8 Geomarketing-Prozess III: Komplexe GIS-Methoden und Umsetzung GBI-Tool

Grundlage vieler im Geomarketing erforderlicher Analysen sind Punktdaten, die gegebenfalls bereits auf einer Fläche verteilt vorliegen (Mikrozellen als Punkte und als Flächen) und vielfach durch *Thiessen-Polygone* dargestellt werden. Viele Marktdaten, die hier herangezogen werden, sind auf Gebäudeebene vorhanden, geometrisch gesehen als Punkt. Von daher eignet sich insbesondere die Anwendung von *Interpolationsverfahren, Point Pattern Analysen, Spatial Interaction Models* und *Gravitationsmodellen* zur Darstellung und Analyse der Phänomene. Für die Anwendung und Kombination dieser Verfahren sind Rasterzellen bzw. Grids hervorragend geeignet (vgl. VOSS & VOSS 2006, S. 35). Es wird hier zunächst auf die unterschiedlichen Möglichkeiten der geometrischen Abbildung zur Berechnung von Daten eingegangen. Schließlich wird dargelegt, warum der Grid-Methodik der Vorzug als Grundlage der Analyseberechnung gegeben wird.

8.1 Konzeption und Entwicklung GBI-Tool: Grid-Methodik und Greenfield-Analyse

8.1.1 Tesselation

Neben der Darstellung und Berechnung von Vektordaten wie Punkten, Linien, Flächen haben sich im Bereich des Geomarketings bisher *Thiessen-Polygone* zu einer wichtigen Analyseebene entwickelt. ARONOFF (1989, S. 218) definiert *Thiessen-Polygone* als diskrete Flächen um jedes Element einer Punktmenge. Die Flächengrenzen werden dabei so konstruiert, dass sie equidistant zu allen benachbarten Punkten liegen. Jede geometrische Position innerhalb eines *Thiessen-Polygons* ist näher zu dem Ursprungspunkt der Fläche als zu jedem anderen Ursprungspunkt der anderen Gebiete. Die Flächen werden auch als *Voronoi-Regionen* bezeichnet. Diese Methode wird in der Praxis genutzt, um z. B. Aussendienstmitarbeiter-Gebiete aufzuteilen ohne das weiße Flecken im Sinne von nicht zugeteilten Gebieten entstehen.

Weitere Möglichkeiten der Abbildung des Raumes stellt TIN (*Triangulated Irregular Network*) für die Interpolation von Höhenpunkten dar. Die verschiedenen Formen der Raumabbildung und Operatoren werden in detaillierter Form im Kapitel *Tesselations* LAURINI & THOMPSON (1992, S. 217ff) beschrieben. Hervorzuheben sind hier im

Gegensatz zu den vektorbasierten, die rasterbasierten Daten. Hierzu gehören *Grids* (*siehe Kapitel 8.1.1*) und *Quadtrees*[1] (LAURINI & THOMPSON 1992, S. 236-243).

Im Zusammenhang mit dieser Untersuchung wird der Grid-Struktur besondere Bedeutung zugemessen. Die zu untersuchenden geometrischen Daten, die hier im Fokus stehen, sind vorwiegend unregelmäßig verstreute Punktdaten (z. B. Post-Standorte, Wettbewerber-, Kunden-, Einzelhandelsstandorte, Haltestellen). Weiterhin wird auf die positiven Erfahrungen der Datenaufteilung in 100 x 100 Meter Grids der Schweizer Post zurückgegriffen[2]. Das *Bundesamt für Landestopografie* (*swisstopo*) bzw. die *Amtliche Vermessung* (AV) der Schweiz veröffentlicht alle erhobenen Daten auf dieser Ebene. Die Kleinräumigkeit der Daten birgt viele Vorteile gegenüber dem großräumigen Ebenen, auf denen in Deutschland Daten veröffentlicht werden. Um dennoch in den Genuß der Vielschichtigkeit der Möglichkeit von Datenanalysen auf dem Grid-Niveau zu kommen, wird hier neben *Filialinfo* ein Tool für Geomarketing konzipiert und implementiert. Das Geomarketing-Tool oder auch *'Geo Business Intelligence'* Tool wird im Weiteren als *GBI-Tool* bezeichnet.

8.1.2 Vergleich Analyse PLZ-Gebiet und Grid-Methodik

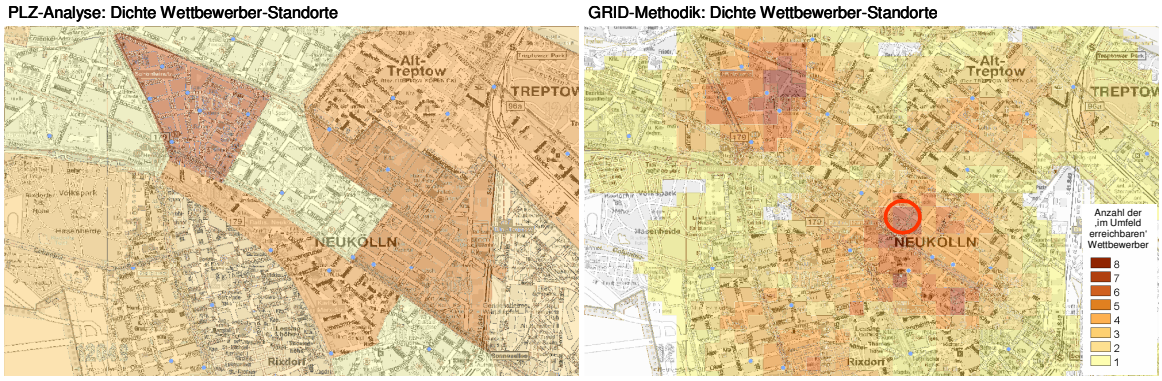
Ein weiterer Grund für die Bearbeitung der Analysen mit der Grid-Methodik wird in diesem Kapitel dargelegt. Hierzu wird ein einfaches Aggregationsverfahren auf PLZ-Ebene einem auf Grid-basierten Verfahren gegenübergestellt.

Ausgewählt wurde anhand einer makroräumlichen Analyse ein Gebiet, das mit Filialstandorten gegenüber den Wettbewerbern unterdurchschnittlich belegt ist. Bezüglich des Cluster liegt das Gebiet in dem 'Innenstadtbezirk West' von Berlin und in dem deutschlandweiten Cluster 14. Die Abbildung *(Abb.* 8-1*)* zeigt die Dichte von *Hermes*-Filialen.

---

[1] Rasterbasierte Grids und Quadtrees können wiederum vektorisiert werden.

[2] In einem von der Autorin durchgeführten Interview zum Thema 'Geomarketing' im Sommer 2004 mit der Schweizer Post (*Herr Pasche*) in Bern wurden die Vorteile des Grid-Vorgehens in der Schweiz erörtert.



Quelle: eigene Berechnung (Filialinfo/ GBI-Tool))

Abb. 8-1: Vergleich: Dichte Wettbewerber-Filialen auf PLZ-Ebene und auf Basis 100 x 100 m Grid.

Die linke Darstellung in der Abbildung gibt nur die Anzahl, der in jedem PLZ-Gebiet vorhandenen Filialen wieder. Je dunkler die Farbgebung wird, desto höhere Filialdichten herrschen vor. Die Wettbewerber-Filialen sind durch blaue Punkte gekennzeichnet. Es wird erkennbar, dass die Klassenzugehörigkeit jedes Gebietes in hohem Maße von den administrativen Gegebenheiten geprägt ist. Die Form und Größe jedes Gebietes nimmt somit Einfluss auf die darin befindliche Anzahl an Filialen. Letztendlich stellt diese Form der Darstellung daher kein besonders exaktes Abbild der Realität dar, was sich durch große Klassensprünge benachbarter Regionen äußert. Auch weitergehende Maßnahmen, bei denen beispielsweise die Anzahl in Relation zur Größe der Gebiete gestellt wird, vermögen den grundsätzlichen Engpass bei dieser Art der Analyse nicht zu umgehen.

Die rechte Darstellung in der Abbildung stellt denselben Sachverhalt auf Grid-Ebene dar. Hier wurde ein Vorgehen gewählt, welches sich auf Einzugsgebiete der Filialen stützt. Ausgehend von diesen Einzugsgebieten wird ein Raster berechnet, welches Zellgrößen von 100 x 100 Metern aufweist. Jede dieser Zellen bekommt den Wert 1 zugewiesen, wenn sich diese innerhalb eines Einzugsgebietes befindet. Für die Zellen, in denen sich die Einzugsgebiete überlagern, werden die Werte aufsummiert.

Durch diese Vorgehensweise wird eine Abhängigkeit von vordefinierten Gebieten vermieden, was zu einer Steigerung des Informationsgehaltes beiträgt. Eine hohe Aussagekraft wird nicht zuletzt dadurch erreicht, dass Nachbarschaftsbeziehungen zwischen den Filialen berücksichtigt werden. Diese drücken sich durch hohe Dichtewerte in solchen Räumen aus, in denen relative Agglomerationen der Filialen herrschen. So ist beispielsweise in der Laubestraße (*siehe roter Kreis in Abb. 8-1)* eine hohe Abdeckung zu verzeichnen, obwohl hier keine Filiale am Standort selbst vorhanden ist, dafür sich aber mehrere Einzugsbereiche in diesem Straßenabschnitt

überdecken. Diese Aussagen könnten sogar noch genauer erfolgen, wenn durch die Methode einer *Inversen Distanzgewichtung* (IDW) die Nähe zu jeder Filiale berücksichtigt wird. Hierbei nimmt der Wert in jeder Einzugsgebietzelle mit zunehmendem Radius ab. Daraus ergibt sich eine höhere Genauigkeit und Realitätsnähe.

Prinzipiell sind alle Geometrien von Vektordaten, also Punkt-, Linien- und Flächendaten, in Raster transformierbar und somit in einem Grid-basiertem System einsetzbar. Um mit den Rasterdaten weitere Analysen vornehmen zu können, müssen diese allerdings quantifizierte Attribute besitzen. Daher eignen sich qualitative Attribute von Objekten, wie etwa Namen oder Kategorien, für den Einsatz in Grid weniger. Vorab klassifizierte Daten sind aus gleichem Grunde ebenfalls nur bedingt geeignet, da hieraus lediglich Häufigkeitsverteilungen zu bilden sind.

Insbesondere Punktlayer im mikroräumlichen Bereich, wie Gebäudedaten oder Mikrozellen, sind für den Einsatz für Grid-Methoden zweckentsprechend. Sozio-demographische Daten lassen sich aus der Geometriestruktur am direktesten in Grids übertragen. Weiterhin werden Standortdaten wie das Filial- und das Wettbewerbernetz als Layer hinzugezogen. Die jeweiligen Dichten können mit Grid berechnet und dargestellt werden.

Die Stärke der Grid-Methodik liegt in der Möglichkeit, mathematische und statistische Operationen großflächig und zudem mit hoher geographischer Genauigkeit durchzuführen. Die Grid-Methodik lässt dem Benutzer viel Freiraum für Analysemethoden. Für die Integration in das GIS-gestützte Geomarketing-System kommen vielfältige statistische Verfahren in Frage. Die automatisierte Umrechnung von Daten auf Grid-Zellen als Voraussetzung für mathematisch, statistische Verfahren, die Einbeziehung der Gewichtungsfaktoren wie Distanzgewichtung anhand eines Attributes sind wichtige Faktoren. Darüber können auch einzelne Attribute für diverse Gewichtungen herangezogen werden. So läßt sich *a priori* der Einfluss jeder einzelnen (Wettbewerber-)Filiale über die Verkaufsfläche o. ä. einbeziehen. Ein Attraktivitätsfaktor kann auch auf Grundlage bereits gemessener Untersuchungen und Erfahrungen zugewiesen werden. Hier kommen die unterschiedlichsten Methoden der *Gravitationsmodelle* zum Einsatz.

Bei der Systemintegration ist die Einbindung von Histogrammen sinnnvoll, damit die Verteilung von Werten im Untersuchungsgebiet im Überblick zu erkennen ist, Datenstrukturen und -muster aufgedeckt werden. Die Berechnung und Darstellung der jeweils wichtigsten statistischen Kenngrößen wie Mittelwert, Median, Modus etc. wird

ebenfalls integriert. Prognostizierende Verfahren wie Trendanalysen aus Zeitreihenvergleichen sind mit dem Grid-Verfahren möglich.

Zudem ergeben sich durch den Einsatz von Grid-Strukturen weiterführende Visualisierungsmöglichkeiten. Es sei hier auch auf die Möglichkeit der Generierung von 3D-Modellen *(siehe auch Kapitel 6.2.2)* hingewiesen, durch welche relevante Sachverhalte eindrucksvoll und anschaulich präsentiert werden können.

8.1.3 Konzeption eines GIS-gestützten Geomarketing-Systems auf Grid-Basis: GBI-Tool

Die Konzeption eines Grid-basierten Tools geht theoretisch zurück auf die *Map Analysis*[3]. In der Regel wird die kartographische Algebra auf der Grundlage von Rasterdaten diskutiert. Insbesondere bei neueren Implementierungen rasterorientierter Systeme wird auf diese Systematik eingegangen. Von daher darf die Erläuterung hier vorweg nicht fehlen, da in den Anwendungsfällen der Filialnetzoptimierung und Kundenanalysen insbesondere mit Grid-Strukturen gearbeitet wird.

Die Grundlagen der Entwicklung der *Map Analysis* gehen auf Arbeiten am *Harvard Lab* wie GRID und IMGRID-Software zurück (CARL STEINITZ & DAVID SINTON). In umfassender Form systematisiert sind diese Ideen des *Map Algebra* von DANA TOMLIN (1990). Zu dem Grundschema der *Map Algebra* gehört die Einteilung nach Operatoren:

- Lokale Operatoren transformieren oder kombinieren Geodaten an einer bestimmten Stelle im Raum, ohne jegliche Berücksichtigung anderer Positionen im Untersuchungsgebiet.
- Fokale Operatoren berücksichtigen neben dem aktuellen Standort dessen unmittelbaren Umgebung (Kernel, Maske). Eine Anwendung als Filter (z. B. Glättung) nach dem Prinzip der gleitenden Fenster (‘Moving Window’) ist aus der Bildverarbeitung bekannt.
- Zonale Operatoren ziehen als Kriterium für die Aggregation von Werten weiterer Themen Zonen heran. Als Zonen werden im Sinne von Klassen potentiell auch disjunkte Gebiete mit gleichartigen Eigenschaften bezeichnet.
- Globale Operatoren berücksichtigen potentiell das gesamte Untersuchungsgebiet, ein Beispiel dafür sind Distanzanalysen.

