

CLAUS HEIDEMANN

Die Vorausberechnung des Pkw-Bestandes in den Großstädten unter Verwendung des Zipfschen Gesetzes

1 Allgemeines

Zur Vorausberechnung des Pkw-Bestandes in der BRD sind in der letzten Zeit mehrere Veröffentlichungen erschienen (vgl. Lit. bei [1]). Im Gegensatz zum amtlichen Verfahren [2] haben sich die meisten Autoren für die Verwen-

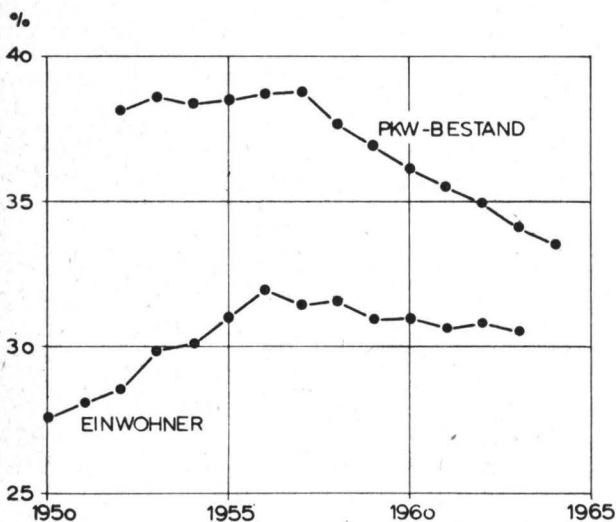


Abb. 1: Anteil der Gruppe der Großstädte an der Einwohnerzahl und am Pkw-Bestand der Bundesrepublik Deutschland (Großstädte: alle Städte der Bundesrepublik mit mehr als 100 000 Einwohnern)

dung der Logistischen Funktion entschieden. In mehreren Arbeiten des Instituts für Stadtbauwesen der TH Braunschweig wurde diese Methode auch für die Vorausberechnung des Pkw-Bestandes kleinerer Gebietseinheiten verwendet ([3], [4]). Zur Kontrolle der Einzelprognosen wurden dabei stets Prognosen größerer Gebiete herausgezogen. Dabei zeigte sich in vielen Fällen, daß nicht nur die Teilbereiche in festen Verhältnissen zum Großbereich stehen, sondern die Elemente der Teilbereiche auch untereinander durch eine stabile Ordnung verknüpft sind.

Für einen besonders interessierenden Teil der Bundesrepublik – die Gruppe der Großstädte (Städte mit mehr als 100 000 Einwohnern) – soll im folgenden die Vorausberechnung des Pkw-Bestandes durchgeführt werden. Wegen der innerhalb dieser Gruppe herrschenden Gesetzmäßigkeit können dann aus der Gruppenprognose auch Vorhersagewerte für Einzelstädte abgeleitet werden.

2 Vorausberechnung des Pkw-Bestandes für die Gruppe der Großstädte

Der Anteil der in den Großstädten der Bundesrepublik (ohne Berlin) gemeldeten Pkw am Gesamtbestand der BRD sinkt seit 1958 stetig (Abb. 1). Es ist damit zu rechnen, daß diese Entwicklung anhalten wird, bis der Anteil der Pkw etwa dem Anteil der Einwohner entspricht. Als Asymptote

dürfte für diesen Fall ein Wert von etwa 30% durchaus realistisch sein. Legt man einen für 1980 in der BRD zu erwartenden Bestand von 16 870 000 Pkw + Kombi zugrunde [1], so entspricht dies einer Zahl von etwa 5 100 000 Pkw + Kombi in den Großstädten.

Ein ähnlicher Wert (5 167 000 Pkw + Kombi) ergibt sich auch bei einer Vorausberechnung mit Hilfe der Logistischen Funktion

$$B_t = \frac{5\,312\,000}{1 + 13.961 \cdot \exp(-0.22184 \cdot t)}$$

mit

B_t: Bestand an Pkw + Kombi in den Großstädten der BRD;
t = 0 für 1952.

