht möglich. Es wäre denkbar, dass die Potentialberechnung nachts in der Datenbank durchgeführt wird und die Daten am folgenden Tag als Layer zur Verfügung stehen.

Sollte jedoch die Potentialberechnung als Mittel zur flächenhaften Darstellung einer Straftatenverteilung eingesetzt werden, stellt sich die Frage, welcher Zeitraum in die Berechnung des Potentials eingehen soll: ein ganzes Jahr, ein Monats, eine Woche, ein Tag oder frei definierbare Zeiträume? Ein weiteres Problem hängt an der zeitlichen Veränderbarkeit der Straßenabschnittsdaten. Die Straßenabschnitte sind zeitlich nicht konstant, sondern können ihre räumliche Lage im Laufe der Zeit verändern. Es kommen neue Straßen(-abschnitte) hinzu. Straßen(-abschnitte) verschwinden oder verlagern ihre Position. In diesem Fall bedarf es einer zeitliche Versionsverwaltung der Straßenabschnitte und der darauf bezogenen Kriminalitätsdaten.

Eine vergleichende Bewertung der jeweiligen Ergebnisse anhand statistischer Kriterien ist leider nicht möglich. Die meisten statistischen Tests vergleichen die Ergebnisse einer Methode im Verhältnis zu einer theoretischen räumlichen Zufallsverteilung der Daten. Es ist statistisch nicht möglich, verschiedene Verfahren gegeneinander zu vergleichen und anhand eines Berechnungskriteriums das beste Verfahren herauszusuchen.

5 Lokale Autokorrelation und Clusterbildung

Im ersten Teil dieser Arbeit wurden bereits die LISA (Local Indicator of Spatial Association) Methoden vorgestellt, die von Anselin (1995) und Getis & Ord (1992) entwickelt wurden. Das Programm ArcView der Firma ESRI stellt nun in seiner Standardausprägung der Version 9.x verschiedene Methoden zur Berechnung räumlich statistischer Parameter (Spatial Statistics) bereit. Darunter finden sich auch die Methoden zur Berechnung des Local Moran I (Anselin, 1995) und des G_i^* -Index (Getis & Ord, 1992).

Die räumliche Verteilung der Straftaten (Einbrüche sowie Diebstähle) wurde deshalb auch durch den Einsatz dieser Methoden untersucht. Mit Hilfe dieser Verfahren sollen diejenigen Datensätze gefunden werden, die ähnliche Werteausprägungen besitzen und räumlich eng benachbart sind. Es handelt sich bei diesen Methoden also um Verfahren zur Aufdeckung räumlicher Cluster bzw. Hotspots.

In der Literatur ist der Einsatz der LISA-Statistiken im Umfeld von Crime-Mapping Untersuchungen bereits gut dokumentiert. Da Attributwerte in die Berechnung der LISA-Statistiken eingehen, wurden diese Untersuchungen bisher entweder an aggregierten Daten (vgl. Murray et al. (2001), S. 325) oder zur Optimierung einer vorhergehenden Kernel-Density-Estimation durchgeführt (vgl. Chainey et al. (2002), S. 33f., Ratcliffe & McCullagh (1999), S. 387ff.). In der vorliegenden Arbeit geht die ermittelte Anzahl der Straftaten pro Straßenabschnitt in die Berechnung der lokalen Indizes der räumlichen Autokorrelation ein. Um diese Berechnung durchführen zu können, wurden mit Hilfe eines Spatial Join die Diebstahlereignisse und die Einbruchereignisse des Jahres 2003 für jeden Straßenabschnittspunkt ermittelt. Die Anzahl der Diebstähle bzw. Einbrüche

pro Straßenabschnittspunkt gingen dann als Parameter in die Berechnung der jeweiligen Indizes ein. In ArcView sind für beide Berechnungsalternativen verschiedene Parameter auswählbar. Diese sind in der folgenden Tabelle 5 aufgelistet.

Tabelle 5: Eingabeparameter zur Berechnung des Local Moran I und des G_i^*

Auswahl	Ausdruck	Bedeutung
Ermittlung der Gewichte	Fixed Distance Band	Alle Daten, die sich innerhalb der definierten Distanz befinden, werden (gleichwertig) berücksichtigt. Alle Daten, die sich außerhalb dieser Distanz befinden, werden ausgeschlossen
	Inverse Distance	Der Einfluss eines Ereignisses auf ein anderes Ereignis nimmt im Verhältnis $1/d$ mit der Entfernung ab
	Inverse Distance Squared	Der Einfluss eines Ereignisses auf ein anderes Ereignis nimmt im Verhältnis $1/d^2$ zur Entfernung ab
	Zone of Indifference	Kombination aus „Fixed Distance Band“ und „Inverse Distance“: Alle Daten, die sich innerhalb einer definierten Distanz zum Ereignis befinden, wirken auf dieses ein.
	Get spatial weights from file	Räumliche Beziehungen sind in einer Datei abgelegt. Der Name der Datei wird hier angegeben.
Berechnung der Distanz	Euclidean Distance	Euklidische Distanz (Luftlinie)
	Manhattan Distance	City Block – die Distanz zwischen zwei Punkten wird entlang senkrechter Linien gemessen (in einem regelmäßig angeordneten Punktraster) und aufsummiert.
Standardisierung der Gewichte	None	Es wird keine Standardisierung der räumlichen Gewichtungen vorgenommen
	Row	Eine Standardisierung erfolgt bezogen auf die Zeile der Gewichtungsmatrix (jedes Gewicht wird durch die Zeilensumme geteilt)
	Global	Eine Standardisierung erfolgt global (jedes Gewicht wird durch die Summe aller Gewichte geteilt)
	Weights_matrix_Datei	Der Pfadname zu der Datei, welche die Gewichte für die räumlichen Beziehungen enthält.

Um die Ergebnisse der beiden Verfahren vergleichen zu können, wurden die Berechnungen in allen Fällen mit dem Distanzparameter „Inverse Distance“ und einer Entfernung von 3000m durchgeführt. Zur Parameterberechnung eines Punktes wurden nun alle Punkte, die innerhalb dieser Entfernung lagen, einbezogen. Punkte, die sich weiter entfernt befanden, wurden nicht mehr berücksichtigt. Die Entfernung zwischen zwei Punkten wurde mit Hilfe der euklidische Distanz ermittelt. Eine Standardisierung der räumlichen Gewichte fand nicht statt.

5.1 Diebstähle 2003

Es wurden beide LISA-Statistiken berechnet. Dies geschah mit dem Ziel, die räumlichen Beziehungen hinter der Verteilung der Diebstähle deutlicher erkennen zu können. Die Ergebnisse der Berechnungen sind in der Abbildung 24 gegenübergestellt.

G_i*

Die linke Grafik stellt die Ergebnisse der Getis/Ord-G_i*-Statistik dar. Nachdem die Berechnung abgeschlossen war, besaß jeder Punkt des Datensatzes einen neuen Wert G_i*. Dieser Wert beschreibt die „... statistical significance of clustering for a specified distance“ (vgl. ESRI (2004)). Ein hoher G_i*-InvDist-Wert für einen Punkt zeigt an, dass seine Nachbarn ebenfalls hohe Attributwerte aufweisen, an diesen Straßenabschnitten also viele Diebstähle festgestellt wurden. Ein geringer G_i*-InvDist-Wert zeigt an, dass seine Nachbarn niedrige Attributwerte aufweisen, hier also wenige Diebstähle ermittelt wurden. Je höher der Ergebniswert, desto stärker ist diese Beziehung. Ein G_i*-InvDist-Ergebnis nahe 0 (von -0,5 bis 0,5) zeigt an, dass die Attributwerte der Nachbarpunkte nicht eindeutig in eine Richtung gehen, sondern variieren. Die Legende in der Abbildung 25 unterscheidet darüber hinaus diejenigen Werte, die statistisch signifikante Ergebnisse liefern, von solchen, die für das Signifikanzniveau $p=0,05$ keine statistisch verwertbaren Aussagen aufweisen. Alle Werte $<-1,96$ oder $>1,96$ können als statistisch signifikant betrachtet werden (vgl. ESRI (2004)).

Das Bild der G_i*-InvDist-Verteilung über die untersuchten Straßenabschnitte stellt für den größten Teil des Stadtgebietes Werte von -0,5 bis 0,5 dar. Diese Werte zeigen an, dass hier keine Beziehung zwischen den betrachteten Punkten und ihren Nachbarn festgestellt wurden. Im Stadtzentrum – sowohl westlich als auch östlich der Alster – befindet sich ein Ring, für den G_i*-InvDist-Werte zwischen 0,5 und 1,96 ermittelt wurden. Diese Werte zeigen an, dass die Nachbarstraßenabschnitte ähnlich hohe Diebstahl-Werte besitzen, jedoch ist diese Beziehung statistisch nicht signifikant ($< 1,96$ für $p = 0,05$). Statistisch signifikant hohe G_i*-InvDist-Werte ($> 1,96$) findet man in dieser Darstellung nur vereinzelt. Im Innenstadtbereich westlich der Alster tritt eine kleine Kette solcher Straßenabschnitte auf, ansonsten sind die Punkte mit signifikant hohen

G_i^* -InvDist-Werten über das Stadtgebiet verteilt. Diese Verteilung erscheint recht zufällig. Da um diese Punkte herum kein Ring mit nicht-signifikant hohen Werten zu erkennen ist, scheinen diese Hotspots inselförmig im Stadtgebiet aufzutreten.

Obwohl die überwiegende Zahl der Straßenabschnitte keine lokalen Beziehungen zu den Nachbarpunkten aufweisen, so ist es dennoch erstaunlich, dass es keine Gebiete gibt, die eine geringere Diebstahlhäufigkeit als erwartet aufweisen. Der Legende entsprechend soll es einige Punkte geben, die eine solche nicht-signifikante Beziehung beschreiben (Werte zwischen -1,96 und -0,5 für $p = 0,05$), jedoch sind sie in der Darstellung nicht auszumachen. Obwohl sich also die Diebstähle an einigen Punkten konzentrieren (vgl. Kapitel 3.2.2 und Kapitel 3.2.3), sind sie doch ubiquitär über das gesamte Stadtgebiet verteilt.

Local Moran I

In der rechten Grafik (vgl. Abbildung 24) werden die Ergebnisse des Local Moran I dargestellt. Die in ArcView 9.0 eingebaute Methode berechnet zwei Werte: zum einen den tatsächlichen lokalen Index I und zum anderen einen z-Wert, der die statistische Signifikanz des Index-Wertes repräsentiert (vgl. ESRI (2004)). Dieser Wert beschreibt, ob die vermeintliche (Un-)Ähnlichkeit zwischen dem betrachteten Punkt und seinen Nachbarn größer ist als es zu erwarten gewesen wäre. Darüber hinaus zeigt ein hoher z-Wert, dass die umliegenden Punkte ähnliche (sowohl hohe als auch niedrige) Werte besitzen als der betrachtete Punkt selber. Ein geringer z-Wert deutet auf starke Unterschiede in der Nachbarschaft dieses Punktes hin (vgl. ESRI (2004)).

Auch in dieser Grafik wurden die Klassen so gewählt, dass die dargestellten Werte die Signifikanz für $p=0,05$ abbilden. Auch hier zeigen die Werte, die unterhalb der Klassengrenze von -1,96 und oberhalb der Klassengrenze von 1,96 liegen, eine signifikante Abweichung von der Annahme einer zufälligen Ergebnisausprägung. Das Bild, das in der rechten Hälfte der Abbildung 24 zu erkennen ist, zeigt eine völlig andere räumliche Werteverteilung als im linken Bild zu beobachten war. In dieser Abbildung zeigen hohe Werte an, dass in der Nachbarschaft dieses Punktes ähnliche Verhältnisse herrschen. Damit kann sowohl eine hohe als auch eine geringe Verteilung der

Diebstähle auf die Straßenabschnitte gemeint sein. Insbesondere im Innenstadtbereich dominiert die Tendenz zu ähnlichen Werten – also benachbarte Straßenabschnitte mit hohen oder benachbarte Straßenabschnitte mit geringen Diebstahlereignissen. Dennoch ist die Anzahl derjenigen Werte, die eine signifikante räumliche Ähnlichkeit bzw. Unähnlichkeit aufweisen, sehr gering.

Obwohl sich die Verteilung des Local Moran I im Vergleich zur Werteverteilung des G_i^* -InvDist wesentlich kleinräumiger darstellt, lassen sich auch hier Bereiche identifizieren, die durch ähnliche Werteausprägungen auffallen. Diejenigen Punkte mit einem signifikant hohen z-Wert finden sich vor allem im Innenstadtbereich. Aus den bisher vorgestellten Überlegungen liegt die Vermutung nahe, dass hier die höchsten Diebstahlvorkommen konzentriert sind, hier also die Hotspots der Diebstahlverteilung zu finden sind. Die signifikant geringen z-Werte zeigen keine räumliche Häufung. Sie treten über das gesamte Stadtgebiet verteilt auf. Auch diese Werte sollten jedoch intensiv betrachtet werden, denn sie können darauf hindeuten, dass sich in einem Gebiet mit relativ geringen Diebstahlvorkommen ein Straßenabschnitt durch besonders hohe Werte auszeichnet. Auf diese Weise lassen sich lokale Hotspots, also hoch belastete Straßenabschnitte, in einem ansonsten durch geringe Diebstahldichten gekennzeichneten Gebiet herausfiltern.

5.2 Einbrüche 2003

In der Abbildung 25 wird die lokale Beziehung der Einbrüche – auch hier wieder an jedem Straßenabschnitt aufaddiert – dargestellt.

