

Energiecluster deutscher Städte – Clusteranalyse deutscher Städte anhand sozio-energetischer Indikatoren

Wilhelm Wall, Hermann-Josef Wagner, Marco K. Koch

Ruhr-Universität Bochum, Lehrstuhl Energiesysteme und Energiewirtschaft,
Universitätsstraße 150, 44801 Bochum, +49 (0)234 32-25986, wall@lee.rub.de,
<http://www.lee.rub.de/>

Kurzfassung: In diesem Beitrag wird eine sozio-energetische Clusterung aller deutschen kreisfreien Städte erarbeitet. Die Städteclusterung kann eine Hilfestellung geben, um Kommunen zu identifizieren, in denen die Ausgangslage sowie die Herausforderungen, in Bezug auf die Klimaschutzziele, ähnlich sind. Die aus drei Schritten bestehende Analyse zeigte, dass die 107 betrachteten Städte 12 sinnvollen Clustern zugeteilt werden können.

Keywords: Energieindikatoren, Clusteranalyse, Städte, Energiecluster

1 Einleitung und Zielsetzung

Energie und deren Verwendung stellt in unserem heutigen Leben ein Grundbedürfnis dar, das derzeit größtenteils durch fossile Energieträger gedeckt wird. Bei ihrer Verbrennung werden große Mengen des treibhausrelevanten Klimagases Kohlenstoffdioxid emittieren. Um dem daraus resultierenden anthropogenen Treibhauseffekt abzubremsen wurden in Deutschland Energieeinspar- und Klimaschutzziele verabschiedet, die eine Reduktion von 80-95 % der Treibhausgase, gegenüber dem Bezugsjahr 1990, vorsehen [BMU2011]. In diesem Zusammenhang werden auch Städte und Kommunen mit in das Wirken einbezogen und gefördert. Ein Beispiel dafür ist der Wettbewerb 'Energieeffiziente Stadt' in dem fünf deutsche Städte bei der Umsetzung innovativer Maßnahmen zur Steigerung der Energieeffizienz gefördert werden [WET2015].

Die Vielzahl der Maßnahmen deckt eine Bandbreite ab und sie sind oft durch den innovativen Ansatz der umsetzenden Stadt sehr „stadtspezifisch“. Die Clusteranalyse kann an dieser Stelle ansetzen, indem Städtecluster ermittelt werden, die sich auf Basis sozio-energetischer Kennwerte ähneln. Strukturell ähnliche Kommunen könnten hinzugezogen werden, um erprobte Maßnahmen in weiteren Städten anzuwenden, die ähnlichen Herausforderungen begegnen müssen und identische Rahmenbedingungen vorweisen. Das Ziel ist somit der Versuch eine im energetischen Bereich bisher kaum verwendete Methode auf diesen anzuwenden sowie erste Ergebnisse zu erhalten, die bei der Entscheidungsfindung im kommunalen Umfeld unterstützend wirken sowie Energieeffizienzmaßnahmen ausgewählter Städte Kommunen übergreifend nutzen zu können.

In diesem Beitrag werden erste Ergebnisse der durchgeführten Clusteranalyse deutscher kreisfreier Städte diskutiert. Zu Beginn erfolgt eine kurze Darstellung der Grundlagen, des Untersuchungsrahmens sowie der Methodik der durchgeführten Clusteranalyse. Daran anschließend erfolgt die Darstellung der durchgeführten Clusteranalyse, die abschließend diskutiert und zusammengefasst wird.

2 Grundlagen: Untersuchungsrahmen und Methodik

Dieses Kapitel befasst sich mit den grundlegenden Definitionen, Annahmen sowie Erläuterungen, die vor der Durchführung der Clusteranalyse festgelegt und diskutiert werden. Neben der Wahl des Untersuchungsrahmens wird die für die Durchführung der Clusteranalyse notwendige Methodik festgelegt sowie kurz erläutert.

2.1 Untersuchungsrahmen

2.1.1 Betrachtete Sektoren

Der städtische Energieverbrauch lässt sich auf vier Wesentliche Energieverbrauchssektoren aufteilen. Die vier nachfolgend aufgelisteten Sektoren sind in den Städten verschieden stark ausgeprägt:

- Private Haushalte,
- Verkehr,
- Industrie
- Gewerbe, Handel und Dienstleistungen (GHD).

So sind z. B. in größeren Städten vermehrt die Sektoren *Industrie* und *Gewerbe Handel und Dienstleistungen (GHD)* stärker ausgebildet. Zeitgleich lassen sich die Grenzen zwischen den Sektoren nicht uneindeutig ziehen, z. B. die Zuordnung des Treibstoffverbrauchs für private Pkw zu dem Sektor *private Haushalte* oder zu dem Sektor *Verkehr*, wodurch die separate Betrachtung eines einzelnen Sektors nur teilweise erfolgen kann und soll. [Eic2014]

Das Leben der Bewohner einer Stadt organisiert sich in *privaten Haushalten*, womit verschiedene Haushaltsformen (Einpersonenhaushalte, Familien, etc.) einhergehen und Einfluss auf die Stadtform und den Energieverbrauch nehmen. Dementsprechend relevant ist dieser Sektor für eine Stadt. Der *Verkehr* als zweiter energieintensiver Sektor ist, besonders im industrialisierten Deutschland, für das tägliche Leben sowie das Wirtschaften unumgänglich. Dieser Sektor kann in zwei grobe Bereiche geteilt werden, den Personen- und den Warenverkehr. Darüber hinaus werden diese beiden Bereiche in Kategorien eingeteilt, die den Straßen-, Schienen- und Luftverkehr sowie den Verkehr auf Wasserwegen umfassen. Letztere beiden sind für die kommunale Betrachtung eher irrelevant, da diese für die Betrachtung von Ländern belastbare Ergebnisse liefern und bei Städten große Unsicherheiten bewirken, da die Verteilung von Wasserwegen als auch Flughäfen in Deutschland inhomogen ist. [DES2013] In Bezug auf Städte wird unter Personenverkehr im Wesentlichen die innerstädtische Mobilität der Stadtbewohner verstanden. Der Güterverkehr ist trotz seiner Energieintensität bei der indikatorbasierten Beschreibung der Städte nur schwer und ungenau einer Stadt zuzuordnen, da der Großteil des Transportaufkommens über Stadtgrenzen hinweg erfolgt. Die *Industrie* stellt den dritten Verbrauchssektor dar, zu dem das produzierende und das verarbeitende Gewerbe gezählt werden [DES2013]. Die Produktion als auch die Verarbeitung erfordern einen höheren Energieeinsatz (ca. 30 % des deutschen Energieverbrauchs [BUN2016]) als der betrachtete vierte Sektor GHD. Im *GHD-Sektor* erfolgt anders als bei der *Industrie* kein materieller Output als Produkt, jedoch ist für beide Sektoren die Datenlage stark limitiert, da in diesen Sektoren das Interesse der Datengeheimhaltung aufgrund der Wahrung von Unternehmensinteressen geschützt wird. Zusätzlich werden Daten für