---

[3] Synonym zu *Map Analysis* werden die Begriffe *Map Algebra*, *Cartographic Modeling* oder *Kartographische Algebra* verwendet.

Die Grundidee liegt ähnlich der mathematischen Algebra in der Allgemeinheit der Verwendung von Variablen: räumliche Themen werden mit Namen bezeichnet und als Variable in Ausdrücken verwendet. Operatoren wirken somit auf ein oder mehrere Themen. Die Ergebnisvariable bezeichnet wiederum ein räumliches Thema. Ausdrücke der kartographischen Algebra werden in formelähnlicher Notation geschrieben. Operatoren können dabei einfach wie Addition, Substraktion oder auch spezifisch räumliche Operatoren sein, wie Flächengröße, Distanzmessung oder geometrische Formdeskriptoren. Jeder dieser Operatoren kann ein oder mehrere Eingabe- sowie ein Ausgabeobjekte aufweisen, die jeweils thematische Informationsschichten sind. Im folgenden bei Geomarketing spielen hier vor allem fokale (z. B. bei der Kerneldichteberechnung), zonale (Cluster) und globale Operatoren (Erreichbarkeits-analysen) eine Rolle. Diese wurden bei der Konzeption und Implementierung des *GBI-Tool* berücksichtigt.

Algorithmen, die den Operatoren zugeordnet werden können, wie *Kernel-Dichtefunktion, Gravitationsmodelle* und weitere, die in den nächsten Kapiteln diskutiert werden, sind in die Konzeption des Tools eingeflossen und bereits in einigen Varianten implementiert. Im Folgenden werden die methodischen Grundlagen zum Thema Grid- und *Greenfield-Analyse* erläutert.

Standortbezogene Analysen basieren in der Regel auf administrativen Einheiten, wie etwa Gemeinden, PLZ-Gebiete oder aber flächenbezogene Marktzellen. Jedes dieser Gebiete kann bestimmte Informationen in Form von Attributen transportieren. Je kleiner diese Einheiten sind, desto näher bewegen sich deren inhaltlichen Aussagen an der Realität. Oder anders formuliert: Mit einer Generalisierung, etwa durch das Zusammenfassen von Standorten oder kleinerer Gebiete zu größeren Einheiten, geht immer ein Informationsverlust einher.

Dieser Umstand führt insbesondere dann zum Verlust von Informationen, wenn es sich bei den Ausgangsdaten um Punktdaten, wie z. B. Filialstandorte, Gebäude o. ä. handelt. Um diese Daten vielen im Geomarketing gängigen Analyseverfahren zugänglich zu machen, müssen diese oftmals zunächst in flächenhafte Informationen umgewandelt werden.

Durch den Einsatz von Grid-basierten Analysemethoden kann auf vordefinierte Bezugseinheiten verzichtet werden, indem punkthafte Daten durch Interpolation oder Modellierung in Flächendaten umgewandelt werden können. Statt einfache Aggregationen anzuwenden, kommen hierbei zum Teil aufwendige mathematische Verfahren zum Tragen. Auf diese Weise werden diskrete Daten zu kontinuierlichen

Informationsträgern, welche in nachfolgenden Schritten für GIS- und Statistikverfahren Verwendung finden können. Die folgende Abbildung (*Abb. 8-2*) gibt einen Überblick über generelle Vor- und Nachteile Grid-basierter Analyseverfahren.

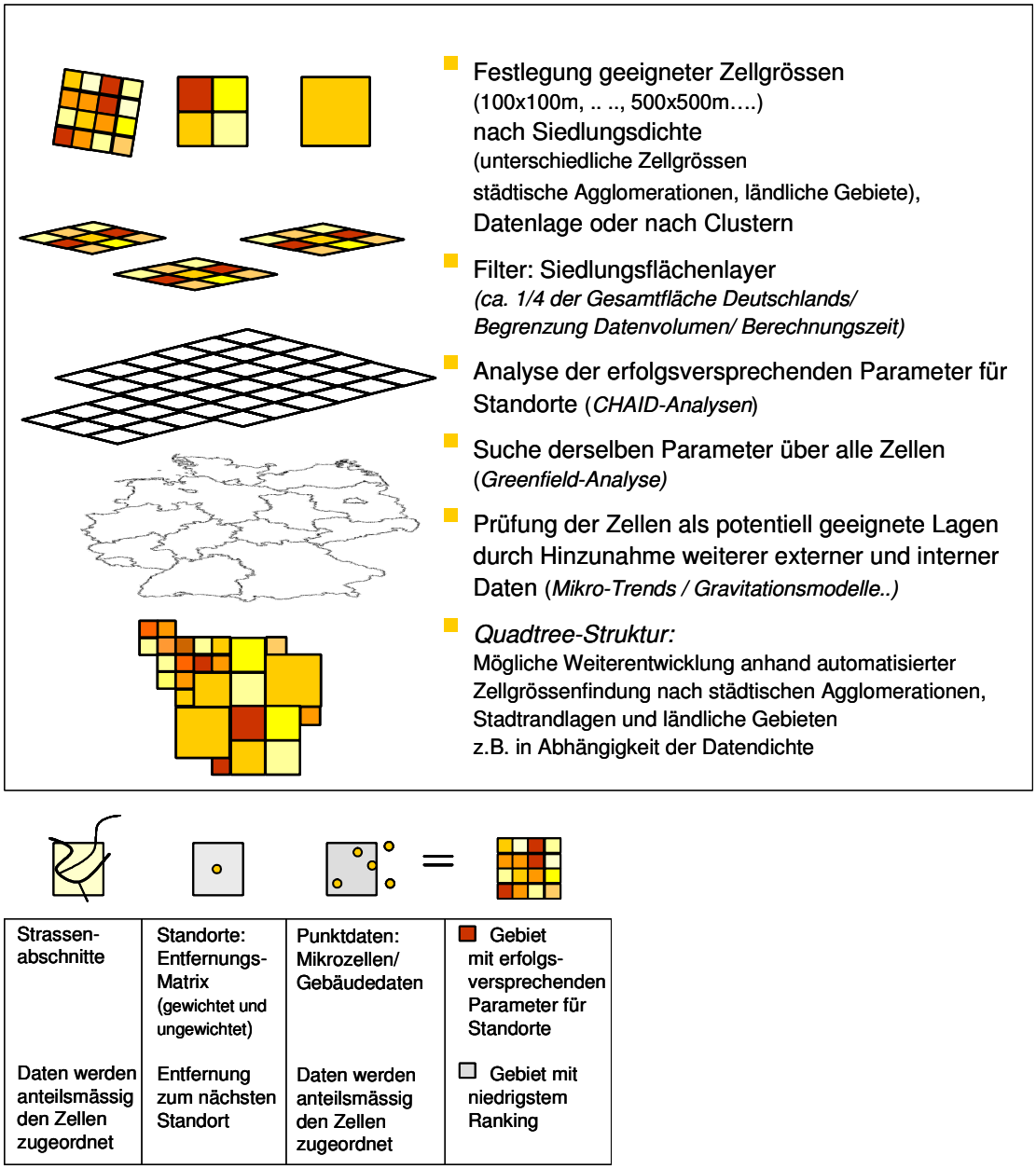
Vorteile	Nachteile
Flächendeckende Auswertungsverfahren bei geringem Informationsverlust Möglichkeit der Anwendung verschiedenster und individueller mathematischer Verfahren Unabhängigkeit von administrativen Einheiten bei der Datenaggregation Anschauliche Visualisierungsmöglichkeiten Einfache Einbindung in vorhandene GIS durch Vektor-to-Raster und Raster-to-Vektor Funktionen	z.T. hoher Speicher- und Rechenbedarf

Quelle: eigene Zusammenstellung

Abb. 8-2: Vor- und Nachteile Grid-basierter Analyseverfahren

Auf Grid-basierende Geoinformationen definieren alle geographischen Objekte als quadratische Zellen oder Pixel, welche zusammen ein Raster bilden. Jede dieser Zellen ist in der Regel gleich groß, wobei die Größe beim Erstellen des Rasters frei wählbar ist. Je feinmaschiger das Raster erstellt wird, desto genauer können die relevanten Daten räumlich abgebildet werden. Die Zellgröße ist dabei jedes Mal auch abhängig von der Skalierung der jeweiligen Analyse. Bei mikrogeographischen Verfahren, die gerade in städtischen Gebieten wichtig sind, ist die Größenordnung von 100 x 100 Metern für jede Zelle empfehlenswert. Jede Zelle kann für sich bestimmte Werte annehmen. In Abhängigkeit von der Art der zu analysierenden Daten können diese sowohl alphanumerisch (Namen oder Kategorien) als auch numerisch ausgeprägt sein. Letztere Eigenschaft eignet sich insbesondere dazu, geographisch kontinuierlich ausgeprägte Daten räumlich abzubilden. Zu solchen Daten zählen etwa die Einwohnerdichte, Kaufkraft oder Mietpreise.

Die Abbildung (*Abb. 8-3*) Grid Verfahren als Basis für *Greenfield-Analysen* zeigt die Vorgehensweise bei der Berechnung: Festlegung geeigneter Zellgrößen nach Agglomerationsbereichen und weniger besiedelten Gebieten, die Möglichkeit durch einen Siedlungslayer (*besiedelte Flächen anhand der ADAC-Karten 1:20:000 im Umriss digitalisiert und fortgeführt*) die Berechnungsfläche der Grids nur auf die relevanten Fläche zu begrenzen, die Erfolgsfaktoren (CHAID) über alle Zellen auszuwerten, Hinzunahme weitere Parameter und die Überlegungen hinsichtlich der Einbindung der *Quadtree*-Struktur.



Quelle: eigene Darstellung

Abb. 8-3: Grid-Verfahren als Basis für Greenfield-Analysen

Ein besonderer Vorteil der Nutzung der Grid-Struktur in methodischer Hinsicht ist die Möglichkeit, mehrere Informationsebenen zu bündeln und effektiv analysieren zu können. So kann theoretisch jede beliebige Rechenoperation auf Rasternetzen durchgeführt werden, wobei auch komplexe Ausdrücke verwendet werden können. Als

einfaches Beispiel einer Rechenoperation sei hier die Subtraktion von Einzugsgebieten zwei verschiedener Filialen genannt, mit welcher die Differenz des Filial-Einflusses für jede Zelle ermittelt werden kann.

Durch das einfache Verknüpfen von Datenschichten in dem Grid-Verfahren eignet sich das Verfahren für die Anwendung der *Greenfield-Analyse*. Hier geht es darum, erfolgsversprechende Parameter, die einen Filialstandort ausmachen, in weiteren Gebieten ausfindig zu machen.

8.1.4 Entwicklung eines GBI-Tools

Die Zielsetzung ist es ein Geomarketing-System zu entwickeln, welches auch in der Praxis Anwendung findet. Für die Umsetzung der Algorithmen und Modelle wurden Programmierroutinen für das Geographischen Informationssystem *Filialinfo* und Grid-basierte Funktionen in einem Zusatzmodul realisiert. Im Rahmen dieser Arbeit wurde dieses Zusatzmodul als Pilot erstellt - ein GIS-basiertes Geomarketing-Tool auf Basis von Grids – das *GBI-Tool*. Viele der im folgenden beschriebenen und für die Anwendung getesteten Algorithmen sind in dem Tool integriert. Sie können im nächsten Schritt verfeinert, erweitert oder auch ergänzt werden. In Kombination mit der GIS-Applikation *Filialinfo* besteht damit ein relativ umfassendes Instrumentarium für Geomarketing-Anwendungen.

In dem *GBI-Tool* sind Routingalgorithmen[4] integriert. Ein Daten- und Metadatenmodell ist erstellt, Karten und soziökonomische Daten sind integriert (*siehe Kapitel 5*).

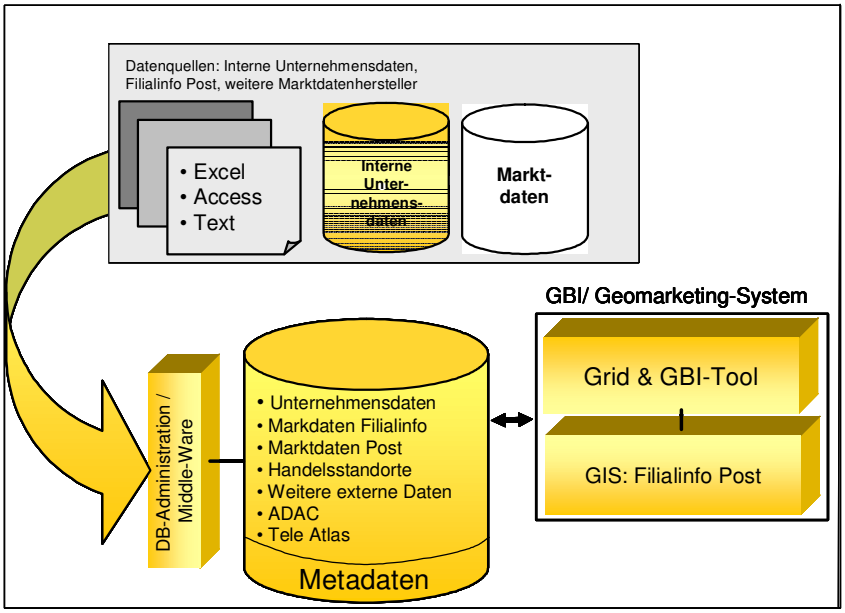
Die Erstellung des *GBI-Tools*[5] erfolgte anhand eines ausgereiften *Software-Engineering-Prozesses* (DENERT & SIEDERSLEBEN 1991), dabei ist die Modularität ein wichtiges Qualitätsmerkmal, die vornehmlich durch die Datenkapselung und Objektorientierung erzielt wurde. Es wurde auf Grundlage der hier gemachten Erkenntnisse im iterativen Vorgehen ein Anforderungsdokument, darauf aufbauend ein designtechnisches und Datenbank-Dokument erstellt (IVU 2005/2006) aus denen die weiteren softwaretechnischen Details hervorgehen, die aber hier nicht Gegenstand der Untersuchung sind.

---

[4] Konrad-Zuse Zentrum/IVU

[5] Der Programmcode ist zu Teilen im Besitz der Deutschen Post und darf hier nicht veröffentlicht werden.

Das Gesamtsystem wird *Geo Business Intelligence*-System oder synonym Geomarketing-System genannt *(Abb.* 8-4). Als Beispiel der Implementierung werden Systemmasken des *GBI-Tools* mit einer berechneten Erreichbarkeitsanalyse und einige Grid-Funktionen als Menuemasken dargestellt.

