

Akademie für Raumforschung und Landesplanung

 ARBEITSMATERIAL DER ARL

Räumliche Konsequenzen des demographischen Wandels

Teil 8

# Demographischer Wandel in ausgewählten Regionstypen Nordrhein-Westfalens

Herausforderungen und Chancen für regionale Politik

Rainer Danielzyk, Heiderose Kilper (Hrsg.)

## **Gemeindetypisierung Nordrhein-Westfalens nach demographischen Merkmalen**

### **Gliederung**

- 1 Einleitung
- 2 Erster, induktiver Analyseansatz
- 3 Zweiter, hypothesengeleiteter Analyseansatz
- 4 Methodenkritische Schlussbemerkungen

Literatur

Anhänge

### **1 Einleitung**

Der aktuelle demographische Wandel in Nordrhein-Westfalen hat ohne Zweifel auch eine räumliche Dimension – aber welche? Schon ein erster Blick auf die räumlichen Ausprägungen von Fertilität, Mortalität, Wanderungen usw. zeigt, dass die regionalen Muster – auch in Nordrhein-Westfalen – immer weniger den überkommenen landesplanerischen Gebietskategorien entsprechen. In den 1950er und 1960er Jahren ließen sich die landesplanerischen Problemstellungen noch weitgehend nach bevölkerungs- und siedlungsräumlichen Struktur- und Entwicklungsräumen differenzieren: Während die Ballungsräume wuchsen, verloren die ländlichen Räume an Bevölkerung. In den 1970er und 1980er Jahren verlagerten sich die wanderungsbedingten Zuwächse auf Kosten der Kernstädte in die Ballungsrandzonen, während sich die Bevölkerungsentwicklung in den ländlichen Räumen zunehmend stabilisierte. Seit den 1990er Jahren wurde die Situation immer unübersichtlicher: Das Wachstum der Ballungsrandzonenstädte flachte ab, in den Ballungsgebieten entwickeln sich die Kernstädte uneinheitlich und in den ländlichen Regionen liegen wachsende Gemeinden teilweise unmittelbar neben stagnierenden und schrumpfenden, sodass es schwer fällt, einheitliche regionale Trends auszumachen.

In dieser Situation ist die Regionalforschung aufgerufen, einen genaueren Blick auf und hinter die Daten zu werfen, um zu prüfen, inwieweit sich neue regionale Muster und Entwicklungstrends identifizieren lassen. Immer mehr setzt sich die Erkenntnis durch, dass die noch im Landesentwicklungsplan NRW von 1995 enthaltene tradierte Gliederung des Landes nach der „siedlungsräumlichen Grundstruktur“, die zwischen Ballungsgebieten, „solitären Verdichtungsgebieten“ und „Gebieten mit überwiegend ländlicher Raumstruktur“ unterschied, die landesplanerische Problemlage heute nicht mehr angemessen beschreibt. Aber welche Raumtypisierung oder Regionalisierung könnte stattdessen geeignet sein? Lässt sich eventuell aus der Bevölkerungsstruktur und -entwicklung eine neue Typisierung oder Regionalisierung ableiten?

Diesen Fragen geht der folgende Beitrag in der Form eines Werkstattberichtes nach. Referiert werden die Ergebnisse von zwei Clusteranalysen zur Typisierung der nord-

rhein-westfälischen Gemeinden nach ihrer Bevölkerungsstruktur und -entwicklung.<sup>1</sup> Die Clusteranalyse ist ein mathematisches Klassifikationsverfahren, das Objekte nach ihrer Ähnlichkeit in einem Merkmalsraum zu möglichst homogenen Gruppen („Cluster“) zusammenfasst. Die besondere Leistungsfähigkeit besteht darin, dass eine große Zahl von Merkmalen (Variablen) berücksichtigt werden kann.

Als multivariates Klassifikationsverfahren gehört die Clusteranalyse inzwischen zum Standardwerkzeug der Regionalforschung.<sup>2</sup> Allerdings täuscht die einheitliche Bezeichnung „Clusteranalyse“ insofern, als sich dahinter nicht eine bestimmte standardisierte Methode im Sinne einer einheitlichen Vorgehensweise von bestimmten Schritten nach wohl definierten Regeln verbirgt, sondern ein weites Feld von Klassifikationsmethoden mit zahllosen Varianten. Ihre Anwendung erfordert an mehreren Stellen Entscheidungen über die Vorgehensweise, wobei sich der einschlägigen Fachliteratur keineswegs immer eindeutige Entscheidungsregeln entnehmen lassen. Daraus resultiert, dass Clusteranalysen verschiedener Autoren häufig zu unterschiedlichen Ergebnissen führen, selbst wenn dieselben Daten verwendet wurden.

Dies beginnt bereits bei der Auswahl der Variablen, die einerseits von der Verfügbarkeit der Daten der amtlichen Statistik abhängt, aber andererseits auch davon, ob eine eher induktive oder eine stärker hypothesengeleitete Vorgehensweise gewählt wird. Von erheblichem Einfluss auf die Ergebnisse sind auch die Definition und Skalierung der Variablen, beispielsweise die Frage, ob Absolut- oder Verhältniszahlen verwendet und ob diese eventuell einer Transformation (beispielsweise zur Beseitigung der Verteilungsschiefe und/oder zur Annäherung an die Normalverteilung) unterzogen werden. Weitere Entscheidungen sind bei der Definition des Ähnlichkeitsmaßes, dann bei der Frage, ob zunächst die Korrelationen zwischen den Variablen dadurch beseitigt werden, dass diese durch unkorrelierte Hauptkomponenten bzw. Faktoren ersetzt werden, dann bei der Auswahl des Klassifikationsalgorithmus sowie bei der Bestimmung der Zahl der Cluster fällig. Schließlich sind die Ergebnisse inhaltlich zu interpretieren. Auch der folgende Ergebnisbericht ist insofern nicht ganz frei von einem subjektiven Moment.

Im Einzelnen geht es um die folgenden Fragestellungen:

- Welche demographischen Struktur- und Entwicklungsmuster lassen sich identifizieren?
- Welche Gemeindetypen nach demographischen Merkmalen lassen sich unterscheiden?
- Welche Gemeinden gehören den ermittelten Typen an?
- Lassen sich räumliche Muster der Zugehörigkeit zu diesen Typen identifizieren?
- Lassen sich aus der Typisierung Hinweise auf eine mögliche neue landesplanerische Regionalisierung ableiten?

Zu diesen Fragen existieren bisher kaum Voruntersuchungen. In methodischer Hinsicht hilfreich sind die älteren Studien zur multivariaten und speziell clusteranalytischen Klassifizierung von Städten insbesondere aus den 1970er Jahren (Forst 1974, Möllers

---

<sup>1</sup> Der Verfasser dankt Markus Jeschke für die wertvolle Unterstützung bei der Zusammenstellung der Daten und der Durchführung der Berechnungen. Verwendet wurde das Programmpaket SPSS.

<sup>2</sup> Einführende allgemeine Darstellungen der Clusteranalyse findet man beispielsweise bei Bortz 2004, Kap. 16, Schlosser 1976 und Sodeur 1974, vertiefende allgemeine Darstellungen bei Backhaus u. a. 2000, Eckes u. Rossbach 1980 sowie Steinhausen u. Langer 1977. Zur Anwendung der Clusteranalyse in der Regionalforschung vgl. insbesondere Bahrenberg u. a. 1992, Eckey u. Klemmer 1975 und Sedlacek 1978.

1977, Sitterberg 1977, Blotevogel 1979; zur Klassifikation nordrhein-westfälischer Städte vgl. insbesondere de Lange 1980). Dieser methodische Ansatz hat sich inzwischen weitgehend als Standard durchgesetzt.<sup>3</sup> Allerdings sind spezielle clusteranalytische Typisierungen anhand demographischer Merkmale bisher selten. Bähr und Gans (2003) haben beispielsweise eine clusteranalytische Regionentypisierung Deutschlands auf der Grundlage von Migrationsdaten vorgenommen. Schlömer (2004) hat unter Verwendung von Daten zur Bevölkerungsprognose bis 2020 eine clusteranalytische Typisierung der 97 Raumordnungsregionen Deutschlands vorgenommen und damit versucht, regionale demographische Problemlagen der kommenden zwei Jahrzehnte zu identifizieren. Untersuchungen, welche die genannten Fragestellungen unmittelbar behandeln, liegen jedoch nicht vor.