diese beiden Sektoren zumeist auf Landes- oder Bundesebene veröffentlicht, wodurch der Umfang der für diese Sektoren erarbeiteten Indikatoren geringer ist als für die beiden erstgenannten Sektoren.

2.1.2 Betrachtete Städte

Die äußerst geläufige Bezeichnung „Stadt“ ist jedem bekannt, dennoch gibt es keine einheitliche Definition für diese strukturelle und organisatorische Form des Zusammenlebens [EU2011]. Es gibt verschiedene Städte, z. B. kreisfreie Städte, Kreisstädte sowie kreisangehörige Städte. Die Unterscheidung dieser Stadttypen erfolgt z. T. historisch (in der Vergangenheit verliehenes Stadtrecht – de jure) sowie Städte, die über den zeitlichen Verlauf um ein regionales Zentrum gewachsen sind und sich mit der Zeit zu einem belebten Raum mit erhöhter Bevölkerungsdichte sowie einer gewissen Funktionalität entwickelt haben, die von Städten per Gesetz kaum unterschieden werden können. Diese Form stellt somit faktisch ebenfalls eine Stadt dar (de facto). Eine weitere Unterteilung der Städte wird wiederum nach der Größe vorgenommen. So wird zwischen Millionenstädten, Großstädten, Mittelstädten sowie kleinen Städten differenziert (s. Tabelle 2-1) [EU2011].

Tabelle 2-1: Definition verschiedener Städte und deren Verteilung im Untersuchungsrahmen [EU2011], [WAL2016]

| Stadt | Einwohnerzahl | Anzahl der Städte im Untersuchungsrahmen |
|----------------|------------------|--|
| Millionenstadt | >1.000.000 | 4 |
| Großstadt | >100.000 | 62 |
| Mittelstadt | 50.000 - 100.000 | 23 |
| Kleinstadt | <50.000 | 18 |

Für diese Untersuchung wurde der Fokus auf die kreisfreien Städte gelegt. Derzeit gibt es in Deutschland 107 kreisfreie Städte, wovon keine Stadt exakt der anderen gleicht. Das zeigt sich ebenfalls durch verschiedenen Merkmalsausprägungen der Städte, die die Indikatoren anzeigen auf die im nachfolgenden Abschnitt eingegangen wird.

2.1.3 Ermittelte Indikatoren, Datenverfügbarkeit und Datenlage

Indikatoren sind Kennwerte, die es erlauben Sachverhalte oder Eigenschaften von Objekten (Städten) quantitativ zu beschreiben. Die Indikatoren bestehen häufig aus Verhältniszahlen, wodurch eine Vergleichbarkeit der untersuchten Städte sichergestellt werden kann. Um die Städte hinreichend genau zu beschreiben wurden die Indikatoren, separat für jeden Verbrauchssektor ermittelt. Für die Durchführung der Clusterung deutscher Städte wurde der Ansatz der freien Verfügbarkeit der Daten verfolgt, die großteils auf statistischen Daten und Fortschreibungen aus den Datenbanken der statistischen Ämter der Länder sowie des Bundes, Zensusdaten, Daten des Kraftfahrtbundesamtes, Daten von Verbänden sowie Unternehmensdaten basieren. Der Vorteil der Verwendung dieser Daten ist neben den nicht anfallenden Kosten oft die Aktualität, da viele Daten in regelmäßigen Abständen flächendeckend erhoben und gleich aufbereitet werden. Aufbauend auf diesen Daten wurden zur Beschreibung der Städte 41 Indikatoren ermittelt (s. Tabelle 2-2). Die ermittelten Indikatoren sind in nachfolgender Tabelle aufgeführt. Einige Indikatoren zeigen ähnliche Aussagen an. Diese wurden dennoch verwendet, da diese in verschiedenen Untersuchungen relevante Aussagen erlauben. Eine detailliertere Beschreibung sowie die genaue Erläuterung der verwendeten Daten und Quellen der einzelnen Indikatoren erfolgt in [WAL2016].

Tabelle 2-2: Übersicht der für die betrachteten Verbrauchssektoren erhobenen Indikatoren