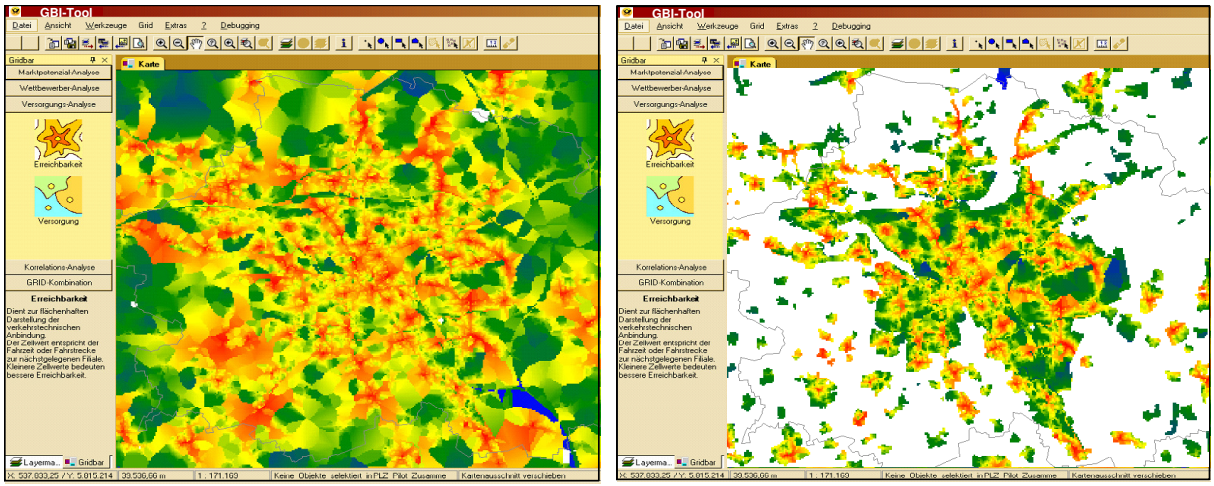


Quelle: eigene Darstellung

Abb. 8-4: Architektur Geomarketing-System (vereinfacht)

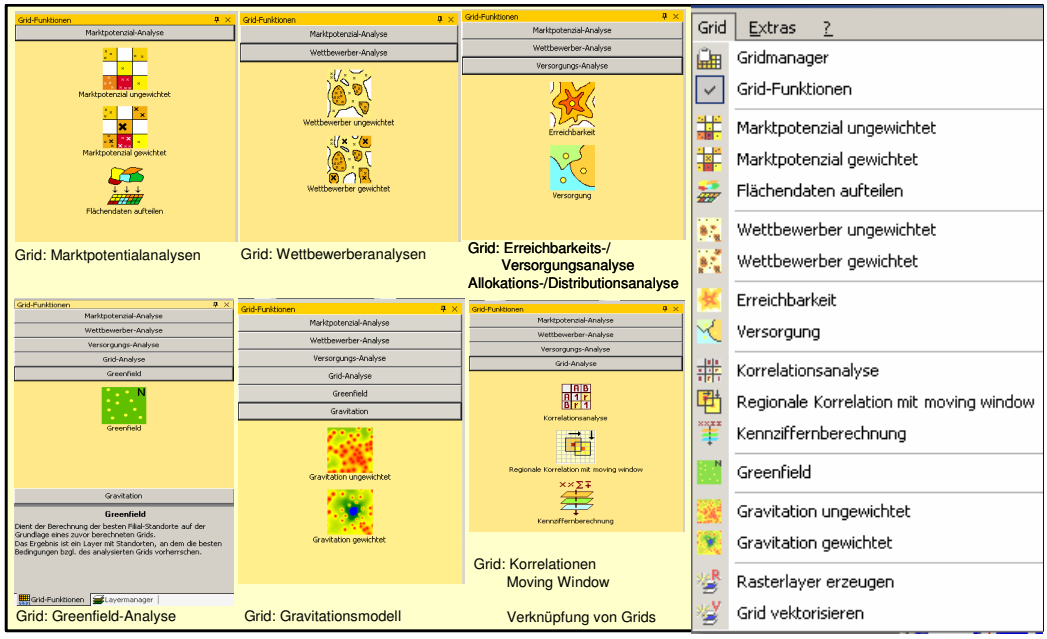
Der Einsatz von GIS und dem *GBI-Tool* erlaubt während der gesamten Untersuchung ständige Plausibilitätskontrollen in tabellarischer und visueller Form. Weiterhin ermöglicht der Einsatz dieser Systeme die effiziente Datenaufbereitung und die Integration und Verknüpfung verschiedenster Datenbestände, die allein über ihren räumlichen Bezug miteinander verbunden sind.

Die Abbildung (*Abb.* 8-5) zeigt beispielhaft eine Analyse mit Erreichbarkeitsalgorithmen (Interpolationsverfahren), die mit dem *GBI-Tool* durchgeführt wurde. Die roten Gebiete sind diejenigen, von denen aus eine Filiale in kurzer Zeit (berechnet anhand des Straßennetzes) erreichbar ist; die grünen Felder zeigen an, welche Gebiete am weitesten von Filialen entfernt liegen. Das Ergebnis wird mit dem Filter Siedlungsflächenlayer kombiniert, um nur die nachfragerelevanten im Sinne von bewohnten/besiedelten Flächen, zu evaluieren. Damit wird die zu bearbeitende Datenmenge und Berechnungszeit erheblich reduziert.



Quelle: eigene Berechnung/ Interpolation (GBI-Tool)

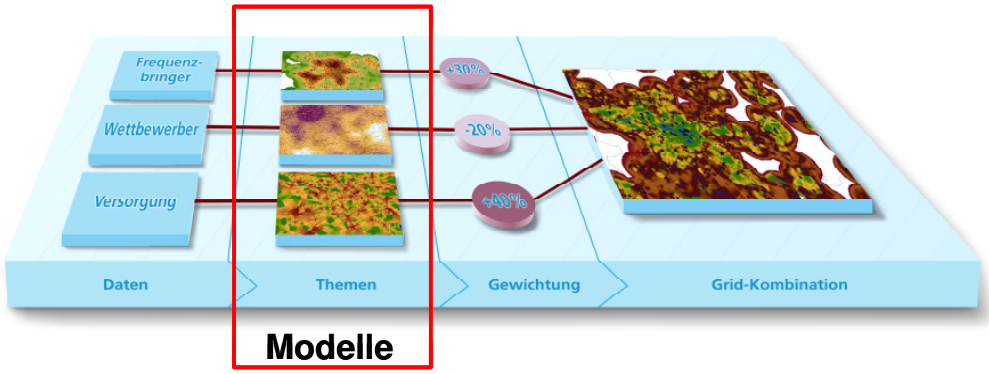
Abb. 8-5: GBI-Tool: Filter: Erreichbarkeit (links) und kombiniert mit Filter: Siedlungslayer (rechts)



Quelle: eigener Entwurf und Zusammenstellung

Abb. 8-6: GBI-Tool: Grid-basiertes Funktionsmenue

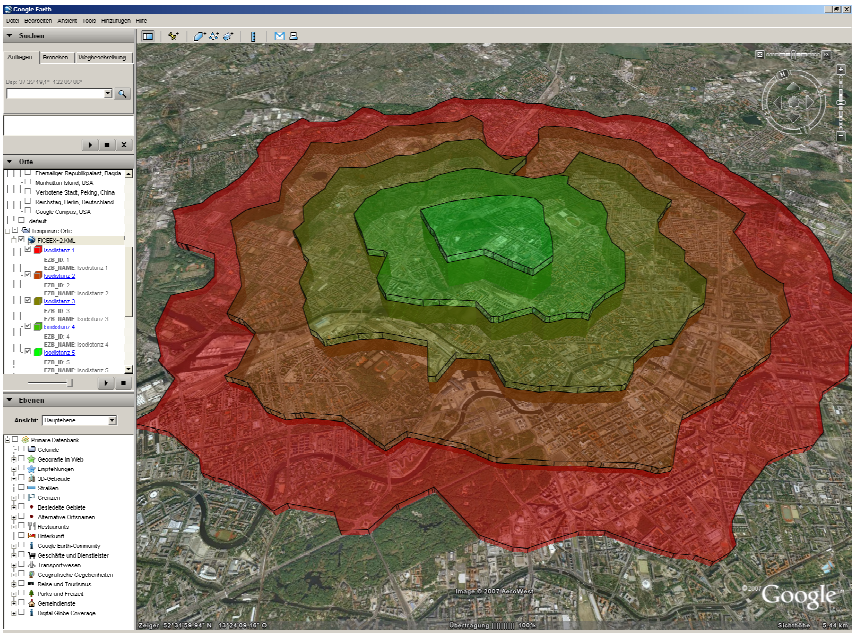
Die weiteren Schritte in der Entwicklung bestehen in dem direkten Aufrufen von Data Mining Routinen wie *Two-Step-Clusteranalyse* und *CHAID-Analyse* aus dem Gesamtsystem heraus. Diese wurden bisher in SPSS und SPSS *Clementine* durchgeführt. Die Konzeption für die Integration ist im Zuge des Vorhabens erarbeitet worden.



Quelle: eigener Darstellung

Abb. 8-7: Modellablauf: Daten, Themen, Gewichtung, Grid-Kombination

Als Ausblick wird für eine weitere Entwicklung die neuesten Erkenntnisse aus dem Bereich der Visualisierung: 3D Stadtmodelle (*Hasso-Plattner-Institut*, *3D Geo GmbH*) für Geomarketinganwendungen berücksichtigt, sowie die Einbindung von Daten und Berechnungen in einem *Google Earth Client*. Gerade die intelligente Visualisierung wird zunehmend wichtiger für eine schnelle Interpretation von Ergebnissen. Die derzeitigen Darstellungen der kombinierten Grids *(Abb. 8-7)* können zwar einerseits noch mit einem Stadtplan hinterlegt werden, aber dennoch ist die Interpretation für den Anwender ungewohnt. Weitere Dimensionen hinzuzunehmen, wird hier in Zukunft zielführend sein. Eine 3D-Visualisierung von Einzugsgebieten eines Standortes dargestellt in *Google Earth* kann eine solche visuelle Erweiterung bieten *(Abb. 8-8)*.



Quelle: eigene Darstellung/ Implementierung in Google Earth, Berechnung Isodistanz/ Routingberechnung auf Teleatlas

Abb. 8-8: 3D-Visualisierung 5 Isodistanzen um einen Standort in Google Earth

8.2 Interpolation, Regionalisierung, Schätzung räumlich kontinuierlicher Daten

Interpolationsmethoden finden dann Anwendung, wenn mehr oder weniger unregelmäßig verteilte, geocodierte Punktdaten Aufschluß über Gebiete geben sollen über die keine Daten vorhanden sind. Interpolation ist ein Prozeß, der aus Punktdaten für ein Untersuchungsgebiet neue zusätzliche Punktwerte berechnet. Es gibt für die Ermittlung der Werte unterschiedliche Verfahren: *Inverse Distance Weighting-Interpolation (IDW Interpolation), Triangulation , Rectangular Interpolation*. Jedes Verfahren birgt für eine bestimmte Anwendung Vorteile. Die *IDW Interpolation* wird vielfach im Bereich des Umweltmonitoring eingesetzt, die *Triangulation* für die Berechnung von Höhenmodellen und die *Rectangular Interpolation*, für gleichmäßig verteilte und dicht beieinander liegende Daten (siehe auch ISSACS, SRIVASTAVA, 1998, S. 250-337). Die Anwendung der Interpolations- und Modellierungsmethoden ist sehr stark von den Qualität und der Verteilungsdichte der dargestellten und untersuchten Attribute abhängig.

Für die Schätzung räumlich kontinuierlicher Daten sind geostatistische Interpolationsverfahren wie z. B. Glättungsverfahren, mit der Berechnung gleitender Mittelwerte oder das *Kriging* von großer Bedeutung. Die Interpolationsmethode beruht auf der Annahme, dass sich räumliche Verteilungen mit demselben Muster fortsetzen. Werte für Variablen an nicht beprobten Orten bzw. dort, wo keine Daten ermittelt wurden, werden aus den erhobenen Daten der Nachbarschaft geschätzt, mit dem Ziel aus den punktuellen Messungen oder Beobachtungen ein möglichst gutes Verständnis der räumlichen Variation thematischer Oberflächen in einem Untersuchungsgebiet zu erlangen.

Kriging ist ein lineares Prognoseverfahren (vgl. CHRISTENSEN 2000), das unter der Annahme eines räumlichen Verteilungsmusters (theoretische Variogramm) über gewichtete räumliche Mittelwertsbildung die Entwicklung der Ausprägung numerischer Daten in unterschiedlichen Entfernungen interpoliert[6]. *Kriging* wird im Zusammenhang mit GIS hauptsächlich verwendet, um Schätzungen von Oberflächen (z. B. Rohstoffvorkommen) oder Emissionsverläufen (z. B. Ausbreitungsmodelle) aus einem Satz einzelner Messpunkte vorzunehmen. Im Gegensatz zu multivariaten Verfahren fließt in *Kriging-Methoden* lediglich eine Variable ein, so dass es als Verfahren für die

[6] Beim *Kriging-Verfahren* wird im Gegensatz zu den linearen Modellen die räumliche Kovarianz anhand des Variogramms modelliert, nicht anhand des Kovariogramms (DRESSMANN 2004, S. 78).

hier geforderten Fragestellung, bei der viele Faktoren in Betracht gezogen werden müssen, nicht geeignet ist.

