

## **.SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis**



Neubauer, Gerhard (2008):

### **Statistik als Werkzeug der strategischen Polizeiarbeit**

SIAC-Journal – Zeitschrift für  
Polizeiwissenschaft und polizeiliche Praxis  
(4), 19-27.

doi: 10.7396/2008\_4\_C

*Um auf diesen Artikel als Quelle zu verweisen, verwenden Sie bitte folgende Angaben:*

Neubauer, Gerhard (2008). Statistik als Werkzeug der strategischen Polizeiarbeit, SIAC-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (4), 19-27, Online: [http://dx.doi.org/10.7396/2008\\_4\\_C](http://dx.doi.org/10.7396/2008_4_C).

© Bundesministerium für Inneres – Sicherheitsakademie / Verlag NWV, 2008

Hinweis: Die gedruckte Ausgabe des Artikels ist in der Print-Version des SIAC-Journals im Verlag NWV (<http://nwv.at>) erschienen.

Online publiziert: 3/2013

# Statistik als Werkzeug der strategischen Polizeiarbeit

Diese Arbeit<sup>1</sup> steht in engem Zusammenhang mit den jüngst in dieser Reihe veröffentlichten Artikeln „Vom Informationsfriedhof zu Führungsinformationssystemen“ und „Das Geografische Informationssystem“ (Marouschek 2008a und 2008b). Seit 2004 wurden in einer Reihe von Projekten mit der Abteilung 4 des Bundeskriminalamtes einerseits die Grundlagen für ein umfassendes System zur strategischen Analyse im Bereich Kriminalität geschaffen, andererseits wurden zwei Module daraus konkret umgesetzt und sind seit 2006 im Einsatz. Im Folgenden wird ein Überblick über diese Arbeit gegeben.

Administrative Daten werden im Sicherheitsbereich seit langem auch zur Berichterstattung verwendet (Kriminalstatistik, Sicherheitsbericht). Seit 2004 existiert mit dem Sicherheitsmonitor (SIMO) ein direktes Meldesystem für Straftaten, das für Österreich flächendeckend aktuelle Daten liefert. Konzepte zur Bekämpfung und Prävention von Kriminalität erfordern strategische Entscheidungen, die durch das Verwenden empirischer Daten unterstützt werden können. Hier sind Lagebericht und Lagebild zu nennen, wobei vor allem letzteres den Zustand und die Entwicklungen der Kriminalität erfassen soll, um darauf aufbauend Bekämpfung und Prävention auszurichten. Dabei ist es von größter Wichtigkeit, die erforderliche Methodik regelrecht einzusetzen. Nur dann kann gewährleistet werden, dass der Informationsgehalt der Daten adäquat genutzt wird. Eine Analyse der Daten soll also zu Aussagen führen, die weder über den Gehalt der Information hinauschießen noch in ihrer Kraft darunter bleiben. Diese Aufgabe ist insbesondere mit den schlussfolgernden Methoden der Statistik zu bewältigen, die konzeptionell durch Wahrscheinlichkeits- und Entscheidungstheorie geprägt sind. In den Kapiteln 1 und 2 werden einige Grundlagen erörtert und in den Kapiteln 3–5 werden drei konkrete Analysemodule vorgestellt.



**GERHARD NEUBAUER,**  
Projektleiter am Institut  
für Angewandte Statistik  
und Systemanalyse der  
JOANNEUM RESEARCH  
Forschungsgesellschaft mbH.

## 1. STRATEGISCHE ANALYSEN – WANN, WIE, WARUM?

Das Interesse an strategischen Analysen wird in vielen Bereichen des öffentlichen und auch des privatwirtschaftlichen Sektors durch zwei Entwicklungen getrieben: 1) einem stärker werdenden Kostenbewusstsein, und 2) der immer größer werdenden Leistungsfähigkeit im Hard- und

Softwarebereich. Die Erhebung und Haltung großer Datenbestände sowie deren Analyse werden dadurch einerseits motiviert und andererseits technisch überhaupt erst möglich.

Allerdings darf dabei nicht übersehen werden, dass die Implementation einer Informationstechnologie für strategische Analysen nur dann erfolgreich sein kann,

wenn auf der Handlungsebene auch Alternativen zur Verfügung stehen. Grundsätzlich können Entscheidungsprobleme durch autoritär-hierarchisch oder demokratisch-kooperativ strukturierte Gruppen bearbeitet werden. Aus der Problemlöseforschung ist bekannt, dass es dabei wesentliche Unterschiede in der Qualität der Lösungen gibt. Der autoritär-hierarchische Stil ist bei einfachen und dringenden Problemen dem demokratisch-kooperativen überlegen. Andererseits führen demokratisch-kooperative Strukturen bei komplexen Problemen und bei genügend Zeit zu qualitativ besseren Lösungen. Der Einsatz empirischer Methoden ist also dann sinnvoll, wenn Handlungsalternativen zur Verfügung stehen, und ausreichend Zeit für Analysen vorhanden ist.