| Haushalte | Verkehr | Industrie | GHD |
|--|---|--|--|
| Bevölkerungsentw. letzte 11 Jahre [%] | Motorisierungsgrad 2014 | Beschäftigtenanteil Industrie 2011 [%] | Beschäftigtenant. GHD 2011 [%] |
| Wohnfläche pro Person [qm] | Pkw pro 1000 EW 2014 | Energieproduktivität Industrie [EUR/GJ] | Energieproduktivität GHD [EUR/GJ] |
| Anteil Wohnungen in 1-2 Fam. Häusern [%] | Dieselfahrzeuge pro Tsd. EW 2014 | Energieintensität Industrie [MJ/EUR] | Energieintensität GHD [MJ/EUR] |
| Durchschnittliche HH Größe [Pers/HH] | Benzinfahrzeuge pro Tsd. EW 2014 | Produktivitätsniveau Industrie [EUR/GJ] | Produktivitätsniveau GHD [EUR/GJ] |
| HH-Dichte [HH/km ²] | Elektrofahrzeuge pro Tsd. EW 2014 | Spez. Energieverbrauch Industrie [MJ/EUR] | Spez. Energieverbrauch GHD [MJ/EUR] |
| Verfügbares Einkommen [€] | Auspenderanteil an soz. Versicherungspflichtigen am Wohnort [%] | Industrie Umsatzsteuer an Gesamtumsatzsteuer [%] | GHD Umsatzsteuer an Gesamtumsatzsteuer [%] |
| Anteil der Siedlungs- und Verkehrsfläche an Gesamtfläche [%] | Anteil zwischen 25 und 65 Jähriger [%] | Entwicklung der Beschäftigtenanteile Industrie 2005-2011 [%] | Entwicklung der Beschäftigtenanteile GHD 2005-2011 [%] |
| Heiztage Mittel der Jahre 2009-2013 | Bevölkerungsdichte [Pers/km ²] | Veränderung Energieintensität 2007-2011 [%] | --- |
| Heizgradtage Mittel der Jahre 1970-2013 | ÖPNV Fahrten pro EW 2012 | --- | --- |
| Gradtagzahlen Mittel der Jahre 1970-2013 | Personenkilometer pro EW 2012 [km] | --- | --- |
| Arbeitslosenquote [%] | Mittlere Reiseweite [km] | --- | --- |
| Anteil Einpersonenhaushalte 2012 [%] | --- | --- | --- |
| Pro-Kopf Stromverbrauch [kWh] | --- | --- | --- |
| Anteil über 65 Jähriger [%] | --- | --- | --- |
| Erdgasverbrauch [kWh/m ² *a] | --- | --- | --- |

Die Datenverfügbarkeit auf Kreisebene ist lückenhaft. Selbst in den Datensätzen der Datenbanken der statistischen Bundes- und Landesämter gibt es Datenlücken. Diese wurden nach Möglichkeit durch gesonderte Recherchen bzw. Anfragen bei den jeweiligen Städten geschlossen. Sofern die Stadt dennoch nach der gesonderten Recherche über alle Indikatoren hinweg noch einige wenige Datenlücken aufwies, so wurden diese durch Mittelwerte aller Indikatorwerte desselben Indikators in dem die Lücke auftrat geschlossen. Insgesamt wurde ein Datensatz mit 4.387 Indikatorwerten erarbeitet.

2.2 Vorgehen und Methodik der Clusteranalyse

Eine Clusteranalyse ist die statistische Betrachtung von Objekten, in dieser Analyse, von deutschen kreisfreien Städten. Für diese Analyse erfolgte dies indem ein Objekt durch quantifizierbare Indikatoren beschrieben wird, die in einem weiteren Schritt zur Gruppierung der Städte herangezogen werden. Dementsprechend wichtig sind die Indikatoren als auch deren Qualität. Die hier durchgeführte Clusteranalyse besteht aus drei aufeinander aufbauenden Schritten (Z-Transformation, Faktoranalyse, Clusteranalyse) die nachfolgend dargestellt werden.

Die nachfolgende Grafik (Abbildung 2-1) zeigt skizzenhaft die drei benannten Schritte, die in den nachfolgenden Unterkapiteln vertieft diskutiert werden. Das hier dargestellte Verfahren erfolgt in dieser Form für die durchgeführte Analyse und ist kein allgemeingültiges, pauschales Vorgehen für sämtliche Clusteranalysen.

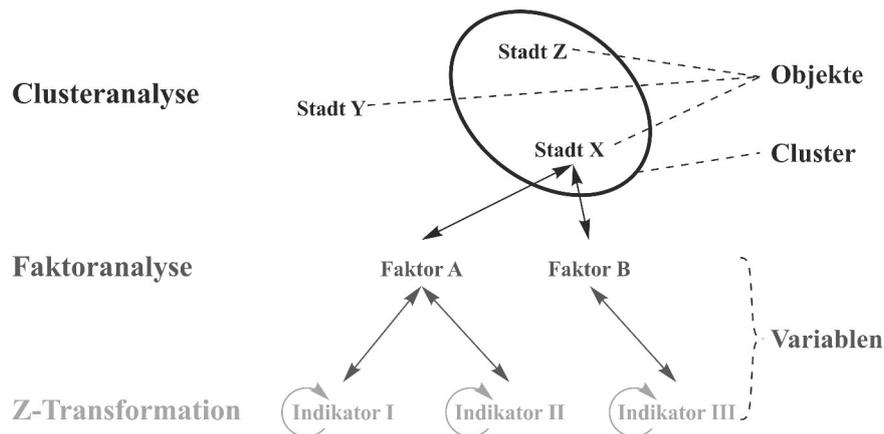


Abbildung 2-1: Struktur des gewählten clusteranalytischen Vorgehens (aus [WAL2014])

2.2.1 Z-Transformation

Der erste Schritt in der durchgeführten Analyse wird durch die Z-Transformation gebildet. Die Z-Transformation schließt an die primäre Datenvorbereitung, die die Erarbeitung des Datensatzes, die Elimination von Datenlücken als auch von Objekten, die sich für die Analyse nicht eignen an. Die erhobenen Daten müssen in diesem Schritt homogenisiert werden. Dies erfolgt durch die Standardisierung der Daten, in der für diese Analyse verwendeten Software SPSS Z-Transformation genannt [IBM2016]. Die Z-Transformation standardisiert unter Zuhilfenahme mathematischer Werte (Minima, Maxima, Mittelwerte, Standardfehler) die Indikatorwerte zu Z-Werten, die in einem Wertebereich zwischen ± 3 liegen. Hierdurch wird die Relationen zwischen den verschiedenen Indikatorwerten gewahrt und deren Aussage kommt im mathematischen Sinn die gleiche Bedeutung zu. [BUE2012]

Grafisch skizziert bilden die Ergebnisse der Z-Transformation das Fundament einer belastbaren Clusteranalyse (s. Abbildung 2-1). Einige wenige „Ausreiser“ unter den Z-Werten gibt es dennoch, die einen Z-Wert außerhalb des benannten Bereichs aufweisen. Diese sind dann zumeist Sonderfälle, die jedoch keinen negativen Einfluss auf die Aussagekraft der Ergebnisse aufweisen.