## **2 Erster, induktiver Analyseansatz**

Es wurden zwei aufeinander aufbauende Analysen mit zwei unterschiedlichen Datensätzen durchgeführt. Zunächst wurden von den 396 Gemeinden des Landes Nordrhein-Westfalen die drei Gemeinden Hemer, Schöppingen und Unna eliminiert, weil Probeberechnungen zeigten, dass die Wanderungsdaten dieser drei Gemeinden in atypischer Weise durch die dort lokalisierten Aufnahmestellen und Übergangseinrichtungen für Aussiedler, Spätaussiedler und Zuwanderer geprägt werden. Es verbleiben damit 393 Gemeinden.

Für diese wurden zwei Datensätze gebildet. Der erste umfasst 65 demographische Variablen; das sind praktisch alle Merkmale, die die amtliche Statistik für den Zeitraum der letzten Jahre (1996 bis 2003) bereitstellt (siehe Anhang). Die Analyse dieses ersten Datensatzes ist induktiv angelegt nach dem Prinzip, möglichst viele Informationen in die Analyse einzubeziehen und dann zu prüfen, ob sich sinnvoll interpretierbare Muster ergeben. Die Variablen enthalten Daten zur Bevölkerungsentwicklung, zur Bevölkerungsdichte, zur Fertilität und Mortalität, zu den Zu- und Fortzügen der Deutschen und der Nichtdeutschen, zu den Pendlersalden (pro Beschäftigten) sowie zur Altersstruktur. Verwendet werden ausschließlich Verhältniszahlen und keine Absolutwerte, um zu vermeiden, dass die unterschiedliche Gemeindegröße die Typisierung in trivialer Weise beeinflusst.

Der zweite Datensatz umfasst lediglich 30 Variablen (siehe Anhang) und ist stärker hypothesengeleitet. Auf der Grundlage der induktiven Ergebnisse der ersten Analysen wurde eine engere Auswahl von 30 Variablen getroffen, die vor allem Merkmale der Altersstruktur, des Wachstums, der Erwerbstätigkeit und des Pendelns umfassen. Die Auswahl wird weiter unten näher begründet.

Die beiden Datensätze umfassen somit:

- a) 393 Gemeinden x 65 Variablen (alle verfügbaren Daten)
- b) 393 Gemeinden x 30 Variablen (hypothesengeleitete Merkmalsauswahl).

Die Aussage vieler Variablen unterscheidet sich teilweise nur minimal, und zwar allein schon deshalb, weil dieselben Merkmale für mehrere Jahre einbezogen wurden. Der statistische Effekt besteht in teilweise hohen Interkorrelationen zwischen diesen Variablen. Würden die Korrelationen nicht beseitigt, erhielten diese Variablen bei der Clusterrung ein unangemessen großes Gewicht und würden damit die Ergebnisse stark verzerren. Deshalb wurde zunächst eine Hauptkomponentenanalyse durchgeführt, um den

---

<sup>3</sup> Vgl. beispielsweise die Städtetypisierungen im Bericht der Enquetekommission „Zukunft der Städte“: Landtag NRW 2004.

vieldimensionalen Merkmalsraum der Variablen auf einen Merkmalsraum mit deutlich weniger, aber unkorrelierten Dimensionen zu reduzieren. Das Ziel der Hauptkomponentenanalyse besteht darin, den 65- bzw. 30-dimensionalen Merkmalsraum so zu reduzieren, dass die in den Ausgangsvariablen enthaltene Information möglichst erhalten bleibt, aber dabei erheblich verdichtet wird, indem hoch korrelierte „Variablenbündel“ durch wenige übersichtliche „Hauptkomponenten“ ersetzt werden.

Das Ergebnis der Hauptkomponentenanalyse des ersten Datensatzes ist in Tab. 1 wiedergegeben. Auf die erste Hauptachse entfallen 29,9% der Gesamtvarianz, auf die zweite 15,0% usw.; die ersten sieben Hauptkomponenten repräsentieren immerhin 78,8% der Ausgangsvarianz. Dieses Ergebnis deutet darauf hin, dass in dem Datensatz zwar einzelne Variablen hoch korreliert sind, dass sich aber im Übrigen die in den Daten enthaltene Ausgangsinformation nur mit erheblichem Informationsverlust auf wenige Hauptachsen reduzieren lässt. Für die weitere Auswertung wird die Lösung mit 7 Hauptkomponenten gewählt. Für diese „Abschneidegrenze“ nach der siebten Hauptkomponente spricht die Verteilung der sog. Eigenwerte (zweite Spalte in Tab. 1): Während die siebte extrahierte (unrotierte) Hauptkomponente mit einem Eigenwert von 2,27 immerhin noch die Information von mehr als zwei Ausgangsvariablen repräsentiert, kommt die achte Hauptkomponente nur noch auf einen Eigenwert von 1,50 und repräsentiert damit nur noch den Informationsgehalt von eineinhalb Variablen.

Tab. 1: Hauptkomponentenanalyse des Datensatzes 1

Erklärte Gesamtvarianz									
Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	19,466	29,948	29,948	19,466	29,948	29,948	9,072	13,957	13,957
2	9,751	15,002	44,950	9,751	15,002	44,950	8,933	13,743	27,700
3	7,325	11,270	56,219	7,325	11,270	56,219	7,407	11,395	39,095
4	5,314	8,175	64,394	5,314	8,175	64,349	7,287	11,211	50,306
5	3,865	5,947	70,341	3,865	5,947	70,341	7,172	11,034	61,340
6	3,201	4,925	75,266	3,201	4,925	75,266	6,745	10,377	71,717
7	2,273	3,497	78,763	2,273	3,497	78,763	4,580	7,046	78,763
8	1,503	2,312	81,076						
9	1,402	2,157	83,233						
10	1,226	1,886	85,119						

Um die Hauptkomponenten inhaltlich zu interpretieren, erfolgt eine Achsenrotation nach dem Varimax-Kriterium. Dadurch wird erreicht, dass die Hauptkomponenten auf die hoch korrelierten „Variablenbündel“ ausgerichtet werden. Das Ergebnis ist im Anhang 2 wiedergegeben. Die sog. Faktorladungen können als Korrelationskoeffizienten zwischen den Hauptkomponenten und den Ausgangsvariablen gelesen werden. Die in fetter Schrift hervorgehobenen hohen Werte ( $r > 0,5$  und  $r < -0,5$ ) deuten darauf hin, dass die Information der betreffenden Variablen in mehr oder weniger hohem Maße durch die entsprechende Hauptkomponente repräsentiert wird. Beispielsweise sind sämtliche Variablen, die sich auf die Zu- und Fortzüge von Deutschen und Nichtdeutschen beziehen, mit der dritten Hauptkomponente hoch korreliert, sodass diese als „Mobilitätsfak-

tor“<sup>4</sup> interpretiert werden kann. In entsprechender Weise sind auch die anderen Hauptkomponenten inhaltlich zu interpretieren.

Die sieben rotierten Hauptkomponenten lassen sich aufgrund der Korrelationen mit den Ausgangsvariablen wie folgt interpretieren (die Prozentzahlen geben an, welchen Varianzanteil des ursprünglichen Datensatzes sie repräsentieren):