### **1.1. DESKRIPTIVSTATISTIK – BESCHREIBEND**

Typischerweise besteht hier das Ziel darin, den Zustand eines Sachgebietes darzustellen. Aus der Sicht der Statistik werden Kennwerte berechnet, Tabellen erstellt und graphische Darstellungen angefertigt. Gliederungen nach definierten Kriterien (z.B. Geschlecht, Regionen, Wirtschaftsklassen) ermöglichen das Durchführen von Vergleichen auf Basis von Augenscheinvalidität. Deskriptive Analysen sind unumgänglich, wenn man einen Gegenstandsbereich neu erschließt und wenig oder gar keine Erfahrung darin hat. Diese so genannte explorative Vorgangsweise ist aber in der Regel der Startpunkt für gezielte Fragestellungen (Hypothesen), die mit den Methoden der schlussfolgernden Statistik bearbeitet werden.

### **1.2. INFERENZSTATISTIK – SCHLUSSFOLGERND**

Das mathematische Hauptinstrument der Inferenzstatistik ist die Wahrscheinlichkeitstheorie. Unter der Annahme, dass

dem beobachteten Phänomen ein gewisses Zufallselement innewohnt, ist die analytische Behandlung mit Methoden der Wahrscheinlichkeitsrechnung unumgänglich. Die Beurteilung, ob ein gewisser Sachverhalt empirisch durch die Daten unterstützt wird – oder eben nicht –, erfolgt nach den Regeln der Entscheidungstheorie. Dazu werden der Sachverhalt und sein Gegenteil als These (Null-Hypothese) und Antithese (Alternativ-Hypothese) formuliert.

***Entscheidungstheoretisch betrachtet kann die These nicht bestätigt, sondern nur verworfen werden.***

Damit ist ein Sachverhalt nicht bewiesen, sondern lediglich „nicht widerlegt“. Theoretisch macht das einen großen Unterschied. In der Forschungspraxis, und mehr noch in der Kommunikation von Forschungsergebnissen, wird dies allerdings oft vergessen. Hoch wahrscheinliche Sachverhalte werden oft als absolute Gewissheiten betrachtet.

### **1.3. METHODIK**

Die Häufigkeit eines Ereignisses ist eine positive diskrete Variable, die im einfachsten Fall durch eine Poissonverteilung charakterisiert werden kann. Dadurch bewegt man sich nicht mehr in der klassischen Statistik der Normalverteilung und den Modelltypen Regression bzw. Zeitreihenmodell. Durch eine Verallgemeinerung des Regressionsansatzes (McCullagh/Nelder 1989) sind Regressionsmodelle für Häufigkeitsdaten möglich. Weitere Entwicklungen dieses Ansatzes erlauben nicht-lineare Regressionsfunktionen (Hastie/Tibshirani 1990) und die Berücksichtigung weiterer stochastischer Elemente, die u.a. auch die Spezifikation serieller Abhängigkeiten ermöglichen (McCulloch/

Searle 2001). Durch die Kombination dieser Methoden ist für Häufigkeitsdaten eine den herkömmlichen Regressions- und Zeitreihenmodellen für metrische Daten (z.B. Börsenkurse) gleichwertige Methodik verfügbar.

## **2. DER SICHERHEITSMONITOR ALS DATENBASIS**

Der SIMO ist eine einzigartige Datenquelle für Analysen auf höchstem Niveau. Im Vergleich zu anderen Datenerfassungssystemen zeichnen sich die SIMO-Daten durch die Charakteristika Aktualität und Vollständigkeit aus. Dadurch ist jederzeit eine bestmögliche analytische Betrachtung zur Kriminalität gewährleistet.

Theoretisch sind den Betrachtungen durch das System keine Grenzen gesetzt. Für praktische Zwecke ist es aber nötig, in den Dimensionen Zeit, Raum und Delikt Kategorisierungen vorzunehmen. Räumlich sind durch die Geo-Kodierung punktgenaue Analysen denkbar, zeitlich kann das kriminelle Ereignis oft im Bereich von Stunden zugeordnet werden, und was die Dimension Delikt betrifft, so ist jedes kriminelle Ereignis durch einen Paragraphen des StGB kategorisiert. Für gewisse Paragraphen ist durch zusätzliche Informationen wie Örtlichkeit, Begehungsart, Gut etc. eine feinere Betrachtung möglich, die unter dem Begriff „Schlagwort“ erfasst wird. Es zeigte sich, dass mit tagesgenauen Analysen in der zeitlichen Auflösung eine Grenze erreicht ist. Ebenso sind zu kleinräumige Betrachtungen nicht sinnvoll. Bei den Delikten wird eine Auswahl von ca. 30 Kategorien aus den Paragraphen und Schlagworten verwendet. Für die hier vorgestellten Analyseansätze werden die Daten auf die Raster (Kalenderwoche X Bezirk X Delikt) aggregiert.