2.2.2 Faktoranalyse

Die Faktoranalyse bildet den zweiten Schritt dieser Untersuchung (Abbildung 2-1). Diese Methode erlaubt es, unter Aufrechterhaltung der Aussagekraft, statistisch sinnvoll Daten zu reduzieren. Dabei werden aus den einzelnen Indikatoren „künstliche“ Variablen (Faktoren) extrahiert, die die einzelnen Indikatoren hinreichend genau beschreiben [MON2008]. Ein Faktor kann somit nicht direkt erhoben werden, sondern muss über den Umweg der Faktorenanalyse ermittelt werden. Ein aussagefähiger Faktor ist thematisch und statistisch sinnvoll, jedoch hat ein Faktor nach seiner Extraktion zunächst keine Bedeutung. Erst mit der Zuordnung der Indikatoren zu den Faktoren lassen sich Schlüsse bzgl. der Aussage eines Faktors ziehen. Die in SPSS durchgeführte Faktoranalyse der 41 Indikatoren erforderte fünf Durchläufe, um belastbare Faktorenlösungen zu erhalten. Aus der Faktoranalyse ging eine Reduktion der Indikatorzahl von 41 auf 34 hervor, die letztlich zu 11 Faktoren zusammengefasst werden. In Tabelle 2-3 sind die elf Faktoren mit den dazugehörigen Indikatoren, den Faktorladungen sowie den Sektoren aufgetragen. Durch die Elimination von sieben Indikatoren wurde dennoch eine belastbare Beschreibung der untersuchten Städte erreicht werden. Der Betrag einer Faktorladung soll möglichst hoch sein jedoch nicht unter $\pm 0,5$ liegen

[BUE2012]. Eliminiert werden Indikatoren, die in keinem Faktor eine ausreichend hohe Faktorladung aufweisen. Eine betragsmäßig hohe Faktorladung zeigt eine sehr gute Übereinstimmung des Indikators mit dem Faktor. Eine negative/positive Übereinstimmung zeigt, dass ein Indikatorwert besonders gering (negativ) bzw. hoch (positiv) ausfällt. Beispielsweise (Faktor 3, Tabelle 2-3) weisen Städte mit einer guten positiven Bevölkerungsentwicklung in den letzten 11 Jahren eine geringe Arbeitslosenquote auf.

Tabelle 2-3: Übersicht der gebildeten Faktoren und den dazugehörigen Indikatoren

| Faktor | Indikator | Faktorladung | Sektor |
|-----------|--|--------------|-----------|
| Faktor 1 | Pkw pro 1000 EW 2014 | 0,912 | Verkehr |
| | Dieselfahrzeuge pro Tsd. EW 2014 | 0,9 | Verkehr |
| | Motorisierungsgrad 2014 | 0,895 | Verkehr |
| | Elektrofahrzeuge pro Tsd. EW 2014 | 0,88 | Verkehr |
| | Benzinfahrzeuge pro Tsd. EW 2014 | 0,715 | Verkehr |
| | Energieproduktivität GHD [EUR/GJ] | 0,578 | GHD |
| Faktor 2 | HH-Dichte [HH/km ²] | 0,906 | Haushalte |
| | Bevölkerungsdichte [Pers/km ²] | 0,901 | Verkehr |
| | Anteil Siedlungs- und Verkehrsfläche an Gesamtfläche [-] | 0,852 | Haushalte |
| | Anteil Wohnungen in 1-2 Familien Häusern [%] | -0,56 | Haushalte |
| Faktor 3 | Anteil über 65 Jähriger [%] | -0,867 | Haushalte |
| | Bevölkerungsentwicklung der letzten 11 Jahre [%] | 0,804 | Haushalte |
| | Anteil zwischen 25 und 65 Jähriger [%] | 0,776 | Verkehr |
| | Arbeitslosenquote [%] | -0,645 | Haushalte |
| Faktor 4 | Heizgradtage Mittel der Jahre 1970-2013 | 0,914 | Haushalte |
| | Gradtagzahlen Mittel der Jahre 1970-2013 | 0,851 | Haushalte |
| | Heiztage Mittel der Jahre 2009-2013 | 0,817 | Haushalte |
| Faktor 5 | Energieintensität Industrie [MJ/EUR] | 0,963 | Industrie |
| | Spezifischer Energieverbrauch Industrie [MJ/EUR] | 0,748 | Industrie |
| | Produktivitätsniveau GHD [EUR/GJ] | 0,721 | GHD |
| Faktor 6 | Beschäftigtenanteil Industrie 2011 [%] | -0,714 | Industrie |
| | Beschäftigtenanteil GHD 2011 [%] | 0,71 | GHD |
| | ÖPNV Fahrten pro EW 2012 | 0,684 | Verkehr |
| Faktor 7 | Energieproduktivität Industrie [EUR/GJ] | 0,965 | Industrie |
| | Produktivitätsniveau Industrie [EUR/GJ] | 0,964 | Industrie |
| Faktor 8 | Wohnfläche pro Person [qm] | 0,693 | Haushalte |
| | Pro-Kopf Stromverbrauch [kWh] | 0,648 | Haushalte |
| | Verfügbares Einkommen [€] | 0,635 | Haushalte |
| Faktor 9 | Industrie Umsatzsteuer an Gesamtumsatzsteuer [%] | 0,835 | Industrie |
| | GHD Umsatzsteuer an Gesamtumsatzsteuer [%] | -0,814 | GHD |
| Faktor 10 | Mittlere Reiseweite [km] | 0,902 | Verkehr |
| | Personenkilometer pro EW 2012 [km] | 0,784 | Verkehr |
| Faktor 11 | Entwicklung der Beschäftigtenanteile Industrie 2005-2011 [%] | -0,892 | Industrie |
| | Entwicklung der Beschäftigtenanteile GHD 2005-2011 | 0,718 | GHD |

2.2.3 Clusteranalyse

Die in dieser Arbeit gewählte Analyse ist die hierarchische agglomerative Clusteranalyse, die den dritten und relevantesten Schritt der gesamten Analyse bildet. Die Abbildung 2-2 stellt exemplarisch eine stark vereinfachte Clusterzusammenführung dar. Anders als bei hierarchisch-divisiven Clusteranalysen, bei denen zu Beginn alle Objekte in einem Cluster vereint sind und im Verlauf der Clusterermittlung in kleinere Gruppen reduziert werden, wird bei der agglomerativen Form der Clusteranalyse nach und nach ein Objekt zum anderen gefügt [BAC2000], bis sich alle Städte in Clustern vereinen, d. h. zu Beginn gibt es so viele Cluster wie Städte (Tabelle 2-2, a)). Erst mit dem Fortschritt der Analyse wachsen die Cluster an und die Clusterzahl verringert sich. Das erfolgt, indem auf Basis der in der Faktoranalyse extrahierten Faktoren Städte ermittelt werden, deren mathematische Distanz möglichst gering ist.