Das einfache Glättungsverfahren mit der Berechnung gleitender Mittelwerte ist in das *GBI-Tool* integriert und wird auf das Untersuchungsgebiet angewendet. Als Prognosewert wird an einem bestimmten Punkt der arithmetischen Mittelwert der Messpunkte aus einer kreisförmigen Umgebung dieses Punktes verwendet. Durch die gleitenden Mittelwerte kann eine Prognosefläche über das gesamte Untersuchungsgebiet bestimmt werden. Es wird ein Raster von 100 x 100 Metern im städtischen Untersuchungsgebieten gelegt. An jeder Stützstelle wird nun ein Prognosewert nach folgender Formel erstellt.

$$\hat{z}(s_o) = \frac{1}{n(s_o, r)} \sum_{s_i \in U(s_o, r)} z(s_i) \quad \textbf{(1)}$$

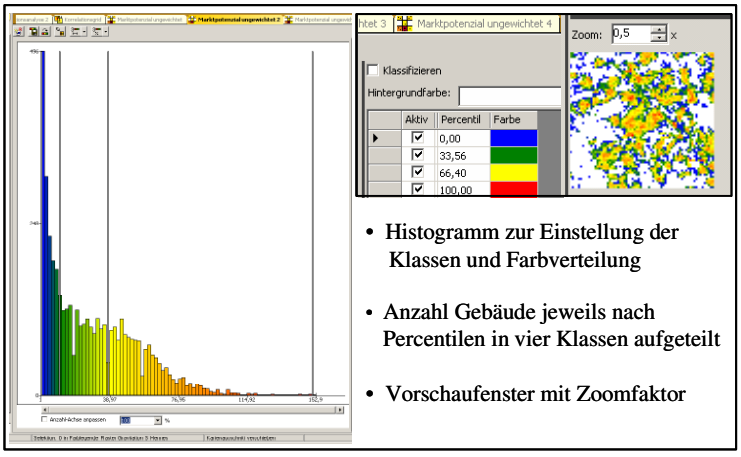
$\hat{z}(s_o)$ = Prognosewert am Punkt S_o
$U(s_o, r)$ = kreisförmige Umgebung um s_o mit Radius r
$n(s_o, r)$ = Anzahl Messpunkte in dieser Umgebung

Dieses Verfahren liefert eine Interpolation des Prozesses durch im Raum liegende Flächenstücke den einzelnen Rasterflächen. Wenn Umgebungen benachbarter Raster dieselben Messpunkte beinhalten, ergibt sich derselbe Prognosewert. Ein Sprung kann dann auftreten, sobald sich die Zusammensetzung der Messpunkte aus der Umgebung ändert. Für die Erkennung von lokalen Trends ist die Fläche möglichst klein zu wählen, eine größere Umgebung ergibt eine glattere Interpolationsfläche. Die Wahl von 100 x 100 Metern in städtischen Gebieten ermöglicht die Berücksichtigung ausreichend genügender Datenpunkte um Interpolationsverfahren anzuwenden. Gleichzeitig gibt sie eine für Planungszwecke ausreichend genauen Trend wieder. Eine Wahl von 50 x 50 Metern ist zwar möglich, da aber bei einer lokalen Betrachtung auf Straßenabschnitten die Einbeziehung regionaler Kenntnisse bzw. auch die 'Vor-Ort Begehung' vorausgesetzt wird, ist eine Raster von 100 x 100 Metern ausreichend, um diese Bereiche ausfindig zu machen. Getestet wurden weiterhin Rastergrößen von 200, 300 und 500 Metern. Ein 200 Meter Raster lässt noch die Identifizierung lokaler Trends zu, bei 300 Metern ist dies allerdings in einigen städtischen eng besiedelten Gebieten nicht mehr ausreichend, da hier zu viele unterschiedliche Regionen zusammengefasst werden. 500 Meter ist eine Umgebungsweite, die im ländlichen Raum durchaus gute Interpretation ermöglicht. Die Ergebnisse der Test werden hier in visueller Form (*Abb.* 8-9) vorgestellt.

Quelle: eigene Zusammenstellung GBI-Tool (Farbskalen nach Percentilen aufgeteilt siehe Abb. 8-10)

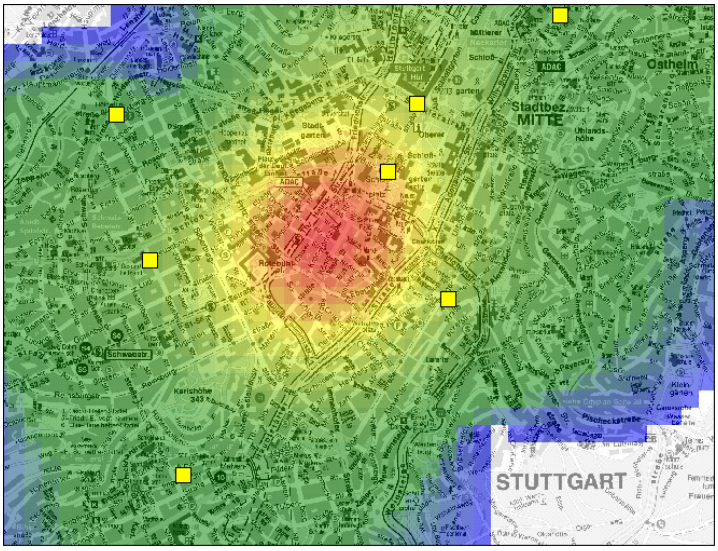
Abb. 8-9: Vergleich städtisches Gebiet mit unterschiedlichen Gridgrößen über Gebäudeanzahl

Eine mögliche Weiterentwicklung ist die automatisierte Berechnung der *Quadtree*-Struktur. Dort wo weniger Daten vorhanden sind, werden die Grids größer angelegt, meist ist das der ländliche Raum. Im städtischen Gebieten bei eng zusammenliegenden Datenpunkten mit unterschiedlicher Ausprägung wird das Maschennetz kleiner. Dies schafft eine schnelle Übersicht über die relevanten Ergebnisse. Weiterhin wird dadurch eine Trennung zwischen ländlichen und städtischen Räumen automatisch anhand der Engmaschigkeit von verfügbaren Daten erkannt. Es wird nicht nach administrativen Gegebenheiten vorgegangen, die oftmals keine genügend genauen Aussage zur Datenverteilung zulassen. Nachteil ist allerdings das Kombinieren von verschiedenen *Quadtreelayer*, aufgrund unterschiedlicher Abgrenzungen, weshalb hier zunächst nur die Grid-Variante mit regelmässig gebildeten Zellen weiter betrachtet wird. Eine unterschiedliche Zellgrößenberechnung wird nach Clustern (Ergebnis der *Two-Step-Clusteranalyse*) definiert.



Quelle: eigene Darstellung (GBI-Tool)

Abb. 8-10: Umsetzung der Algorithmen im GBI-Tool



Quelle: eigene Berechnung (GBI-Tool)

Abb. 8-11: Handelsumsatz mit Gewichtung nach IDW Interpolation im Grid 100 x 100 Meter.

Um ein kleines Untersuchungsgebiet auf sehr detaillierte Trends zu untersuchen, die hier *'Mikro-Trends'* genannt werden, ist es empfehlenswert, die Isodistanz oder Isochrone als Einzugsgebiet zu wählen. Diese gibt letztlich ein exakteres Bild von der zu erreichenden Umgebung als die Anwendung der Radiusbildung. Der Messpunkt muss in einer bestimmten maximal gleich zu erreichenden Entfernung oder zur gleichen Zeit zu erreichenden Entfernung liegen, d. h. die Daten werden innerhalb der über die maximalen Isodistanz/Isochronen entfernt liegende Punkt aufgespannt. In dem so entstandenen Einzugsgebiets (siehe *Kapitel 6.2.3*) werden die Messpunkte 'eingesammelt' und zur Berechnung herangezogen.

Die Umgebungsgröße kann entweder direkt über den Radius *r* gesteuert werden oder die Umgebung wird so eingestellt, dass sie eine feste Anzahl *k* von Messpunkten erhält. Damit variiert der Radius so, dass immer $n(s_o, r) = k$ gilt. Sind die Messpunkte sehr ungleichmäßig verteilt, dann empfiehlt sich die zweite Variante, um so zu vermeiden, dass bei festem Radius an bestimmten Stellen wenig Messpunkte in eine Berechnung eingehen. Die Grundlage der vorangegangenen Berechnung bezieht sich auf sozioökonomischen Daten, die auf Gebäudeebene vorhanden sind. Findet die Berechnungen innerhalb einer besiedelten Fläche statt, ist eine gleichmäßige Verteilung der Beobachtungen anzunehmen. Damit kann ein fester Radius zur Berechnung verwendet werden. Bezieht sich das Untersuchungsgebietes auf die gesamte Fläche von Berlin, ist aufgrund von großen Grün- und Wasserflächen an denen keine sozioökonomischen Daten vorhanden sind, die Variante mit der festen Anzahl an Messpunkten pro Radius empfehlenswert. Wenn bei der Berechnung der Filter Siedlungsflächenlayer im GIS einbezogen wird, dann hat die Wahl des Verfahrens keine Auswirkung auf das Ergebnis.

Eine weitere Variante ergibt sich, wenn man eine Gewichtung vornimmt, so dass der Einfluss der Umgebungswerte auf den Mittelwert mit zunehmenden Abstand abnimmt[7].

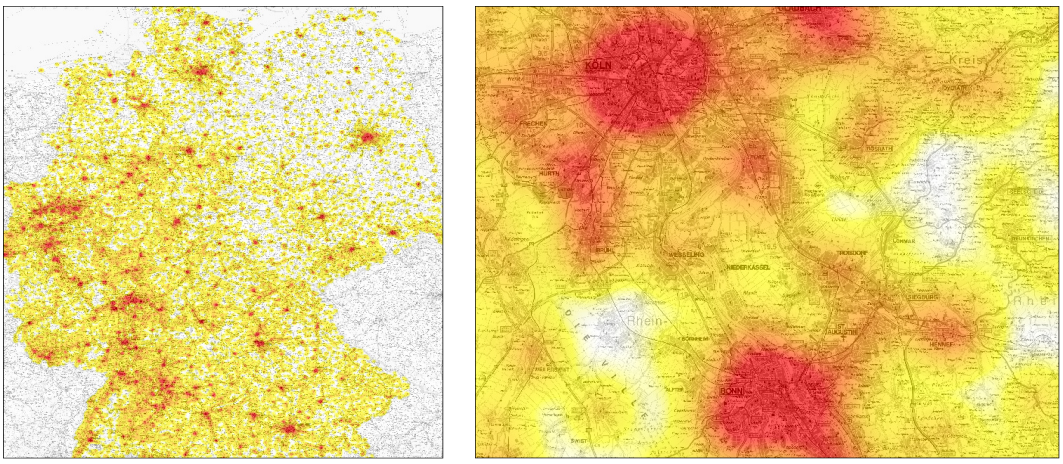
$$\hat{z}(s_o) = \frac{\sum_{s_i \in D} w(d(s_o, s_i)) \cdot z(s_i)}{\sum_{s_i \in D} w(d(s_o, s_i))} \quad \textbf{(2)}$$

Neben der hier dargestellten einfachen monoton fallenden Gewichtung, weist die *'Tricube Funktion'* gute Eigenschaften der Gewichtung auf (DRESSMANN 2004, S. 74):

$$w(d(s_o, s_i)) = (1 - w(d(s_o, s_i)/r)^3)^3 \text{ falls } d(s_o, s_i) < \text{r}$$
$$w(d(s_o, s_i)) = 0 \text{ falls } d(s_o, s_i) >= \text{r}$$

Die Gewichtung wird hier für den Wettbewerbereinfluss berechnet.

[7] Hierzu siehe auch: Lokale Analysen - die Einordnung der Methoden in *'Geographically Weighted Regressions'* (GWR) (FORTHERINGHAM et al 2000, S. 107-114; vgl. FORTHERINGHAM E et al 2002).



Quelle: eigene Darstellung (GBI-Tool)

Abb. 8-12: Wettbewerberdichte Deutschland und Ausschnitt Köln/Bonn mit Gewichtung nach Tricube-Funktion

8.3 Point Pattern Analysis, Punktedichteschätzungen

Generell sind die *Point Pattern Analysis* (PPA) Techniken wichtige weiterführendere Analysemethoden (siehe dazu BOOTS & GETIS 1988, ISSACS & SRIVASTAVA 1998; LEITNER 2001; LORUP 2003; STALLLEIKEN 2005, S. 22ff). Unter PPA werden alle Techniken zusammengefasst, welche die räumliche Verteilung von Ereignissen, die als Punkte dargestellt sind, untersuchen. Das Ziel für die Analyse räumlicher Punktmuster ist der Versuch ein besseres Verständnis für den Prozeß zu entwickeln, welcher diese Punkteverteilung hervorgebracht hat. Aus punktförmigen Daten, z. B. aus umsatzrelevanten Informationen zu den einzelnen Filialen (Umsatz, Vertriebserfolg etc.) kann anhand vorgegebener Kriterien ein System von Punktclustern (PPA) oder Regionen (*Thiessen, Voronoi*) abgeleitet werden. So können z. B. Einzugsgebiete und Werbeverteiler definiert werden, um die richtigen Standorte zu finden, in denen sich Punkte bestimmter Ausprägung häufen (hier z. B. gehobener Einzelhandel, 1-2 Personen-Haushalte etc.). Gleichzeitig muss aber möglicher *Kannibalismus*[8] zu vermeiden versucht werden. Bezogen auf Kundendaten ist PPA eine sehr weit verbreitete Methode, die die Kundendaten als Punkte darstellt (*customer spotting)* und damit Marktgebietsdichte bestimmt (APPLEBAUM 1961), davon abgeleitet können Stern-

[8] Entzug von Marktanteilen durch unternehmenseigene Filialen

und Rubinigraphdarstellungen[9] betrachtet werden, die die Dichte, Kundenwolken und die Beziehungen von Kunden zu Standorten visuell unterstreichen.

Punktprozessdaten werden anhand von *Point Pattern Analysis* (PPA) Verfahren behandelt (vgl. FOTHERINGHAM U.A. 2000, S. 130-161; SULLIVAN & UNWIN 2003, S. 77-133) mit dem Ziel ein Verständnis für den Prozess der Punktverteilung zu erlangen. Die Ursprünge der Techniken, die heute in der statistischen Analyse von Punktmustern eingesetzt werden, sind in den 50er Jahren bereits in der Pflanzenökologie entwickelt und in der Archäologie und Anthropologie (HODDER & ORTON 1976) sowie Astronomie und Werkstoffkunde weiterentwickelt worden. RIPLEY (1981) und DIGGLE (1983, 1993) haben seit den 80er Jahren komplexere PPA-Techniken entwickelt[10].

Wird nur eine Variable untersucht, kann das beobachtete Muster anhand eines übergreifenden räumlichen Trend der Punktdaten (*Kernel density estimation*) oder anhand der Nachbarschaftsanalyse erklärt werden. Bei der Nachbarschaftsanalyse handelt es sich um eine Methode, räumliche Muster lokaler Punktdaten anhand der graphisch gemessenen Distanzen, der jeweils nächsten Nachbarn verglichen mit den theoretisch erwarteten Distanzen, unter Annahme einer räumlichen Zufallsverteilung zu untersuchen. Die räumliche Punktverteilung wird damit definiert. Die Punktedichte-schätzung dient zur Feststellung von Konzentrationen oder Risikobereichen. Generell wird zwischen einfacher Dichteschätzung und Kernel-Dichteschätzung unterschieden. In der einfachen Dichteschätzungen werden Punktdaten, die innerhalb eines vorgegebenen Suchradius liegen, aufsummiert und durch die Fläche des Kreises dividiert (vgl. auch FOTHERINGHAM U.A. 2000, S. 146-161).