Die Übereinstimmung zwischen errechneten und tatsächlichen Werten (Tab. 1) ist um so bemerkenswerter, als in dem allgemeinen Trend der Zunahme des Pkw-Bestandes die Aufnahme weiterer Städte in die Gruppe der Großstädte im Laufe des Beobachtungszeitraumes enthalten ist: so hat sich die Zahl der Großstädte von 49 (1953) auf 55 (1964) erhöht.

3 Vorausberechnung des Pkw-Bestandes für einzelne Großstädte mit Hilfe des Zipfschen Gesetzes

Voraussetzung für die Aufteilung des vorausberechneten Bestandes auf die einzelnen Städte ist das Auffinden eines möglichst stabilen Ordnungsgesetzes für die Elemente der Gruppe. Ein derartiges Gesetz wurde erstmals von Karl Zipf [5] im Rahmen linguistischer Untersuchungen für den Zusammenhang zwischen der Rangzahl R und der Häufigkeit P der nach der Häufigkeit ihres Auftretens geordneten Wörter einer Sprache entdeckt.

Das Hyperbelgesetz

$$R \cdot P^q = \text{const.} = C,$$

das in seiner logarithmierten Form

$$\log R + q \cdot \log P = \log C$$

mit R: Rangziffer eines Elementes in der nach der Größe geordneten Reihe der Elemente

P: Wert des Elements

q: Exponent = Steigungsmaß der Geraden

C: Konstanter Wert = Durchstoßpunkt der Geraden durch die P-Achse

in Potenzpapier leicht als Gerade darzustellen ist, wurde von Zipf selbst bereits für geographische Probleme (Einwohnerzahl und Rangziffer der Gemeinden eines Landes) verifiziert. Als Exponent ergab sich dabei in sehr vielen Fällen q = 1.

Der Exponent q = 1 ist auch mit hinreichender Genauigkeit für die Einwohnerzahl-Rangziffer-Beziehung der Gemeinden ≥ 20 000 EW im Bereich der BRD anzusetzen (Abb. 2). Lediglich die zehn größten Städte weichen stark

Tabelle 1:

Vergleich der tatsächlichen Entwicklung des Bestandes an Pkw + Kombi in den Großstädten der Bundesrepublik (ohne Berlin) mit den nach der Logistischen Funktion errechneten Werten sowie Prognosewerte bis 1980

Jahr	Tatsächlicher Bestand [Pkw + Kombi]	Rechnerischer Bestand [Pkw + Kombi]	Abweichung [%]
1953	435 526	436 007	+ 0,1
1954	535 992	533 424	- 0,5
1955	642 208	649 707	+ 1,2
1956	788 042	787 164	- 0,1
1957	953 705	947 790	- 0,6
1958	1 126 770	1 132 985	+ 0,6
1959	1 319 256	1 343 231	+ 1,8
1960	1 572 805	1 577 761	+ 0,3
1961	1 837 989	1 834 315	- 0,2
1962	2 133 262	2 109 026	- 1,1
1963	2 410 955	2 396 527	- 0,6
1964	2 693 807	2 690 303	- 0,1
1965	—	2 983 000	—
1970	—	4 224 000	—
1975	—	4 896 000	—
1980	—	5 167 000	—

von der Geraden ab; ihre Einwohnerzahlen sind – gemessen an ihren Rangziffern – zu niedrig. Allerdings kann diese Erscheinung, die sich von 1875 bis 1962 fast über 90 Jahre erhalten hat, als Hinweis auf die Stabilität des Ordnungsschemas gewertet werden.