Für die hier durchgeführte Analyse wurde die quadrierte euklidische Distanz verwendet. Der Vorteil dieses Distanzmaßes ist, dass der Berücksichtigung großer Differenzen bei der Distanzberechnung der Objekte eine größere Beachtung zukommt, wodurch auch kleinste Unterschiede in den Ausprägungen einer Stadt gewichtiger in die Clusterzusammenführung einfließen [BUE2012]. Wird ein solches Städtepaar gefunden, so bildet dieses ein erstes Cluster. Dabei bildet der Mittelwert der Faktorwerte dieser beiden Städte das neue Clusterzentrum. In Tabelle 2-2 (b) wird dieses Clusterzentrum als schwarzer Punkt dargestellt. Im nächsten Schritt (Abbildung 2-2, c)) fusioniert das gefundene Clusterzentrum mit einer weiteren Stadt. Letztlich bleibt ein Cluster übrig (Tabelle 2-2, d)) das aus drei Städten besteht. Der verwendete Algorithmus ist der Ward-Algorithmus, der dieses Vorgehen sinnvoll abbildet, da dieser zum einen jede Stadt ausschließlich einem Cluster zuordnet (disjunkt – nicht überlappend), wodurch die Clustergrundidee gewahrt bleibt (Homogenität in den Clustern - Heterogenität zwischen den Clustern). Zum anderen benötigt der Ward Algorithmus keine Vorgabe eine Clusterzahl [BUE2012], was für die Durchführung einer Clusterung deutscher Städte nach sozio-energetischen Indikatoren entgegenkommt, da keine Clusteranzahl im Vorfeld bekannt ist und nur durch das Erraten gewählt werden kann. [BAC2000] In dem nachfolgenden Kapitel werden die Ergebnisse der Clusteranalyse dargestellt.

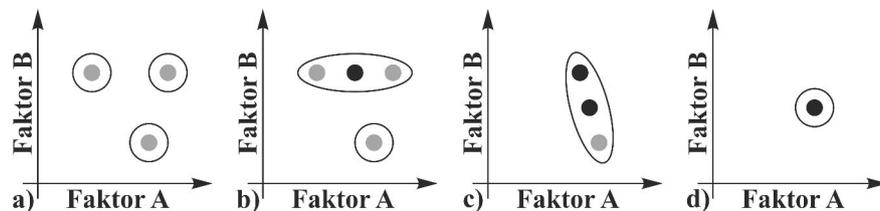


Abbildung 2-2: Beispielhafte schematische Clusterzusammenführung (aus [Wal2014])

3 Clusteranalyse deutscher kreisfreier Städte

Clusteranalysen finden in vielen Wissenschaftsbereichen Anwendung [WAL2016]. Diese Arbeit stellt eine von einigen wenigen Anwendungen dieser Methode auf Problemstellungen mit energetischem Ansatz dar, die in diesem Kapitel umrissen wird. Zunächst werden die Ergebnisse erläutert. Abschließend werden die ermittelten Städtecluster diskutiert.

3.1 Ergebnisse der durchgeführten Clusteranalyse

Die vielfältig erarbeiteten Indikatoren bilden die Städte des Untersuchungsrahmens umfassend ab. Es können für eine tiefgehende Beschreibung von kreisfreien Städten nicht ausschließlich direkte energetische Indikatoren verwendet werden. Deshalb wurden mittelbare als auch nicht mittelbare Indikatoren verwendet. Darauf basierend wurde die Clusteranalyse aufgebaut aus der folgende Ergebnisse hervorgebracht wurden. Die betrachteten 107 kreisfreien Städte können, nach der Kommunengröße unterschieden, vier Städtegruppen zugeordnet werden. Eingeteilt nach den ermittelten 11 Faktoren können wiederum 12 Städtegruppen erfasst werden, die statistisch sinnvoll gebildet werden können.

Bei der Clusteranalyse erfolgt die Berechnung in der Software SPSS 22. Die Ergebnisse der durchgeführten Clusteranalyse in SPSS hingegen sind nicht direkt als endgültiges Ergebnis nutzbar, da diese aus einer Schaar verschiedener Lösungsmöglichkeiten bestehen. Diese müssen durch die Auswahl einer geeigneten Clusteranzahl finalisiert werden. In Tabelle 3-1

sind die relevanten Fusionschritte, dimensionslosen Koeffizienten, berechneten Steigungen, prozentualen Differenzen sowie die daraus ableitbaren Clusteranzahlen dargestellt. Mit diesen Größen lassen sich statistisch sinnvolle Clusteranzahlen bestimmen. Mithilfe der dimensionslosen Koeffizienten lassen sich die Steigungen der jeweiligen Fusionschritte berechnen. Diese wiederum deuten bei starken Schwankungen auf eine Erhöhung der Unordnung im entstehenden Cluster hin, die auf eine Clusterlösung hindeuten. Wird die Steigung grafisch aufgetragen bildet sich eine Kurve, die mit den Fusionschritten anwächst. An dieser Kurve lassen sich die zu betrachtenden Steigungssprünge grafisch darstellen, jedoch macht die grafische Darstellung nur digital Sinn, da die Kurve stark vergrößert betrachtet werden muss. Die Betrachtung wird auch als Ellenbogenkriterium bzw. Scree-Diagramm bezeichnet und ist ein gängiges Vorgehen in der multivariaten Clusteranalyse. Fusionschritte, die einen großen Steigungsanstieg verursachen deuten auf eine unproportionale Erhöhung der Unordnung (Inhomogenität) in einem Cluster hin, wodurch sich die Clusterlösung von der Clustergrundidee entfernt [BAC2002]. Solche Stellen müssen gesondert betrachtet werden, da ideale Clusterlösungen immer einen Schritt vor solchen Stellen vorliegen. Diese Fälle sind in Tabelle 3-1 grau unterlegt.