---

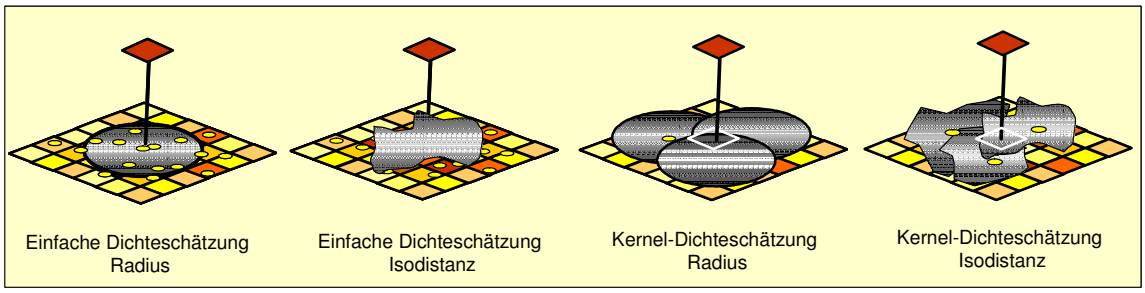
[9] Der Rubini-Graph verbindet die jeweils nächsten merkmalsgleichen Objekte miteinander. So können Adressen von Kunden, die von einem Außendienstmitarbeiter betreut werden, miteinander verbunden werden. Diese Art der Darstellung ist hilfreich, wenn die Kompaktheit der Gebiete beurteilt werden soll. Die Sterndarstellung dient dazu Überlappungsbereiche bzgl. der Kundenversorgung sofort ersichtlich zu machen, dabei werden Standorte von Außendienstmitarbeitern, Service-Centern oder Filialen mit den Standorten der jeweiligen Kunden verbunden.

[10] Weitere Arbeiten zu PPA siehe: GATRELL und BAILEY (1996), GATRELL u.a. (1996), BHOPAL u.a. (1992) sowie DIGGLE und ROWLINGSON (1994).

$$\lambda_p = \frac{N[S \in C(p,r)]}{\pi r^2} \quad \textbf{(3)}$$

N = Anzahl der Punktdaten
S = Ereignis
p = Untersuchungsgebiet
r = Radius
C(p,r) = Kreis des Radius r um einen Punkt x

Eine Verfeinerung lässt sich anwenden, in dem statt der Summe aus den Radien, die Summe der Punktdaten aus definierten Flächen, die durch Isodistanzen[11] oder Isochronen[12] aufgespannt werden, gebildet werden und damit eine genauere Aussage hinsichtlich Erreichbarkeit zugelassen wird.



Quelle: eigene Darstellung erweitert nach McCoy 2002
Abb. 8-13: Verfahren der Kernel-Dichteschätzung

Das Kernel-Dichteverfahren gewichtet die Punkte nahe dem Zentrum des Suchradius einer Rasterzelle höher als weiter entfernt liegende Punkte. Für jeden Punkt wird eine symmetrische Kernel-Funktion berechnet, die eine Normalverteilung darstellt. Die zugrunde liegende Dichtefunktion wird durch das Summieren der individuellen Kernelfunktionen berechnet, um eine kumulative Dichtefunktion zu erhalten.

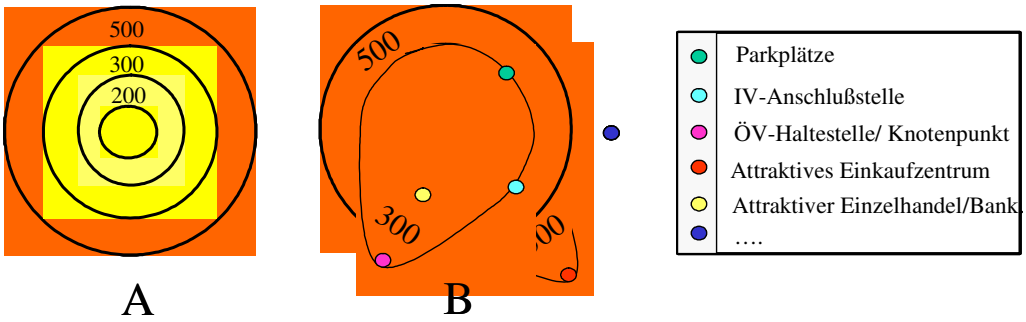
Das Ergebnis ist eine interpolierte geglättete Rasteroberfläche (vgl. BAILEY 1998, MITCHELL 1999). *"Here the smoothed value at any point is essentially estimated by a weighted average of all other values, with the weights arising from a probability distribution centered at that point and referred to as the kernel"* (BAILEY 1998, S. 27).

Es gibt eine Vielzahl von Kernel-Funktionen, die von der Funktion der Normalverteilung (KELSALL & DIGGLE 1995) abweichen: Dreiecksfunktion (*triangular function*) (BURT und BARBER 1996), biquadratische und quadratische Funktion (*biquartic function/ quartic function)* (BAILEY & GATRELL 1995). Bei der angepassten

[11] anhand der Entfernung entlang des Straßennetzes
[12] anhand der zeitlichen Erreichbarkeit entlang des Straßen- oder Fußgängernetzes

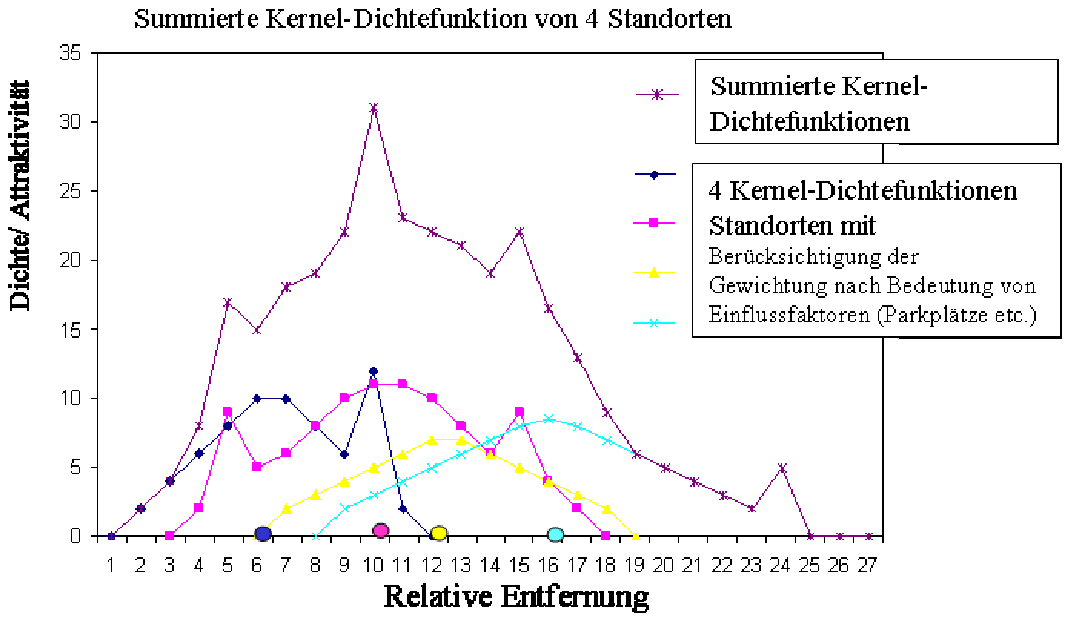
Kernel-Dichteschätzung variiert man den Suchradius entsprechend dem Untersuchungsgebiet, um Effekte weiterer Variablen zu erklären.

Die Erweiterung um die Isodistanz/Isochronen Kernel-Funktion, bei der sich das Einzugsgebiet anhand des Straßennetzes und damit an der realen Entfernung orientiert, ist in der Abbildung *(Abb. 8-13)* dargestellt. Die Hinzunahme eines Attraktivitätsfaktors ausgerichtet nach bestimmten Anziehungspunkten oder möglichweise auch Erfolgsfaktoren, läßt sich folgendermaßen skizzieren (*Abb. 8-14*):



A Standort - Bemessung von Einflussfaktoren anhand Luftline
B Standort - Bemessung von Einflussfaktoren anhand Attraktivitätsfaktor unter theoretischer Berücksichtigung der Isodistanz
Quelle: eigene Darstellung

Abb. 8-14: Bemessung Einflussfaktoren anhand Luftline und nach Gewichtung der Attraktivität



Quelle: eigene Darstellung (ausgehend von *Abb. 8-14 B* bei Berücksichtigung von Gewichtung von Einflußfaktoren)

Abb. 8-15: Summierte Kernelfunktion von 4 Standorten

Die absolute Entfernung gemessen anhand des Radius oder der Isodistanz wird ersetzt durch die 'relative' Entfernung oder einem Index, bei der ein Punkt, der weiter von einem Ausgangspunkt, einer Filiale wegliegt, aber eine hohe Attraktivität besitzt, eine geringere Entfernung zugeteilt wird und damit das gesamte Feld um die Filiale eine höhere Dichte erhält. Die absolute Entfernung anhand der Isodistanz oder Isochrone wird bei der Indexbildung berücksichtigt.

In dem für diese Untersuchung entwickelten Geomarketing-Tool (*GBI-Tool*) ist mit den unterschiedlichen *Kernel-Density-Funktionen* versehen, so dass Vergleichsanalysen getätigt und visualisiert werden können. Zur Anwendung kommt diese Methodik bei Wettbewerberanalysen. Hier werden sowohl die Dichte der unterschiedlichen Wettbewerber als auch die Dichte der eigenen Standorte untersucht, um somit mögliche Kannibalisierungseffekte zu berechnen.

Deutlich wird aus oben genannten Ergebnissen, dass eine reine Distanzfunktion nicht genügend Aussagekraft für die Beantwortung von Fragestellungen bezüglich des Standortes oder des Marktgebietes besitzen. Entscheidend ist die Attraktivität. Befragungen von Geschäftskunden (DP AG 2005) zur Attraktivität der Postfiliale bestätigen z. B. das Parkplätze und Öffnungszeiten wichtige Kriterien für den Besuch des Kunden sind. Diese Attraktivitätsfaktoren werden bei *Spatial Interaction Models* und *Gravitationsmodellen* berücksichtigt. Herleitung, Bedeutung und Anpassungsmöglichkeiten der Modelle werden nachfolgend erläutert.

8.4 Spatial Interaction Models und Gravitationsmodelle

Im nordamerikanischen Raum beschäftigt sich erstmals REILLY (1931) mit Anwendungen gravitationstheoretischer Ansätze[13] für den Einzelhandel. Erst viel später finden diese Methoden Eingang im deutschsprachigen Raum[14] (z. B. KOTSCHEDORFF 1976, KLEIN & LÖFFLER 1989, GÜßEFELDT 2003). Seit den 90er Jahren sind weitere

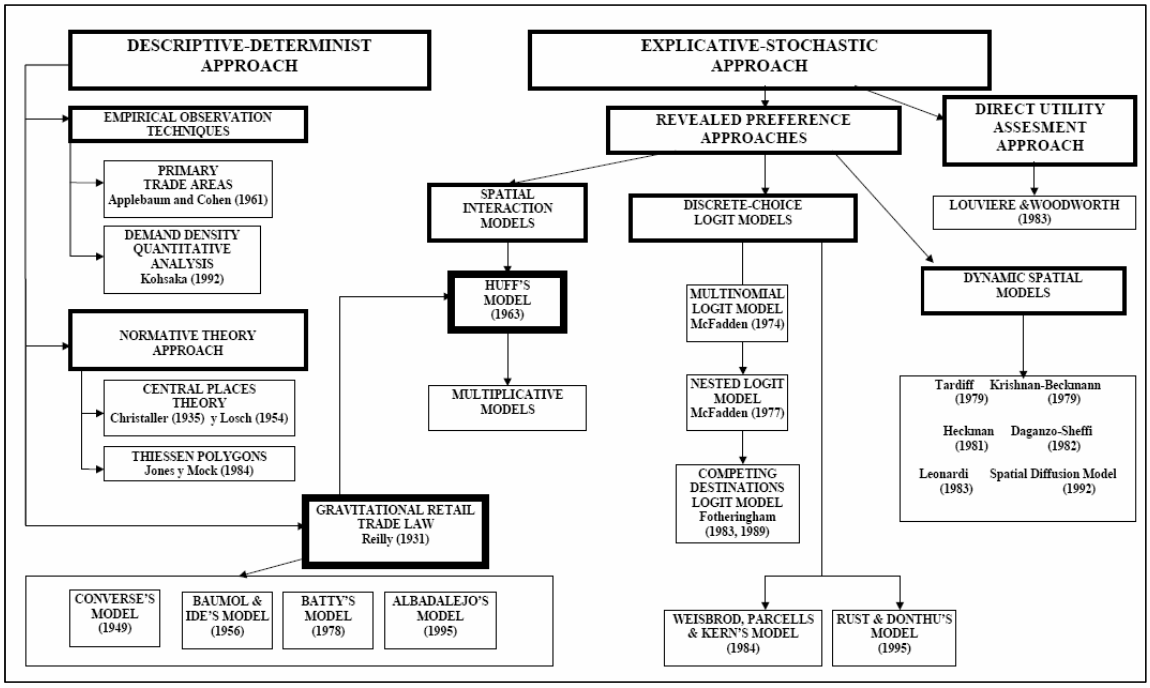
[13] Mit Hilfe der Modelle werden sozialwirtschaftliche Phänomene in Analogie zum Gravitationsgesetz der Physik von NEWTON (1687) beschrieben. Erstmals belegt ist dies durch die Analogie, die CAREY (1858) vorgenommen hat, das sich die Anzahl der Interaktionen zwischen den Bewohnern zweier Städte proportional zu der Anzahl der Bewohner in den Städten und umgekehrt proportional zur Entfernung zwischen den Städten verhält.

[14] Was auch dazu führt, dass Softwareprodukte zur Berechnung dieser zu 'nordamerika-spezifisch'. Es müssen spezielle Anpassungen erfolgen, um Daten wie sie in Deutschland vorzufinden sind, integrieren zu können. Das wiederum bedeutet Individualentwicklungen und hat einen hohen Preis. Im Zuge der Entwicklung von Geomarketing bei der Post ist eine solche Entwicklung eines GBI/ Geomarketing-Tools mit Berücksichtigung von speziellen Methoden ermöglicht worden.

Modelle wie *Point Pattern Analysen*, *Thiessen-Polygone* oder auch Grid-Bildung verstärkt für den Bereich der Vertriebs- und Kundenanalyse *Trade Area Analyses* (JONES; SIMMONS 1993, S. 329-349) eingesetzt worden.