## **3. ZEITLICHE PROGNOSEMODELLE**

Die zeitliche Prognose der Kriminalität ist ein wesentliches Grundelement bei der Unterstützung strategischer Entscheidungsprozesse. Das mit hoher Wahrscheinlichkeit erwartete Ausmaß an krimineller Aktivität ist dabei ebenso wichtig wie ein System, das es ermöglicht, gravierende Änderungen frühzeitig zu erkennen. Für beide Problemstellungen Prognosesicherheit und Trendmonitoring ist die adäquate Erfassung des zugrunde liegenden zeitlichen Prozesses von zentraler Bedeutung. Die hier vorgestellte Methodik ist im sogenannten Trendmonitoring-System (TMS) des Bundeskriminalamtes umgesetzt. Die vollständige Betrachtung aller Kombinationen von Bezirk X Delikt ergibt eine Menge von ca. 3.600 Modellen, wovon ca. 90 % irrelevant sind. Das TMS umfasst daher nur jene Kombinationen von Bezirk und Delikt, in denen auch regelmäßig ausreichend Fälle zu beobachten sind.

### **3.1. PROGNOSEMODELLE**

In Anlehnung an Modelle für ökonomische Zeitreihen wird ein Komponentenansatz verwendet. Es wird angenommen, dass der zeitliche Verlauf eines kriminellen Geschehens die Komponenten langfristiger Trend, Saisonalität und Kalendereffekte enthält. Unter Kalendereffekten versteht man regelmäßig wiederkehrende Ereignisse, die möglicherweise kein festes Datum haben, wie etwa Weihnachten, Ostern, Ferienzeiten. Abbildung 1 (siehe Seite 22) zeigt oben die Wochenhäufigkeiten zum Fahrraddiebstahl der Stadt Salzburg 2004–2007 (Punkte) und das Prognosemodell (Linie). Darunter sind die Komponenten Trend, Saison und Kalendereffekt abgebildet.

Grafik: Neubauer

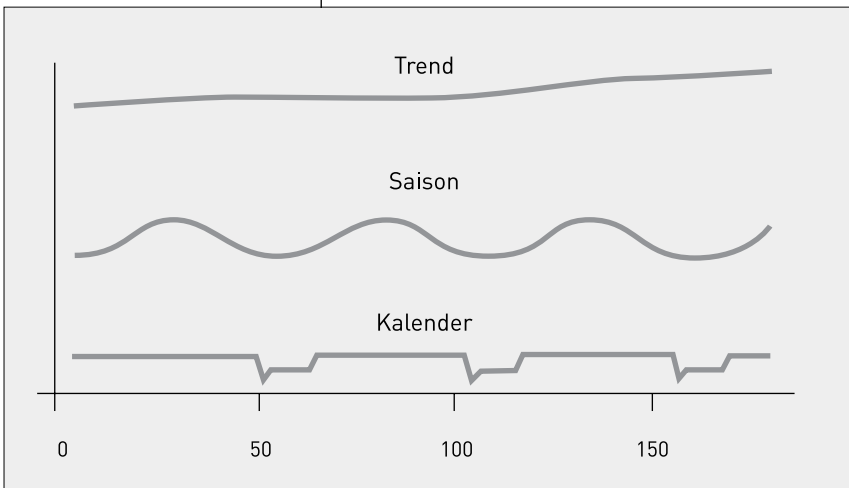
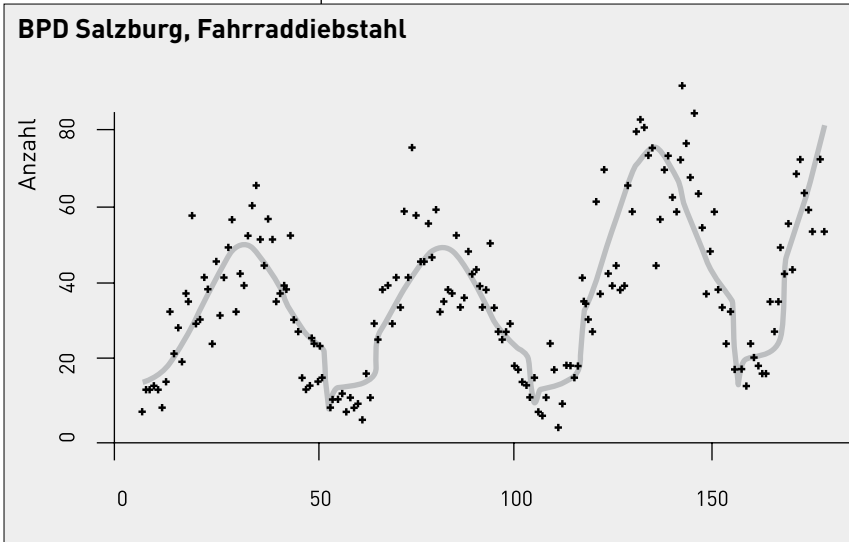


Abb. 1: oben – Daten und Prognosemodell; unten – Komponenten des Prognosemodells

### 3.2. TRENDMONITORING

Das wesentliche Ziel der Modellierung ist hier eine Art Frühwarnsystem für Kriminalitätsentwicklungen. Das Vorbild „Klassische Qualitätskontrolle in der Produktion“ (Schlagwort Six Sigma) ist nicht anwendbar, weil – anders als in der Produktion – der beobachtete Prozess nicht ausreichend kontrollierbar ist.