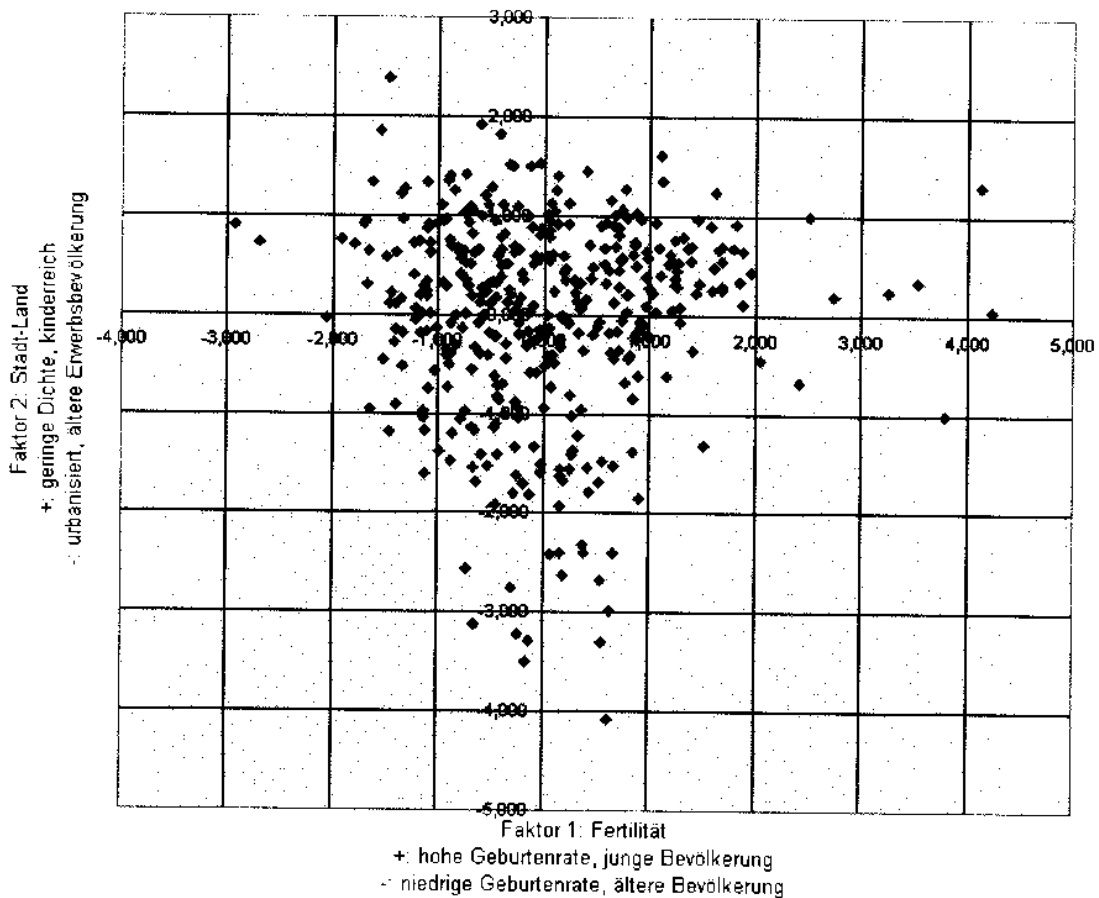
Faktor 1: Fertilität	14,0%
Faktor 2: Stadt-Land	13,7%
Faktor 3: Wanderungsintensität	1,4%
Faktor 4: Wachstum	1,2%
Faktor 5: Arbeitsmarktzentralität	11,0%
Faktor 6: Erwerbstätigkeit/Rentner	10,4%
Faktor 7: Altersspezifische Mortalität	7,0%

Die Ausprägungen der 393 Gemeinden auf den 7 Hauptkomponenten sind die sog. Faktorwerte (dimensionslos, z-standardisiert). Die Faktorwerte treten bei der Clusterung an die Stelle der Ausgangswerte der Variablen. Graphisch darstellen lassen sich die Faktorwerte nur im zwei- bzw. allenfalls dreidimensionalen Merkmalsraum. Als Beispiel zur Veranschaulichung sind in Abb. 1 die 393 Gemeinden im Merkmalsraum der ersten beiden Hauptkomponenten wiedergegeben. Während die als Abszisse dargestellte Hauptkomponente 1 als Fertilitätsfaktor gedeutet werden kann (positive Faktorwerte: Gemeinden mit hoher Fertilität und hohen Anteilen junger Erwachsener und Kinder; negative Faktorwerte: geringe Fertilität, hoher Anteil älterer Bevölkerung), bildet die als Ordinate dargestellte zweite Hauptkomponente eine Stadt-Land-Differenz ab (positive Faktorwerte: Gemeinden mit geringer Bevölkerungsdichte und junger Bevölkerung; negative Faktorwerte: Gemeinden mit hoher Bevölkerungsdichte und älterer Bevölkerung). Das Streudiagramm zeigt einen ziemlich amorphen Punkteschwarm ohne klare Clusterung, wobei selbstverständlich von diesem Bild nicht ohne Weiteres auf die Datenstruktur im mehrdimensionalen Merkmalsraum sämtlicher Hauptkomponenten geschlossen werden darf.

---

<sup>4</sup> Streng genommen darf nur bei einer Faktorenanalyse und nicht bei einer Hauptkomponentenanalyse von „Faktoren“ die Rede sein. Die beiden Analysen unterscheiden sich jedoch nur geringfügig (Schätzung der Kommunalitäten bei der Faktorenanalyse) und sind bei großen Datensätzen i. d. R. im Ergebnis weitgehend identisch. Da der Stichprobencharakter der Ausgangsdaten zumindest zweifelhaft ist, wird die Analyse hier mit einer deskriptiven Fragestellung durchgeführt. Aus diesem Grund wird nicht eine Faktorenanalyse, sondern eine Hauptkomponentenanalyse (Hauptachsentransformation) durchgeführt. Da es sich um einen verhältnismäßig großen Datensatz handelt, sind die Unterschiede jedoch zu vernachlässigen.

Abb. 1: Die 393 Gemeinden im Merkmalsraum der ersten beiden Hauptkomponenten (Analyse 1)



Als nächster Schritt folgt die eigentliche Clusteranalyse, also die Klassifikation der 393 nordrhein-westfälischen Gemeinden nach ihrer Ähnlichkeit. Als Merkmalsraum dient nun der um Interkorrelationen bereinigte sieben-dimensionale Raum der Hauptkomponenten. In die Clusteranalyse geht die Matrix der Faktorwerte ein, d.h. der Ausprägungen der 393 Gemeinden auf den sieben extrahierten varimax-rotierten Hauptachsen.

Als Ähnlichkeitsmaß dient die üblicherweise gewählte quadrierte euklidische Distanz. Von den zahlreichen Cluster-Algorithmen wird die nichthierarchische Cluster-Zentren-Methode (varianzorientiertes Zentroid-Verfahren, auch „K-Means“-Verfahren genannt) gewählt (Prozedur QUICK CLUSTER im Programmpaket SPSS) (vgl. Diehl u. Staufenbiel 2001, Brosius 2002). Im Unterschied zu den meist verwendeten hierarchischen Verfahren werden die Cluster hier nicht schrittweise durch den Zusammenschluss ähnlicher Objekte gebildet, sondern durch eine Partitionierung der Ausgangsmenge der Objekte nach einer vorgegebenen Clusteranzahl, und zwar unter der Bedingung, dass die Quadratsummen innerhalb der Cluster minimiert werden. Gegenüber einer hierarchischen Clusterung ist das Verfahren nicht nur aus rechentechnischen Gründen für große Datenmengen besser geeignet, sondern es erzeugt in der Regel auch etwas bessere Ergebnisse, d.h. homogenere Cluster.

Allerdings ist eine Vorentscheidung über die Zahl der zu bildenden Cluster erforderlich. Da für diese Entscheidung kein formales Kriterium existiert, wurde iterativ vorgegangen, indem probeweise Berechnungen mit 3, 4 ... 10 Clustern durchgeführt und die

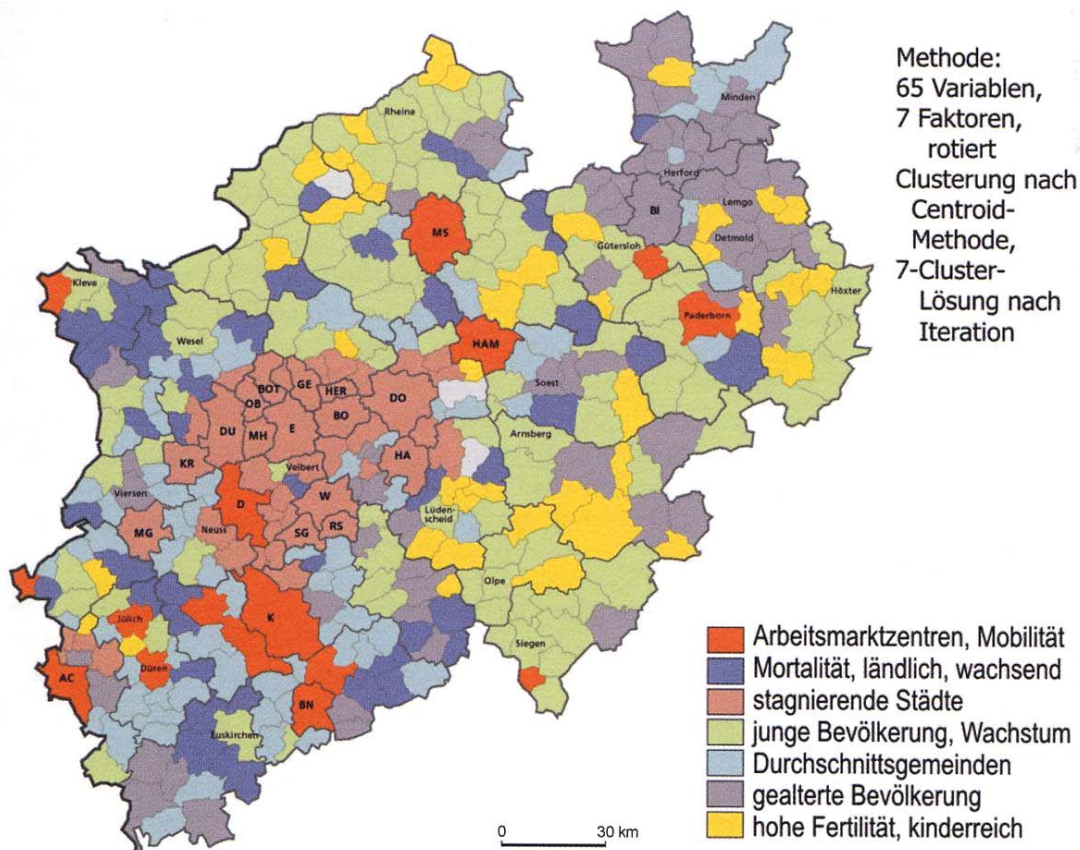
Clusterbildung miteinander verglichen wurde. Inhaltlich am besten interpretierbar erscheint die Lösung mit sieben Clustern.