Die Anpassung und Anwendung der Modelle auf Beispiele in städtischen Gebieten Deutschlands soll die Übertragbarkeit deutlich machen und die Vorteile dieser räumlichen Methoden gegenüber einfacheren GIS-Standardverfahren darstellen.

Marktgebiete *(trade area)* sind im ökonomischen Sinne genauso Regionen wie auch die rein territorialen Einteilung in Gemeinden und Postleitzahlen. Sie stellen keine amtlich gültige Einteilung dar, sondern eine Gebietsordnung, die hinsichtlich spezieller Fragestellungen optimiert ist.



Quelle: Chasco Yrigoyen & Otero 1998, o.S.

Abb. 8-16: 'Spatial' Modelle, Methoden zur Berechnung von Einzugsgebieten für Einzelhandel

1931 war REILLY der erste, der das Problem der Marktgebietseinteilung diskutierte. Basierend auf dem NEWTON´SCHEN *Gravitationsmodell* hat er die 'Gravitations-Idee' auch im räumliche Modelle übertragen *('Reilly´s Law of Retail Gravitation')*: statt der physikalisch magnetischen Anziehungskraft wird die 'psychologische' bzw. 'verhaltensorientierte Anziehungskraft' in den Mittelpunkt gerückt. Gravitation wird durch Attraktivität ersetzt. Das Attraktivitätsmaß wird hier durch die alternativen Einkaufsmöglichkeiten der unmittelbaren Umgebung abgebildet. Für die verhaltensorientierte ausgelegte Funktion wird hier zur Unterscheidung von den

deterministischen Modellen (CHRISTALLER 1933, LÖSCHE 1954, APPLEBAUM 1961, THIESSEN POLYGONE – JONES & MOCK 1984 etc.) der Begriff *'behavior gravitation function'* eingeführt.

$$\frac{R_A}{R_B} = \left(\frac{P_A}{P_B}\right)^{a} * \left(\frac{D_B}{D_A}\right)^{\lambda}$$

R_A, R_B: Kaufkraft/ Umsatz des Einzelhandels eines Ortes/ Filialclusters A und B mit einem zwischen ihnen liegenden Standort
P_A, P_B: Attraktivität, *hier:* Bevölkerung von Ort A und Ort B
D_A, D_B: Entfernung des 'zwischenliegenden' Ortes/ Filialclusters zum Ort A und Ort B
α Gewichtungsfaktor für die Attraktivität
λ Gewichtungsfaktor für die Entfernung

Quelle: Location Strategies for Retail and Service Firms (Gosh u.a. 1987) nach CHASCO YRIGOYEN & OTERO (1998) o.S.

Abb. 8-17: Reilly´s Law of Retail Gravitation

Die Attraktivität wird von REILLY als Einwohnerzahl des Ortes definiert. Der Gewichtungsfaktor für die Attraktivität α hat den Wert 1 und für die Entfernung λ schlägt er den Gewichtungsfaktor 2 vor (KOTSCHEDOFF 1976, S. 87). Mit Hilfe des Modells läßt sich die Kaufkraft eines Nachfrageortes auf zwei Angebotsstandorte aufteilen. Die Bestimmung der Marktgrenze ermöglicht erst die *'breaking-point-formula'* von Converse (1949). Der *'breaking-point'* ist der Punkt, der zwischen den beiden Zentren ausfindig zu machen, ab dem der Kunde eher zur Filiale A oder zu B tendiert.

$$D_A = \frac{D_{AB}}{1 + \sqrt{\frac{P(B)}{P(A)}}}$$

where: D_A: Distance from city A to breaking point
D_{AB}: Distance between cities A and B
P (A): Population of city A
P (B): Population of city B.

Quelle: Location Strategies for Retail and Service Firms
(CHASCO YRIGOYEN & OTERO 1998 nach GHOSH ET AL 1987, S.68)

Abb. 8-18: Reilly (1931)-Converse' s (1949) 'Breaking- Point Formula'

Der in Abbildung (*Abb. 8-18*) dargestellte '*breaking point'* ist dort, wo die relative Attraktivität eines Ortes, einer Filiale gleich ist, unabhängig von der tatsächlich

gemessenen Entfernung, die zurückzulegen ist, um die jeweilige Filiale zu erreichen. Der *'breaking point'* ist der Grenzpunkt zwischen den Einzugsbereichen zweier Angebotsstandorte. An diesem Punkt sind die Nachfrager gegenüber den beiden Standorten indifferent. Errechnet man mit Hilfe der Formel von CONVERSE (1949) die *'breaking points'* zwischen einem Angebotsstandort A und allen umliegenden Wettbewerberzentren und verbindet diese miteinander, lässt sich das ungefähre Absatzgebiet dieses Angebotsstandortes A ermitteln (HUFF 1964, S. 35).

Es läßt sich allerdings nur die Grenzen zwischen jeweils zwei Angebotsstandorten ermitteln, wendet man dieses Verfahren für mehrere, benachbarte Angebotsstandorte an so kann dies zu Überschneidungen der Einzugsbereiche und somit zu über- und unterversorgten Gebieten führen. Diese Gebiete wiederum müssten separat betrachtet und im iterativen Schritten entsprechend aufgeteilt werden. Ein weiterer zu variierender Aspekt ist der Distanzexponent, der bei REILLY den Durchschnittswert 2 annimmt. Ebenso der Exponent 1 für die Gewichtung der Attraktivität. Diese beiden Parameter sind für jede Untersuchung neu zu ermitteln. Sie haben keine Allgemeingültigkeit.

Der Idee des verhaltensorientierten Ansatz des *Gravitationsmodells* sind viele Geographen gefolgt und haben damit einen wichtigen Weg für die Wirtschaftsgeographie und für Geomarketing eingeleitet: HUFF (1963), NAKANISHI & COOPER (1974) *'Multiplicative Competive Interaction Model'* MCI, FORTHERINGHAM & O´KELLY (1989) '*Competing Destinations Model'* CDM, WEISBROD, PARCELLS & KERN´S Model (1984), JONES & MOCK (1994) mit der Einführung der *Thiessen-Polygone*, RUST & DONTHU´S Model (1995) u.a.

8.5 Gravitationsmodell nach Huff

Vielfach zitiert wird das HUFF Modell (1963), das auf REILLY´S *Law* basiert. HUFF definiert die Wahrscheinlichkeit des Besuchs einer Filiale als die, dass mehrere Filialen in dem Einzugsgebiet die Chance haben, besucht zu werden. Das Attraktivitätsmaß ist dabei die Größe der Verkaufsfläche der Filiale, bei REILLY ist es die Einwohnerzahl. Distanz und Größe der Filiale sind die entscheidenden Parameter bei HUFF.

Das HUFF-Modell ist ein Spezialfall allgemeiner Interaktionsmodelle und bezieht sich auf die Gravitationstheorie ISAAC NEWTONS nach der sich zwei Objekte um so stärker anziehen, je näher sie beisammen liegen und je größer deren Masse ist. Das HUFF-Modell und die HUFF-Gleichung ist wiederum ein Spezialfall dieses *Gravitationsmodells* und damit der NEWTONS'schen Gravitationsgleichung für die

Vorhersage von Kundenverhalten, Kundenströme im Einzelhandel. Die HUFF-Gleichung wird herangezogen, um die Wahrscheinlichkeit abzuschätzen, nach der ein potentieller Kunde eine oder mehrere Filiale aufsucht. Es gelten folgende Annahmen: Die Interaktionen zwischen Standort *i* und *j* nehmen mit zunehmender Entfernung (D_{ij}) ab und die Interaktionen zwischen Standort *i* und *j* nehmen mit wachsenden Standortmassen von Quellort *i* bzw. Zielort *j* zu; d. h. die Wahrscheinlichkeit, dass ein Kunde die Filiale aufsucht, steigt mit der Attraktivität (vgl. HUFF 1963)[15].

$$P_{ij} = \frac{U_{ij}}{\sum_{j=1}^{n} U_{ij}} = \frac{A_j^{\alpha} D_{ij}^{-\lambda}}{\sum_{j=1}^{n} A_j^{\alpha} D_{ij}^{-\lambda}} \quad \textbf{(4)}$$

P_{ij}	Wahrscheinlichkeit, daß ein Kunde am Ort i eine Filiale j (oder Ort j) aufsucht
Uij	(Utility) Nützlichkeit/ Mehrwert einer Filiale j (oder Ortes j) für den Kunden am Ort i
A_j	Attraktivitätsmaß, bei Huff Flächengröße (Size); *hier:* weitere Faktoren und Kombinationen
D_{ij}	Distanz zwischen dem Kunden am Ort i und der Filiale am Ort j
α	Attraktivitäts-Parameter geschätzt auf Basis empirischer Beobachtungen (Reilly´s Law α=1)
λ	Distanz Decay[16] Parameter geschätzt auf Basis empirischer Beobachtungen (Reilly´s Law λ = 2)
n	Anzahl der Standorte inklusive Standort j

HUFF (1963) befasste sich als Erster mit der Dynamik des Wettbewerbs im Einzelhandel und wie Kunden zwischen einem alternativen Angebot an Einkaufsmöglichkeiten wählen. Der Kunde wählt in erster Linie nach der gesamten Nützlichkeit der Filiale für seine Bedürfnisse und weniger nach dem Ort, wo diese liegt. Dies bedeuted, dass die Kombination von Produkten in größeren Geschäften weit aus attraktiver sind, als die in kleinen Filialen, auch wenn diese günstig gelegen sind. Es ist anzunehmen, dass wenn Filialen mit anderen Filialen in der unmittelbaren Umgebung liegen, der Kunde die Besuche kombiniert und mehrere Filialen aufsucht. Jede Filiale, innerhalb eines bestimmten Distanzraumes hat eine Chance besucht zu werden, in Abhängigkeit der Größe und der genauen Distanz.

Der entscheidende Unterschied der Modelle besteht darin, dass REILLY und CONVERSE den Einfluss zweier Absatzzentren auf das räumliche Konsumverhalten berücksichtigen und den Punkt bestimmen, an dem die Einwohner indifferent gegenüber den zwei Zentren sind. HUFF definiert hingegen den Einfluss alle konkurrierenden Zentren. Die Attraktivität des jeweils untersuchten Zentrums wird in Beziehung zur Summe der

---

[15] "*Die Interaktionswahrscheinlichkeit P_{ij} zwischen zwei beliebigen Orten i und j ergibt sich als Quotient der mit der inversen Distanz zwischen i und j gewichteten Attraktivität A_i des Zielortes i und der Summe der distanziell gewichteten Attraktivitäten A_j aller Orte*" (GÜßEFELDT 2002, S. 354).

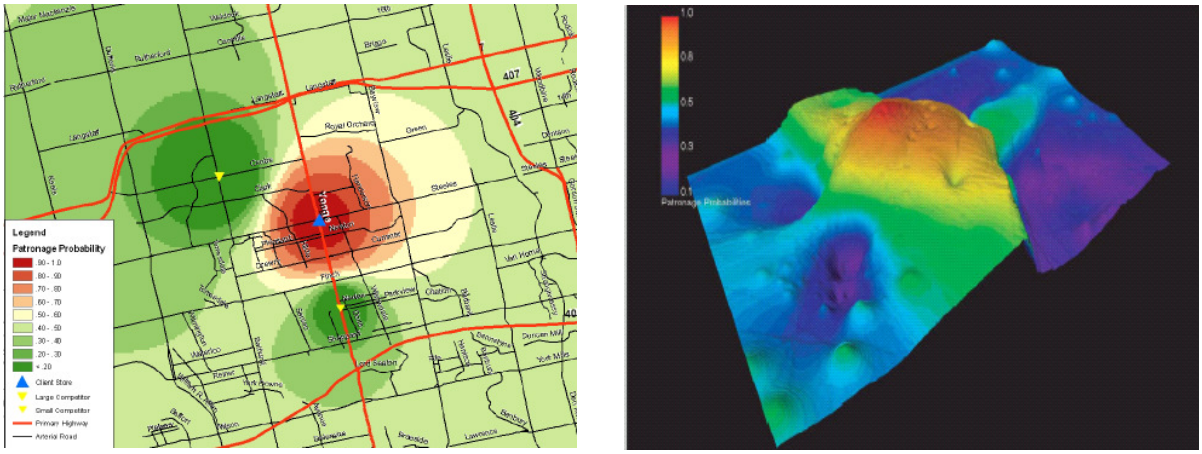
[16] Der *Distance Decay Parameter* gibt die Art der Steigung/ 'Abklingen' der Kurve bei zunehmender Entfernung an.

Attraktivität aller umgebenden Zentren gesetzt. Damit ist eine Abgrenzung der Einzugsgebiete sowohl für die eigenen Standorte als auch gegenüber der Wettbewerberstandorte möglich. Jedem Ort (hier jedem Grid) wird ein Wert zugewiesen mit welcher Wahrscheinlichkeit unter Berücksichtigung der Attraktivität der anderen Zentren der Besucher zu einem bestimmten Standort tendiert. Dies lässt eine Abschätzung des Absatzpotenzials einzelner Zentren zu.

Weiterhin hervorzuheben ist, dass HUFF den Parameter λ flexible einsetzt im Gegesatz zu REILLY, der von dem konstanten Durchschnittswert 2 ausgeht. Bei HUFF schwankt λ in Abhängigkeit der Nachfragefristigkeit von Gütern. In der Praxis ist die Bestimmung des Parameterwertes allerdings eine der größten Schwächen des Modells.

Die Größe der Filiale als ein Attraktivitätsmaß zu gewichten, ist aufgrund der Änderung der Filiallandschaft insbesondere bei dem Einzelhandel den mit Größe zusammenhängenden Erfolg von Filialen ein in den 70er bis 90er Jahren (SCHÜSSLER 2000, S. 87ff) gutes Indiz für Attraktivität. Allerdings im Zeitalter der schnellebigen Mobilitäts-, Kommunikations- und Konsumveränderungen ist der Faktor Größe einer Filiale alleine als Attraktivitätsmaß in keinem Falle ausreichend für die Erklärung von Einflüssen.

Die Bedeutung von HUFF ist insbesondere in den USA und Canada im Bereich von Geomarketing groß. Jüngere Literatur macht deutlich, dass dieses Gebiet weiter entwickelt wird. Das HUFF Modell wird weiterhin für komplexe Analysen im Bereich Vertriebsgebietanalyse und Standortoptimierung herangezogen. So beschreibt DRAMOWICZ (2005) die Anwendung von HUFF unter verschiedenen Parametern. Das Modell wird variiert mit der Anzahl der Parameter, die die Attraktivität einer Filiale ausmachen können, wie Parkplätze, Anzahl von Einzelhändler in der Umgebung. Weiterhin wird die Anzahl der Filialen variiert, die sich in der Einflußzone befinden, es werden Kundendaten zu Ausgabeverhalten für verschiedene Produkte hinzugezogen und Angaben von Kundenbefragungen zur Kalibrierung des Modells berücksichtigt (siehe auch HUFF 2003, 2005). Insofern werden hier auch Aspekte der Modelle von MCFADDEN 1974 und FORTHERINGHAM 1989 (*siehe Kapitel 8.6*) verwendet, ohne dass auf diese Bezug genommen wird.