G_i^*

Wie gerade für die Diebstählen beobachtet, zeigt sich auch hier an den meisten Straßenabschnitten der Stadt keine nachbarschaftliche Beziehung. Die meisten der dargestellten Straßenabschnitte weisen Werte um 0 herum aus. Im Gegensatz zur Verteilung der Diebstähle sind alle Daten, die auf hohe Werte in der Nachbarschaft hinweisen, räumlich verteilt. Dies gilt sowohl für die Punkte, die eine nicht-signifikant hohe Beziehung zu den Nachbarpunkten anzeigen, als auch für die Punkte, die signifikant hohe

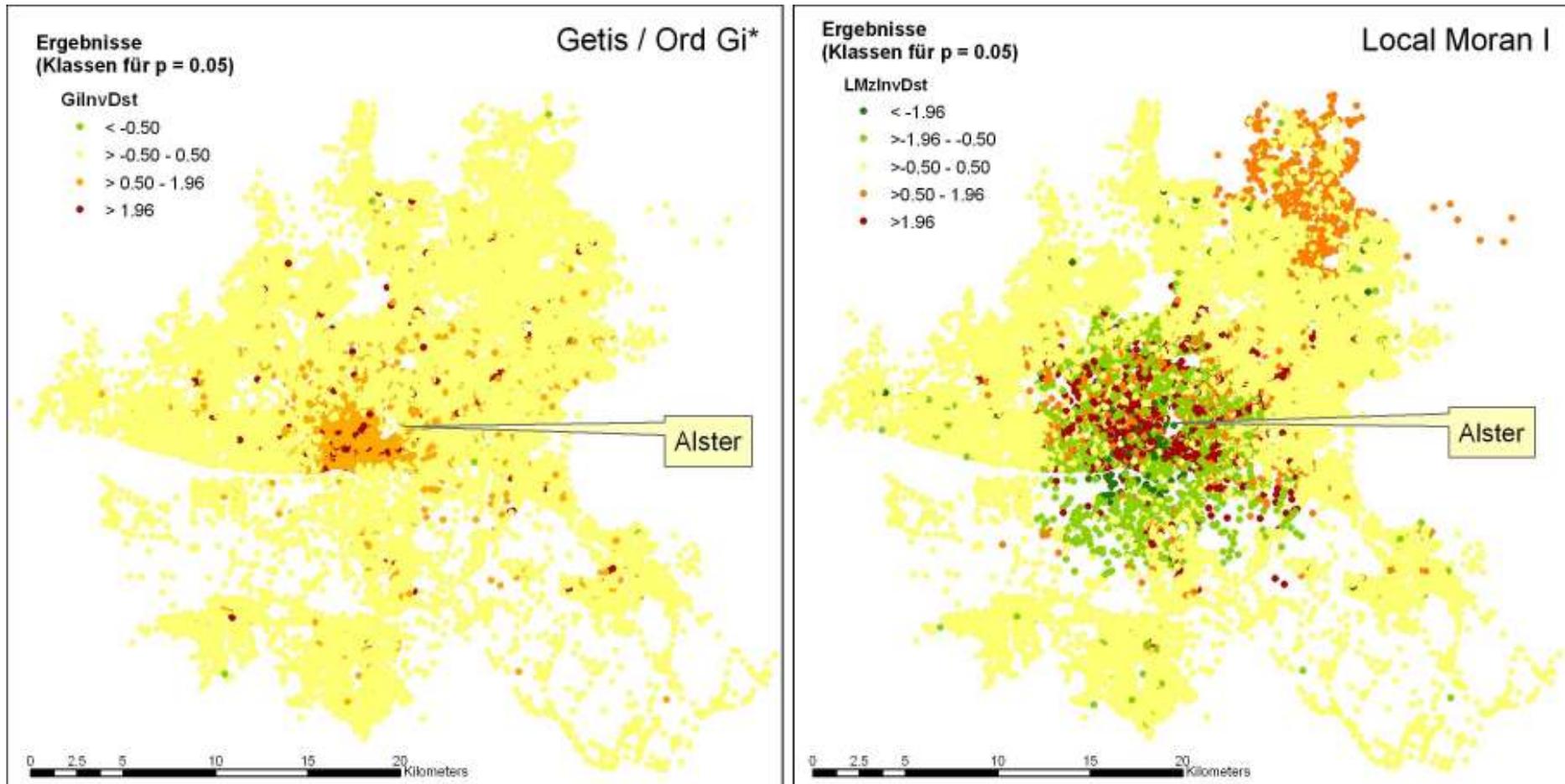


Abbildung 24: Diebstähle 2003 - Ergebnisse der LISA-Statistiken

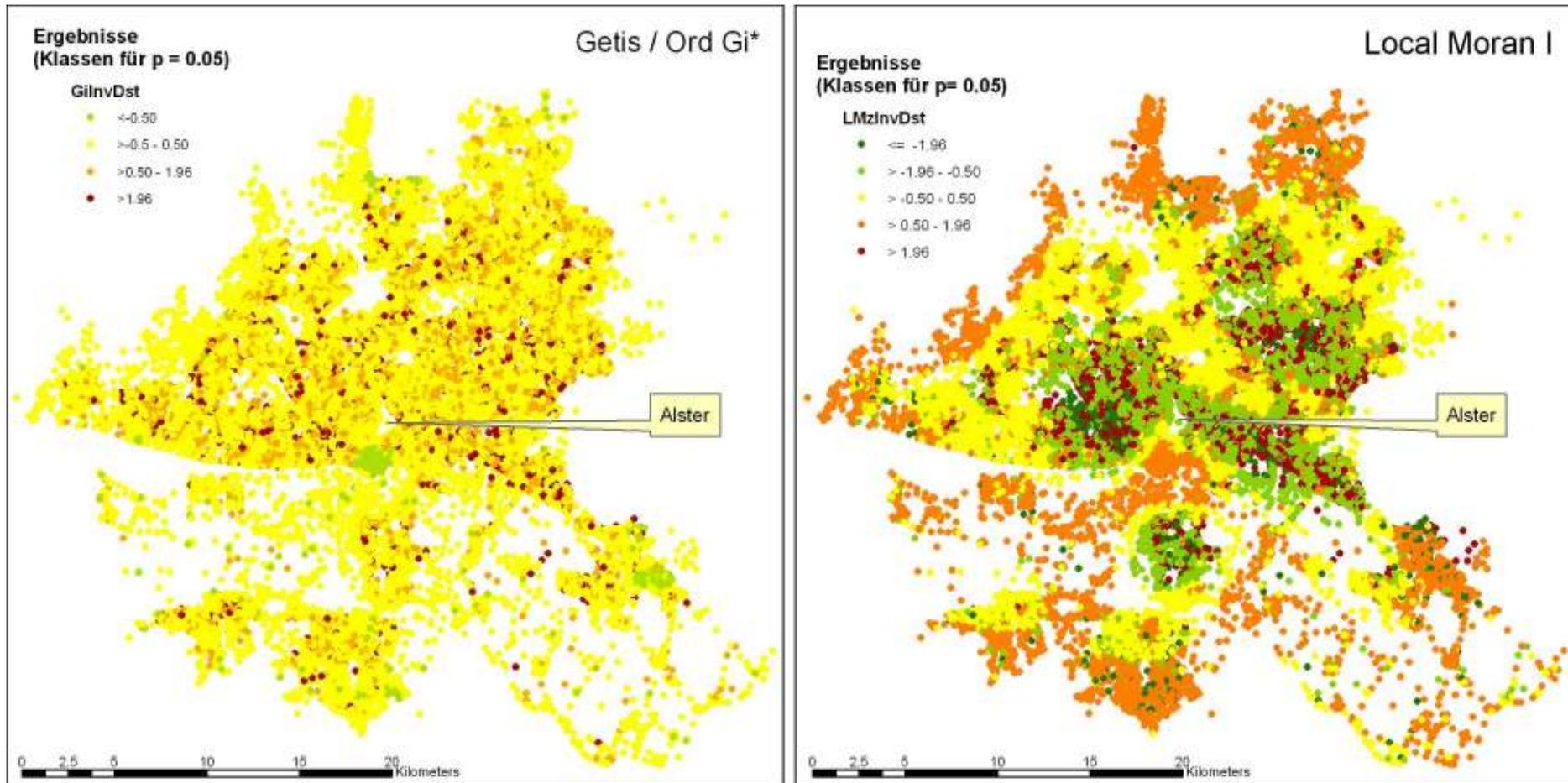


Abbildung 25: Einbrüche 2003 - Ergebnisse der LISA-Statistiken

Werte anzeigen. Letztere sind – wiederum im Gegensatz zu den Diebstählen – deutlicher zu erkennen, da sie nicht einzeln, sondern sehr oft in kleinen Gruppen auftreten. Dies deutet darauf hin, dass Einbrüche im allgemeinen – also bezogen auf das Hamburger Stadtgebiet als Ganzes – verstreut auftreten, sich jedoch kleinräumig konzentrieren.

Neben den Hotspots lassen sich sogar zwei (nicht signifikante, da die Werte zwischen -1,96 und -0,5 liegen für $p=0,05$) Cold-Spots erkennen. Unter Cold-Spots versteht man solche Gebiete, an denen eine unterdurchschnittliche Häufigkeit der untersuchten Ereignisse zu beobachten ist (vgl. Levine (2002), S. 289). Im vorliegenden Fall wären hier also weniger Einbrüche geschehen, als theoretisch zu erwarten gewesen wären. Einer dieser Cold-Spots befindet sich im Bereich der Altstadt, zwischen Innenalster, Großen Bleichen und der Speicherstadt, der zweite Cold-Spot ist im südöstlichen Teil Bergedorfs anzutreffen.

Local Moran I

Wiederum im rechten Teil der Abbildung 26 ist das Ergebnis der Berechnung des Local Moran I dargestellt. Diese Darstellung zeigt einen (nicht-signifikanten, da Werte zwischen 0,5 und 1,96 für $p = 0,05$) Ring ähnlicher Werte, der fast den gesamten Stadtrand umschließt. Lediglich im östlichen Teil ist dieser Ring durchbrochen. Im Südwesten schiebt sich ein Keil dieser Klasse um die Elbinsel herum bis in den Innenstadtbereich vor.

Westlich und östlich der Alster zeigen sich große Bereiche, die signifikant ähnliche Werteausprägungen besitzen (hier sind die Werte $> 1,96$ für $p = 0,05$). Solche Bereiche sind auch auf der Elbinsel zu finden. Diese werden umschlossen von Bereichen, die auf – ebenfalls signifikant – unähnliche Werteverteilungen ($< -1,96$ für $p = 0,05$) in ihrer Nachbarschaft hindeuten. Die Verteilung dieser Werte könnte auf das Vorhandensein von Clustern hindeuten: ähnliche Werte stellen den Cluster dar. In der Übergangszone zeigen sich dann Variationen der Werteausprägungen. Diese Gebiete sollten isoliert und genauer untersucht werden.

5.3 Methodenkritik

Die im gesamten Kapitel 5 vorgestellten Methoden aus dem Anwendungsumfang des neuen Softwarepaktes ArcGIS 9.0 der Firma ESRI haben gezeigt, dass sie ein hohes Potential zur räumlichen Auswertung der Straftatenverteilung besitzen. Zu einer solchen Bewertung des G_i^* und Local Moran kamen bereits (Ratcliffe & McCullagh (1999), S. 397), die die Ergebnisse der LISA-Statistiken für Crime-Mapping-Anwendungen untersuchten. Die dargestellten Bilder zeigen neue Zusammenhänge auf und geben Anlass für weitere Untersuchungen zur Überprüfung der dargestellten Ergebnisverteilungen und zur Suche nach Erklärungen (Hypothesen) für diese Verteilungen.

Da der Distanzparameter bei der Anwendung der ArcGIS-Funktionen in vielfältiger Weise modifiziert werden kann, sollte untersucht werden, mit welchen Einstellungen die besten Ergebnisse erzielt werden können. Interessant wäre in diesem Zusammenhang auch ein Test, der untersucht, ob es möglich wäre, anstatt der aktuell gewählten euklidischen Distanz andere Distanzwerte (z. B. die tatsächlichen Entfernungen zwischen den jeweiligen Straßenabschnitten) in Form einer Gewichtungsmatrix in die Berechnungen einzubeziehen. Auf diese Weise könnte z. B. der Einfluss natürlicher (z. B. Elbe) und anthropogener (z. B. Flughafen) Hindernisse stärker in die Berechnung und die anschließende Darstellung der Werte einbezogen werden.

Für die Interpretation der Werte stellt sich die Frage, inwieweit Randeffekte zum Tragen kommen, welche die Werteverteilungen des G_i^* , bzw. des Local Moran I beeinflussen. Diese dürfen nicht unberücksichtigt bleiben.

6 Raumzeitliche Verteilung der Straftaten

Neben der räumlichen Verteilung interessiert auch die zeitliche Verteilung der Straftaten über den Tag, die Woche, den Monat oder das Jahr. GISPOL stellt bereits Abfragemöglichkeiten nach der zeitlichen Verteilung von Straftaten bereit. Die Ergebnisse dieser Abfragen stellen die zeitliche Verteilung der Daten als Balkendiagramme dar. Darüber hinaus wirkt die zeitliche Anfrage als Filter für die im Viewer dargestellten Punktsymbole. Es fehlt jedoch eine raum-zeitliche Darstellungsvariante, mit deren Hilfe die

räumliche Verteilung der Straftaten zu bestimmten Zeitpunkten in Form einer flächenhaften Darstellung simuliert werden kann.

6.1 Qualität der Zeitangaben

Ebenso wie die räumlichen Koordinaten sich auf die Darstellung der räumlichen Verteilung der Straftaten auswirken und das dargestellte Bild bestimmen, so wirkt sich der festgestellte Tatzeitpunkt auf die Darstellung der zeitlichen Straftatenverteilung aus. Die Problematik, die hinter der Erfassung und Auswertung des Tatzeitpunktes steht, wurde bereits im ersten Teil dieser Arbeit (vgl. Teil I, Kapitel 3.2.1) erläutert. An dieser Stelle sei darauf verwiesen.

Die Hamburger Straftatenanalyse in GISPOL stellt ein Datum zur Auswertung bereit. Das angegebene Datum stellt den frühestmöglichen Zeitpunkt dar, an dem die Straftat hätte begangen werden können – also den Zeitpunkt, an dem der Geschädigte das Objekt zuletzt unversehrt gesehen hat. In Abhängigkeit von der Art des Deliktes werden entweder nur die Tage (Diebstähle) oder auch die Uhrzeiten (Einbrüche) erfasst.