Ein Analogon zur Rangziffer-Einwohner-Beziehung ist die Beziehung zwischen Rangziffer und Pkw-Bestand (Abb. 3) der Großstädte im Bereich der BRD. Der Zusam-

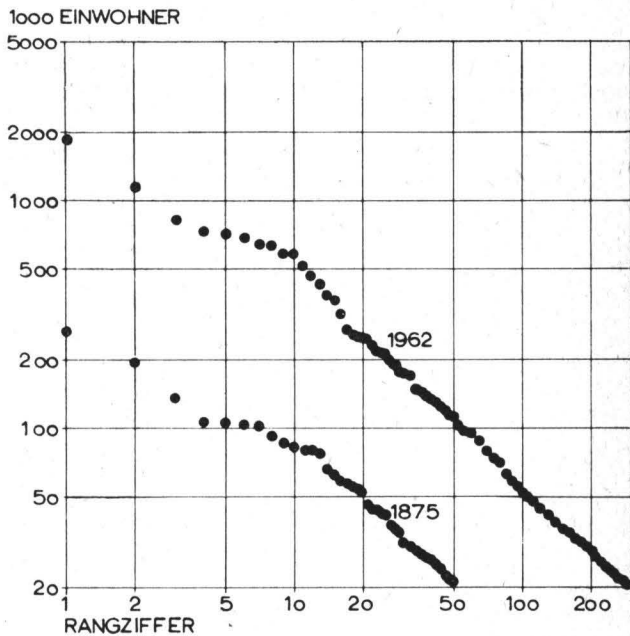


Abb. 2: Beziehung zwischen Stadtgröße und Rangziffer der Städte mit mehr als 20 000 Einwohnern im Bereich der Bundesrepublik 1875 und 1962

menhang ist durchweg nicht so straff wie in Abb. 2, doch ist ein allmähliches Angleichen an die Gerade erkennbar. Das typische Verhalten der Städte mit Rangziffern < 10 der Einwohner-Rangziffer-Beziehung wiederholt sich für die Pkw-Bestände. Auffällig ist ferner das Zurückbleiben des letzten Bereichs (Rangziffern $> N - 5$), doch ist die zeitliche Tendenz zum Angleichen an die Gerade deutlich (Vergleich 1928 bis 1964). Das Aufholen der Städte hoher Rangziffer ist ein Kennzeichen der Homogenisierung der räumlichen Pkw-Verteilung, die auch in den seit 1958 gegenüber den Großstädten höheren Zuwachsraten der Kleinstädte und Landkreise festzustellen ist. Es darf angenommen werden, daß die Pkw-Verteilung sich im Laufe der Zeit der Einwohnerverteilung (Abb. 2) anpassen wird.

Die Verwendung des Zipschen Gesetzes zu Zwecken der Vorausberechnung setzt jedoch voraus, daß folgende drei Bedingungen erfüllt sind:

(1) Die Rangstruktur muß ausreichend stabil sein.

- (2) Die Abweichungen der Einzelwerte von der Ausgleichsgeraden müssen ausreichend klein sein.
 (3) Die Lage der Ausgleichsgeraden muß aus typischen Gruppenwerten (Totalwert, Mittelwert usw.) abschätzbar sein.

3.1 Stabilität der Rangstruktur

Als Maß für die Konstanz der Reihenfolge kann der Rangkorrelationskoeffizient nach *Spearman* benutzt werden [6]. Danach ist

$$\text{Rho} = 1 - \frac{6 \cdot \sum_{i=1}^N d_i^2}{N(N-1)(N+1)}$$

mit Rho: Rangkorrelationskoeffizient, $-1 \leq \text{Rho} \leq 1$
 d_i : Differenz der Rangziffern des Wertes i
 N : Anzahl der Werte.

Rho = 1 bedeutet dann eine im Laufe der Jahre unveränderte Reihenfolge, Rho = 0 das Fehlen eines Zusammenhangs und Rho = -1 die Umkehrung der ursprünglichen Reihung.

Für den Zeitraum 1928–1938, der besonders durch Rangverschiebungen der Ruhrgebietsstädte infolge größerer Eingemeindungen gekennzeichnet ist, errechnet sich $\text{Rho} = 0,87$ ($N = 32$). Für den Zeitraum 1952–1964 ergibt sich $\text{Rho} = 0,88$ ($N = 44$), und bei einem Vergleich über 36 Jahre (1928 bis 1964) ist immer noch ein $\text{Rho} = 0,83$ ($N = 32$) vorhanden. Da alle Werte auf dem 99,9%-Niveau gesichert sind, kann die Rangstruktur als ausreichend stabil angesehen werden.