Quelle: Lea (2005), S. 28f CSCA, Ryerson Universität Toronto

Abb. 8-19: Huff Gravitationsmodell als Interpolations- und als 3D Darstellung

Das HUFF-Modell wird in dem Kapitel "*Trade Area Analysis*" PICK (2005) ausführlich diskutiert und als den anderen Methoden überlegende für das Anwendungsfeld dargestellt. Ebenfalls wird es in der Stadt- und Regionalforschung weitergehend behandelt[17].

Zur Anwendung des Modells werden Daten zum Umsatz bzw. Kundenaufkommen in den Zielorten, Attraktivitätsfaktoren, Nachfragevolumen der Quellorte[18] und Distanzen in Form von Luftlinien, Straßenkilometer, Fahr- oder Laufzeiten benötigt.

Es zeigt auf, dass der Einsatz des HUFF-Modells ermöglicht, verschiedene Aspekte mit unterschiedlicher Gewichtung für die Untersuchung von Standorte und Kundenströmen hinzuzuziehen in Abhängigkeit der Datenlage, mit der man das Modell 'füttern' kann. Eine Variante des HUFF-Modells ist in dem *GBI-Tool* implementiert und im Folgenden angewendet worden[19].

Zum Vorgehen für eine GIS-technische Umsetzung wird hier zunächst ein regelmäßiges Raster erstellt, wobei jede Rasterzelle (Grid) einen potentiellen Kunden darstellt. Nun

---

[17] z. B. an der Technischen Universität in Wien (GIFFINGER, KRAMER 2001) und der Universität Köln in Zusammenhang mit der Untersuchung des Einflusses der Verkehrslenkung (MÜLLER-HAGEDORN u.a.1997).

[18] Die Daten sollten so feinräumig wie möglich vorliegen. Eine Auswertung auf Gemeindeebene erbringt keinen Mehrwert, sondern läßt viele Gebiete unberücksichtigt und verfälscht somit das Ergebnis.

[19] Implementierung des Modell finden sich sowohl bei *Mapinfo Vertical Mapper* als auch bei dem *Business Analyst* von *ESRI*. Auf die Möglichkeit der Modellkalibrierung, wie sie DRAMOWICZ (2005) beschreibt, ist verzichtet worden, da sie in der Praxis sehr komplex, und es kostenintensiv ist, die notwendigen Daten für die Kalibrierung zu erheben (siehe auch HUFF 2004, S. 2).

wird mittels der HUFF-Gleichung für jede Zelle ermittelt, mit welcher Wahrscheinlichkeit ein Kunde an dieser Position eine Filiale aufsucht. Dabei hängt die Wahrscheinlichkeit des Kundenbesuchs zum einen von der Attraktivität der untersuchten Filiale ab und zum anderen von der der umliegenden Filialen. Diese Attraktivität wird den Filialen in Form von Bewertungen von Variablen zugewiesen, z. B. Größe der Verkaufsfläche, Ausstattungsqualität, Parkplätze im Umfeld, attraktiver Einzelhandel (hochwertiger) oder Bankstandorte im Umfeld. Die Wahl eines oder mehrerer kombinierter Attraktivitätsvariablen erfolgt aus der Berechnung der Erfolgsfaktoren, wobei eine Unterscheidung in interne Unternehmensfaktoren (Öffnungzeiten, Fläche etc.) für die Angebotsseite und externe Daten (sozioökonomische Daten, die das Umfeld des Standortes bestimmen) für die Nachfrageseite zu unterscheiden sind. Den internen Faktoren kann das Unternehmen direkt beeinflussen, die externen nur gegebenenfalls auf die Veränderung dieser wirken. Es kann daraus ein Gesamtindex für die Attraktivität aufgestellt werden[20].

Die auf diese Weise mit Werten gefüllte Rasterfläche kann als idealisierte Darstellung der Angebotssituation interpretiert werden. Für die Integration der Nachfragesituation in das Modell wird anhand von punktbezogenen Kundendaten z. B. in Form geocodierten Kundenadressen, Einwohnern, Beschäftigte oder Verkehrsfrequenzen eine oder mehrere weitere Rasterebene erstellt werden. Es läßt sich demnach für eine Filiale zunächst das Gebiet ermitteln, in welchem die Kaufwahrscheinlichkeit bei x % liegt. Anschließend kann diese Zone entweder vektorisiert und die darin befindlichen Kundendaten können in einem vektoriellen GIS ausgewertet werden, oder die Rasterfläche selbst dient als Ausgangspunkt für weitere Analysen.

Mit der Anwendung des HUFF-Modells in Grid-Methodik und der Kombination mit weiteren Attraktivitätsfaktoren als GIS-Informationslayern ist hier Angebots- und Nachfrageseite berücksichtigt.

Die ‘Wirkung’ der Attraktivitität steht wiederum umnmittelbar mit der Distanz in Zusammenhang. Von daher sind die weiterführenden Modelle zu betrachten.

---

[20] unter der Anwendung von CHAID- und Clusteranalysen und Integration in eine LDSC- siehe hierzu auch *Kapitel 9*

8.6 Competitive Destination Model und Sättigungsnindex

Mit zunehmender Zeit spielen wesentlich mehr Parameter eine Rolle für Attraktivität als bei dem HUFF-Modell ursprünglich vorgesehen ist. Dazu zählen insbesondere Parkplätze im Umfeld der Filiale, Öffnungszeiten, attraktive Angebote etc. (siehe auch Befragung DP AG 2005) als interne und z. B. hochwertige Einzelhändler als externe Daten. Von daher liefern die bereits seit 1974 aufgestellten Modelle berechtigte Vorschläge weitere Variablen zu integrieren; NAKANISHI & COOPER (1974) MCI *Multiplicative Competitive Interaction Model*, *Gautschi´s Model* (GAUTSCHI 1981) die z. B. weitere Distanzfaktoren als Modell-Parameter berücksichtigen.

Das *'Competing Destinations Model'* (CDM) (FORTHERINGHAM 1983), dass als Weiterentwicklung der Modells von MCFADDEN 1974, ('*Multinomial Logit Model'*) bezeichnet wird, unterscheidet sich dadurch, dass es stärker an der räumlichen Ebene orientiert ist und mit der Einführung des Zentralitätsmasses C_j dieses ergänzt.

$$P_{ij} = \frac{C_j^{\theta} \exp(V_{ij})}{C_j^{\theta} \sum_{j=1}^{J} \exp(V_{ij})} \quad \textbf{(5)}$$

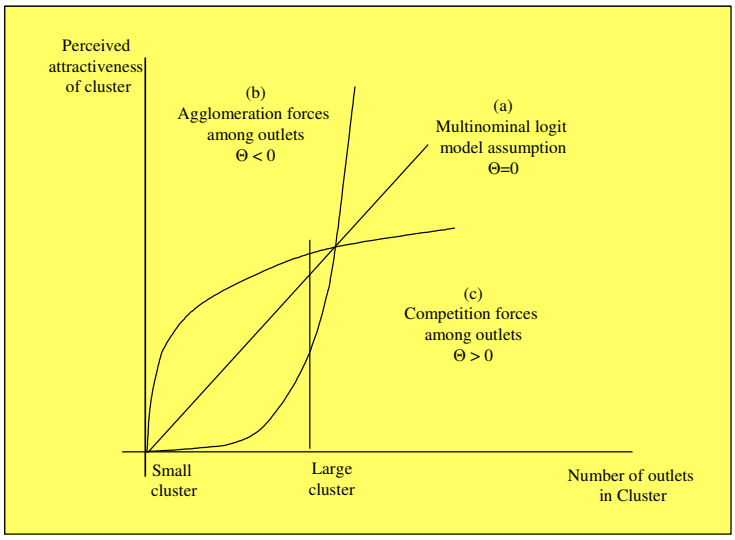
θ Attraktivitäts-/ Sensibilitäts-Parameter

C_j Zentralitätsmaß. In Anlehnung an BORGERS & TIMMERMANS (1987): $C_j = \frac{\sum_{j' \neq j} d_{j'j}}{J-1}$

Vij repräsentiert die Attraktivität, die den Besuch des Kunden der Filiale j am Ort i ausmacht.

Dem Modell '*Competing Destinations Model*' (CDM) liegt die Annahme zugrunde, dass jedes Individuum nur begrenzt eine Vielfalt von Informationen aufnehmen kann. Von daher ist auch die räumliche Wahl ein Ergebnis eines hierarchischen Informationsprozess bei dem zuerst ein Cluster von Alternativen gewählt wird.

Wenn θ > 0: Die Filialen weiten ihren Marktanteil aus, dadurch das sie isoliert von den Wettbewerbern betrachtet werden *('competition forces')*.
Wenn θ < 0: Die Filialen weiten ihren Marktanteil aus, dadurch das sie in unmittelbarer Nähe zu anderen Einzelhändlern/ Banken platziert sind *('agglomeration forces')*.
Wenn θ = 0: Alle Alternativen werden gleichwertig gewichtet.



Quelle: Fortheringham & Kelly (1989)

Abb. 8-20: Zusammenhang zwischen der angenommenen Attraktivität eines Clusters von Einzelhändlerm und der Anzahl der Filialen im Cluster

Die Abbildung *(Abb. 8-20)* zeigt auf, dass bis zu einem bestimmten Punkt ein Cluster von Attraktivitätsfaktoren (hier bestimmte berechnete Erfolgsfaktoren) eine Gesamtattraktivität für den Besuch einer Filiale steigert *('agglomeration forces')*. Übertragen auf die ermittelten Erfolgsfaktoren heisst dass, dass der Marktanteil ausgeweitet werden kann, wenn bestimmter Einzelhandel oder z. B. Bankstandorte in der unmittelbaren Umgebung liegen und damit auch die Attraktivität des eigenen Standortes gesteigert ist. Das optimale Verhältnis zu Erreichung eines Marktanteils und der Anzahl der Filialen ist begrenzt, da hier der *'breaking point'* einsetzt. Bei Vorhandensein anderen Einzelhändler, haben diese ebenfalls die Chance besucht zu werden und möglicherweise werden die Produkte, die von der Post angeboten werden, dort gekauft. Die Höhe der Wahrscheinlichkeit steht in Abhängigkeit der Anzahl der gesamten Einzelhandelsläden des Clusters.

In dem Falle, dass es keinen Wettbewerber in der Umgebung gibt, weitet sich der Martkanteil in Abhängigkeit der Expansion der Filialen aus. Genau diesem Prinzip ist *Hermes* gefolgt. *Hermes* zeigt einerseits dort Präsenz, wo sich die Post zurückgezogen hat, andererseits aber platziert sich *Hermes* gleichzeitig direkt in der Nähe von Post-Clustern (*siehe Kapitel 7.2.4*). Der Marktanteil von *Hermes* ist mit dieser Wettbewerbsstrategie zu ungunsten der Post gestiegen.

8.7 Modifizierte Potenzialmodelle von Klein, Löffler und Güßefeldt

Um eine bessere Abbildungsqualität des HUFF Modells zu erreichen, entwickelten KLEIN & LÖFFLER (1988, 1989) und KLEIN (1988, S. 29) das Zentrums- und Hierarchiefilter, die dazu dienen sollen, die Anzahl der Interaktionen auf die aus Ihrer Sicht relevanten zu reduzieren. Der Einsatz der Filter führte bei den Untersuchungen zu einer Verbesserung der Schätzgüte des Modells. Diese beziehen sich aber immer nur auf das jeweilige Gesamtuntersuchungsgebiet.

Der Hierarchiefilter schränkt die relevanten Zielorte ein, indem er lediglich Interaktionen zwischen Orten zulassen, bei denen der Zielort eine größere Attraktivität besitzt als der Quellort. Dieser Filter ist allerdings in Frage zu stellen, wenn Daten nur auf Gemeindeebene betrachtet werden, da es durchaus rational ist, dass Kunden zu einem großen Angebotsstandort in einer kleineren Gemeinde zum Einkauf fahren, weil dieser besser erreichbar ist als das Einkaufszentrum in der eigenen Gemeinde[21].

Der Zentrumsfilter beruht darauf, dass Kunden innerhalb eines Untersuchungsgebietes nicht bereit sind, eine größere Distanz zurück zu legen als zu dem Ort, an dem sie das größte Warenangebot vorfinden. Alle Orte, die weiter entfernt liegen als der attraktivste Angebotsort, bleiben deshalb unberücksichtigt. Hierbei werden dann allerdings Spezialanbieter, die eventuell weiter als der attraktivste Ort liegen, nicht berücksichtigt.

Anstelle von Hierarchie- und Zentrumsfilter wird in dieser Untersuchung anhand der Einbindung des Straßennetzes, der Isodistanzen/Isochronen ein Erreichbarkeitsfilter genutzt, das die realen Gegebenheiten berücksichtigt. Weiterhin wird mit dem Filter Siedlungslayer für die Nachfrageseite gearbeitet, um somit die Datenmengen zu reduzieren.