6.2 Die raumzeitliche Verteilung der Einbrüche und der Diebstähle

Die Darstellung der raumzeitlichen Verteilung von Straftaten kann auf verschiedene Weise erfolgen. Einerseits kann die zeitliche Verteilung im Jahresverlauf untersucht werden. Jedoch dürfte auch eine vergleichende Betrachtung der Verteilung von Straftaten im Verlauf einer Woche, bzw. im Verlauf mehrerer Tage interessante Ergebnisse für die Straftatenanalyse liefern. Aus diesem Grund werden sowohl die räumlichen Verteilungen der Straftaten im Jahresverlauf (im Monatsvergleich) als auch im Verlauf einer Woche, durch den Vergleich der Wochentage, untersucht.

Die Verteilung der Straftaten im Verlauf eines Jahres wird dabei mit Hilfe der Einbruchsdaten untersucht. Zur Untersuchung von Veränderungen im Verlauf einer Woche werden die Diebstahlsdaten herangezogen. In beiden Fällen werden die Ergebnisse der Kernel-Density-Estimation den Ergebnissen der Potentialanalyse gegenübergestellt.

6.2.1 Zeitliche Veränderungsdaten der Einbrüche auf Monatsbasis

Der zeitliche Vergleich der Einbrüche im Verlauf des Jahres 2003 beschreibt die räumliche Verteilung der Einbrüche für jeden einzelnen Monat des Jahres und stellt diese einander gegenüber. Die Verteilungen werden anhand der Kernel-Density-Estimation sowie der Potentialanalyse ermittelt. Für die Kernel-Density Methode wurde ein Suchradius von einem Kilometer definiert. Die Größe der Rasterzellen wurden aus Gründen der Vergleichbarkeit zum Potentialansatz mit 300 x 300 m relativ groß gewählt. Die Rasterzellen der Potentialanalyse erhielten anschließend dieselbe Größe, Auch für die Potentialanalyse wurde in diesem Fall ein Suchradius definiert. Dies geschah mit dem Ziel, die Rechenzeit der Potentialberechnung zu verkürzen, in dem nicht mehr alle Datensätze in die Potentialberechnung einbezogen wurden, sondern nur noch solche, die innerhalb des definierten Radius liegen. Um jedoch genügend Punkte in die Berechnung einzubeziehen, wurde hier der Suchradius auf 3000 m gesetzt.

Ein Vergleich der beiden Methoden kann dabei nur bedingt erfolgen. Nicht vergleichbar sind die quantitativen Angaben der beiden Methoden. Deshalb wird ein Vergleich ausschließlich auf die dargestellte räumliche Schwerpunkt-Verteilung reduziert. Sollten sich Unterschiede ergeben, werden diese benannt. Die Quelle dieser Unterschiede kann jedoch nicht ermittelt werden.

Kernel-Density-Estimation

Die Abbildungen 26 und 27 beschreiben die räumliche Verteilung der Einbrüche für jeden einzelnen Monat des Jahres 2003. Bereits der erste Blick auf die zwölf Einzelbilder macht deutlich, dass die räumliche Verteilung der Einbrüche in den einzelnen Monaten stark variiert. Dennoch lassen sich auch hier räumliche Schwerpunkte erkennen. So bil-

det sich der Bereich St. Pauli in fast allen Monaten als Schwerpunkt heraus. Auch im Süden Hamburgs stellen die beiden Stadtteile auf der Elbinsel (Wilhelmsburg und Veddel), Harburg, Bergedorf und insbesondere Neugraben (im Süd-Westen) diejenigen Be-

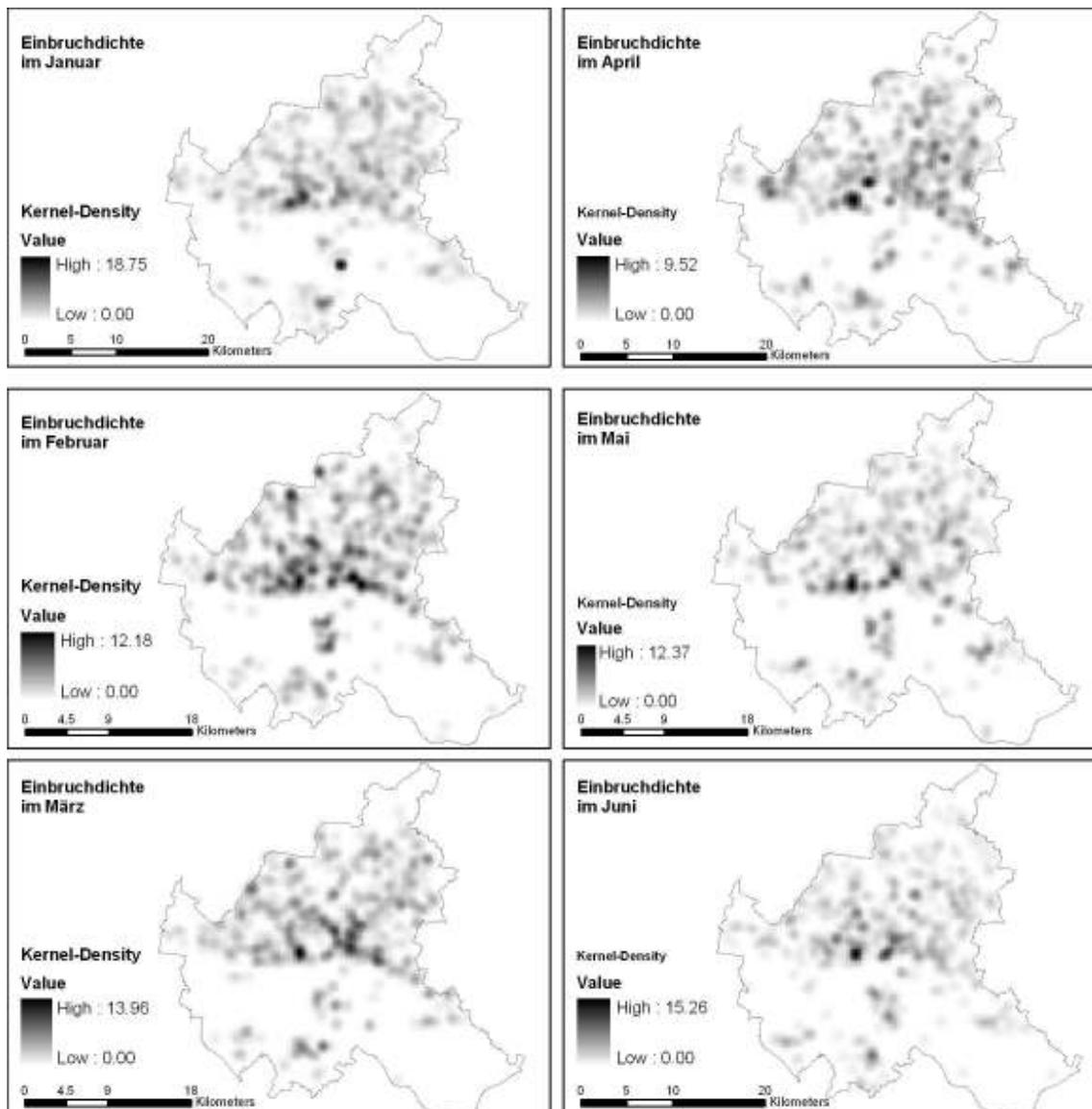


Abbildung 26: die raum-zeitliche Verteilung der Einbrüche des Jahres 2003 (Januar - Juni) ermittelt durch die Kernel-Density-Estimation

reiche dar, die häufig hohe Einbruchsdichten im Jahresverlauf aufweisen. Darüber hinaus gibt es – insbesondere im Osten der Stadt – einige Gebiete, die nur in einem einzigen Monat hohe Einbruchsdichten anzeigen (vgl. Abbildung 27: September, November oder Dezember).

In den meisten Monaten bildet sich aus der Verteilung der Einbrüche ein Straßenverlauf deutlich erkennbar heraus. Hier handelt es sich um den Straßenzug, der von der östlichen Stadtgrenze, bzw. von der Auffahrt auf der A24 am Horner Kreisel von Südosten nach Nordwesten in die Hamburger Innenstadt hineinführt.

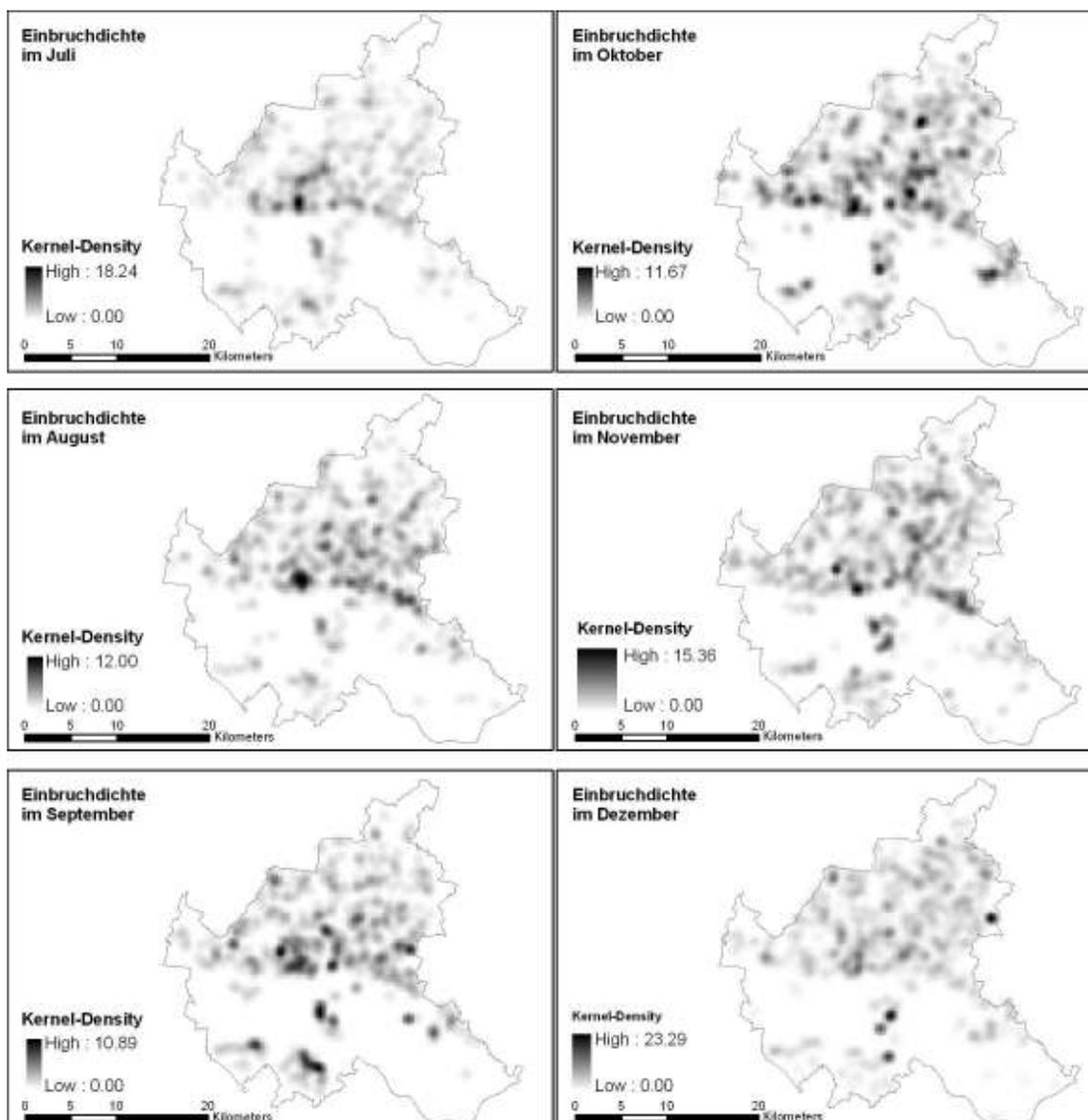


Abbildung 27: die raum-zeitliche Verteilung der Einbrüche des Jahres 2003 (Juli- Dezember) ermittelt durch die Kernel-Density-Estimation

Potentialanalyse

Auch die Potentialanalyse wurde zur Beschreibung der raum-zeitlichen Einbruchsverteilung in Hamburg herangezogen. Die Ergebnisse werden in den Abbil-

dungen 28 und 29 vorgestellt. Im Gegensatz zu den vorhergehend beschriebenen Ergebnissen der Kernel-Density-Estimation, die in jedem Monat mindestens einen Schwerpunkt ausweisen, zeigen die Ergebnisse der Potentialanalyse andere Verteilungen. Im Januar und im Februar stimmt die räumliche Darstellung des Einbruchspotentials mit der Darstellung der KDE überein. Es gibt hohe Werte westlich der Alster, auf der Elbinsel sowie östlich der Alster. Weitere kleinere Flächen mit hohen Potentialwerten sind im Stadtgebiet nördlich der Elbe verteilt. Die Darstellung für März und April zeigt kaum Gebiete, die besonders hohe Potentialwerte aufzeigen. Betrachtet man jedoch die Legende für diese Monate, dann zeigen sich besonders hohe Werte am Ende der Skala, und es stellt sich die Frage, ob diese Extremwerte die Darstellung möglicherweise beeinflusst haben und deshalb keine eindeutigen Schwerpunkte erkennbar sind.

Für die Einbrüche im Mai und im Juni zeigt die Legende ähnliche Werte in ihrer Skala. Dennoch zeigt die räumliche Verteilung der Ereignisse im Monat Mai mehr Gebiete mit einem hohen Einbruchspotential als der Juni. Diese hohen Einbruchspotentiale befinden sich zum einen westlich, zum anderen östlich der Alster in relativer Nähe zum Stadtzentrum. In den Sommermonaten Juli bis September (insbesondere im August/September) dehnen sich die Flächen mit höheren Potentialwerten weiter aus. Dies könnte sich damit erklären lassen, dass in den Sommerferien viele Häuser und Wohnungen über einen längeren Zeitraum leer stehen und somit die Einbruchgelegenheiten größer werden.