Die sieben Cluster lassen sich durch die Position ihrer Mittelwerte (Zentroide) im Merkmalsraum der sieben Hauptkomponenten (Anhang 3) inhaltlich interpretieren. Anhand dessen lassen sich die sieben Cluster wie folgt charakterisieren:

- Cluster 1 (21 Gemeinden): Arbeitsmarktzentren, hohe Mobilität
- Cluster 2 (48 Gemeinden): altersspezifische Mortalität, ländlich, wachsend
- Cluster 3 (43 Gemeinden): stagnierende, insb. größere Städte
- Cluster 4 (103 Gemeinden): junge Bevölkerung, Wachstum
- Cluster 5 (74 Gemeinden): durchschnittliche Wohngemeinden
- Cluster 6 (64 Gemeinden): gealterte Bevölkerung
- Cluster 7 (40 Gemeinden): hohe Fertilität, kinderreich

Die Zugehörigkeit der einzelnen Gemeinden zu den Clustern geht aus Abb. 2 hervor:<sup>5</sup>

Abb. 2: Gemeindetypisierung 1: 65 Variablen, 7 Faktoren, 7 Cluster



Sind die Ergebnisse geeignet, um die eingangs gestellten Fragen zu lösen? Die Antwort fällt ambivalent aus, denn die ermittelte Typisierung der Gemeinden stellt mindestens so viele Fragen, wie sie beantworten kann. Die Charakterisierung der Cluster und die Zugehörigkeit der einzelnen Gemeinden zu diesen Typen erscheint teils plausibel, aber vor dem Hintergrund der Kenntnis der lokalen Verhältnisse teilweise auch irritierend. Für die Typenbildung wirksam sind insbesondere die Merkmale Dichte, Fertilität

<sup>5</sup> Der Verfasser dankt Herrn Dipl.-Ing. Harald Krähe für die Erstellung der beiden Karten.

und Altersstruktur. Allerdings werden daraus Cluster gebildet, die sich teilweise nur schwer eindeutig charakterisieren lassen.

Immerhin lassen sich hinsichtlich der regionalen Verteilung der Clusterzugehörigkeiten einige Regionalisierungstendenzen erkennen. So dominieren im Ruhrgebiet die Gemeinden des Typs 3, in den meisten ländlichen Regionen Westfalens die Gemeinden des Typs 4 und in Minden-Ravensberg und Lippe die Gemeinden des Typs 6. Hingegen sind im Rheinland kaum regionale Muster auszumachen; allenfalls fallen hier die häufigen Gemeinden des Typs 5 auf. Insgesamt gesehen lässt sich daraus allerdings keine klare Regionalisierung ableiten, schon gar nicht eine Regionalisierung, die sich zur Beschreibung der landesplanerischen Problemlandschaft eignen könnte. Allerdings bestätigt die räumliche Verteilung der Clusterzugehörigkeiten, dass sie nicht den räumlichen Mustern der traditionellen landesplanerischen Gebietskategorien entspricht. Diese scheinen sich vielmehr ausdifferenziert zu haben: die Ballungskernstädte in die Typen 1 und 3 und die ländlichen Gemeinden vor allem in die Typen 4, 6 und 7. Eine einheitliche Zuordnung der Ballungsrandzonengemeinden lässt sich nicht erkennen; allenfalls lässt sich eine gewisse Tendenz zu Typ 5 erkennen.

Dieser Analyseansatz ist erstens induktiv und zweitens beschreibend. Als Variablen gingen praktisch alle Merkmale in die Analyse ein, die die amtliche Bevölkerungsstatistik bereitstellt. In der methodologischen Literatur wird eine solche induktive Vorgehensweise eher skeptisch gesehen und allenfalls für wenig erforschte Untersuchungsgebiete akzeptiert, wenn elaborierte Theorien fehlen und die induktive Typisierung dazu genutzt werden kann, Hypothesen über die strukturbildenden Merkmale und Zusammenhänge zu bilden.

Aber lassen sich aufgrund der Ergebnisse des ersten Analyseansatzes tatsächlich weiterführende Hypothesen generieren? Die bisher referierten Ergebnisse geben eher Anlass zur Skepsis. Die ermittelte Hauptkomponenten- und Clusterstruktur erwies sich zwar insofern als robust, als alternative Analysen desselben Datensatzes mit leicht variierten methodischen Optionen (Ähnlichkeitsmaß, Klassifikationsalgorithmus, Anzahl der Cluster) zu sehr ähnlichen Ergebnissen führten, sodass die Hauptkomponenten- und Clusterstruktur nicht als reines Methodenartefakt gelten muss, sondern tatsächlich ein Strukturmuster der Realität abbildet. Aber die dennoch verbleibende Skepsis gründet sich auf zwei Argumente:

Erstens deutet die Struktur der Ausgangsdaten darauf hin, dass sie sich nicht ohne großen Informationsverlust auf einige wenige Hauptkomponenten und Cluster verdichten lassen. Anders formuliert: Die Struktur der Ausgangsdaten-Matrix zeigt einen weitgehend amorphen Punkteschwarm sowohl im Variablen- als auch im Faktorenraum, wie die Streudiagramme mit ausgewählten Variablen und Hauptachsen zeigen (vgl. Abb. 1). Sie ist deshalb nur wenig geeignet für eine Faktorisierung und Clusterung. Zwar führen Hauptkomponentenanalyse und Clusteranalyse stets zu irgendwelchen Ergebnissen, aber bei einer solchen Struktur der Ausgangsdaten ist die aufgezeigte Faktoren- und Clusterstruktur nur von begrenzter Aussagekraft.

Zweitens machen die ermittelten Hauptkomponenten- und Clusterstrukturen nur insoweit „Sinn“, als wir sie im Licht von Theorien über die räumliche Ausdifferenzierung von demographischen Prozessen „interpretieren“, also sinnhaft verständlich machen können. Bei dem gegenwärtigen sehr begrenzten Stand der Hypothesen- und Theoriebildung gelingt dies aber nur in sehr eingeschränktem Maße.

### 3 Zweiter, hypothesengeleiteter Analyseansatz

Um die Schwächen des ersten Analyseansatzes wenigstens teilweise auszugleichen, wurde in einem zweiten Analyseansatz eine hypothesengeleitete Auswahl von Variablen vorgenommen. Ein Interpretationsproblem des ersten Untersuchungsansatzes besteht darin, dass durch die Vielzahl der simultan berücksichtigten Merkmale und der darin enthaltenen heterogenen räumlichen Muster eine relativ große Zahl von Hauptachsen gebildet wird. Wenn diese dann in die Clusteranalyse eingehen, werden die Gemeinde-Cluster durch hoch komplexe Merkmalskombinationen gebildet. Es werden zwar viele Informationen verarbeitet, aber die Kehrseite dieses Vorgehens besteht darin, dass die Cluster inhaltlich nicht klar profiliert sind. U. a. liegt dies daran, dass die Information der Ausgangsvariablen auf Grund der komplexen Interkorrelationen auf mehrere Hauptkomponenten „verteilt“ wird.

Solche Defizite lassen sich zumindest teilweise beheben, wenn von vornherein nur solche Variablengruppen in die Analyse eingehen, von denen erwartet werden kann, dass sie auf Grund bekannter oder vermuteter Zusammenhänge zur Typenbildung beitragen.