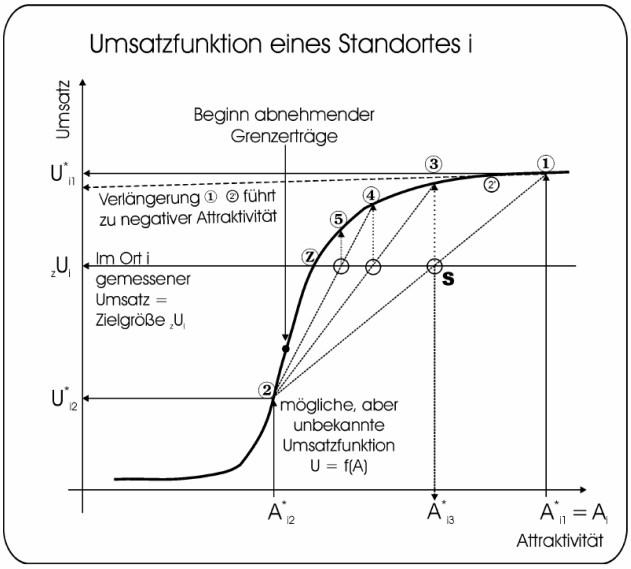
Weiterhin nutzen KLEIN & LÖFFLER neben dem Distanzexponenten einen weiteren Exponenten zur Gewichtung der Attraktivität. Hier ist anzumerken, dass die Verwendung eines einheitlichen Wertes für den Exponenten in einem Gebiet mit mehreren Standorten unterschiedlicher Attraktivität dazu führen kann, dass für das gesamte Gebiet zwar eine akzeptable Güte des Modells erreicht wird wie KLEIN belegt (1988, S. 47 ff und 1992, S. 70f), aber für die einzelnen Standorte die Werte anders aussehen. Weiterhin addieren KLEIN &LÖFFLER (1989 S. 406, *Translation der Distanzfunktion*) zu jeder Distanz den Wert 1, um den am Wohnort verbleibenden Ausgabeanteil berechnen zu können. Das heisst aber, dass unabhängig von der

[21] Anhand eines einfachen Routings auf Straßennetzdaten (hier *TeleAtlas*) ist die Entfernung zwischen Kundenstandort und Einkaufsorten messbar und vergleichbar.

flächenhaften Ausdehnung einer Gemeinde, die Zeit oder Entfernung vom Ortsrand ins Zentrum immer dieselbe sei. Insofern ist diese Modellierung der Innerortsdistanz realitätsfern.GÜßEFELDT (2002) schlägt ein Verfahren zur Bestimmung der Innerortsdistanz vor: aus der halben euklidischen Nächst-Nachbar-Distanz schätzt er die Fahrzeiten bei einer Durchschnittsgeschwindigkeit von 50 km/h. Eine Definition für 'Innerorts' gibt es allerdings nicht. Eine Durchschnittsgeschwindkeit von 50 km/h in ländlichen Orten kann realistisch sein. Dieser Durchschnittswert ist aber nicht für Großstädte anzuwenden. Entweder ist hier aufgrund von Erreichbarkeiten durch angrenzende Stadtautobahnen eine schnellere Geschwindigkeit realistisch oder aufgrund von Ampeln, Verkehrsdichte und zunehmenden 30 km/h Zonen eine wesentliche geringere Durschnittgeschwindkeit anzusetzen. Wo liegt nun die Grenze zwischen den 50 km/h Parameter und dem 30 oder 40 km/h Parameter? Wo hört der Raum 'Innerorts' auf? Wenn man die Speckgürtel der Städte einbezieht, gehören diese noch zu 'Innerorts'? Auch hier muss eine Unterscheidung getroffen werden, sonst handelt es sich wiederum lediglich um einen Austausch von einem festen Parameter durch einen anderen, welches damit keine tiefgreifende Modifikation des HUFF-Modells darstellt.

Andererseits strebt GÜßEFELDT (2002) an, die Ungleichgewichte der Märkte zu modellieren und nicht alleine die Abbildungsqualität des HUFF-Modells zu verbessern. Denn er geht davon aus, dass Einzelhandelsunternehmen Maßnahmen ergreifen zur Verbesserung ihrer Produktionsfunktion, die sich nur marginal in Strukturindikatoren bemerkbar machen. Bei der Quantifizierung der Attraktivitäten sind diese Unvollkommenheiten der Einzelhandelsmärkte zu berücksichtigen, wenn sie mit Hilfe des HUFF Modells abgebildet werden (GÜßEFELDT 2002, S. 356). GÜßEFELDT stellt eine theoretische Ertrags- und Umsatzfunktion eines Angebotsstandorts dar (*Abb. 8-21*). Die Abzisse entspricht der Attraktivität, die Ordinate dem Umsatz. Besitzt man Daten über das Nachfragevolumen in den Quellorten, so lässt sich Umsatzhöhe am Angebotsstandort schätzen. Es wird angenommen, dass der so geschätzte Umsatz weitaus höher liegt als der gemessene. Ziel ist es die Attraktivität des Zielortes zu verändern, damit der geschätzte Umsatz mit dem gemessenen nahe übereinstimmt.

Beachtet werden muss bei diesem Verfahren, dass die Optimierung immer für einen Zielort durchgeführt werden kann, bei Betrachtung mehrerer Orte müssen zunächst die Zielorte nach Attraktivitätswerten sortiert werden und dann iterativ vorgegangen werden.



Quelle: GÜßEFELDT (2002, S. 356) nach FITTKAU (2004, S.88)

Abb. 8-21: Umsatzfunktion eines Standortes in Abhängigkeit der Attraktivität

FITTKAU (2004) setzt das GÜßEFELDT Modell erstmals für einen Anwendungsfall der Beeinflussung regionaler Kaufkraftströme durch einen Autobahnlückenschluss ein und resümiert, dass es sich im ein realitätsnah entwickeltes Verfahren handelt (FITTKAU 2004, S. 92)[22]. Anhand von empirischen Untersuchungen von FITTKAU (2004) erweist sich die GÜßEFELDT Variante des HUFF Modells als vorteilhaft: *"Das von Güßefeldt entwickelte Modell stellt eine erhebliche Verbesserung der Abbildungsqualität des ursprünglichen Huff-Modells sowie dessen Abwandlungen durch Klein / Löffler dar. Beleg für das Funktionieren des Verfahrens ist in erster Linie, das sich nach dessen Anwendung die Abweichungen zwischen modellierten und statistischen Zielgrößen, in diesem Fall die Einzelhandelsumsätze in allen Gemeinden des Untersuchungsgebietes, innerhalb zuvor definierter Schwankungsbreiten bewegen"* (FITTKAU 2004, S. 198). Gleichzeitig merkt FITTKAU an, dass weiterer Forschungsbedarf bei den Parametern des Modells bestehen und das aufgrund entsprechend fehlenden Datenmaterials für die Maßstabsebene unterhalb der Gemeindeebene die Methodik nur eingeschränkt geprüft werden konnte.

Es bleibt auch hier anzumerken, dass die Parametrisierung der Funktion erhebliche Schwierigkeiten bereitet, hängt sie doch vom Verhalten der Unternehmen und der Wettbewerber am Standort ab. Den Einfluss des Konsumentenverhaltens auf die Umsatzfunktion klammert GÜßEFELDT aus.

Von daher ist der in dieser Untersuchung gewählte Ansatz einer Vorab-Clusterung (*Two-Step-Clusteranalyse*) eine mögliche Voraussetzung für die Anwendung von *Gravitationsmodellen*. Anhand der Cluster, die bereits vielfältige Merkmale der Gebiete in Betracht ziehen und bewerten, werden die Distanz-Parameter unterschiedlich zu wählen sein. Dies bietet eine aufwendigere Anpassung des Modells, aber eine

[22] Ein Verfahren den möglichen Umsatz an Angebotsstandorten abzubilden ist hier nicht Gegenstand dieser Untersuchung, sondern ein möglicher nächster Schritt.

realitätsnähere Abbildung. In Hinblick der heutigen technischen Möglichkeiten von GIS und *Spatial* Data Mining, zur Verfügung stehenden Routingalgorithmen und der Berücksichtigung der Datenverfügbarkeit auf Gebäudeebene bietet dies neue Dimensionen der Evaluierung von Modellen. Dies kann hier nur ansatzweise ausgeführt werden.

Beispielsweise finden in bisher keinem Modell Produktpreise Berücksichtigung. Es gibt im Marketing vielfache Untersuchung zum Einzelhandel und Unternehmenspolitik u.a. von Billigmarktanbietern wie *ALDI* oder *Lidl*. Den Einfluß, der die Nähe eines bestimmten Klientels von Handelsunternehmen als Attraktivitätsfaktor ausmacht oder nach FORTERHINGHAM als Agglomerationsvorteil, ist bisher im wirtschaftsgeographischen Bereich in dem Zusammenhang mit dem HUFF-Modell nicht ausreichend bearbeitet worden. Ein nachfrageorientiertes Modell wird entsprechend diese Effekte berücksichtigen müssen und in die Attraktivitätsparameter einbeziehen. Attraktivitätsfaktoren werden dann wiederum mit Hilfe von *CHAID-Analysen* definiert.

8.8 Weiterentwicklung der probabilistischen Modelle

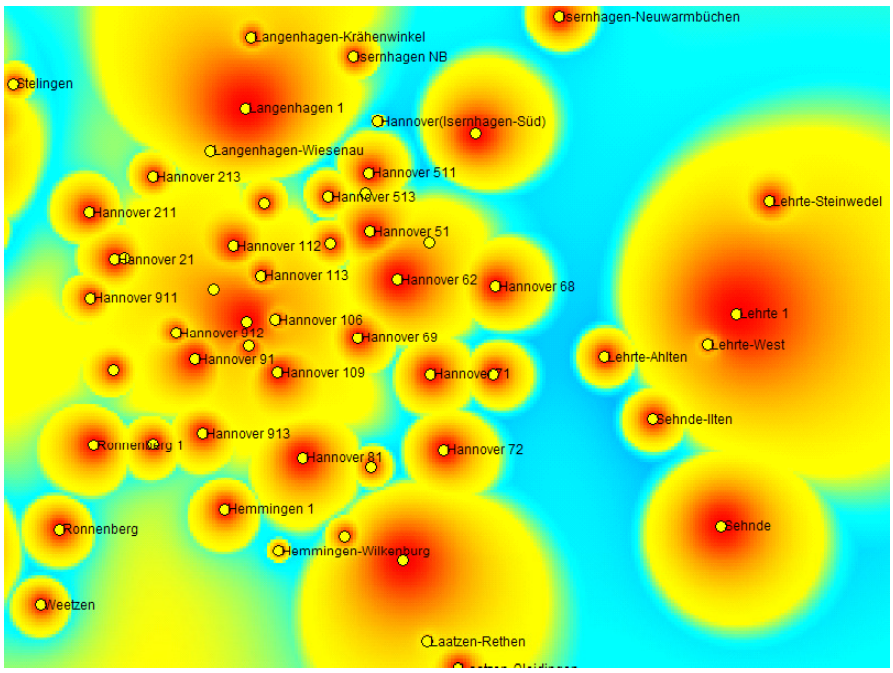
Im Gegensatz zu den deterministischen Interaktionsmodellen (*aggregierte Modelle*) stehen die verhaltensorientierte oder probabilistischen Modelle, so auch das HUFF-Modell und je nach Interpretation das Modell von REILLY. Konkrete Aussagen z. B. über Verkehrsbelastungen oder Fußgängerfrequenzen einzelner Straßenabschnitte können im HUFF-Modell auch probabilistisch formuliert werden (TAAFFE, GAUTHIER & O'KELLY 1996, S. 336 ff.). Probabilistische Modelle treffen Aussagen über die Wahrscheinlichkeit eines bestimmten Verhaltens, z. B. der Wahl des Besuchs eines Einkaufszentrums. Die wichtigste Form der Operationalisierung dieser Modellklasse sind *Logit-Modelle* (MCFADDEN 1974, 1977; FORTHERINGHAM 1983, 1989 U.A.), die auf Verfahren der Logistischen Regression beruhen. In diesen Modellen werden die z. B. aus Befragungen gewonnenen Erkenntnisse über tatsächliche Entscheidungssituationen auf Merkmale der Alternativen z. B. anhand der Entfernung bezogen und so ein Entscheidungsmodell für das betrachtete Phänomen generiert. Der verhaltensorientierte Ansatz hat vor allem bei der Betrachtung kleinräumiger Verkehrsverhalten und in der Verkehrsursachenforschung eine große Bedeutung (vgl. VOGEL 2000, S. 25f). Dieser Ansatz gewinnt durch die weitere Zusammenführung von GIS mit Data Mining und der Verfügbarkeit der entsprechenden kleinräumigen soziökonomischen Daten auch im Geomarketing weiter an Bedeutung.

Der Fortschritt in den probabilistischen Modellen besteht gegenüber den deterministischen Verfahren zwar weniger in der Genauigkeit von Hochrechnungen als in dem umfassenden Einblick in die komplexen Strukturen, die das Mobilitätsverhalten und die Entscheidungsfindung zur Wahl eines Besuchs von Standorten bestimmen.

Der eingeführte Parameter Attraktivität α, der bei REILLY mit Einwohner belegt ist und bei HUFF mit Ladenfläche, lässt sich abwandeln und unterschiedlich definieren je nach Region, Cluster, Branche, aber auch nach Kulturkreis. Die Entfernung λ unterliegt ebenfalls unterschiedlicher Kriterien bzw. Gewichtungsfaktoren (siehe auch Kapitel 9). Damit lässt sich das HUFF-Modell in Zusammenhang mit dem Möglichkeiten des Einsetzen von Filtern (Erreichbarkeiten, Siedlungsflächen, Sättigungsindex, Wettbewerberstandorte etc.) als ein sehr flexible einsetzbares Modell gestalten.

Die Kombination von HUFF mit den Überlegungen des Modells von FORTHERINGHAM, der Betrachtung von Zentralität, Einzelhandels- und Wettbewerberdichte unter Berücksichtigung eines Sättigungsindices (Filter) sind in die Untersuchungen eingegangen.

In das Konzept *GBI-Tool* ist das *Gravitationsmodell* von HUFF unter der Berücksichtigung des Modells von FORTHERINGHAM für diese Untersuchung integriert worden. Die Sättigungsindices werden als separates Grid-Layer anhand der bestehenden Anzahl von Standorten pro Einzelhändler berechnet analog kann bei Vorliegen von Kundendaten die bereits erreichte Anzahl an Kunden pro Einwohner berechnet werden.



Quelle: eigene Berechnung; Daten: Deutsche Post AG (GBI-Tool)

Abb. 8-22: Huff Gravitationsmodell: Filialen Hannover, Attraktivitätsmaß: Zeitanteil den Mitarbeiter mit Kunden verbringen

In der Abbildung *(Abb. 8-22)* wird die beispielhafte Anwendung des HUFF-Modell auf Filialen der *Deutschen Post* berechnet, wobei als Attraktivitätsmass die 'KGrAZ' (KundenGrundArbeitszeit)[23] als synonym für die Größe einer Filiale eingegangen ist. Zu erkennen sind die Einzugsbereiche von Filialen (Lehrte 1), die andere Filialen (Steinwedel, West) einschließen. Im nächsten Schritt werden beispielsweise Haushaltezahlen und Filter hinzugenommen. Auch hier wird in einem iterativen Verfahren versucht eine realitätnahe Abbildung zu generieren.

[23] Die KGrAZ ist eine spezielle Kennziffer der *Deutschen Post*, die sich aus der Anzahl der Arbeitszeit, die die Mitarbeiter mit dem Kunden verbringen (Schalter- und Beraterzeiten) ergibt. Sie lässt sich hier für eine beispielhafte Berechnung nach HUFF anwenden, geht man davon aus, dass ein hoher Zeitanteil, der mit Kunden verbracht wird, als Ausdruck der Attraktivität für die Filiale zulässig ist.