Tabelle 3-1: Auswahl aus der Zuordnungsübersicht der Fusionschritte

| Fusions-schritt | Fusions-koeffizienten | Steigung | Differenz zwischen den Schritten | Clusteranzahl |
|------------------------|------------------------------|-----------------|---|----------------------|
| 1 | 0,364 | 0,364 | 0,623 | 106 |
| 2 | 0,949 | 0,584 | 0,731 | 105 |
| 3 | 1,748 | 0,800 | 0,991 | 104 |
| ... | ... | ... | ... | ... |
| 83 | 311,640 | 12,407 | 0,984 | 24 |
| 84 | 324,255 | 12,615 | 0,862 | 23 |
| 85 | 338,897 | 14,642 | 0,980 | 22 |
| 86 | 353,834 | 14,937 | 0,970 | 21 |
| 87 | 369,232 | 15,398 | 0,832 | 20 |
| 88 | 387,735 | 18,503 | 0,980 | 19 |
| ... | ... | ... | ... | ... |
| 94 | 522,557 | 25,853 | 0,953 | 13 |
| 95 | 549,694 | 27,136 | 0,846 | 12 |
| 96 | 581,751 | 32,057 | 0,929 | 11 |
| 97 | 616,254 | 34,503 | 0,894 | 10 |
| 98 | 654,848 | 38,594 | 0,882 | 9 |
| 99 | 698,626 | 43,778 | 0,839 | 8 |

Als mögliche Lösungen gingen aus der Clusteranalyse Clusteranzahlen von 106, 105, 101, 53, 46, 31, 23, 20, 12 sowie 8 hervor. Eine gekürzte Zuordnungsübersicht mit einer Auswahl möglicher Lösungen ist in vorangestellter Tabelle 3-1 gezeigt, wobei zwar die möglichen aber unrelevanten Clusterlösungen (101, 53, 46, 31) eliminiert wurden. Die ersten beiden Clusterlösungen (106, 105) in obiger Tabelle dienen der Übersicht und können ausgeschlossen werden, da der Clustergedanke, nur wenige Städtegruppen zu finden, nicht erfüllt wird. Die idealste Clusterlösung wäre, wenn jede Stadt ein Cluster bildet, denn nur dann wird jede Stadt bestmöglich abgebildet. Die sinnvollen Clusterlösungen 23, 20, 12 sowie 8 konnten als aussagekräftigste Cluster identifiziert werden (Strichpunktlinien in Abbildung 3-1). Aus dieser Anzahl denkbarer Clusteranzahlen muss mithilfe des nachfolgenden Dendrogramms eine Clusterlösung qualifiziert werden.