In den Monaten Oktober bis Dezember zeigt sich eine ähnliche Situation, wie im Juni. Betrachtet man jedoch ein Detail im Monat November etwas genauer, so zeigt sich ein interessantes Bild im südöstlichen Stadtgebiet. Hier wird ein lokaler Schwerpunkt ausgewiesen, der sich als dunkler Punkt von seiner Nachbarschaft abhebt. Interessanterweise stellt diese Darstellung ein Ergebnis dar, welches von dem Bild abweicht, das durch die Kernel-Density-Estimation für diesen Monat und diesen Bereich dargestellt wurde. Die KDE ließ Einbruchschwerpunkte erst weiter östlich von diesem Punkt erkennen.

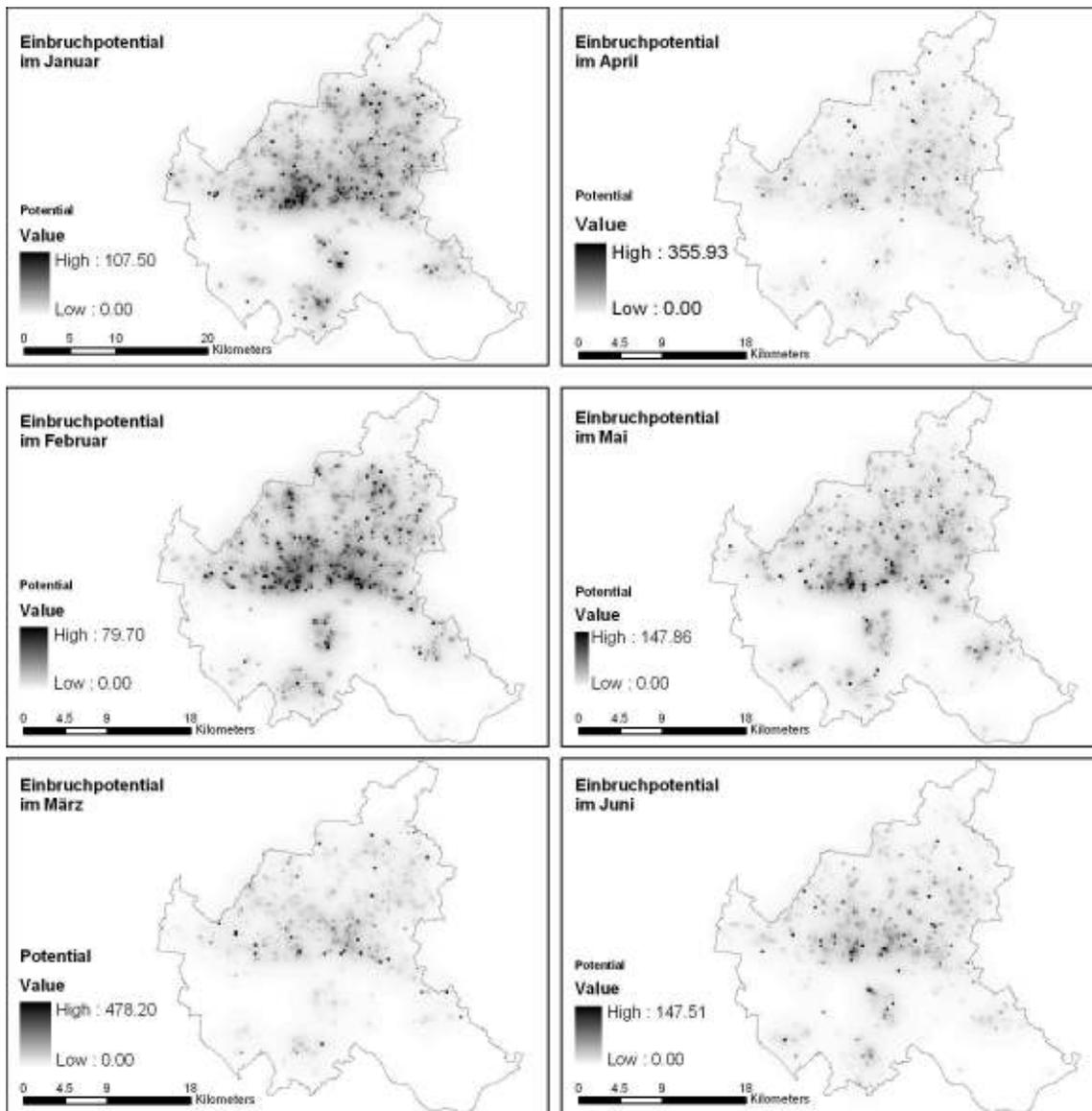


Abbildung 28: raum-zeitliche Verteilung der Einbrüche im Jahr 2003 (Januar - Juni) ermittelt durch die Potentialanalyse

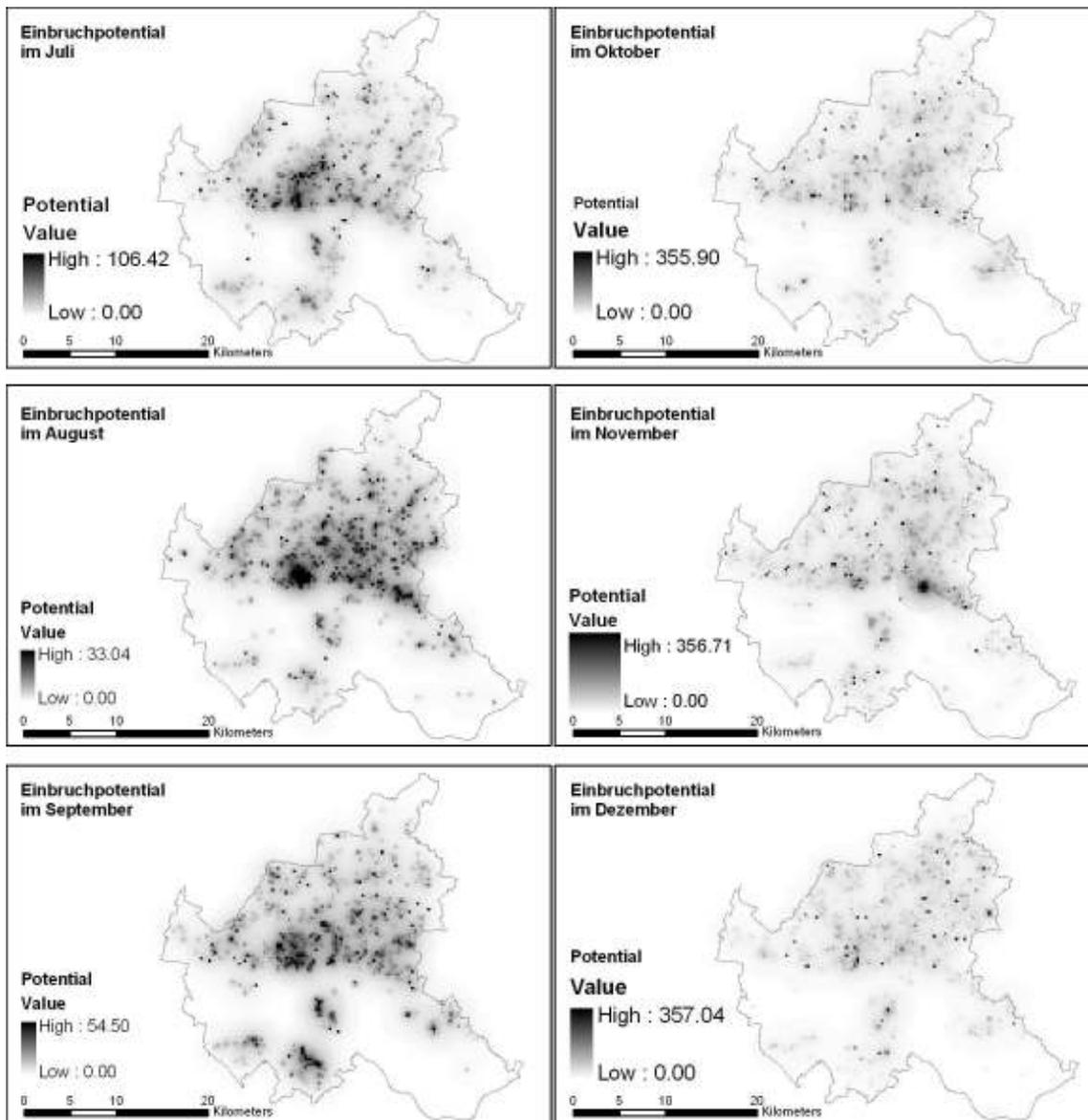


Abbildung 29: die raum-zeitliche Verteilung der Einbrüche des Jahres 2003 (Juli - Dezember) ermittelt durch die Potentialanalyse

6.2.2 Zeitliche Veränderungsrate der Diebstähle im Wochenverlauf

Die raumzeitliche Verteilung von Straftaten im Verlauf einer Woche wurde anhand der Diebstahldaten des Jahres 2003 untersucht. Zu diesem Zweck wurde der Wochentag aus den vorhandenen Datumsangaben ermittelt und in einer zusätzlichen Spalte abgelegt. Anschließend wurde die räumliche Verteilung der Diebstahlereignisse für jeden einzelnen Wochentag aus den Gesamtereignissen des Jahres 2003 berechnet. Es wurden

wieder beide Verfahren mit den im Kapitel 6.2.1 vorgestellten Parametern eingesetzt. Die Ergebnisse dieser Berechnungen sind den Abbildungen 30 und 31 zu entnehmen.

Kernel-Density Estimation

Die Verteilung der Diebstähle im Wochenverlauf zeigt relativ stabile Verhältnisse in der räumlichen Verteilung der Diebstahlschwerpunkte an den Werktagen. Lediglich die Wochenenden (Samstag und Sonntag) zeigen veränderte räumliche Strukturen. Der Freitag nimmt eine Übergangstellung ein, die sowohl die Verteilung der Werktage als auch die Verteilung der Wochenenden beinhaltet.

Die höchsten Diebstahldichten finden sich innerhalb der Woche im Innenstadtbereich, südlich der Alster. Dieser Bereich stellt sowohl die innerstädtischen Haupt-Einkaufsstraßen (um die Mönkebergstraße, den Neuen Wall und das Hanse-Viertel), als auch das Gebiet um den Hauptbahnhof dar. Dieses Gebiet ist an allen Tagen als Schwerpunkt erkennbar. Bis auf den Sonntag lassen sich in diesem Gebiet regelmäßig die höchsten Diebstahldichten feststellen.

Westlich von diesem Hot-Spot befindet sich ein weiteres Gebiet, welches sich insbesondere an den Wochenenden (Freitag, Samstag und Sonntag) als Diebstahl-Schwerpunkt herausbildet. Für den Sonntag ist dieses Gebiet als alleiniger Schwerpunkt ausgewiesen. Dieses Gebiet befindet sich in St. Pauli. Dort ist um die Reeperbahn das Hamburger Vergnügungsviertel angesiedelt. Dieses wird insbesondere an den Wochenenden stark besucht. Offenbar ergeben sich dann viele Gelegenheiten für Diebstähle. Wie es scheint, finden auch am Sonntag noch ausreichend viele Menschen ihren Weg in dieses Gebiet.

Viele Gebiete, in denen sich erhöhte Diebstahldichten zeigen, die allerdings nicht mehr an die innerstädtischen Werte heranreichen können, stellen Stadtteilzentren und/oder Haltestellen des öffentlichen Personennahverkehrs (insbesondere U-Bahn und S-Bahn) dar. Auch der Flughafen (im Norden der Stadt) ist als Gebiet mit hohen Diebstahldichten an allen Wochentagen erkennbar.

Potentialanalyse

Die Ergebnisse der Potentialanalyse (s. Abbildung 31) liefert für die Werktage der Woche (Montag bis Freitag) ein relativ einheitliches Bild. Es gibt eine Zone, die besonders hohe Diebstahlpotentiale aufweist. Dieses Gebiet stellt den Innenstadtbereich dar, welcher seinerseits von einer Wolke geringerer Potentiale umgeben ist. Die Konzentration dieser Wolke nimmt mit zunehmender Entfernung von Zentrum ab. In den Stadtzentren von Harburg (im Süden) und Bergedorf (im Südosten) zeigen sich weitere Regionen erhöhter Diebstahlpotentiale, die sich deutlich von den Potentialen des Umlandes abheben.

Am Freitag scheint sich im Stadtteil St. Pauli ein zweiter Schwerpunkt zu etablieren, an dem sich die Diebstahldelikte konzentrieren. Die von diesem Schwerpunkt eingenommene Fläche ist mit 4 benachbarten Rasterzellen größer als an den übrigen Tagen in diesem Bereich. Er ist jedoch noch nicht so ausgeprägt, wie am Samstag oder wie auf der Darstellung der KDE zu vermuten. Das Entstehen eines zweiten Schwerpunktes in St. Pauli lässt sich eindeutig für den Samstag feststellen. An diesem Tag verschmelzen beide Schwerpunkte miteinander.

Überraschend hingegen fällt die Darstellung des Diebstahlpotentials für den Sonntag aus. Im Zentrum Hamburgs zeigt sich eine sehr große Fläche, für die hohe Potentiale berechnet wurden. Die Wolke zeigt eine weitaus größere Ausbreitung in die umliegenden Flächen, als an den anderen Tagen. Die räumliche Verteilung des Potentials am Sonntag darf jedoch nicht ohne weiteres mit den Potentialwerten der übrigen Wochentage verglichen werden. Der Maximalwert, der Potentiale am Sonntag fällt wenigstens um das Sechsfache geringer aus als die Potentialwerte der übrigen Wochentage. Die Darstellung der Potentiale ergibt sich jedoch ausschließlich aus den Potentialwerten des jeweiligen Rasterdatensatzes. Vor diesem Hintergrund ist die Darstellung des Potentials für den Sonntag nicht mit den Potentialwerten der übrigen Wochentage vergleichbar. Um einen Vergleich aller Wochentage herbeizuführen, müsste eine einheitliche Klassifizierung über alle Datensätze erfolgen. Dies ist in der Darstellung der Abbildung 31 jedoch nicht der Fall.