Beispielsweise lassen sich Zusammenhänge erwarten:

- a) zwischen dem Bevölkerungswachstum und der Altersstruktur (Gemeinden, die in den letzten Jahren stark gewachsen sind, verfügen über eine relativ junge Bevölkerung und umgekehrt)
- b) zwischen Pendlerbilanz und Altersstruktur (Auspendlerwohngemeinden verfügen als Ergebnis von Suburbanisierungswanderungen über eine insgesamt jüngere Bevölkerung).

Aufgrund solcher Überlegungen wurde ein zweiter Datensatz gebildet, in den insgesamt 30 Variablen zur Bevölkerungsentwicklung, zur Pendlerbilanz und zur Altersstruktur eingehen (vgl. Anhang 1).

Die weiteren Untersuchungsschritte entsprechen dem ersten Analyseansatz. Zunächst wird eine Hauptkomponentenanalyse durchgeführt, um die Interkorrelationen zwischen den Variablen zu beseitigen und die Ausgangsinformationen durch die Reduktion auf wenige orthogonale Hauptachsen zu verdichten. Auf Grund der gezielten Variablenauswahl ist nicht erstaunlich, dass bereits mit drei Hauptkomponenten 76,4% der gesamten Ausgangsinformation reproduziert werden kann, wie aus Tab. 2 hervorgeht.

Tab. 2: Hauptkomponentenanalyse des Datensatzes 2

Erklärte Gesamtvarianz									
Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	13,713	45,709	45,709	13,713	45,709	45,709	11,062	36,875	36,875
2	5,631	18,770	64,479	5,631	18,770	64,479	6,164	20,547	57,422
3	3,583	11,944	76,423	3,583	11,944	76,423	5,700	19,001	76,423
4	1,738	5,795	82,218						
5	1,210	4,033	86,251						
6	0,975	3,251	89,502						
7	0,659	2,198	91,700						

Nach der Varimax-Rotation der drei Hauptkomponenten lassen sich diese anhand der Faktorladungen (Korrelationen zwischen Hauptkomponenten und Variablen) wie folgt inhaltlich interpretieren:

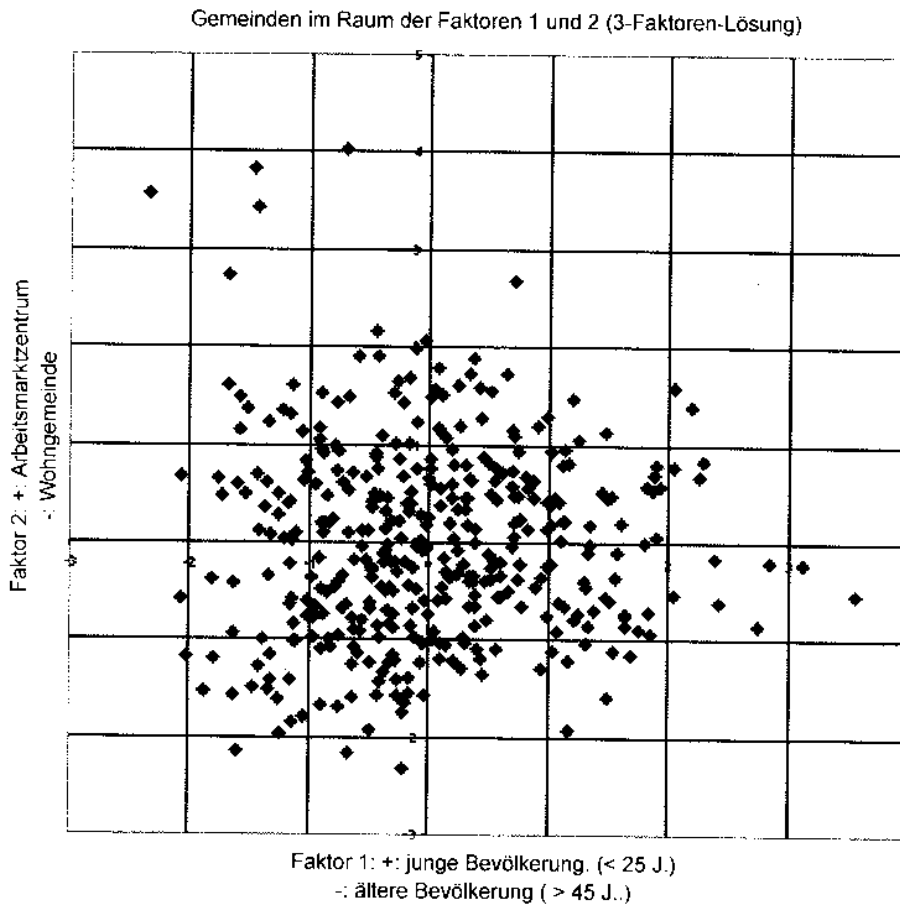
- Faktor 1: Junge (0–25 J.) / alte (>45 J.) Bevölkerung
- Faktor 2: Arbeitsmarktzentrum/Wohngemeinde
- Faktor 3: Bevölkerungswachstum und jüngere Erwerbsbevölkerung/hoher Rentneranteil

Damit ist die Faktorialstruktur einfacher und klarer als beim ersten Analyseansatz. Dies erleichtert die inhaltliche Interpretation der folgenden Clusteranalyse. Zu den drei Hauptkomponenten sei angemerkt, dass sie nicht einfach die Informationen der drei Variablengruppen Altersstruktur, Pendlerbilanz und Bevölkerungsveränderung enthalten, indem sie diese drei „Variablenbündel“ repräsentieren. Berücksichtigt werden auch Zusammenhänge zwischen diesen Variablengruppen. So repräsentiert der dritte Faktor beispielsweise den engen Zusammenhang zwischen hohen Anteilen der jüngeren Erwerbspersonen (25- bis 45-Jährige) einerseits und dem Bevölkerungswachstum (zwischen 1994 und 2002) andererseits. Der erste Faktor wird primär durch die Altersstruktur der Gemeinden gebildet, wobei interessant ist, dass dieser nur schwach mit den Variablen zur Bevölkerungsveränderung korreliert. Dasselbe gilt für den zweiten Faktor, der primär die Variablen zur Pendlerbilanz repräsentiert: Hohe Einpendlerintensität geht tendenziell mit einer negativen Bevölkerungsentwicklung einher.

Beispielhaft wird auch hier ein Streudiagramm wiedergegeben, das die Position der 393 Gemeinden im Merkmalsraum der ersten beiden Hauptachsen zeigt. Auch hier wird wieder deutlich, dass die Datenstruktur keine klare Clusterung aufweist.

Anschließend erfolgt die Clusterung der 393 Gemeinden im Merkmalsraum der drei rotierten Hauptkomponenten, wobei wiederum das nichthierarchische Zentroidverfahren gewählt wird. Vergleichende Berechnungen mit unterschiedlichen Clusterzahlen ergaben, dass die Lösung mit fünf Clustern die am besten interpretierbaren Ergebnisse bringt. Das bedeutet konkret, dass die Cluster einerseits klar inhaltlich profiliert, d. h. nicht zu groß und damit zu heterogen sind, aber andererseits auch nicht durch zu spezielle Merkmale gebildet werden, d. h. nicht zu klein sind. Diese Anforderung lässt sich als Abwägung zwischen der Forderung nach möglichst großen Distanzen zwischen den Cluster-Mittelwerten und möglichst geringen Streuungen innerhalb der Cluster formulieren.