Dieses für die Gruppe der Großstädte gültige Ergebnis besagt jedoch nicht, daß die Rangziffer der Einzelstadt nicht starken Schwankungen unterliegen kann. Das gilt insbesondere für einige Ruhrgebietsstädte (Gelsenkirchen: $R(1928) = 27$, $R(1938) = 34$, $R(1964) = 18$; Oberhausen $R(1928) = 32$, $R(1918) = 39$, $R(1964) = 23$), während Städte wie Braunschweig ($R(1928) = 16$, $R(1938) = 18$, $R(1964) = 19$) und Karlsruhe ($R(1928) = 15$, $R(1938) = 19$, $R(1964) = 16$) relativ stetige Rangziffern aufweisen. Für die Vorausberechnung ist es also erforderlich, für jede Stadt die mutmaßliche künftige Rangziffer im Kollektiv der Großstädte abzuschätzen. Gerade das Einordnen der Einzelstadt in die Gruppe dürfte andererseits wertvolle Rückschlüsse auf die mögliche Einzelentwicklung erlauben.

3.2 Aufsuchen der optimalen Ausgleichsgeraden

Nimmt man die Steigung der Geraden mit $q = 1$ als gegeben an, so reduziert sich die Aufgabe auf das Bestimmen des Durchstoßpunktes C. Wegen der systematischen Abweichungen in den Bereichen $R < 8$ und $R > N - 5$ ist es nicht möglich, die übliche Abschätzung von

$$C = \sum_{i=1}^N P_i / \sum_{i=1}^N 1/R_i$$

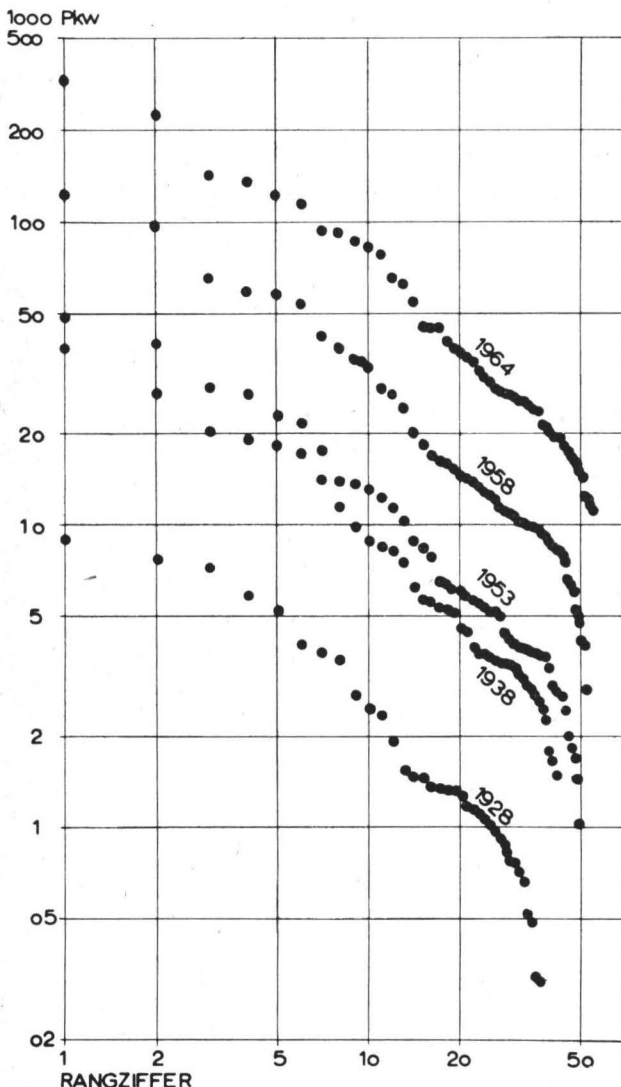


Abb. 3: Beziehung zwischen Pkw-Bestand und Rangziffer für die Großstädte im Bereich der Bundesrepublik

Tabelle 2:

Mittelwert (obere Zahl) und Streuung (untere Zahl) der Abweichungen der tatsächlichen Werte von den nach verschiedenen Verfahren berechneten Werten im Bereich $8 \leq R \leq N-5$