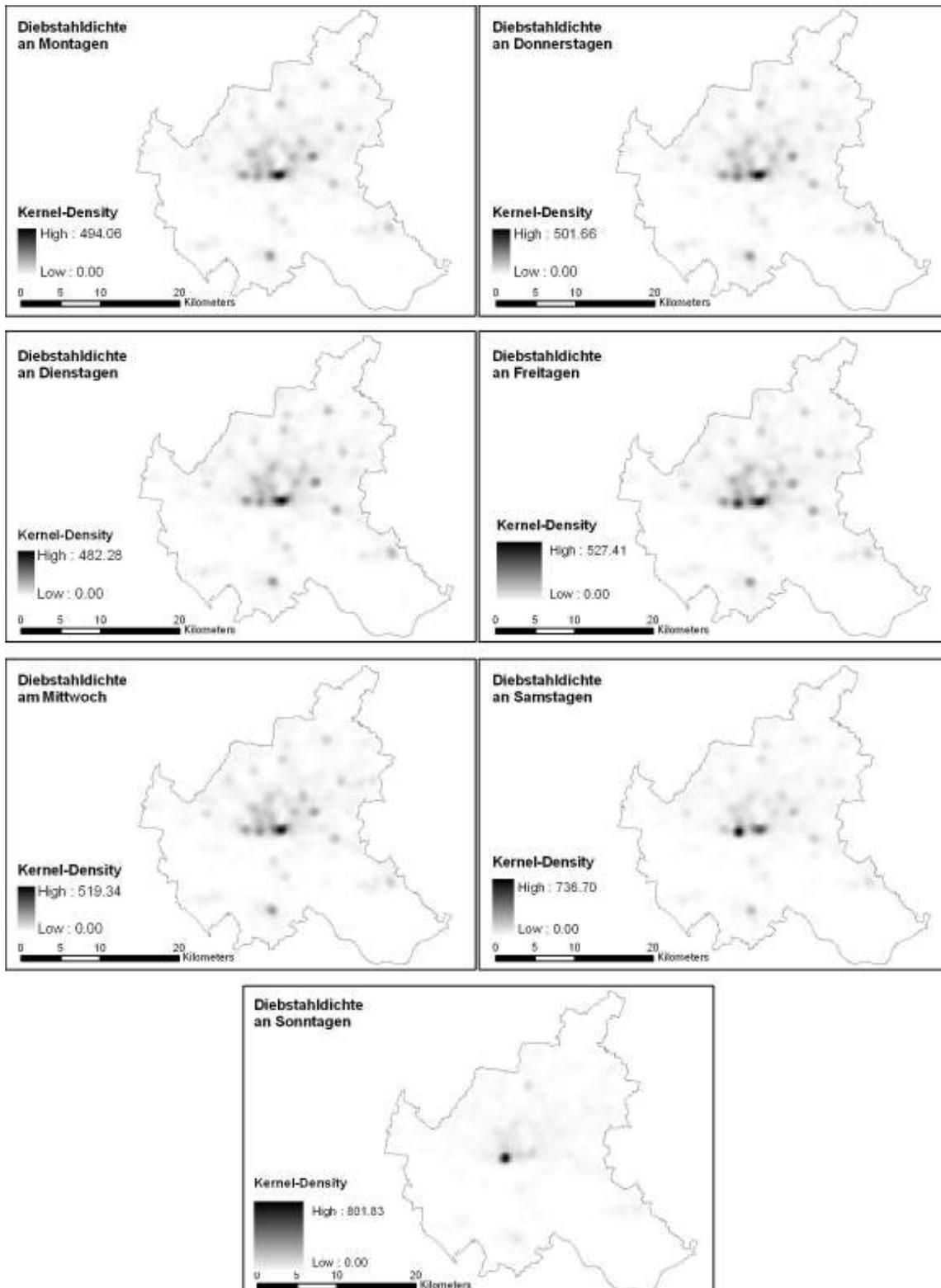


Abbildung 30: die raum-zeitliche Verteilung der Diebstähle im Wochenverlauf des Jahres 2003 (Kernel-Density Estimation)

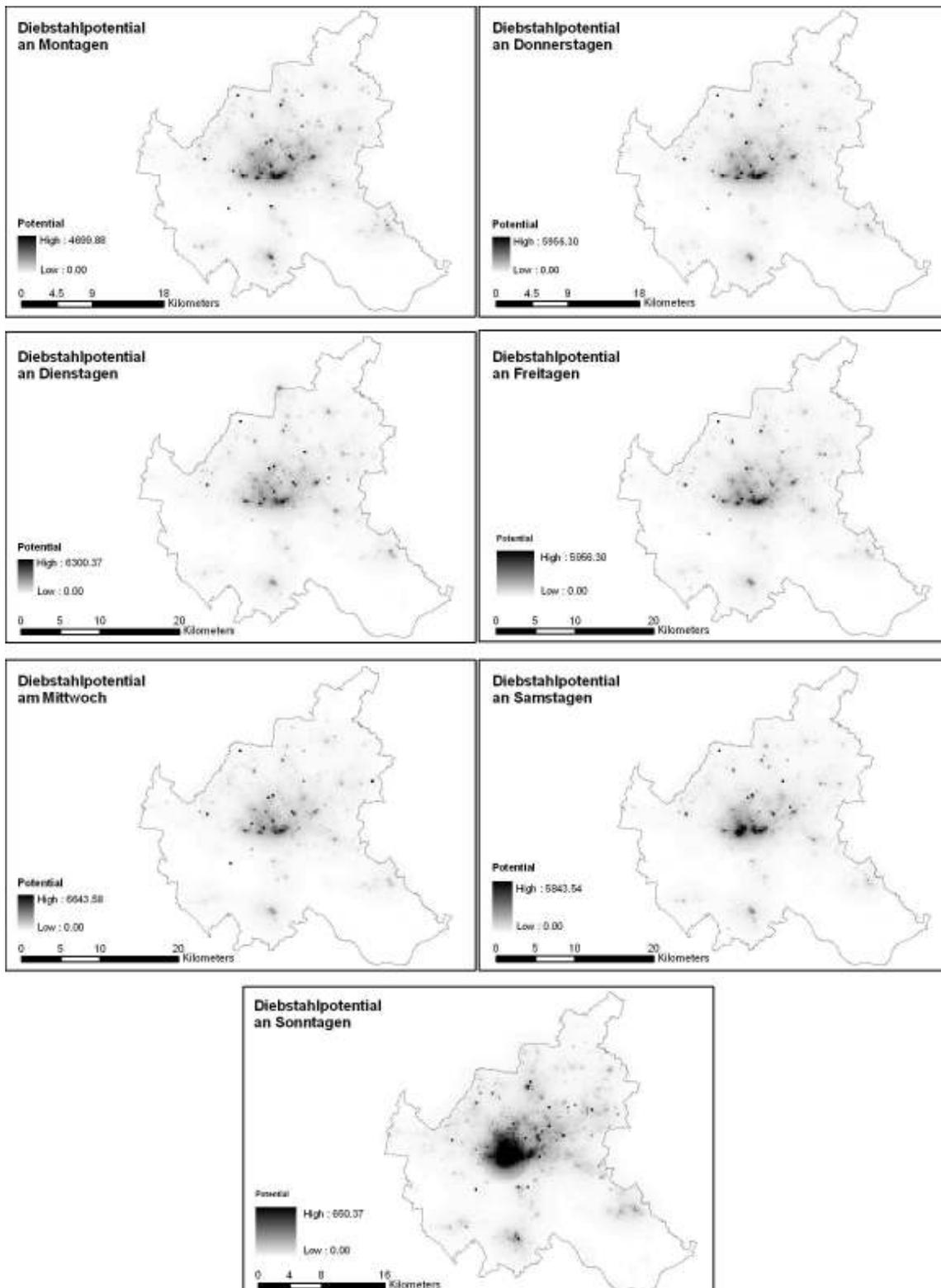


Abbildung 31: die raum-zeitliche Verteilung der Diebstähle im Wochenverlauf des Jahres 2003 (Potentialanalyse)

6.2.3 Vergleichende Bewertung der vorgestellten Methoden

Die beiden vorhergehenden Kapitel 6.2.1 und 6.2.2 haben Alternativen zur Betrachtung der Veränderungsraten von Einbrüchen und Diebstählen im Verlauf des Jahres 2003 bzw. im Verlauf des Wochenmittels des Jahres 2003 vorgestellt. Aus der Zusammenschau der Einzelbilder lässt sich ein Eindruck über die raum-zeitliche Verteilung der Dichten bzw. der Potentiale gewinnen.

Die Darstellung der raum-zeitlichen Straftatenverteilung ergibt sich aus den Zeitangaben, die in den Polizeidaten gepflegt sind. Diese geben nicht den echten Tatzeitpunkt an, sondern den frühestmöglichen Zeitpunkt, an dem das Ereignis stattgefunden haben könnte. Während man bei den Diebstahldaten davon ausgehen kann, dass diese Zeitpunkte dem Tatzeitpunkt relativ nahe kommen, muss dies für viele Einbruchsdaten bezweifelt werden. Insofern ist die raumzeitliche Darstellung dieser Daten kritischer zu hinterfragen. Eine quantitative Aussage zu möglichen Darstellungsfehlern kann leider nicht vorgenommen werden.

Raumzeitliche Ergebnisdarstellung als Animation

Die Bilder, die in den Abbildungen 27 - 31 aufgereiht sind, können zu einzelnen Bilderserien zusammengestellt und den Benutzern als Animation angeboten werden. Diese Animation könnte z. B. in einem eigenen Fenster der Web-GIS-Anwendung (GISPOL) präsentiert werden. Die Ablaufgeschwindigkeit sollte durch den Anwender bestimmbar sein, so dass mindestens eine Taste zum Abspielen/Anhalten der Zeitreihe vorhanden sein sollte. Diese Forderungen würden z. B. durch die Nutzung der Film-Datei-Formate (mpeg, avi, Quicktime, etc.) erfüllt werden können (vgl. Ott & Swiaczny (2001), S. 150).

Die vergleichende Betrachtung der Abbildungen 27 - 31 lässt jedoch ein Problem offenbar werden, welches es zu lösen gilt, bevor eine Animation erstellt werden kann. Die Legenden der Einzelbilder beziehen sich ausschließlich auf die Daten, die in der aktuellen Darstellung präsent sind. Ein bildübergreifender Ergebnisvergleich fällt daher schwer. Dies gilt besonders für die Darstellung der Potentialanalyse, da diese mit Hilfe der Standardabweichung klassifiziert ist. Mittelwert und Standardabweichung verändern

sich jedoch erheblich zwischen den jeweiligen Bildern. Dies wird besonders deutlich durch die Darstellung der Diebstahlsverteilung am Sonntag. Um Bilder besser miteinander vergleichen zu können, sollte deshalb eine einheitliche Legende (mit einheitlicher Klassifizierung) für alle Bilder erstellt werden.

Eine Lösung dieses Problems könnte in der raum-zeitlichen Organisation der Daten liegen, wie sie von Openshaw et al. (1994, S. 136f.) am Beispiel einer Animation von Polizei-Einsatzdaten aus Northumbria (England) vorgestellt wird. Die Autoren schlagen vor, dass die dargestellten Zeiträume sich zeitlich überlagern sollen, damit Veränderungstendenzen für den Betrachter nachvollziehbar werden und nicht plötzlich entstehen, wie dies z. B. in den monats- und wochenbasierten Darstellungen der Kapitel 6.2.1 und 6.2.2 erkennbar sind. Weiterhin verwenden sie eine raum-zeitlich operierende Kernel-Funktion (räumlicher Radius des Kernels 1,5km, zeitlicher Radius des Kernels 0,5h) zur Erzeugung der einzelnen Bilder (vgl. Openshaw, et al. (1994), S. 137).

Um die Darstellung der Kriminalitätsdaten ein wenig „fließender“ zu gestalten, könnte jedoch im ersten Ansatz dazu übergegangen werden, „Zeitscheiben“ zu erzeugen, die einen Zeitraum von 4 Wochen beschreiben, sich jedoch im Abstand von 14 Tagen überlagern, so dass die Hälfte der Daten immer in zwei Darstellungen enthalten wären. Auf diese Weise sollten sich auch die Legenden der einzelnen Bilder einander annähern.

6.3 Suche nach raumzeitlichen Clustern mit LISA

Im vorherigen Kapitel 6.2 wurde die Zeit als Filter eingesetzt, um für einen bestimmten Monat oder einen Wochentag das Muster der Straftatenverteilung zu ermitteln. In diesem Kapitel wird die Zeit als Parameter eingesetzt, der in die Funktionen zur Berechnung der LISA-Statistiken eingehen wird. Das Ziel dieser Untersuchung besteht darin, diejenigen Gebiete zu finden, in denen Straftaten sowohl räumlich als auch zeitlich konzentriert auftreten. Es besteht die Hoffnung, auf diese Weise Serienstraftaten (frühzeitig) entdecken zu können.

Die Berechnung der LISA Statistiken wurde wiederum mit Hilfe von ArcView 9.0 durchgeführt. Da dieses Programm als Eingangsparameter jedoch ausschließlich Zahlenwerte akzeptiert, das Datum systemintern aber auch als Zahlenwert abgebildet wird,

wurde das Erfassungsdatum der Straftat vor dem Aufruf der Funktion mit Hilfe eines VBA-Scriptes in einen Zahlenwert (Long) umgewandelt.

Im Gegensatz zu dem Verfahren, das im Kapitel 5 (Teil II) vorgestellt wurde und wo die Anzahl der Straftaten pro Straßenabschnitt ausgezählt wurde und die räumliche Verteilung der Straßenabschnitte mit dem Attribut „Anzahl Straftaten“ in die Wertevermittlung einging, werden hier die Original-Datensätze eingesetzt. Diese Daten zeigen das Problem, dass sich verschiedene Straftatereignisse überlagern. Es stellt sich nun die Frage, ob LISA-Statistiken diesem Phänomen Rechnung tragen können. Vor diesem Hintergrund und den Erläuterungen zu den beiden Verfahren Local Moran I (Anselin 1995) und G_i^* (Getis/Ord 1995) ist zu erwarten, dass die Berechnung des G_i^* geeigneter sein wird als die Berechnung der Local Moran Index, denn die Berechnung des G_i^* erfolgt gerade unter der Bedingung dass die Punkte i und j identisch sein können (s. Teil I, Kapitel 4.3.3.2). In der Tat konnte mit Hilfe des G_i^* eine Darstellung erzeugt werden in welcher alle übereinander liegenden Punkten dieselben Werte aufweisen und deshalb durch den Betrachter interpretierbar ist.