Als Alternative wurde ein System implementiert, das monatlich ein aktuelles Prozessmodell liefert und auf dieser Basis Prognosen für vier zukünftige Monate abgibt. Die Modellprognosen aus zwei Be-

reichen der Vergangenheit werden dann mit den Zukunftsprognosen verglichen. Als Vergleichszeitraum 1 (VZ 1) dienen die sechs Monate unmittelbar vor der Zukunftsprognose und als Vergleichszeitraum 2 (VZ 2) derselbe Zeitraum ein Jahr zuvor. Abbildung 2 (siehe Seite 23) zeigt als Beispiel die Prognose der Fahrraddiebstähle für die Stadt Salzburg. Im Vergleich zu VZ 1 wird ein erwarteter Anstieg von 45–58 Fällen pro Woche prognostiziert. Für VZ 2 beträgt der erwartete Anstieg 14–28 Fälle pro Woche. Fahrraddiebstahl ist offensichtlich ein saisonales Phänomen und daher ist der Vergleich mit VZ 2 adäquat.

Das implementierte System bewertet die Situation anhand des Vergleiches mit VZ 1. Wird ein Anstieg prognostiziert, so erhält die Abbildung einen roten Rahmen. Im Fall eines prognostizierten Rückgangs erhält die Abbildung einen grünen Rahmen. Kann keine Veränderung festgestellt werden, wird der Rahmen grau. Diese Farbe ist auch in einer Übersichtstabelle enthalten, wodurch eine schnelle Orientierung möglich ist. Zusätzlich kann die Geschichte aller Prognosen dargestellt werden.

### 3.3. QUALITÄT DER MODELLE

Ein statistisches Modell kann durch ausreichend flexible Gestaltung beliebig genau an vorhandene Daten angepasst werden. Ein derartiges Modell ist in der Regel sehr komplex, und es verliert die Kraft der Verallgemeinerung. Es ist spezifisch an einen Datensatz adaptiert. Deswegen wird in der Modellbildung versucht, mit dem einfachsten Modell die beste Anpassung zu erzielen. Dieses Prinzip ist im TMS umgesetzt. Es sind zwar die Komponenten Trend, Saison und Kalender in den Modellen vorgesehen, allerdings wird in jeder Aktualisierung überprüft, ob sie auch wirklich notwendig sind. So ist es auch möglich, dass sich Modelle über die Zeit ändern – also an veränderte Verhält-

nisse anpassen. Eine weitere Art der Qualitätskontrolle ist durch Kennwerte innerhalb der Modelle gegeben. Die Erfüllung verschiedener technischer Voraussetzungen wird durch geeignete Testverfahren überprüft. Zu diesen Voraussetzungen gehören Annahmen über die Verteilungseigenschaften der Residuen.

### 3.4. QUALITÄT DER PROGNOSEN

Bei der Konzeption des Systems wurde besonderes Augenmerk darauf gelegt, die Qualität der abgegebenen Prognosen beurteilen zu können.

***Daher wurde ein Kreuzvalidierungsansatz verwendet, in dem, mit dem zur damaligen Zeit vorhandenen Datenmaterial, der Echtbetrieb simuliert wurde.***

Konkret wurden die Daten der aktuellsten vier Monate (1/2006–4/2006) als Validierungsdaten deklariert. Mit den verbleibenden Daten (1/2004–12/2005) wurden Prognosemodelle (M1) und Prognosen (P1) erstellt. Danach wurde das System um ein Monat durch die Zeit bewegt und die Monate 2/2006–4/2006 dienten als Validierungsdaten, während die Daten 1/2004–1/2006 für Modellierung (M2) und Prognose (P2) verwendet wurden. Diese Vorgangsweise konnte noch zweimal wiederholt werden und lieferte insgesamt vier Prognosen für die Daten 4/2006, drei Prognosen für 3/2006, zwei für 2/2006 und eine für 1/2006. Daraus werden Kennwerte berechnet und somit eine erste Abschätzung der Prognosequalität geleistet. Mit dem Betrieb des Systems liegen derartige Informationen als Nebenprodukt vor und eine weitere Evaluierung wurde im Juli 2007 durchgeführt. Es zeigte sich, dass in ca. 75 % der Modelle die Prognosen von ausreichend guter Qualität sind.