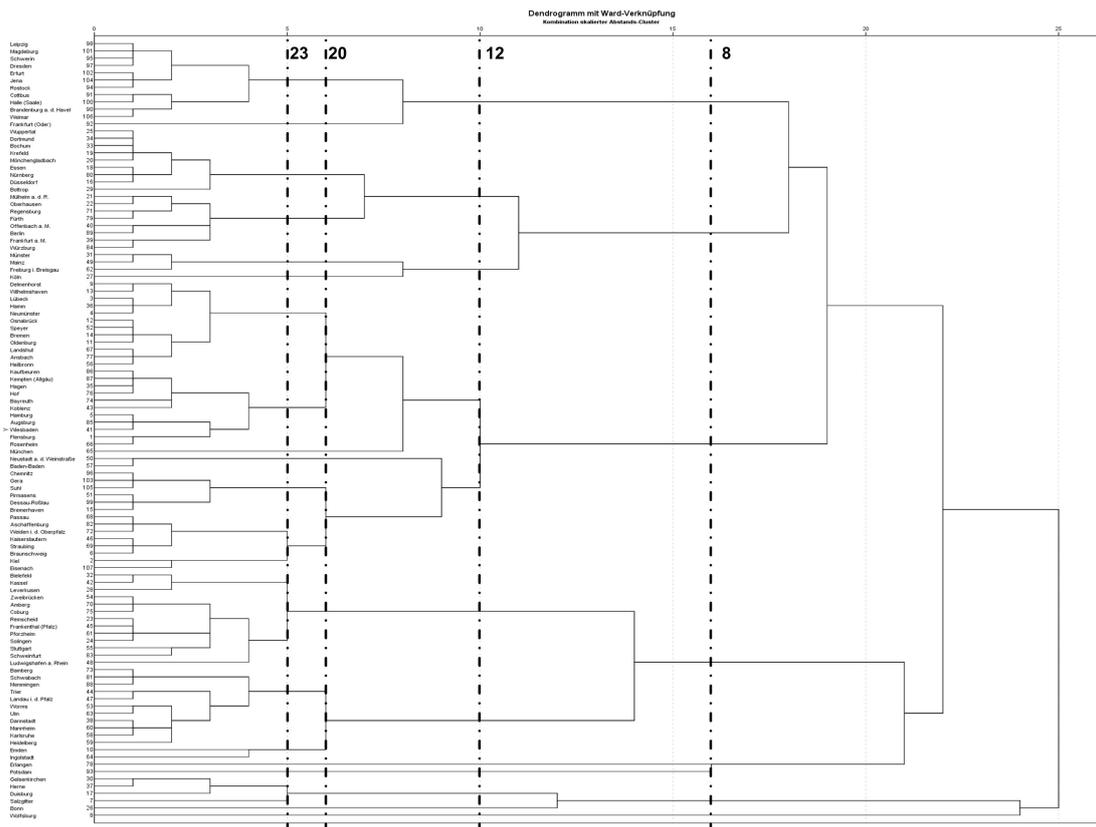


Abbildung 3-1: Dendrogramm der Fusionsschritte

Ein Dendrogramm stellt die Fusionsschritte der Clusterzusammenführung grafisch dar. Dabei stellen vertikale Linien Clusterfusionen und horizontale Linien die Städte bzw. Cluster dar. Auf der Ordinate sind die Städte aufgetragen. Die Abszisse zeigt ein dimensionsloses Distanzmaß. Nach einer Fusion bildet die entstehende horizontale Linie das gebildete Cluster. Die Länge der horizontalen Linien (Clusterlinien) bis zur Fusion sollte möglichst gering sein. Lange horizontale Linien deuten auf eine große Inhomogenität in den Clustern hin, da Cluster entstehen die aus heterogenen Objekten (Städten) bestehen. Die 8 Clusterlösung kann aus diesem Grund ausgeschlossen werden, da heterogene Cluster generiert werden (viele lange horizontale Linien). Die 23 und 20 Clusterlösung liegen aufgrund der geringen Differenz in der dimensionslosen Distanz nah beieinander. Das deutet auf eine „Ausfransung“ der Clusterergebnisse hin. Die 12 Clusterlösung stellt in dieser Analyse einen sinnvollen Kompromiss dar, da zum einen die Cluster ineinander, unter einer vertretbaren Inhomogenität, eine gute Heterogenität zwischen den Clustern realisieren. Dennoch lässt sich, wie bei jeder Datenreduktion, auch bei der Clusterung ein Informationsverlust nicht vermeiden. Trotz des Informationsverlusts werden bestimmte Erkenntnisse gewonnen, die auf anderem Wege nicht ermittelt werden können.

3.2 Diskussion der Ergebnisse

Die Tabelle 3-2 zeigt die einzelnen Cluster, die Anzahl der Städte aus denen diese bestehen sowie die den Clustern zugeordneten Städte. Es zeigt sich, dass unter den 12 identifizierten Clustern vier Cluster aus jeweils nur einer Stadt bestehen. Das größte identifizierte Cluster besteht aus 24 Städten.

Tabelle 3-2: Ermittelte Cluster sowie die zugeordneten Städte

| Cluster | Zugeordnete Städte |
|------------|---|
| Cluster 1 | Flensburg, Lübeck, Neumünster, Hamburg, Delmenhorst, Oldenburg, Osnabrück, Wilhelmshaven, Bremen, Hagen, Hamm, Wiesbaden, Koblenz, Speyer, Heilbronn, München, Rosenheim, Landshut, Bayreuth, Hof, Ansbach, Augsburg, Kaufbeuren, Kempten (24 Städte) |
| Cluster 2 | Kiel, Braunschweig, Bremerhaven, Kaiserslautern, Neustadt a.d.W., Pirmasens, Baden-Baden, Passau, Straubing, Weiden i.d.O., Aschaffenburg, Chemnitz, Dessau-Roßlau, Gera, Suhl, Eisenach (16 Städte) |
| Cluster 3 | Salzgitter, Duisburg, Gelsenkirchen, Herne (4 Städte) |
| Cluster 4 | Wolfsburg (1 Stadt) |
| Cluster 5 | Emden, Darmstadt, Trier, Landau i.d.P., Worms, Karlsruhe, Heidelberg, Mannheim, Ulm, Ingolstadt, Bamberg, Schwabach, Memmingen (13 Stadt) |
| Cluster 6 | Düsseldorf, Essen, Krefeld, Mönchengladbach, Mülheim a.d.R., Oberhausen, Wuppertal, Bottrop, Bochum, Dortmund, Frankfurt a.M., Offenbach a.R., Regensburg, Fürth, Nürnberg, Würzburg, Berlin (17 Städte) |
| Cluster 7 | Remscheid, Solingen, Leverkusen, Bielefeld, Kassel, Frankenthal (Pfalz), Ludwigshafen a.R., Zweibrücken, Stuttgart, Pforzheim, Amberg, Coburg, Schweinfurt (13 Städte) |
| Cluster 8 | Bonn (1 Stadt) |
| Cluster 9 | Köln, Münster, Mainz, Freiburg i.B. (4 Städte) |
| Cluster 10 | Erlangen (1 Stadt) |
| Cluster 11 | Brandenburg a.d.H., Cottbus, Frankfurt (Oder), Rostock, Schwerin, Dresden, Leipzig, Halle (Saale), Magdeburg, Erfurt, Jena, Weimar (12 Städte) |
| Cluster 12 | Potsdam (1 Stadt) |

Mithilfe der tabellarischen Darstellung der Clustermittelwerte (Tabelle 3-3) können die Zuordnungen zu den Clustern nachvollzogen werden. Als Beispiel wird der Faktor 3 betrachtet. Dort fällt der niedrigste Mittelwert im Cluster 3 an. Das bedeutet, dass die Ausprägungen der Städte in Cluster 3 einen besonders hohen Anteil an Personen über 65 Jahren, eine stark negative Bevölkerungsentwicklung in den letzten Jahren, eine hohe Arbeitslosenquote sowie einen geringen Anteil der Bevölkerung im Alter zwischen 25 und 65 Jahren. Weitere Beispiele lassen sich unter der Betrachtung der Clustermittelwerte sowie der Faktorladungen und Indikatorzuordnung (Tabelle 2-3) durchführen.