Da die Betrachtung über das gesamte Jahr 2003 eine übermäßig dichte Punktwolke entstehen ließ und die Werteverteilung nicht mehr klar zu interpretieren war, wurde der G_i^* exemplarisch für zwei ausgewählte Monate (Diebstähle - Januar und Einbrüche - August) betrachtet. Die Ergebnisse dieser Betrachtung zeigt die Abbildung 32.

Die Verteilung des G_i^* für die Diebstähle und die Einbrüche in den beiden Monaten Januar und Juli zeichnet sich insgesamt dadurch aus, dass keine räumlichen Konzentrationen signifikant hoher G_i^* -Werte vorliegen. Diese Werte tauchen vereinzelt und zufällig verteilt im Stadtgebiet auf. Es zeigt sich aber, dass die besonders hohen ($> 1,96$) oder besonders niedrigen ($< 1,96$) G_i^* -Werte an solchen Straßenabschnitten errechnet wurden, an denen mehrere Straftaten übereinander liegen. Darüber hinaus lagen diese Straftaten auch zeitlich nahe beieinander (eine Woche). Sehr oft wiesen die Straftaten sogar das gleiche Datum auf.

An jedem Tag des Jahres werden Straftaten erfasst und die Datumswerte, die zur Berechnung des G_i^* -Wertes herangezogen werden, begrenzen die maximale Differenz zwischen den einzelnen Tagen (zwischen dem ersten und dem letzten Tag eines Monats

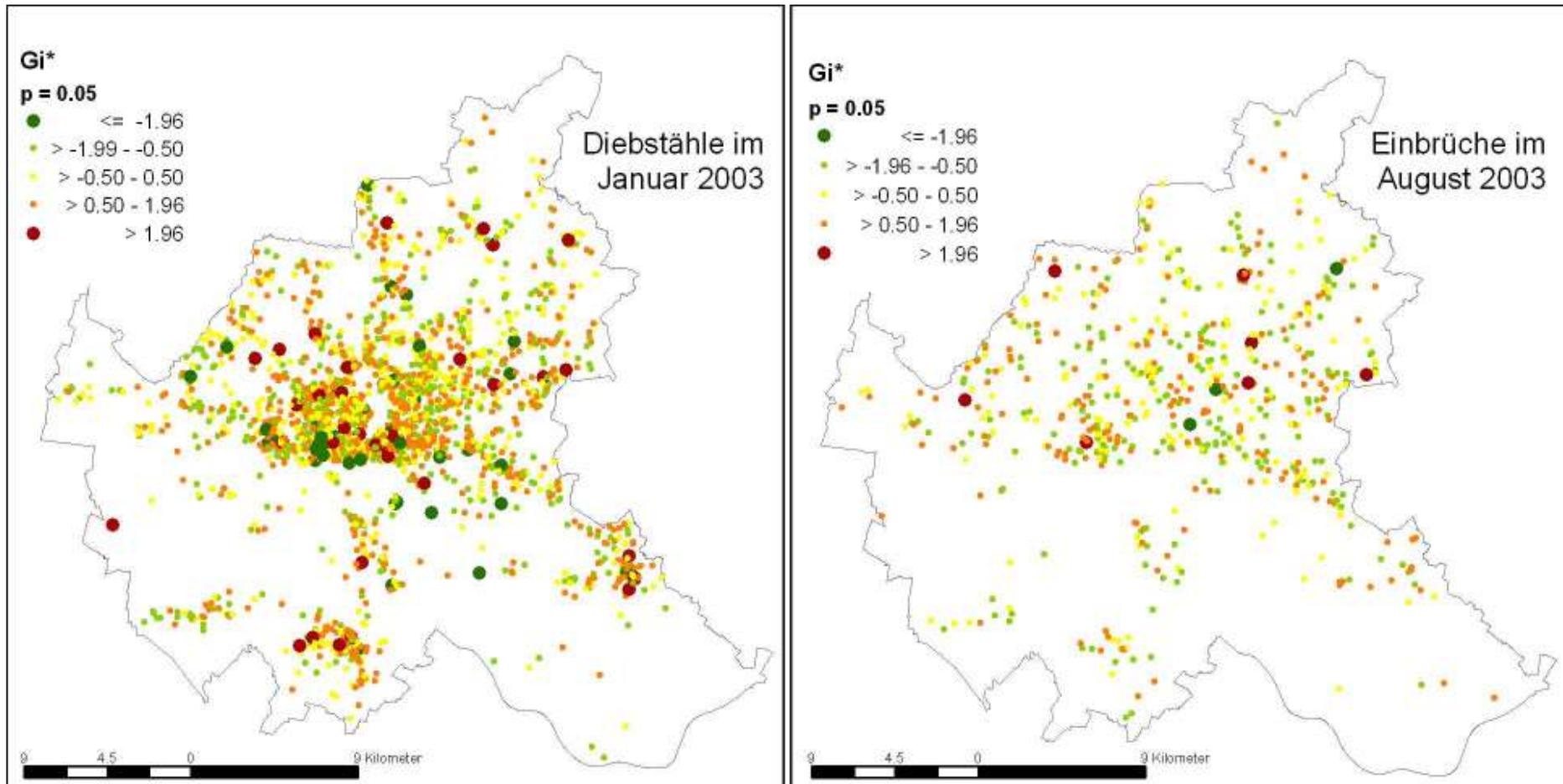


Abbildung 32: Lassen sich mit Hilfe des G_i^* zeitliche Schwerpunkte im Auftreten von Diebstählen und Einbrüchen feststellen?

liegen maximal 30 Tage). Darüber hinaus sorgt die Interpretation des G_i^* (ein großer G_i^* steht für eine räumliche Konzentration hoher Ausgangswerte, ein kleiner G_i^* weist auf die Konzentration geringer Ausgangswerte hin) dafür, dass nur solche raum-zeitlichen Cluster angezeigt werden, die am Monatsanfang (geringe Werte, große grüne Punkte in der Legende) oder am Monatsende (hohe Werte, große rote Punkte in der Legende) aufgetreten sind. Die Tage, die in der Monatsmitte liegen, werden bei dieser Suche nach raum-zeitlichen Beziehungen nicht berücksichtigt.

Insofern eignet sich dieses Verfahren nur eingeschränkt, um Hotspots oder Coldspots der raum-zeitlichen Straftatenverteilung zu entdecken. Sollte diese Methoden für die Suche nach zeitlichen Zusammenhängen genutzt werden, dann müssen die zu untersuchenden Daten in entsprechende zeitliche Sequenzen aufgeteilt werden, so dass auch die Daten, die aktuell im Zentrum des Untersuchungszeitraumes anzutreffen waren, an den Rand des Untersuchungszeitraumes geraten.

7 Einsatz von Methoden des Spatial Data Mining

Da die Straftatendatenbank der Hamburger Polizei bereits jetzt eine große Datensammlung darstellt und die dort abgelegten Datenmengen stündlich größer werden, wird die Suche nach Mustern in den Daten immer schwieriger und die Suche nach Möglichkeiten zur Automatisierung der Mustererkennung immer dringender. Eine Version der automatisierten Suche nach Beziehungen zwischen Daten in einer (oder auch verschiedenen) Datenbanken wird Data Mining genannt. „Unter Data Mining versteht man das systematische (in der Regel automatische oder halbautomatische Entdecken und Extrahieren unbekannter Informationen aus großen Mengen von Daten“ (vgl. Wikipedia (2005)). Um Muster in großen Datenmengen zu entdecken, kommen verschiedene Aufgaben im Data Mining zur Anwendung. Neben dem Aufdecken von Clustern werden z.B. auch Klassifikationen vorgenommen oder Assoziationsregeln aufgestellt und Generalisierungen vorgenommen (vgl. Ester & Sander (2000), S. 4). Es kommen verschiedene statistische Verfahren zur Anwendung.

7.1 Spatial Data Mining – Aufdecken räumlicher Cluster

Die Besonderheiten der Daten, die mit räumlichen Bezügen ausgestattet sind, wurden in den vorherigen Kapiteln bereits deutlich hervorgehoben. Auch im Bereich des Data Mining hat sich gezeigt, dass die allgemeinen Methoden des Data Mining für eine Suche nach Mustern in räumlichen Datenbanken nicht ausreichen (vgl. Peuquet & Guo (2000), S. 1). So hat sich auch hier der Spezialfall des *Spatial* Data Mining entwickelt. Das Spatial Data Mining erfordert, dass die räumlichen Beziehungen der Daten (absolut sowie relativ) betrachtet werden (vgl. Ester et al. (2001), S. 2). Neben dem Aufdecken von Clustern (räumlichen sowie raumzeitlichen Clustern) sollen mit diesen Verfahren auch Ausreißer entdeckt oder Trendanalysen durchgeführt werden (vgl. Han (2003), S. 1). In diversen Projekten ([SPIN!](#), [GeoVistaStudio](#)) werden Anwendungen und Methoden für das Spatial Data Mining entwickelt. Es ist zu erwarten, dass diese Methoden auch zu neuen Erkenntnissen für das Crime-Mapping führen werden.

Die Anwendung CrimeStat hält bereits zwei Methoden zur Suche nach räumlichen Clustern bereit. Hierbei handelt es sich um das Verfahren des „K-means“ und um die Methode des „Hierarchical Nearest Neighbour Clustering“. Ein Ergebnis der „Hierarchical Nearest Neighbour“ Clustering-Methode zur Darstellung der Einbruchverteilung des Jahres 2003 stellt die Abbildung 33 dar. Dieses Ergebnis wurde mit Hilfe der Standardwerte zu diesem Methodenaufruf erzielt (mindestens 10 Punkte für einen Cluster erforderlich, eine Berechnung der Ellipse mit einer Standardabweichung, 0 Simulationen). Um dieses Ergebnis anhand bereits vorgestellter Ergebnisse einschätzen zu können, wurde die nach der Kernel-Density-Estimation ermittelte Einbruchverteilung darunter gelegt.

Aus der ursprünglichen Verteilung der Einbrüche des Jahres 2003 werden die Cluster erster Ordnung ermittelt. Es werden mindestens 10 Punkte zu einem Cluster zusammengefasst, für die die Nächste-Nachbar-Distanz zueinander geringer ist als zu den übrigen Punkten im Untersuchungsgebiet (Cluster 1. Ordnung). Nachdem alle Cluster 1. Ordnung erzeugt wurden, werden die Cluster 2. Ordnung gebildet. Hier werden diejenigen Cluster 1. Ordnung – ebenfalls unter dem Gesichtspunkt der Minimierung der Nächsten-Nachbar-Distanzen zu Clustern 2. Ordnung zusammengefasst.

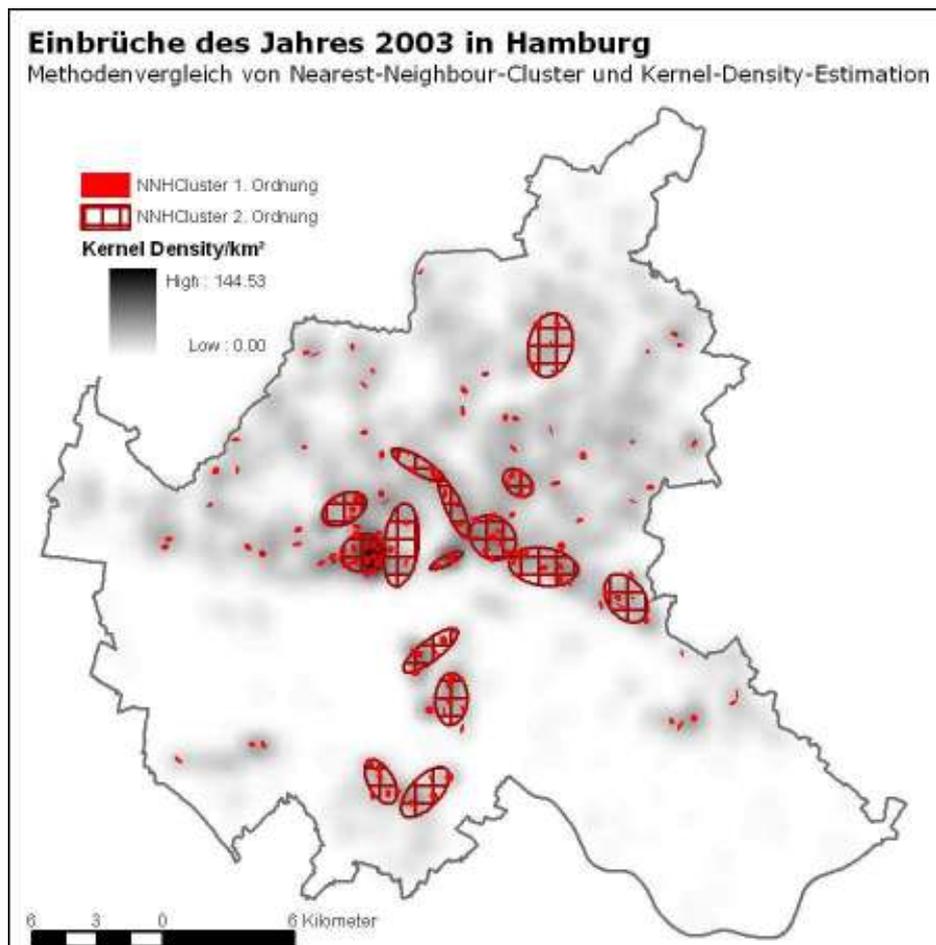


Abbildung 33: Cluster erster und zweiter Ordnung der Einbruchverteilung ermittelt mit CrimeStat (Nearest Neighbour Clustering)

Der Vergleich der räumlichen Verteilung der Cluster mit der Verteilung der Kernel-Density-Estimation zeigt, dass die Cluster erster Ordnung sehr häufig in der Nähe eines Dichtezentrums der KDE liegen. Weiterhin weisen diejenigen Bereiche, die in der KDE durch hohe Einbruchdichten aufgefallen sind, die größte Menge an Clustern erster Ordnung auf. Auch die Anordnung der Cluster 2. Ordnung – insbesondere östlich der Alster – spiegelt einige Besonderheiten dieses Datensatzes wieder (den Straßenverlauf von der A24 in das Zentrum nördlich der Alster).