Grafik: Neubauer

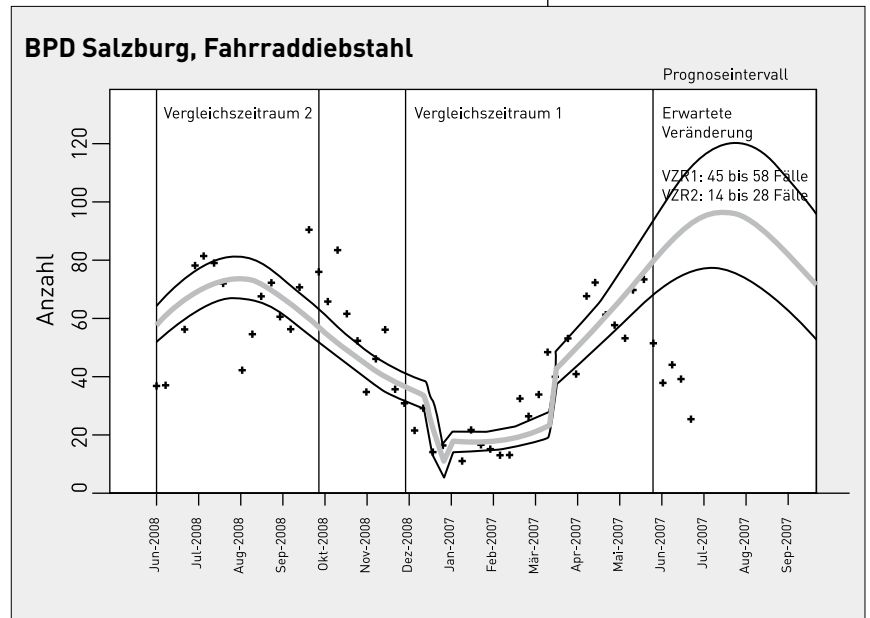


Abb. 2: Beispiel aus dem Tendmonitoring-System

Eine Detailanalyse der verbliebenen 25 % zeigte vor allem, dass in diesen Fällen die Variabilität in den Daten durch die bisher verfügbaren Informationen nicht erfasst werden kann. Durch eine Erweiterung der Datenbasis um lokale Informationen zu besonderen Ereignissen ist hier eine Verbesserung zu erwarten. Zu diesen Ereignissen zählen lokale Feste, Sportveranstaltungen, Kulturevents, aber auch strukturelle Veränderungen, wie die Eröffnung (Schließung) von Einkaufszentren oder Diskotheken oder die Anbindung an Schnellstraßen.

### 4. IDENTIFIKATION UND QUANTIFIZIERUNG VON UNTERSCHIEDEN

Die folgenden Beispiele umfassen Varianten eines Themas, nämlich die Identifikation und Quantifizierung von Unterschieden in der Kriminalität. Diese Unterschiede können durch zeitliche Veränderungen innerhalb eines Gebietes auftreten, die Folge von Maßnahmen oder Events sein, oder durch regionale Andersartigkeit bedingt sein.

1. Zeitliche Veränderungen: Dieser Aspekt ist besonders wichtig für die

Berichterstattung, da hier häufig der Unterschied zwischen vergleichbaren Zeiträumen (erstes Halbjahr 2006 vs. erstes Halbjahr 2007) dargestellt wird.

2. Evaluierung von Maßnahmen/Ereignissen: Maßnahmen wie Prävention oder Repression können auf ihre Wirksamkeit überprüft und beziffert werden, wenn sie so gesetzt werden, dass sie zeitlich und räumlich klar abgegrenzt sind. Die Auswirkung sportlicher oder kultureller Großereignisse kann durch einen Vorher/Nachher-Vergleich der lokalen (regionalen) Kriminalität anhand der SIMO-Daten bestimmt werden.
3. Regionale Unterschiede: Regionale Unterschiede können durch „belastende Infrastruktur“, wie Grenzübergang, Autobahnabfahrt, Einkaufszentrum etc. verursacht sein. Die Quantifizierung des Belastungsfaktors soll durch „Matching“ von Regionen bzw. Berücksichtigung von regionaler Heterogenität im Modell möglich werden (ceteris paribus Annahme).

Die beschriebene Methodik wurde in der Easy Test Application (ETA) des Bundeskriminalamtes umgesetzt. Sie ermöglicht geschulten Benutzern die selbständige Formulierung und Testung einfacher statistischer Hypothesen. Abbildung 3 (siehe Seite 25) zeigt das Ergebnis einer Analyse über die Wirkung einer polizeilichen Maßnahme. Im Mai 2007 wurde in Salzburg eine Bande von Fahrraddieben verhaftet und es wurde die Frage gestellt, ob diese Maßnahme einen nachweisbaren Einfluss auf die Anzahl der Fahrraddiebstähle hat. Der obere Teil der Abbildung 3 zeigt die Daten und das Trendmodell samt Effekt der Maßnahme. Es ist deutlich zu erkennen, dass die Kurve im Mai 2007 schlagartig absinkt.