	Mittelwert und Streuung der prozentualen Abweichungen bei Verwendung des Verfahrens			
	(1)	(2)	(3)	(4)
	[0/0]	[0/0]	[0/0]	[0/0]
1953	- 3,37 ± 4,91	+ 3,76 ± 4,09	- 5,64 ± 6,07	- 0,94 ± 4,08
1954	- 0,19 ± 3,72	+ 6,97 ± 4,18	- 6,25 ± 6,19	+ 0,72 ± 3,42
1955	- 4,08 ± 5,43	+ 2,86 ± 3,52	- 6,91 ± 6,61	+ 0,72 ± 3,52
1956	- 8,37 ± 7,88	- 1,60 ± 4,84	- 6,38 ± 7,26	+ 0,38 ± 4,06
1957	- 7,14 ± 8,29	- 0,45 ± 4,85	- 5,64 ± 7,54	+ 0,92 ± 4,36
1958	- 6,25 ± 6,59	+ 0,40 ± 3,62	- 2,04 ± 4,42	+ 1,27 ± 3,53
1959	- 5,98 ± 5,87	+ 0,64 ± 3,31	- 0,09 ± 3,40	+ 2,01 ± 3,44
1960	- 3,67 ± 4,67	+ 3,01 ± 3,70	+ 3,64 ± 3,90	+ 2,24 ± 3,47
1961	- 4,93 ± 5,38	+ 1,63 ± 3,36	+ 3,38 ± 3,99	+ 2,37 ± 3,58
1962	- 4,47 ± 4,96	+ 2,06 ± 3,57	+ 4,11 ± 4,47	+ 2,68 ± 3,75
1963	- 5,54 ± 5,23	+ 0,86 ± 3,63	+ 3,68 ± 4,30	+ 2,80 ± 4,06
1964	- 8,53 ± 5,94	- 2,33 ± 3,81	+ 2,90 ± 3,65	+ 2,94 ± 3,67

zu benutzen. Wenn auch in Analogie zur Einwohner-Rangziffer-Beziehung angenommen werden darf, daß im Laufe der Zeit das starke Ausscheren der Punkte mit $R > N-5$ verschwinden wird, so wird die Asymmetrie der Verteilung infolge Abknickens der Punkte mit $R < 8$ erhalten bleiben. Diese Asymmetrie zeigt sich deutlich bei einem Vergleich von arithmetischem Mittel und Median; die Rangziffer des Mittelwertes liegt durchschnittlich um 11 Rangstufen unterhalb der des Medians. Allerdings ist die Beziehung zwischen beiden Werten sehr eng (Abb. 4). Die lineare Regression führt auf

$$x' = 759 + 0,55 \bar{x}$$

mit x' : Median der Verteilung $x' = P[N/2]$
 $\bar{x}$: arithmetisches Mittel der Verteilung $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$

Das Bestimmtheitsmaß ist mit $B = 0,99$ sehr hoch.

Zur Auswahl der „besten“ Ausgleichsgeraden wurden mehrere Verfahren durchgerechnet, um an Hand der geringsten Abweichungen die Entscheidung für eine Prognosegerade treffen zu können. Dabei sollte besonders der Bereich $8 \leq R \leq N-5$ berücksichtigt werden, da Prognose-

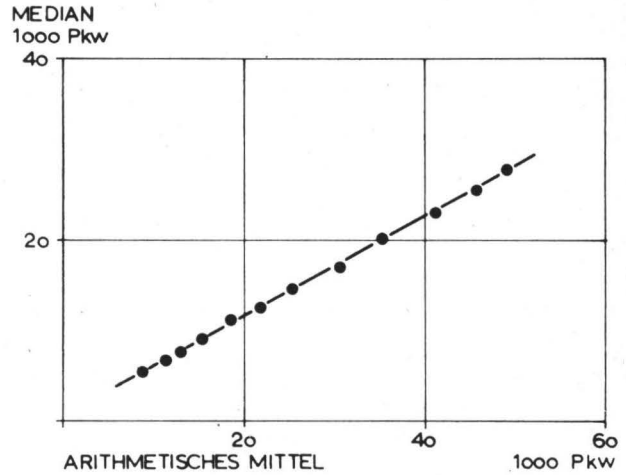


Abb. 4: Beziehung zwischen arithmetischem Mittel ($\bar{x}$) und Median (x') des Pkw-Bestandes in den Großstädten der Bundesrepublik 1953–1964

daten aus der Geradengleichung ohnehin nur für diesen Bereich ableitbar sind.