Abb. 3: Die 393 Gemeinden im Merkmalsraum der ersten beiden Hauptkomponenten (Analyse 2)



Die fünf Cluster lassen sich aufgrund der Position der Cluster-Mittelwerte im Merkmalsraum der drei rotierten Hauptachsen (vgl. Anhang 5) wie folgt inhaltlich interpretieren:

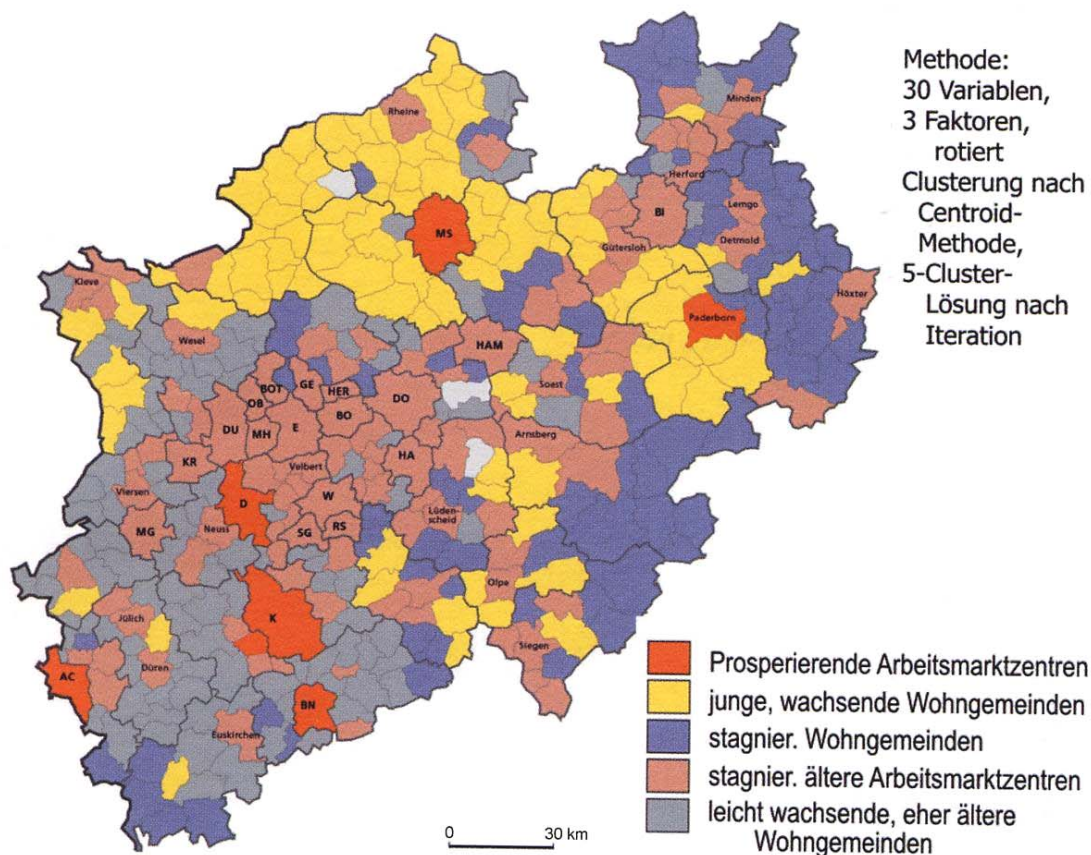
- Cluster 1: prosperierende Arbeitsmarktzentren
- Cluster 2: junge, wachsende Wohngemeinden
- Cluster 3: stagnierende Wohngemeinden
- Cluster 4: stagnierende ältere Arbeitsmarktzentren
- Cluster 5: leicht wachsende, eher ältere Wohngemeinden

Die Zugehörigkeit der 393 Gemeinden zu den einzelnen Clustern geht aus Abb. 4 hervor.

Die Ergebnisse des zweiten Analyseansatzes sind im Großen und Ganzen plausibel und erscheinen insgesamt gesehen besser geeignet, um zumindest in Teilen Antworten auf die eingangs gestellten Fragen zu geben.

Cluster 1 umfasst mit Münster, Paderborn, Düsseldorf, Köln, Bonn, Aachen und Hürth Städte, die nicht nur durch eine hohe Arbeitsmarktzentralität, sondern auch durch einen großen Anteil jüngerer Erwerbspersonen und eine stabile Bevölkerungsentwicklung charakterisiert werden. Es sind vor allem Städte, die durch einen hohen Dienstleistungsbesatz und große Hochschulen geprägt sind.

Abb. 4: Gemeindetypisierung 2: 30 Variablen, 3 Faktoren, 5 Cluster



Cluster 2 umfasst 91 Gemeinden, die schwerpunktmäßig im Münsterland, aber auch im Umland von Paderborn liegen. Es sind besonders kinderreiche ländliche Wohngemeinden mit überwiegend katholischer Bevölkerung, die auf Grund der Geborenenüberschüsse auch überwiegend eine positive Bevölkerungsentwicklung verzeichnen können.

Zu Cluster 3 gehören 77 Gemeinden, die sich vor allem in Ostwestfalen, im östlichen Sauerland, im Wittgensteiner Land und in der Nordeifel konzentrieren. Es handelt sich um meist ländliche Gemeinden mit stagnierender Bevölkerungsentwicklung und einem hohen Rentneranteil. Dies sind nicht nur Bade- und Kurorte, sondern auch ländliche Gemeinden, in denen bereits früher der Geburtenrückgang eingesetzt hat, sodass sie geringere Kinderanteile als die Gemeinden des Clusters 2 aufweisen.

Cluster 4 umfasst 108 Gemeinden, die sich im Ganzen als stagnierende Arbeitsmarktzentren charakterisieren lassen. Zu diesem Typ gehören fast alle Städte des Ruhrgebiets, aber auch viele Gemeinden der anderen traditionellen Industrieregionen wie im Bergischen Land, im Aachener, im Sauer- und Siegerland sowie in Minden-Ravensberg und Lippe. Die stagnierende bis negative Bevölkerungsentwicklung wird vor allem durch die anhaltende Entindustrialisierung der ökonomischen Basis dieser Gemeinden hervorgerufen.

Cluster 5 ist mit 110 Gemeinden das umfangreichste, es ist aber zugleich das inhaltlich am wenigsten profilierte. Am ehesten lassen sich die Gemeinden als leicht wachsende Wohngemeinden mit überdurchschnittlichen Anteilen älterer Bevölkerungsgruppen charakterisieren. Die Gemeinden dieses Typs sind vor allem im rheinischen Lan-

desteil verbreitet, wobei neben den Städten der Ballungsrandzonen auch viele ländliche Gemeinden diesem Cluster zugeordnet sind.

Lassen sich aus dieser Typisierung, die zugleich deutliche Tendenzen einer Regionalisierung aufweist, weiterführende Schlussfolgerungen ableiten? Zunächst lässt sich ein Befund des ersten Analyseansatzes bestätigen: Die tradierte siedlungsräumliche Einteilung in Ballungskerne, Ballungsrandzone und ländliche Räume ist nicht völlig verschwunden, wird aber stark ausdifferenziert und überlagert. Mindestens ebenso wichtig wie die Stadt-Land-Dimension sind die Unterschiede *innerhalb* der großen siedlungsräumlichen Kategorien: die Trennung zwischen den prosperierenden Dienstleistungszentren (Cluster 1) und den stagnierenden, häufig infolge der Entindustrialisierung strukturschwachen Arbeitsmarktzentren andererseits (Cluster 4), aber auch die Unterschiede zwischen den verschiedenen Typen von Wohngemeinden in den ländlichen Räumen und Ballungsrandzonen.

Die Typenzugehörigkeit der Gemeinden zeigt im ganzen Land deutliche regionale Häufungen, d. h. eine deutliche Tendenz zur Regionalisierung. Mit der gebotenen Vorsicht lassen sich daraus mindestens zwei Schlussfolgerungen für eine landesplanerische Regionalisierung ableiten.

Erstens hat die Analyse ergeben, dass der demographische Wandel durchaus eine regionale Dimension hat. Damit ist gemeint, dass Gemeinden gleicher Clusterzugehörigkeit häufig aneinander grenzen und somit regionale Muster bilden. Das ist keineswegs selbstverständlich, denn es wäre auch denkbar, dass bei einer kartographischen Darstellung der Clusterzugehörigkeiten ein bunter Flickenteppich entstünde, der demographische Wandel also gemeindespezifisch und nicht regional ausgeprägt wäre. Abgesehen vom ersten Cluster ist dies offensichtlich nicht der Fall. Die regionale Lage und/oder die regionalen Struktur- und Entwicklungsbedingungen sind offenbar nicht unwichtig. Daraus lässt sich zumindest schlussfolgern, dass ein Ansatz einer landesplanerischen Regionalisierung nicht von vornherein an der Heterogenität lokaler Verhältnisse scheitern müsste.

Zweitens lässt sich aus den Ergebnissen folgern, dass die tradierten siedlungsräumlichen Kategorien wie Ballungskerne, -randzonen und ländliche Räume weniger denn je als homogene Gebietskategorien gelten können. Dies gilt sowohl für die großen Städte als auch in besonderem Maße für die ländlichen Zonen, deren Heterogenität nach Struktur- und Entwicklungsmerkmalen auch hier wieder dokumentiert wird.