Die Evaluierung des Kriminalitätsaufkommens während der EURO 2008 wird

ganz wesentlich auf Analysen mit der ETA beruhen.

## 5. REGIONALE HETEROGENITÄT

Die Entwicklung regional differenzierter Strategien der Bekämpfung und der Prävention erfordert eine regionalisierte Analyse der Kriminalität. Durch die zeitliche Prognose auf Bezirksebene ist dafür bereits ein wichtiges Element geschaffen. Darüber hinaus sind jedoch kriminogene Faktoren von besonderem Interesse. Modelle zur Identifikation und Quantifizierung solcher Faktoren lassen sich auch aus der Beschäftigung mit Kennzahlen der Kriminalität ableiten, was im Folgenden ausführlich dargestellt wird.

### 5.1. KENNZAHLEN DER KRIMINALITÄT

Eine Kennzahl der Kriminalität ist die Kriminalitätsrate, definiert als „the ratio of crimes in an area to the population of that area“.<sup>2</sup> Die Relation einer Häufigkeit zur Größe der Bevölkerung ist aus Demographie und Epidemiologie bekannt. So genannte „vital rates“ (Geburten-, Sterbe-, Erkrankungsraten etc.) werden als Maßzahlen von Prozessen in der Bevölkerung nach dem obigen Schema berechnet. Die Standardisierung in der Epidemiologie verfolgt den Zweck, den Effekt „trivialer“ Einflussfaktoren (z.B. des Alters) auf das Erkrankungsrisiko zu eliminieren. Dadurch soll der reine Effekt einer krankmachenden Einflussgröße sichtbar werden. Analog dazu soll eine Kennzahl der Kriminalität nicht-triviale regionale Unterschiede sichtbar machen. Das Problem der Analogie besteht nun darin, dass Kriminalität aus der Sicht der Exekutive primär ein Phänomen des Gebietes und nicht der Bevölkerung ist. Anders als bei der Erforschung von Populationseigenschaften haben wir hier die Eigenschaften eines nur teilweise an die lokale Population gebun-

denen Prozesses vorliegen. Das hat zur Konsequenz, dass – zumindest theoretisch – eine Kennzahl wie die Kriminalitätsrate größer als 1 werden kann, nämlich genau dann, wenn die Anzahl der Delikte in einem Zählgebiet die Anzahl der dort ansässigen Personen überschreitet. Als Beispiel denken wir uns einen Serienbetrüger, der in einem dünn besiedelten Gebiet ansässig ist und von dort aus eine große Anzahl von Delikten begeht. Das Ratenkonzept ist nur dann methodisch sauber anwendbar, wenn ausschließlich in Österreich ansässige Personen als Täter oder als Opfer Gegenstand der Betrachtung sind. Damit würde aber ein wesentlicher Teil der im Bundesgebiet stattfindenden Kriminalität aus der Analyse fallen.

### 5.1.1. KONSEQUENZEN FÜR DIE METHODIK

Um die Konsequenzen für die statistische Bearbeitung der Problematik zu verdeutlichen, vergegenwärtigen wir uns noch einmal die Situation für Analysen von „vital rates“. Das statistische Erhebungsmodell ist das Bernoulli-Experiment (z.B. Münzwurf). In einer Reihe identer Wiederholungen, deren Anzahl wir mit  $N$  bezeichnen, wird der Ausgang jedes Wurfes aufgezeichnet und binär kodiert. Also, wenn Kopf dann 1 und wenn Adler dann 0. Den Ausgang eines Experimentes können wir daher als

$$Y_i = \begin{cases} 1 & \text{wenn das Ereignis (Kopf) eintritt} \\ 0 & \text{sonst} \end{cases}$$

schreiben. Über alle  $N$  Wiederholungen gibt die Summe  $S = \sum_i Y_i$  genau die Anzahl der günstigen Ereignisse wieder. Da  $0 \leq S \leq N$  gilt, muss mit  $R = S/N$  auch  $0 \leq R \leq 1$  gelten. Für „vital rates“ wird dieses Modell nun auf Bevölkerungen so angewandt, dass man in einem definierten Zeitintervall den Bestand der Population