Die vier geprüften Verfahren zur Abschätzung von C sind:

- (1) Berechnung aus dem Mittelwert $\bar{x}$:

$$C1 = \bar{x} \left(\frac{N}{2} - 10 \right)$$

- (2) Berechnung aus dem Mittelwert $\bar{x}$:

$$C2 = \bar{x} \left(\frac{N}{2} - 11 \right)$$

- (3) Berechnung aus dem Median x' :

$$C3 = x' \cdot \frac{N}{2}$$

- (4) Berechnung aus dem mittleren Kurvenverlauf im Bereich $10 \leq R \leq 30$:

$$C4 = \frac{1}{21} \sum_{i=10}^{30} R_i P_i$$

Ein Vergleich der Rechenergebnisse mit den tatsächlichen Werten an Hand von Mittelwert und Streuung der prozentualen Abweichungen (Tab. 2) zeigt, daß sich die mittleren Abweichungen insbesondere bei den Verfahren (2), (3) und (4) innerhalb zulässiger Bereiche bewegen. Die Entscheidung für eines dieser Verfahren ist leichter zu treffen, wenn die Varianzanalyse die statistisch gesicherte Ungleichwertigkeit der Verfahren nachweist. Wegen der zeitlichen Komponente in der Datenmenge wird ein verteilungsfreier Test (Friedman-Test, [6]) benutzt. Danach kann für Mittelwerte ($V = 15,7$) und Streuungen ($V = 22,0$) die Nullhypothese (Gleichwertigkeit aller Verfahren) auf dem 99,5%-Niveau ($\chi^2 = 12,8$ für $f = 3$) abgelehnt werden. Aus den Werten der Berechnung ergeben sich die Verfahren (2) und (4) als die gleichwertig besten.

PARAMETER C2

1000 Pkw

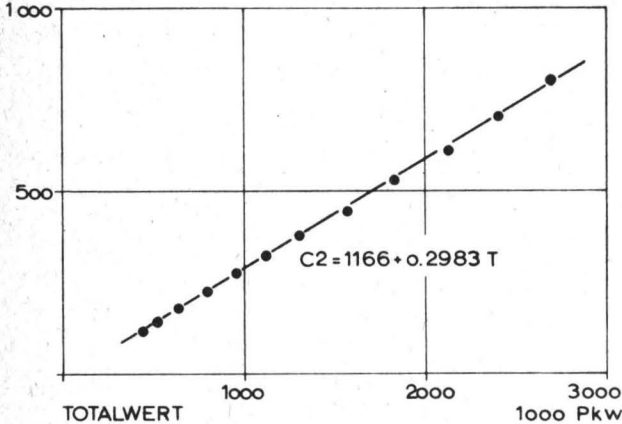


Abb. 5: Beziehung zwischen dem Parameter C2 und dem Gesamtbestand an Pkw in den Großstädten der Bundesrepublik 1953–1964

3.3 Abschätzen des Parameters C aus Gruppenwerten

Der Wert von C2 ist aus dem Totalwert der Gruppe dann unschwer zu bestimmen, wenn über die Veränderung von N ebenfalls Annahmen gemacht werden können:

$$C2 = \bar{x} \left(\frac{N}{2} - 11 \right) = \left(0,5 - 11/N \right) \cdot \sum_{i=1}^N P_i$$

Dies bedeutet jedoch eine zusätzliche Schwierigkeit, obwohl andererseits ein geringer Fehler von N keine großen Auswirkungen hat.