#### **4 Methodenkritische Schlussbemerkungen**

Die den vorgestellten Ergebnissen zu Grunde liegende Clusteranalyse mit vorgeschalteter Hauptkomponentenanalyse ist im Kern eine beschreibende Klassifikation, nicht mehr und nicht weniger. Damit lassen sich weder Hypothesen über Ursache-Wirkungs-Zusammenhänge überprüfen noch etwaige regionale Muster mit Allzweckcharakter aufdecken. Auch die Bezeichnung „Hauptkomponenten-“ bzw. „Faktorenanalyse“ sollte nicht zu der Annahme verleiten, durch eine Hauptachsentransformation ließen sich gleichsam automatisch Kausalzusammenhänge identifizieren. Was die Clusteranalyse leisten kann, ist eine Typisierung (Klassifikation) von Objekten (hier: Gemeinden) nach der Ähnlichkeit bestimmter Merkmalsausprägungen (Variablen).

Dass die Ergebnisse der Clusteranalyse vor allem von der Auswahl der Variablen abhängen, ist im Grunde trivial und muss dennoch ausdrücklich betont werden. Wie der erste Analyseansatz gezeigt hat, führt eine weitgehend wahllose Verwendung von Variablen nach dem Prinzip, möglichst alles in die Analyse einzugeben, was die amtliche Bevölkerungsstatistik hergibt, zu wenig befriedigenden Ergebnissen. Zielführender er-

scheinen dagegen Typisierungen mit hypothesengeleiteten Variablenauswahlen, wie im zweiten Analyseansatz versucht wurde. Aber so oder so ist die Erwartung ungerechtfertigt, dass damit Typisierungen oder gar Regionalisierungen mit Allzweckcharakter aufgedeckt werden könnten.

Schließlich sei noch einmal darauf hingewiesen, dass die Clusteranalyse sowohl bei Verfahrensentscheidungen als auch bei der Interpretation der Hauptkomponenten und Cluster unvermeidlich ein subjektives Moment enthält. Dies gilt in besonderem Maße dann, wenn wie in unserem Fall die Struktur der Ausgangsdaten-Matrix eher einen amorphen Punkteschwarm als eine klare Clusterstruktur aufweist. In solchen Fällen ist besondere Vorsicht gegenüber zu weit reichenden Interpretationen der Ergebnisse geboten.

Diese methodenkritischen Bemerkungen sollen nicht zu der Schlussfolgerung führen, Untersuchungen wie diese würden nur Methoden-Artefakte erzeugen und seien deshalb wertlos. Auch wenn die eingangs gestellten Fragen nicht in vollem Umfang befriedigend beantwortet werden können, sollte dieser Beitrag eher ein Ansporn für weiterführende Untersuchungen über die räumliche Dimension des demographischen Wandels sein.

## Literatur

- Backhaus, K. u. a. (2000): *Multivariate Analysemethoden*. 9. Aufl. Berlin.
- Bähr, J.; Gans, P. (2003): Regionale Typen ähnlicher Migrationsdynamik in Deutschland. In: *Zs. f. Bevölkerungswiss.* 28, S. 233–242.
- Bahrenberg, G. u. a. (1992): *Statistische Methoden in der Geographie*. Bd. 2: *Multivariate Statistik*. 2. Aufl. Stuttgart.
- Blotvogel, H. H. (1979): Methodische Probleme der Erfassung städtischer Funktionen und funktionaler Städtetypen anhand quantitativer Analysen der Berufsstatistik 1907. In: *Voraussetzungen und Methoden geschichtlicher Städteforschung*. Hg. v. W. Ehbrecht. Köln. S. 217–269. = *Städteforschung*, Reihe A, Bd. 7.
- Bortz, J. (2004): *Statistik für Human- und Sozialwissenschaftler*. 6. Aufl. Berlin.
- Brosius, F. (2002): *SPSS 11*. Bonn.
- de Lange, N. (1980): Städtetypisierung in Nordrhein-Westfalen im raum-zeitlichen Vergleich 1961 und 1970 mit Hilfe multivariater Methoden. Paderborn. = *Münstersche Geogr. Arb.* 8.
- Diehl, J. M.; Staufienbiel, Th. (2001): *Statistik mit SPSS, Version 10.0*. Eschborn.
- Eckes, Th.; Rossbach, H. (1980): *Clusteranalysen*. Stuttgart.
- Eckey, H.-F.; Klemmer, P. (1975): Die Clusteranalyse und ihre Anwendungsmöglichkeiten im Rahmen der Regionalforschung. In: *Methoden der empirischen Regionalforschung* (2. Teil). Hannover, S. 145–155. = *Forschungs- und Sitzungsberichte der ARL* Bd. 105.
- Forst, H. Th. (1974): Zur Klassifizierung von Städten nach wirtschafts- und sozialstatistischen Strukturmerkmalen. Würzburg. = *Arb. z. angew. Statistik* 17.
- Landtag NRW (Hg.) (2004): *Zukunft der Städte*. Bericht der Enquetekommission des Landtags von Nordrhein-Westfalen. Düsseldorf.
- Möllers, H. (1977): *Infrastrukturausstattung und Entwicklung von Städten*. Münster. = *Beitr. z. Siedl.- u. Wohnungswesen u. z. Raumplanung* 42.
- Schlömer, C. (2004): Bestimmungsfaktoren der künftigen räumlich-demographischen Entwicklung in Deutschland. In: M. T. W. Rosenfeld; C. Schlömer (Hg.): *Räumliche Konsequenzen des demographischen Wandels* Teil 4. Hannover. S. 1–36. = *Arbeitsmaterial der ARL* Nr. 312.
- Schlosser, O. (1976): *Einführung in die sozialwissenschaftliche Zusammenhangsanalyse*. Reinbek.
- Sedlacek, P. (Hg.) (1978): *Regionalisierungsverfahren*. Darmstadt. = *Wege d. Forsch.* 195.
- Sitterberg, G. (1977): *Multivariate Analyse der Struktur und Entwicklung von Städten*. Münster. = *Beitr. z. Siedl.- u. Wohnungswesen u. z. Raumplanung* 41.
- Sodeur, W. (1974): *Empirische Verfahren zur Klassifikation*. Stuttgart. = *Studienskripten zur Soziologie*.
- Steinhausen, D.; Langer, K. (1977): *Clusteranalyse*. Berlin.