Grafik: Neubauer

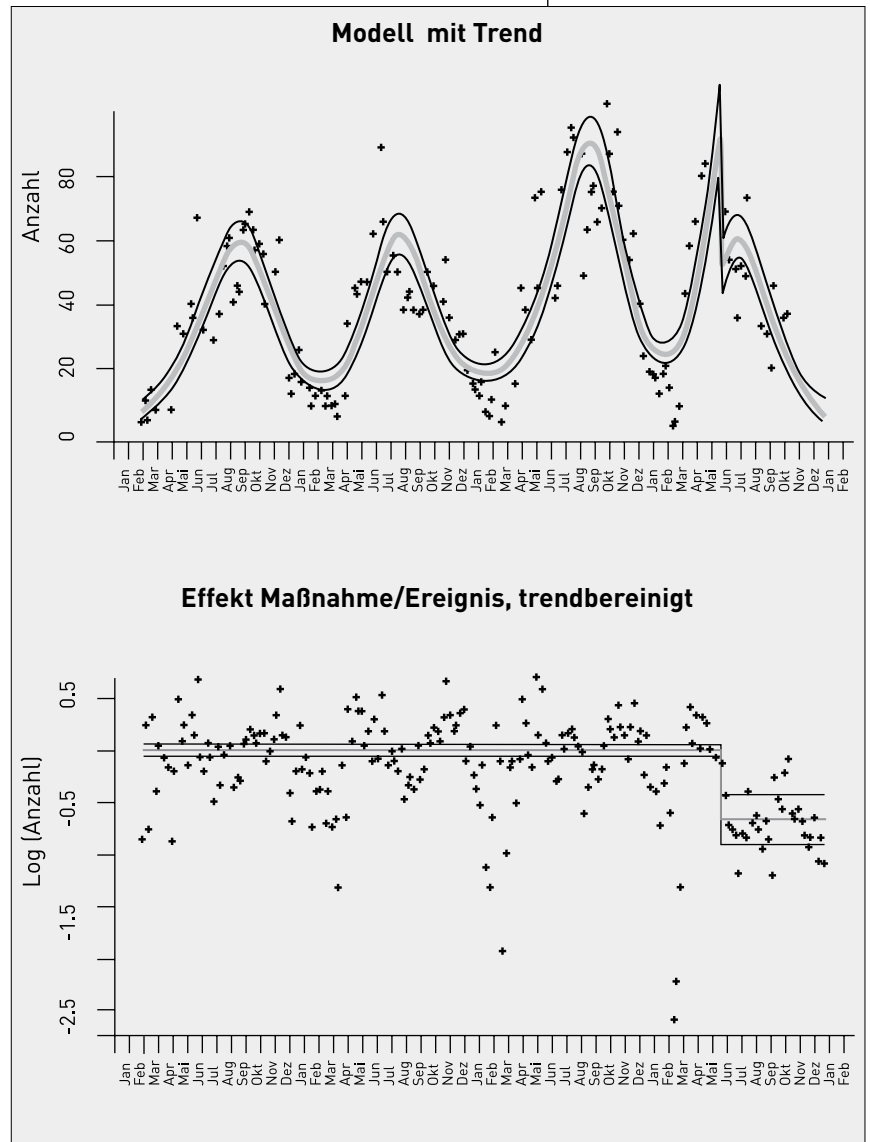


Abb. 3: Ergebnis einer Hypothesenprüfung mit der Easy Test Application

als konstant betrachtet und die Häufigkeit eines Ereignisses erfasst ( $S$ ). Unter der Annahme, dass sich die Individuen nicht systematisch unterscheiden, hat man  $N$  Wiederholungen.

Bei Übertragung dieses Konzeptes auf die Kriminalität wird sofort deutlich, dass hier die Grundgesamtheit nicht gegeben ist, wenn wir kriminelle Vorfälle betrachten. Natürlich ist zumindest eine physische Person an einem kriminellen Vorfall beteiligt – nämlich der Täter. Aber bereits bei



den Geschädigten können ausschließlich juristische Personen vorkommen. Im einfachsten Fall haben wir einen Täter und ein Opfer. Jedoch sind auch Fälle von 1:m, n:1 und n:m Relationen bekannt und geläufig. Eine Kennzahl, die die Anzahl der Delikte in Relation zur Größe der Wohnbevölkerung stellt, trifft implizit die Annahme, dass die Intensität des lokalen kriminellen Prozesses proportional zur Größe der Wohnbevölkerung ist. Das ist nur in idealisierten Populationen realistisch, die – quasi von der Außenwelt abgeschlossen – eine „endemische“ Kriminalität aufweisen.

Daher ist ein Ratenansatz nur bedingt zweckdienlich. Das Konzept der Standardisierung von Raten in der Epidemiologie lässt sich methodisch allgemein in Form von Modellen erfassen, die die Quantifizierung von Einflussgrößen ermöglichen. Also eine Reihe von potentiell Krankheit verursachenden Effekten, die simultan in einem Regressionsmodell berücksichtigt werden. Dieser Ansatz zur Erfassung und Beschreibung regionaler Unterschiede eignet sich auch für die Daten, mit denen eine Ratenrechnung nicht sinnvoll ist, also auch für die Daten der Kriminalität. Die Kriminalitätsrate wird ersetzt durch regionalisierte kriminogene Faktoren. Sie sind als jener Anteil der regionalen Kriminalität interpretierbar, der durch einen definierten Einfluss verursacht ist.