Es muss jedoch darauf geachtet werden, dass die Grenzen der Cluster, die hier die Form von Ellipsen darstellen, nicht als scharfe Grenzen interpretiert werden. Vielmehr stellen sie das Ergebnis einer bestimmten Art der Grenzberechnung dar, die sich in der realen Welt selten in dieser Form antreffen lassen. Chainey et al. (2002, S.22) weisen auf die

Problematik hin, die aus der Darstellung von Clustern in Form von räumlichen Ellipsen resultieren. Aus diesem Grund setzen z.B. Estivill-Castro & Lee (2002) das Verfahren der Delaunay-Triangulation (vgl. Teil I, Kapitel 4.2.1.3) ein, um potentielle Grenzen von Clustern unmittelbar aus der Verteilung der Punkte heraus ermitteln zu können. Jedoch setzt dieses Verfahren voraus, dass jeder Punkt nur einmal im Datensatz vorkommt und deshalb wäre dieses Verfahren für die hier vorliegenden Datensätze der Straftatenverteilung unbrauchbar (vgl. Brimbicombe (2003), S. 3). Brimbicombe selber schlägt ein Verfahren vor, das er Geo-Pro-Zones (GPZ) nennt und welches durch „... a recursive decomposition of space, similar to quadtrees, but allowing variable decomposition ratios and rectangular cells“ (vgl. Brimbicombe (2003), S. 4) Cluster abgrenzen soll. Dieses Verfahren arbeitet darüber hinaus ohne Parametereingabe durch die Anwender.

7.2 Entdecken raum-zeitlicher Cluster durch „Spatio-Temporal“ Data Mining

Die bisher in den vorhergehenden Kapiteln vorgestellten Untersuchungen bezogen sich überwiegend auf die Untersuchung der räumlichen Verteilung der Daten. Auch die bisher im Kapitel 6 durchgeführten Untersuchungen zur raum-zeitlichen Variabilität der Straftatenverteilung betrachtete primär die räumliche Verteilung der Daten zu unterschiedlichen Zeitpunkten und untersuchte somit die Zeit eher als eine Variable des Raumes, anstatt eine gleichberechtigte Untersuchung von Raum und Zeit vorzunehmen. Die Betrachtung eines Tatortes als „...discrete location in space and time...“ (vgl. Brantingham & Brantingham (1981) S. 7) muss jedoch die räumliche und die zeitliche Komponente gleichermaßen berücksichtigen. Vor diesem Hintergrund besteht die Hoffnung, dass mit Hilfe raum-zeitlicher Clusterverfahren oder anderer (raum-zeitlicher) Verfahren des Data-Minings eine gleichberechtigte Untersuchung der räumlichen sowie zeitlichen Beziehungen der Straftaten zueinander, bzw. zu anderen räumlichen oder zeitlichen Mustern möglich wird und sich neue Erkenntnisse ergeben.

8 Methodendiskussion

Die Suche nach Schwerpunktgebieten bestimmter Straftaten (Crime Hotspots) wird in vielen Orten zu einem zentralen Thema der polizeilichen Arbeit. Die Lokalisierung dieser Hotspots führt dazu, dass die der Polizei, bzw. den jeweiligen Verantwortlichen in den Revieren, zur Verfügung stehenden Ressourcen gezielt in diesen Schwerpunktgebieten eingesetzt werden können (vgl. Ratcliffe (2004c), S. 5). Aus der Analyse der räumlichen und zeitlichen Verteilung von Straftaten in einem Gebiet gibt Ratcliffe (2004c, S. 17) dann in einer sog. Hot-Spot Matrix Empfehlungen für die Einsatzplanung der Polizeikräfte.

Auch die Hamburger Polizei hat eine Crime-Mapping Anwendung im Einsatz. Ein Ziel dieser Intranet-Anwendung besteht darin, Schwerpunkte frühzeitig zu erkennen und geeignete Maßnahmen zu entwickeln. Die Darstellung der Straftaten als Punktdaten erschweren jedoch die Interpretation der Darstellung. Vor dem Hintergrund dieser Ausgangsbedingungen werden im Rahmen dieser Arbeit verschiedene Methoden im Hinblick auf ihre Aussagekraft und Anwendbarkeit im Rahmen des Projektes GISPOL der Hamburger Polizei untersucht. Es wurden Methoden betrachtet, die im Crime Mapping bereits etabliert sind (Quadrat-Methode, Kernel-Density Estimation) und andere Methoden, die noch nicht, bzw. erst selten eingesetzt wurde (Potentialanalyse, aber auch die LISA-Statistiken, wie sie hier zum Einsatz kommen). Die Untersuchungen beschränken sich dabei nicht auf eine Darstellung der räumlichen Verteilung sondern werden auf die Darstellung raum-zeitlicher Verteilungen ausgedehnt.

Die abschließende Methodenbewertung geschieht vor dem Hintergrund, dass die ausgewählten Methoden in eine Intranet-Anwendung integriert werden sollen. Die Antwortzeiten der einzelnen Berechnungen sollten daher möglichst kurz sein. Diese Forderung lässt sich jedoch selbst für die Kernel-Density Estimation, die am schnellsten Ergebnisse präsentieren konnte, nicht befriedigend umsetzen, da auch die KDE – an Anhängigkeit von der Größe der zu berücksichtigenden Datensätze einige Sekunden bis zu einigen Minuten benötigt und dieses Antwortverhalten von den Anwendern im Intranet nicht toleriert wird. Vor diesem Hintergrund wäre denkbar, dass verschiedene Berechnungen automatisiert ablaufen und die Ergebnisse zu bestimmten Zeiten

bereitgestellt werden oder aber, dass zwar Berechnungen sozusagen – on demand – ausgeführt werden können, die Ergebnisse jedoch nicht unmittelbar zu erwarten sind.

Weiterhin ist zu berücksichtigen, dass die wenigsten Anwender Erfahrungen in der Anwendung und Interpretation räumlicher Analysen haben. Es sind deshalb Verfahren vorzuziehen, die mit wenigen Eingabeparametern auskommen oder bei denen sich die Eingabeparameter im Hintergrund programmtechnisch mitgeben lassen. In diesem Fall muss gewährleistet sein, dass die Parameter bei variierenden räumlichen Datenverteilungen zu qualitativ gleichwertigen Ergebnissen führen werden.

Der Methodenvergleich im Kapitel 4.2.4 (des II. Teils) ergab, dass die Potentialanalyse mit den wenigsten Parametern auskam. Die Anzahl der Parameter war dabei abhängig von der eingesetzten Bezugsfläche (Thiessen-Polygone oder Rasterzellen). Dabei wurde die Darstellung der Straftatenverteilung, die sich auf die Thiessen-Polygone bezog, favorisiert, da hier die Flächenbezugspunkte identisch waren mit den Deliktspunkten und sich deshalb der Einfluss des Zufalls auf die Darstellung der Straftatenverteilung eliminieren ließ. Dieses Verfahren beansprucht (aktuell) jedoch die längste Rechenzeit. Sollte dieses Verfahren zum Einsatz kommen, muss der verwendete Algorithmus beschleunigt werden.

Ein Methodenvergleich (Potentialanalyse vs. Kernel-Density Estimation) der raumzeitlichen Verteilung und die Überlegungen bezüglich der möglichen Darstellung zeigte dann jedoch Vorteile der Kernel-Density-Estimation in Bezug auf die Klassifizierung der Ergebnisse und deren vergleichende Bewertung im Zeitverlauf. Die für die Kernel-Density-Estimation eingesetzte Klassifizierungsmethode der „Gleichen Intervalle“ bewirkt, dass die Ergebnisse einfach zu vergleichen und zu interpretieren sind.

Auch die LISA-Statistiken lieferten Ergebnisse, die sowohl die räumliche als auch die raumzeitliche Verteilung der Straftaten erklären. Da diese Methoden jedoch nicht nur die räumliche Verteilung der Punkte im Raum, sondern auch die räumliche Verteilung eines bestimmten Attributs (wie z.B. hier die Häufigkeiten der Straftaten pro Straßenabschnitt), können diese Methoden nur dann eingesetzt werden, wenn die Straftaten weiterhin auf die Straßenabschnitte bezogen werden. Hiervon hat die Polizei Hamburg im Verlauf des Jahres 2004 jedoch Abstand genommen: Die Straftaten werden

jetzt auf die Koordinaten der Adressen bezogen, die im Verlauf der Straftatenerfassung eingegeben wurde (mdl. Mitteilung Kondoch, PP Hamburg). Ob die LISA-Statistiken weiterhin ein hohes Erklärungspotential besitzen, müsste vor dem Hintergrund dieser Entwicklung erneut getestet werden.

Abschließend wurde das Themengebiet des Spatial-Data Mining gestreift. Dieses Thema wird in letzter Zeit intensiv erforscht und in der Fachliteratur diskutiert. Es werden verschiedene Verfahren ausprobiert, jedoch hat sich bisher noch kein Verfahren als das bestgeeignete Verfahren für die Untersuchung raumbezogener Fragestellungen herausgestellt. Möglicherweise wird dies aufgrund der Vielfalt der zu untersuchenden Phänomene auch nicht möglich sein. Die eingesetzten Clusterverfahren sollten dabei möglichst nach an den Datenquelle – also der Datenbank, bzw. dem Datawarehouse ansetzen. Mit räumlichen Clusterverfahren ließe sich dann eine horizontale Auswertung der Daten innerhalb eines Layers durchführen. Zusätzlich dazu können Assoziationsmethoden eingesetzt werden, die dabei helfen, vertikale Beziehungen in Daten aufzudecken, die in verschiedenen Layern abgelegt sind (vgl. Estivill-Castro & Lee (2001)).

In Bezug auf die Intranet-Anwendung GISPOL würde dieses Vorgehen – insbesondere vor dem Hintergrund der Antwortzeiten auf benutzerdefinierte Anfragen – eine interessante Alternative bedeuten. Die Entwicklungen auf dem Feld des Spatial Data Mining sollten deshalb weiter verfolgt werden.

Literaturverzeichnis

ANSELIN, L. (1995): Local Indicators of Spatial Association – LISA. In: Geographical Analysis 27 (2), S. 93 – 115.

ANSELIN, L. & BAO, S. (1997): Exploratory Data Analysis Linking SpaceStat and ArcView In: FISCHER, M. M. & GETIS, A. (ed.) (1997) Recent Developments in Spatial Analysis. Berlin, Heidelberg, New York: Springer, S. 35-59.

ARIZONA CRIMINAL JUSTICE COMMISSION (2002): Crime Mapping in Arizona Report [online]. Erreichbar unter: http://azcjc.gov/pubs/home/092502_CrimeMapping.pdf [zuletzt besucht am 10. 04. 2005].

BAILEY, L. (1994) A review of statistical spatial analysis in geographical information systems In: FOTHERINGHAM, S. & ROGERSON, P. (ed.) (1994) Spatial analysis and GIS, London, Bristol: Taylor & Francis, S. 13-44.

BAILEY, T. C. & GATRELL, A. C. (1995): Interactive Spatial Data Analysis. Harlow, Essex: Longman.

BLOCK, C. R. (1995): STAC Hot Spot areas [online]. Erreichbar unter: www.icjia.state.il.us/public/pdf/stac/Hotspot.pdf [zuletzt besucht am 29.03.2005].

BRANTINGHAM, P. J. & BRANTINGHAM, P. L. (1982): Environmental Criminology. Beverly Hills: Sage.

BRIMBICOMBE, A. J. (2003): A Variable Resolution Approach to Cluster Discovery in Spatial Data Mining. In: Lecture Notes in Computer Science 2669, S. 1-11.

BRISACH, C. E.; ULLMANN, R.; SASSE, G.; HÜBNER, E.; DESCH, R. (2001): Planung der Kriminalitätskontrolle. Stuttgart: Boorberg.

CATALANO, P. (2000): Applying geographical analysis to serial crime investigations to predict the location of future targets and determine offender residence [online]. Konferenzbeitrag für: Crime Mapping – Adding Value to Crime Prevention and Control. Australian Mineral Foundation, Adelaide. 21-22 September 2000. Erreichbar unter: www.aic.gov.au/conferences/mapping/catalano.pdf [zuletzt besucht am 02. 04. 2005].

CHAINEDY, S.; REID, S.; STUART, N. (2002): When is a Hotspot a Hotspot? A procedure for creating statistically robust hotspot maps of crime. In: KIDNER, C.; HIGGS, G.; WHITE, S. (2002) Socio-Economic Applications of Geographic Information Science, London.

CLARKE, D. & MADDEN, M. (2001): Regional Science on business. Berlin, Heidelberg, New York: Springer.

ECRI (2001): How does Geographic Profiling work [online]? Erreichbar unter: www.geographicprofiling.com/geopro/howdoesitwork.html [zuletzt besucht am 24. 04. 2005].

ESRI (2004): Online-Hilfe zu ArcView 9.0.

ESRI (2005): Law Enforcement and the GIS Enterprise [online]. Erreichbar unter: www.esri.com/industries/lawenforce/business/law_gis_enterprise.html [zuletzt besucht am 24. 04. 2005].

ESTER, M. & SANDER, J. (2000): Knowledge Discovery in Databases: Techniken und Anwendungen. Berlin, Heidelberg, New York. Springer.

ESTER, M., KRIEGEL, H.-P., SANDER, J. (2001): Algorithms and Applications for Spatial Data Mining [online]. Erreichbar unter: www.dbs.informatik.uni-muenchen.de/Publikationen/Papers/Chapter7.revised.pdf [zuletzt besucht am 16. 04. 2005].

ESTIVILL-CASTRO, V. & LEE, I. (2001): Data Mining Techniques for Autonomous Exploration of Large Volumes of Geo-referenced Crime Data [online]. Erreichbar unter: www.cs.helsinki.fi/u/gionis/seminar_papers/estivillcastro01data.pdf [zuletzt besucht am 23. 05. 2005].