Einfacher dürfte die Abschätzung von C2 und C4 aus dem Totalwert der Gruppe sein, da beide Werte mit der Gruppensumme durch eine einfache lineare Beziehung verbunden werden können (Abb. 5 und 6). Damit ist auch die letzte Voraussetzung für die Verwendung des Zipfschen Gesetzes als Verteilungsmodell für Prognosezwecke erfüllt.

PARAMETER C4

1000 Pkw

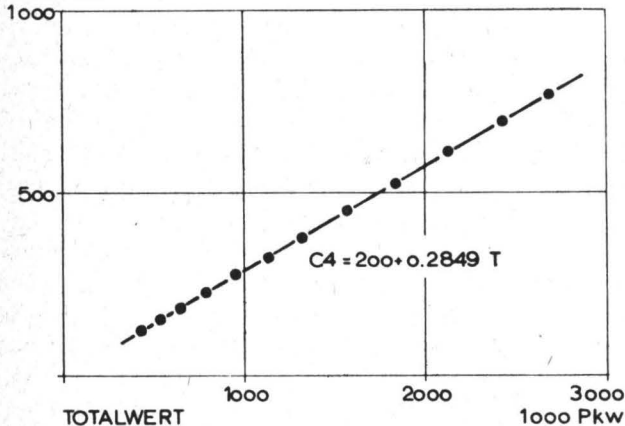


Abb. 6: Beziehung zwischen dem Parameter C4 und dem Gesamtbestand an Pkw in den Großstädten der Bundesrepublik 1953–1964

Tabelle 3:

Prognosewerte der Parameter C2 und C4 zur Abschätzung der Einzelwerte

Jahr	Pkw-Bestand der Großstädte [Pkw + Kombi]	Parameter des Zipfschen Gesetzes	
		C 2 [Pkw + Kombi]	C 4 [Pkw + Kombi]
1965	2 983 000	891 000	850 000
1970	4 224 000	1 261 000	1 204 000
1975	4 896 000	1 462 000	1 395 000
1980	5 167 000	1 542 000	1 472 000

3.4 Berechnung der Prognosewerte für Einzelstädte

Aus den vorausberechneten Werten der Tabelle 1 lassen sich die Parameter C2 und C4 nach den Gleichungen der Abb. 5 und 6 errechnen (Tab. 3). Da keine der beiden Schätzmethode als eindeutig überlegen anzusehen ist, werden die mutmaßlichen Geraden durch den zwischen ihnen liegenden Bereich ersetzt (Abb. 7). Die Bereichsgrenzen können dann als untere und obere Grenze der Prognosewerte interpretiert werden. Für die Städte mit Rangziffern $R < 8$ wurde in Abb. 6 die Punktreihung der letzten Jahre, die zwar stark von der Geraden abweicht, in sich aber durchaus stabil ist, auch für die Prognosejahre eingetragen.

Die Ableitung von Bestandsprognosen für Städte mit Rangziffern $8 < R < 50 \dots 60$ kann so vor sich gehen, daß die Rangziffer einer Stadt für den Prognosezeitraum an Hand der bisherigen Rangzifferentwicklung oder einer Ein-

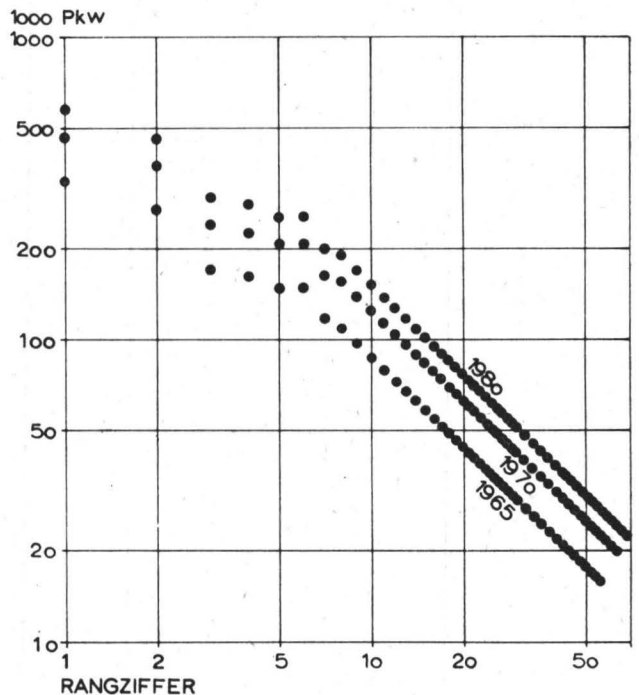


Abb. 7: Prognosewerte der Beziehung Pkw-Bestand – Rangziffer für die Großstädte der Bundesrepublik