## Anhang 1: Liste der 65 bzw. 30 Variablen

			1. Datensatz	2. Datensatz
RVE9496	Entwicklung der Bevölkerung 1994-96	%	x	x
RVE9698	Entwicklung der Bevölkerung 1996-98	%	x	x
RVE9800	Entwicklung der Bevölkerung 1998-2000	%	x	x
RVE0002	Entwicklung der Bevölkerung 2000-02	%	x	x
EQ96	Bevölkerungsdichte (Ew. pro qkm) 1996		x	-
EQ98	Bevölkerungsdichte (Ew. pro qkm) 1998		x	-
EQ00	Bevölkerungsdichte (Ew. pro qkm) 2000		x	-
EQ02	Bevölkerungsdichte (Ew. pro qkm) 2002		x	-
GEB01	Geborene pro 1000 Ew. 2001		x	-
GEB02	Geborene pro 1000 Ew. 2002		x	-
GEB03	Geborene pro 1000 Ew. 2003		x	-
GBE15451	Geborene pro 1000 Frauen zwischen 15 und 45 Jahren 2001		x	-
GBE15452	Geborene pro 1000 Frauen zwischen 15 und 45 Jahren 2002		x	-
GBE15453	Geborene pro 1000 Frauen zwischen 15 und 45 Jahren 2003		x	-
GEST01	Gestorbene pro 1000 Ew. 2001		x	-
GEST02	Gestorbene pro 1000 Ew. 2002		x	-
GEST03	Gestorbene pro 1000 Ew. 2003		x	-
GSE71801	Gestorbene pro 1000 Ew. im Alter von 71-80 Jahren 2001		x	-
GSE71802	Gestorbene pro 1000 Ew. im Alter von 71-80 Jahren 2002		x	-
GSE71803	Gestorbene pro 1000 Ew. im Alter von 71-80 Jahren 2003		x	-
GSE81901	Gestorbene pro 1000 Ew. im Alter von 81-90 Jahre und älter 2001		x	-
GSE81902	Gestorbene pro 1000 Ew. im Alter von 81-90 Jahre und älter 2002		x	-
GSE81903	Gestorbene pro 1000 Ew. im Alter von 81-90 Jahre und älter 2003		x	-
ZUZE01	Zuzüge pro 1000 Ew. 2001		x	-
ZUZE02	Zuzüge pro 1000 Ew. 2002		x	-
ZUZE03	Zuzüge pro 1000 Ew. 2003		x	-
FOZE01	Fortzüge pro 1000 Ew. 2001		x	-
FOZE02	Fortzüge pro 1000 Ew. 2002		x	-
FOZE03	Fortzüge pro 1000 Ew. 2003		x	-
ZUZNE01	Zuzüge der Nichtdeutschen pro 1000 Ew. 2001		x	-
ZUZNE02	Zuzüge der Nichtdeutschen pro 1000 Ew. 2002		x	-
ZUZNE03	Zuzüge der Nichtdeutschen pro 1000 Ew. 2003		x	-
FOZNE01	Fortzüge der Nichtdeutschen pro 1000 Ew. 2001		x	-
FOZNE02	Fortzüge der Nichtdeutschen pro 1000 Ew. 2002		x	-
FOZNE03	Fortzüge der Nichtdeutschen pro 1000 Ew. 2003		x	-
Entw9202	Bevölkerungsentwicklung 1992-2002 in%		x	-
Entw0102	Bevölkerungsentwicklung 2001-2002 pro 1000		x	-
Mver	Mittelwertveränderung		x	-
saldsvpf	Pendlersaldo pro 1000 sozialversicherungspflichtige Beschäftigte		x	-
PjeTB98	Pendlersaldo je 1.000 Besch. am Wohnort 1998		x	x
PjeTB99	Pendlersaldo je 1.000 Besch. am Wohnort 1999		x	x
PjeTB00	Pendlersaldo je 1.000 Besch. am Wohnort 2000		x	x
PjeTB01	Pendlersaldo je 1.000 Besch. am Wohnort 2001		x	x
PjeTB02	Pendlersaldo je 1.000 Besch. am Wohnort 2002		x	x
a6_92	Anteil der unter 6-Jährigen 1992	%	x	x
a616_92	Anteil der 6-16-Jährigen 1992	%	x	x
a1618_92	Anteil der 16-18-Jährigen 1992	%	x	x
a1825_92	Anteil der 18-25-Jährigen 1992	%	x	x
a2545_92	Anteil der 25-45-Jährigen 1992	%	x	x
a4565_92	Anteil der 45-65-Jährigen 1992	%	x	x
a_65um_92	Anteil der 65-Jährigen und Ältere 1992	%	x	x
a6_96	Anteil der unter 6-Jährigen 1996	%	x	x
a616_96	Anteil der 6-16-Jährigen 1996	%	x	x
a1618_96	Anteil der 16-18-Jährigen 1996	%	x	x
a1825_96	Anteil der 18-25-Jährigen 1996	%	x	x
a2545_96	Anteil der 25-45-Jährigen 1996	%	x	x
a4565_96	Anteil der 45-65-Jährigen 1996	%	x	x
a65um_96	Anteil der 65-Jährigen und Ältere 1996	%	x	x
a6_02	Anteil der unter 6-Jährigen 2002	%	x	x
a616_02	Anteil der 6-16-Jährigen 2002	%	x	x
a1618_02	Anteil der 16-18-Jährigen 2002	%	x	x
a1825_02	Anteil der 18-25-Jährigen 2002	%	x	x
a2545_02	Anteil der 25-45-Jährigen 2002	%	x	x
a4565_02	Anteil der 45-65-Jährigen 2002	%	x	x
a65um_02	Anteil der 65-Jährigen und Ältere 2002	%	x	x

## Anhang 2: Faktorladungen der 7 rotierten Hauptkomponenten (Analyse 1)

Rotierte Komponentenmatrix

	Komponente						
	1	2	3	4	5	6	7
RVE9496				0,63			
RVE9698				0,70			
RVE9800				0,73			
RVE0002				0,83			
EQ96		-0,86					
EQ98		-0,86					
EQ00		-0,86					
EQ02		-0,86					
GEB01	0,83						
GEB02	0,83						
GEB03	0,80						
GBE15451	0,83						
GBE15452	0,83						
GBE15453	0,78						
GEST01					-0,59	0,63	
GEST02					-0,57	0,62	
GEST03					-0,55	0,64	
GSE71801						0,82	
GSE71802						0,80	
GSE71803						0,83	
GSE81901						0,63	
GSE81902						0,62	
GSE81903						0,66	
ZUZE01			0,66	0,59			
ZUZE02			0,67	0,61			
ZUZE03			0,65	0,54			
FOZE01			0,78				
FOZE02			0,80				
FOZE03			0,77				
ZUZNE01			0,82				
ZUZNE02			0,79				
ZUZNE03						0,76	
FOZNE01						0,82	
FOZNE02						0,82	
FOZNE03						0,80	
entw9202							0,81
ent0102							0,79
mver							0,88
saldsvf							0,82
PjeTB98							0,92
PjeTB99							0,93
PjeTB00							0,93
PjeTB01							0,93
PjeTB02							0,93
a6_92		0,51	0,59				
a616_92			0,70				
a1618_92		0,51	0,54				
a1825_92		0,58					
a2545_92							0,77
a4565_92		-0,50	-0,61				
a_65um_92							-0,80
a6_96		0,58	0,54				
a616_96			0,72				
a1618_96			0,63				
a1825_96		0,61					
a2545_96							0,75
a4565_96		-0,56	-0,56				
a65um_96							
a6_02		0,77					
a616_02			0,64				
a1618_02			0,71				
a1825_02							
a2545_02							0,68
a4565_02		-0,68					
a65um_02							

## Anhang 3: Clusterzentren im Hauptkomponentenraum (Analyse 1)

Clusterzentren im Raum der 7 Hauptkomponenten

Komponente	Cluster						
	1	2	3	4	5	6	7
1	-0,69	-0,19	-0,24	-0,08	-0,63	0,23	1,50
2	-0,78	0,36	-1,80	0,65	-0,12	0,05	0,38
3	2,28	0,03	-0,55	-0,42	-0,02	0,15	0,25
4	0,22	0,83	-0,86	-0,27	0,56	0,22	-0,89
5	1,24	-0,23	0,10	0,58	-1,04	0,45	-0,77
6	0,97	0,15	0,21	0,49	0,05	-1,35	-0,10
7	-0,15	1,60	0,27	-0,27	-0,68	-0,22	0,17

#### Anhang 4: Faktorladungen der drei rotierten Hauptkomponenten (Analyse 2)

	Komponente		
	1	2	3
RVE9496	0,349	-0,340	0,499
RVE9698	0,218	-0,449	<b>0,558</b>
RVE9800	0,203	-0,351	<b>0,589</b>
RVE0002	0,074	-0,193	<b>0,650</b>
PjeTB98	-0,055	<b>0,957</b>	-0,138
PjeTB99	-0,048	<b>0,965</b>	-0,123
PjeTB00	-0,055	<b>0,967</b>	-0,112
PjeTB01	-0,053	<b>0,965</b>	-0,107
PjeTB02	-0,055	<b>0,961</b>	-0,113
a6_92	<b>0,834</b>	-0,225	0,254
a616_92	<b>0,871</b>	-0,318	0,187
a1618_92	<b>0,833</b>	-0,138	0,026
a1825_92	<b>0,738</b>	0,146	0,175
a2545_92	-0,080	0,016	<b>0,854</b>
a4565_92	<b>-0,853</b>	-0,036	-0,210
a_65um_92	-0,405	0,273	<b>-0,656</b>
a6_96	<b>0,815</b>	-0,249	0,289
a616_96	<b>0,867</b>	-0,339	0,171
a1618_96	<b>0,850</b>	-0,137	0,006
a1825_96	<b>0,813</b>	0,240	0,051
a2545_96	0,056	0,065	<b>0,913</b>
a4565_96	<b>-0,854</b>	-0,047	-0,188
a65um_96	-0,478	0,255	<b>-0,695</b>
a6_02	<b>0,769</b>	-0,161	0,340
a616_02	<b>0,816</b>	-0,388	0,200
a1618_02	<b>0,839</b>	-0,242	0,047
a1825_02	<b>0,743</b>	0,352	0,042
a2545_02	0,124	0,180	<b>0,873</b>
a4565_02	<b>-0,807</b>	-0,134	-0,152
a65um_02	<b>-0,547</b>	0,215	<b>-0,742</b>

Extraktionsmethode: Hauptkomponentenanalyse

Rotationsmethode: Varimax mit Kaiser-Normalisierung

Die Rotation ist in 5 Iterationen konvergiert.

Faktorladungen >0,5 und <-0,5 sind durch fette Schrift hervorgehoben.

#### Anhang 5: Clusterzentren im Hauptkomponentenraum (Analyse 2)

##### Clusterzentren im Raum der 3 Hauptkomponenten

		Cluster				
		1	2	3	4	5
Komponente	1	-1,22	1,16	0,40	-0,55	-0,62
	2	3,11	0,06	-0,29	0,87	-0,90
	3	2,27	0,66	-1,21	-0,38	0,53