## 5.2. MODELLANSATZ – KRIMINOGENE FAKTOREN

Die Erfassung, Beschreibung und Quantifizierung regionaler Unterschiede in der Kriminalität führt zu Modellen mit erklärenden Variablen (kriminogenen Faktoren). Aus unterschiedlichen theoretischen Ansätzen können Bereiche identifiziert werden – wie Demographie, wirtschaftliche Situation, soziale Verhältnisse, Psychologie, Infrastruktur – die kriminogen wirken

können. Im Sinne eines explorativen Vorgehens verwenden wir unterschiedliche Indikatoren zur Erklärung regionaler Variabilität. Dabei bleiben aufgrund der verwendeten Datenquellen Täter- und Opfercharakteristiken unberücksichtigt. Vielmehr wird hier der Versuch unternommen, eine Region zu charakterisieren und nicht die dort ansässige Bevölkerung. Merkmale einer Region sind z.B.:

- a) Wohnbevölkerung, Alter, Ausbildung,
- b) Einkommen, Arbeitslosenrate,
- c) Wohnsituation, Kindergartenplätze,
- d) Anteil ausländischer Beschäftigter, Angestelltenanteil, Betriebsgröße,
- e) verbaute Fläche, Infrastrukturfläche, Autobahnkilometer, Anzahl der Autobahnabfahrten, Anzahl der Raststätten.

## 5.3. MODELLVARIANTEN

Auf den Trendmodellen aufsetzend versuchen wir die regionalen Unterschiede durch erklärende Variablen zu erfassen. Dazu ist es nötig, die bisher separaten Trendmodelle in ein Ganzes zu integrieren. Für die simultane Modellierung von Trendfunktionen gibt es mehrere Möglichkeiten:

- 1) Verketteten der Trendmodelle durch regionale Variablen,
- 2) Clustern der Trendfunktionen und
- 3) Aggregation der Zeitreihen.

Die Behandlung durch Raum-Zeit-Modelle wird hier nicht in Betracht gezogen, weil dadurch regionale Unterschiede nicht ursächlich erklärt werden.

Mit dem Ansatz 3) wurde eine Pilotstudie durchgeführt, in der gezeigt werden konnte, dass kriminogene Faktoren identifizierbar sind und gut interpretiert werden können. Der praktische Nutzen kann durch die Verwendung der regionalen Einflussgrößen in einem „What-If“-Tool erfolgen. Mit einem derartigen Werkzeug können, wie in einem Planspiel, die Auswirkungen von Veränderungen in den

regionalen Bevölkerungs- und Strukturmerkmalen abgeschätzt werden. Wiederrum liegt hier die Bedeutung des Analysemodules in der strategischen Orientierung hinsichtlich krimineller Entwicklungen.

## 6. AUSBLICK

Die vorgestellten Analysemodule sind als Teile eines Austrian Crime Information System (ACIS) konzipiert worden, dessen modularer Aufbau jederzeit Erweiterungen ermöglicht. Zwei Themen werden hier seit Beginn der Arbeit ganz unterschiedlich erörtert. Es sind dies die Themen Dunkelfzifferschätzung und Ausbreitungsmodelle.

Das Thema Ausbreitungsmodelle ist nach wie vor im Ideenstadium. Es ist getragen von der Vorstellung, dass es zwischen der raum-zeitlichen Verbreitung von

Kriminalität und der raum-zeitlichen Ausbreitung von Infektionskrankheiten eine gewisse Ähnlichkeit gibt. Für die Ausbreitung von Krankheiten gibt es sehr gute Simulationsmodelle und generell sind so genannte Punkt-Prozess-Modelle geeignet, die räumliche Verteilung von Ereignissen über die Zeit zu erfassen. Sollte es gelingen, gewisse Formen der Kriminalität durch derartige Modelle zu beschreiben, so müsste eine Prognose über zukünftige Tatorte möglich sein. Das klingt zugegebenermaßen etwas utopisch, trifft aber den Kern der Sache recht gut. Wiederrum war es die hervorragende Beschaffenheit des Datenbestandes SIMO, die hier der Idee auf die Sprünge half. Nur mit Geo-kodierten Daten ist an eine derartige Analyse überhaupt zu denken.

<sup>1</sup> Diese Arbeit ist meinem Kollegen und Freund Herbert Poltnig gewidmet. Sein Engagement und seine Begeisterung für die analytische Betrachtung realer Phänomene waren Auslöser meiner Beschäftigung mit strategischer Kriminalanalyse.

<sup>2</sup> Zum Beispiel <http://www.hyperdictionary.com/>.

### Quellenangaben

Hastie, T./Tibshirani, R. (1990). *Generalized Additive Models*, New York.

Marouschek, P. (2008a). Vom Informationsfriedhof zu Führungsinformationssystemen, *.SIAK-Journal* (2), 15–25 ff.

Marouschek, P. (2008b). Das Geografische Informationssystem, *.SIAK-Journal* (3), 88–98.

McCullagh, P./Nelder, J. A. (1989). *Generalized Linear Models*, New York.

McCulloch, C. E./Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*, New York.