ESTIVILL-CASTRO, V. & LEE, I. (2002): Multi-Level Clustering and its Visualization for Exploratory Spatial Analysis. In: GeoInformatica 6 (2), S. 123-152.

FISCHER, M. M. & GETIS, A. (ed.) (1997): Recent Developments in Spatial Analysis. Berlin, Heidelberg, New York: Springer.

FOTHERINGHAM, S. & ROGERSON, P. (ed.) (1994): Spatial analysis and GIS, London, Bristol: Taylor & Francis.

FOTHERINGHAM, A. S. & ZHAN, F. B. (1996): A Comparison of Three Exploratory Methods for Cluster Detection in Spatial Point Patterns. In: Geographical Analysis 28 (3). S. 200-218.

FOTHERINGHAM, A. S., BRUNSDON, C., CHARLTON, M. (2000): Quantitative Geography: Perspectives on Spatial Data Analysis. London, Thousand Oaks, New Delhi: SAGE Publications.

FRANCIS, C.; LALISHA, H.; MCCORMICK, J.; PALMER, N. (2002): The design and development of WEBCAT: An online crime analysis tool [online]. Erreichbar unter: www.sys.virginia.edu/capstone/past/cap2002/2002-14.pdf [zuletzt besucht am 06. 04. 2005].

GATRELL, A.C.; BAILEY, T. C.; DOGGLE, P.J.; ROWLINGSON †, B. S. (1996): Spatial point pattern analysis and its applications in geographical epidemiology. In: Transactions of the Institute of British Geographers N. S. (21), S. 256-274.

GETIS, A. & ORD, J. K. (1992): The Analysis of spatial Association by use of Distance Statistics. In: Geographical Analysis 24 (3). S. 189-206.

GRIFFITH, D.A. & AMRHEIN, G. (1991): Statistical Analysis for Geographers. London, Sidney, Toronto: Prentice-Hall.

HAGGETT, P.; CLIFF, A. D.; FREY, A. (1977): Locational Analysis in Human Geography. 2. Auflage. London: Arnold.

HAINING, R. (1990): Spatial data analysis in the social and environmental sciences. Cambridge: Cambridge University Press.

HAKE, G.; GRÜNREICH, D.; MENG, L. (2002): Kartographie. 8. Auflage. Berlin, New York: de Gruyter.

HAN, J. (2003): Mining Spatiotemporal Knowledge: Methodologies and Research Issues. [online]. Erreichbar unter: www.ucgis.org/Visualization/whitepapers/Han-gis03.pdf [zuletzt besucht am 23. 04. 2005].

HARRIES, K. (1999): Mapping Crime: Principles and Practice [online]. Research Report. National Institute of Justice. Crime Mapping Research Center. NCJ 178919. Erreichbar unter: www.ncjrs.org/html/nij/mapping/pdf.html. [zuletzt besucht am 02. 04. 2005].

HEARNSCHAW, H. M. & UNWIN, D. J. (1994): Visualization in Geographical Information Systems. Chichester: Wiley.

HENRY, L. M. & BRYAN, B. A. (2000): Visualising the spatio-temporal patterns of motor vehicle theft in Adelaide, South Australia. [online]. Konferenzbeitrag für: Crime Mapping – Adding Value to Crime Prevention and Control. Australian Mineral Foundation, Adelaide. 21-22 September 2000. Erreichbar unter: www.aic.gov.au/conferences/mapping/henry.pdf [zuletzt besucht am 02. 04. 2005].

HIRSCHFELD, A. & BOWERS, K. (ed.) (2001): Mapping and Analysing Crime Data. London: Taylor & Francis.

HIRSCHFELD, A.; YARWOOD, D.; BOWERS, K. (2001): Crime Pattern Analysis, Spatial targeting and GIS: The Development of New Approaches for use in Evaluating Community Safety Initiatives. In: CLARKE, D. & MADDEN, M. (2001) Regional Science on business. Berlin, Heidelberg, New York: Springer. S. 323-342.

ICJIA (2005): STAC Facts [online]. Erreichbar unter: www.icjia.state.il.us/public/index.cfm?metasection=Data&metapage=StacFacts [zuletzt besucht am 10. 04. 2005].

KIDNER, C.; HIGGS, G.; WHITE, S. (2002): Socio-Economic Applications of Geographic Information Science, London.

KRAAK, M.-J. & ORMELING, F. (2003): Cartography: Visualization of Geospatial Data. 2. Auflage. Harlow, England: Pearson Education.

LEVINE, N. (1999): CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations [online]. Konferenzbeitrag für: GeoComputation99. Mary Washington College, Fredericksburg, VA, USA. 25 – 28. Juli 1999. Erreichbar unter: www.geovista.psu.edu/sites/geocomp99/about.htm. [zuletzt besucht am 02. 04. 2005].

LEVINE, N. (2002): CrimeStat II: A Spatial Statistics Program for the Analysis of Crime Incident Locations, Chapter 7 – 'Hot Spot Analysis II' [online]. Erreichbar unter: www.icpsr.umich.edu/NACJD/crimestat.html [zuletzt besucht am 13. 04. 2005].

LEYKAUF, J. (1983): Erfassung und Darstellung von Bevölkerungsverteilungen durch Potentiale und deren Trendflächen. In: Petermanns Geographische Mitteilungen : PGM. 3/83. Gotha: Klett-Perthes. S. 161-165.

LKA HAMBURG (2003): Lageauswertung und Kriminalitätsdarstellung in Hamburg 2003. Internes Dokument.

LONGLEY, P. A.; GOODCHILD, M. F.; MAGIURE, D. J.; RHIND, D. W. (2001): Geographic Information Systems and Science. Chichester: Wiley.

MURRAY, A. T., MCGUFFOG, I., WESTERN, J. S. , MULLINS, P. (2001): Exploratory Spatial Data Analysis Techniques for Examining Urban Crime. In: British Journal of Criminology Vol. 41, S. 309-329.

OKABE, A.; BOOTS, B.; SUGIHARA, K. (1992): Spatial Tesselations: Concepts and Applications of Voroni Diagrams. Chichester: Wiley.

OKON, G. & WEINREICH, G. (1999): Darstellung der Kriminalitätslage unter Verwendung von GIS [online]. Konferenzbeitrag für: 14th European ESRI, 7th ESRI German User Conference, ESRI Germany 20th Anniversary. Munich. 15 – 17. 11. 1999. URL: <http://gis.esri.com/library/userconf/europroc99/html/vortraege/v05/head05.html>. [zuletzt besucht am 02. 04. 2005].

OKON G. & DOERFFEL, G. (2003): Arresting times in Munich. In: GEO:connexion. March 2003, S. 44-45.

THE OMEGA GROUP (2002): CrimeView®: Complete GIS Solutions in Crime Analysis Mapping [online]. Erreichbar unter: www.theomegagroup.com/downloads/slideshow_cv2002.zip [zuletzt besucht am 03. 04. 2005].

O'SULLIVAN, D. & UNWIN, D. J. (2003): Geographic Information Analysis. Hoboken, New Jersey: Wiley.

ORD, K. J. & GETIS, A. (1995): Local Spatial Autocorrelation Statistics: Distributional Issues and an Application In: Geographical Analysis 27 (4). S. 286-306.

OPENSHAW, S.; CHARLTON, M.; WYMER, C.; CRAFT, A. (1987): A Mark 1 Geographical Analysis Machine for the automated analysis of point data sets. In: International Journal of Geographical Information Systems 1 (4). S. 335 – 358.

OPENSHAW, S.; WAUGH, D.; CROSS, A. (1994): Some ideas about the use of Map Animation as a Spatial Analysis Tool. In: HEARNSCHAW, H. M. & UNWIN, D. J. (1994) Visualization in Geographical Information Systems. Chichester: Wiley. S. 131-138.

OTT, T. & SWIACZNY, F. (2001): Time-Integrative Geographic Information Systems: Management and Analysis of Spatio-Temporal Data. - Berlin, Heidelberg, New York. Springer.

PANDLER, T. (2002): Das Intranet-GIS der Hamburger Polizei. In: arcaktuell 4/2002. S. 22-23.

PEUQUET, D. J. & GUO, D. (2000): Mining Spatial Data Using an Interactive Rule-Based Approach [online]. Erreichbar unter: www.giscience.org/GIScience2000/papers/232-Peuquet.pdf [zuletzt besucht am 23. 04. 2005].

PHILIPS, P. D. (1972): A prologue to the geography of crime. In: Proceedings, Association of American Geographers 4. S. 86-91.

POLIZEI HAMBURG LANDESKRIMINALAMT (LKA) (2003): Polizeiliche Kriminalstatistik 2003. [online]. Freie und Hansestadt Hamburg Behörde für Inneres. Erreichbar unter: http://fhh.hamburg.de/stadt/Aktuell/behoerden/inneres/polizei/broschueren/pks2003-pdf_property=source.pdf [zuletzt besucht am 02. 04. 2005].

POOLER, J. (1987): Measuring Geographical Accessibility: a Review of Current Approaches and Problems in the Use of Population Potentials. In: Geoforum 18 (3). S. 269-289.

RATCLIFFE, J. H. (2000): Visualising crime hotspots and making sense of high volume crime. Konferenzbeitrag für: Crime Mapping – Adding Value to Crime Prevention and Control. Australian Mineral Foundation, Adelaide. 21-22 September 2000. Erreichbar unter: www.aic.gov.au/conferences/mapping/ratcliffe.pdf [zuletzt besucht am 03. 04. 2005].

- RATCLIFFE, J. H. (2002): Aoristic Signatures and the Spatio-Temporal Analysis of High Volume Crime Patterns. In: Journal of Quantitative Criminology 18 (1), S. 23-43.
- RATCLIFFE, J. H. (2004a): Geocoding crime and a first estimate of a minimum acceptable hit rate. In: International Journal of geographical Information Science 18 (1), S. 61-72.
- RATCLIFFE, J. H. (2004b): Crime Mapping and the Training Needs of Law Enforcement. In: European journal on criminal policy and research 10(1), S. 65-84.
- RATCLIFFE, J. H. (2004c): The Hotspot Matrix: A Framework for the Spatio-Temporal Targeting of Crime Reduction [online]. Erreichbar unter: [http://j.ratcliffe.net/papers/Ratcliffe%20\(2004\)%20The%20Hotspot%20Matrix.pdf](http://j.ratcliffe.net/papers/Ratcliffe%20(2004)%20The%20Hotspot%20Matrix.pdf) [zuletzt besucht am 16. 04. 2005]:
- RATCLIFFE, J. H. & McCULLAGH, M. J. (1999): Hotbeds of crime and the search for spatial accuracy. In: Journal of Geographical Systems 1. S. 385-398.
- RATCLIFFE, J. & McCULLAGH, M. J. (2001): Crime, repeat victimisation and GIS. IN: HIRSCHFELD, A. & BOWERS, K. (ed.) (2001) Mapping and Analysing Crime Data. London: Taylor & Francis. S. 61-92.
- SCHIEFER, K. (2002): Von der Steckkarte zur Animation. In: arcaktuell 4/2002. S. 16-17.
- SCHUHMACHER, U. (1989): Zur Anwendung des Potentialansatzes in der Geographie. In: Geographische Berichte 131 (2). S. 129-137.
- SCHWIND, H.-D. (2003): Kriminologie: eine praxisorientierte Einführung mit Beispielen. Heidelberg: Kriminalistik-Verlag.
- SILVERMAN, B. W. (1986): Density Estimation for Statistics and Data Analysis. London, New York: Champan & Hall.
- STATISTISCHES AMT FÜR HAMBURG UND SCHLESWIG-HOLSTEIN (2004): RBS-Information, internes Dokument.
- TETRAD (2005): Crime Analysis [online]. Erreichbar unter www.tetrad.com/new/crime.html [zuletzt besucht am 24. 04. 2005].
- TURTON, I. & OPENSHAW, S. (2001): Methods for automating the geographical analysis of crime incident data. In: HIRSCHFELD, A. & BOWERS, K. (ed.) (2001) Mapping and Analysing Crime Date. London: Taylor & Francis. S. 11-26.
- UPTON, G. & FINGLETON, B. (1985): Spatial data analysis by example. Vol. 1: Point Pattern and Quantitative Data. Chichester: Wiley.

VANN, I. B. & GARSON, G. D. (2003): Crime Mapping: New Tools for Law Enforcement. New York, Washington: Lang.

WALFORD, N. (1996): Geographical data analysis. Nachdruck. Chichester: Wiley.

WIKIPEDIA (2005): Data Mining [online]. Erreichbar unter:

www.wikipedia.org/wiki/datamining [zuletzt besucht am 16. 04. 2005].

WILLIAMSON, D.; MCLAFFERTY, S.; GOLDSMITH, V.; MOLLENKOPF, J.; MCGUIRE, P. (1999): A Better Method To Smooth Crime Incident Data [online]. In: ArcUser 01/99. Erreichbar unter: www.esri.com/news/arcuser/0199/crimedata.html.

Weitere Web-Adressen

Crime Mapping

CrimeStat: <http://www.icpsr.umich.edu/NACJD/crimestat.html> [zuletzt besucht am 24. 04. 2005]

Illinois Criminal Justice Information authority: <http://www.icjia.org/public/indec.cfm> [zuletzt besucht am 24. 04. 2005]

International Association of Crime Analyst: <http://www.iaca.net> [zuletzt besucht am 24. 04. 2005]

National Institute of Justice: <http://www.ojp.usdoj.gov/nij> [zuletzt besucht am 24. 04. 2005]

MAPS-Programm: <http://www.ojp.usdoj.gov/nij/sciencetech/ma.htm#crimemapping>
[zuletzt besucht am 24. 04. 2005]

CMAF: <http://www.crimeanalyst.net> [zuletzt besucht am 24. 04. 2005]

Policefoundation: <http://www.policefoundation.org> [zuletzt besucht am 24. 04. 2005]

Vera Institute of Justice: <http://www.vera.org> [zuletzt besucht am 24. 04. 2005]

Spatial Data Mining/Cluster Detection

SPIN!: www.ccg.leeds.ac.uk/spin [zuletzt besucht am 16. 04. 2005]

GeoVistaStudio: www.geovistastudio.psu.edu [zuletzt besucht am 17. 04. 2005]