Tabelle 4:
Prognosewerte für die Stadt Braunschweig

Jahr	Bestand an Pkw + Kombi nach Schätzung	
	(a)	(b)
1965	42 500–44 500	43 500
1970	60 300–63 100	61 400
1975	69 800–73 200	71 000
1980	73 700–77 100	75 000

(a) Hochrechnung der Gruppe der Großstädte nach der Logistischen Funktion und Verteilung nach dem Zipfschen Gesetz
(b) Hochrechnung des Pkw-Bestandes nach der Logistischen Funktion [4]

wohnerprognose abgeschätzt wird. Division der C-Werte in Tab. 3 durch die Rangziffer ergibt dann den wahrscheinlichen Bestand an Pkw + Kombi. Für die Stadt Braunschweig z. B., deren Rangziffer seit 1953 leicht steigende Tendenz zeigt, ergeben sich für $R = 20$ die Schätzwerte der Tab. 4. Die Übereinstimmung der Ergebnisse darf als befriedigend angesehen werden.

4 Zusammenfassung

Ausgehend von der Beschreibung des Zusammenhanges zwischen Einwohnerzahl und Rangziffer der Städte > 20 000 Einwohner im Bereich der Bundesrepublik durch das Zipfsche Gesetz, kann nachgewiesen werden, daß dieses Gesetz auch für die Beziehung zwischen Pkw-Bestand und Rangziffer der Städte > 100 000 Einwohner gilt. Die wesentlichen Bedingungen zur Anwendung des Gesetzes für Zwecke der Vorausberechnung, nämlich

- (1) Stabilität der Rangstruktur,
- (2) Einhalten eines zulässigen Fehlerbereiches,
- (3) Abschätzen des Lageparameters aus Gruppendaten,

können als hinreichend erfüllt gelten. Aus der Prognose des Gesamtbestandes an Pkw in der BRD bis 1980 können damit Prognosewerte des Pkw-Bestandes einzelner Großstädte abgeleitet werden.

Diese Methode erlaubt die Prognose für eine Stadt auch dann, wenn eindeutige Ausgangsdaten bzw. die der Entwicklung aus der Vergangenheit nicht vorliegen. Üblicherweise wird man die Voraussage der künftigen Entwicklung auf gemessenen Daten aufbauen und mit eigenen mathematischen Methoden ermitteln ([1], [2]). Aber auch dann hat die Anwendung des Zipfschen Gesetzes in vorstehend beschriebenen Verfahren noch Bedeutung, indem auf diese Weise eine unabhängige Kontrolle der errechneten Werte möglich ist.

Literatur

[1] Heidemann, C.: Zur Vorausberechnung des Bestandes an Pkw in der Bundesrepublik bis 1980. In: Straßenverkehrstechnik. 9. Jg. (1965) H. 9/10, S. 157–160.
[2] Richtlinien für die Anlage von Landstraßen. I. Teil RAL-Q. Hrsg.: Forschungsgesellschaft für das Straßenwesen. – Bielefeld 1956.
[3] Habekost, H.; C. Heidemann; P. Kirchhoff: Verkehrsgutachten Nordhorn. – Institut für Stadtbauwesen, TH Braunschweig 1965. (Typoskript.)
[4] Raumplanungsgutachten Südniedersachsen. Hrsg.: F. Zimmermann. Monographie Verkehr. – Braunschweig. (Im Erscheinen.)
[5] Zipf, G.K.: Human behavior and the principle of least effort. – Cambridge/Mass. 1949.
[6] Pfanzagl, J.: Allgemeine Methodenlehre der Statistik. – Berlin 1962.