Tabelle 3-3: Clustermittelwerte der Faktoren

| Cluster | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Faktor 1 | -0,14 | -0,11 | -0,24 | 8,62 | 0,07 | -0,04 | 0,16 | 0,56 | -0,19 | -0,05 | -0,50 | -0,88 |
| Faktor 2 | 0,11 | -0,50 | -1,30 | 0,25 | -0,67 | 1,05 | 0,09 | -0,26 | -0,36 | 0,18 | -0,54 | -0,44 |
| Faktor 3 | -0,06 | -0,80 | -0,93 | -0,83 | 1,21 | 0,05 | -0,08 | 1,28 | 1,10 | 0,80 | -0,34 | 0,81 |
| Faktor 4 | 0,93 | 0,35 | -1,08 | -0,28 | -0,24 | -0,84 | -0,31 | 1,88 | -1,30 | -0,05 | 0,18 | -0,69 |
| Faktor 5 | -0,35 | 0,21 | -0,34 | 0,38 | -0,26 | -0,08 | 0,13 | -0,29 | -0,98 | 2,04 | 0,30 | 7,59 |
| Faktor 6 | 0,07 | 0,06 | -1,07 | -0,20 | -0,86 | 0,27 | -0,37 | 2,93 | 1,83 | -1,97 | 0,37 | 0,42 |
| Faktor 7 | 0,56 | 0,33 | -0,52 | -1,62 | 0,01 | 0,22 | 0,30 | 0,68 | -0,61 | 0,86 | -1,77 | -0,71 |
| Faktor 8 | -0,16 | -0,25 | 2,96 | 0,13 | -0,20 | -0,14 | -0,18 | 6,81 | -0,61 | -0,05 | -0,15 | 0,77 |
| Faktor 9 | -0,19 | -0,64 | -0,25 | -0,63 | 0,09 | -0,44 | 1,92 | 0,13 | -0,31 | -0,43 | -0,07 | 0,00 |
| Faktor 10 | -0,37 | 0,36 | 0,23 | 0,02 | -0,56 | -0,15 | 0,26 | -0,16 | 1,13 | 6,85 | -0,14 | -1,05 |
| Faktor 11 | 0,60 | -0,37 | -0,07 | 0,57 | -0,14 | -0,28 | -0,13 | 0,51 | 0,95 | 0,21 | -0,52 | 1,01 |

Auch aus einer statistischen Betrachtung der Städte des Untersuchungsrahmens gehen einige Städte als nicht zuteilbar hervor. Diese wurden in dieser Analyse vorerst weiter

berücksichtigt, damit die Gesamtheit der kreisfreien Städte Deutschlands in die Betrachtung einfließt. Die Lösungen sind nicht ideal, da die vier Cluster die je nur eine Stadt beinhalten die Clustergrundidee verwerfen. Eine Möglichkeit wäre es die Clusteranzahl zu verringern, um diese „Einzelcluster/-städte“ in andere größere Cluster zu fusionieren, jedoch würde dadurch die Clustergrundidee nicht weniger verletzt (homogene Städte in einem Cluster).

Bei der Deutung der Cluster auf die an dieser Stelle nicht eingegangen wird bietet die Möglichkeit beispielsweise Energie- oder Klimaschutzkonzepte zwischen den Städten, basierend auf den ermittelten Städtegruppen, abzugleichen und dadurch eine primäre Vorauswahl für die Identifikation möglicher Umsetzungsstädte zu geben. Die Deutung der Cluster kann durchaus Hilfestellung sein, jedoch ist genau das der Schritt bei dem eine subjektive Abschätzung in die bisher mathematisch-statistische Analyse einfließt, wodurch ein Absinken der Belastbarkeit der Ergebnisse nicht auszuschließen ist.

4 Zusammenfassung und Ausblick

In dieser Untersuchung wurden die vier Verbrauchssektoren der deutschen kreisfreien Städte durch frei verfügbare Daten Indikator basiert beschrieben. Der ermittelte Datensatz wurde in einem dreistufigen Analyseprozess, bestehend aus der Z-Transformation zur Vereinheitlichung der Indikatoren, der Faktoranalyse zur Zusammenfassung und Reduktion der Indikatorzahl sowie der letztendlichen Clusteranalyse der Städte untersucht. Die Analyse hat gezeigt, dass sich die 107 deutschen kreisfreien Städte des Untersuchungsrahmens idealerweise auf 12 verschieden große Cluster aufteilen lassen. Die Cluster vereinen somit Städte auf statistisch ähnliche Städtegruppen, die für weitere Untersuchungen herangezogen werden können. Die Ergebnisse der Clusteranalyse erheben keinen Anspruch darauf, dass eine Übertragbarkeit bereits umgesetzter Maßnahmen in Modellstädten (z. B. den Städten des Wettbewerb 'Energieeffiziente Stadt') bereits im Vorfeld überprüft sowie beschieden werden kann, ob die Maßnahme in anderen Städten erfolgreich umgesetzt werden kann. Die Ergebnisse der Clusteranalyse können jedoch als Hilfestellung für die Identifikation in Frage kommender Kommunen dienen.

In der Fortführung der Anwendung der Clusteranalyse werden zunächst die Ausreißerstädte eliminiert, um die Ergebnisse (Clusteranzahl) zu verschlanken und dadurch die Qualität der Clusterlösungen zu steigern. Darüber hinaus wird die Clusteranalyse fokussiert auf lediglich einen Verbrauchssektor angewendet, um daran andere Anwendungsbereiche und Hilfen durch die Clusteranalyseergebnisse zu erproben (z. B. Potenzialabschätzungen).

5 Literatur

- [BAC2000] BACKHAUS, K.; ERICHSON, B.; PLINKE, W. ET AL.: *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. 9. Auflage. Berlin, Springer, 2000. (Springer-Lehrbuch). ISBN: 978-3-662-08893-7.
- [BAC2002] BACHER, J.: *Clusteranalyse. Anwendungsorientierte Einführung*. 2. Auflage. München: Oldenbourg, 2002. ISBN: 3-486-23760-8.
- [BUE2012] BÜHL, A.: *SPSS 20. Einführung in die moderne Datenanalyse*. 13., aktualisierte Aufl. München: Pearson, 2012. ISBN: 9783868941500.

- [BUN2016] BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: *Gesamtausgabe der Energiedaten - Zahlen und Fakten, Energiedaten. Nationale und internationale Entwicklung*, 2016. <http://www.bmwi.de>, zuletzt aktualisiert am 12.10.2015. Abruf: 08.01.2016.
- [DES2013] STATISTISCHES BUNDESAMT: *Zahlen und Fakten - Wirtschaftsbereiche*. 2016. <https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/Wirtschaftsbereiche.html>. Abruf: 07.01.2016.
- [EIC2014] EICKELKAMP, T.; WALL, W.; KOCH, M. K.: *Indikatoren zur energetischen Charakterisierung und Clusterung von Städten*. In: H.-J. Wagner und J. Görres (Hg.): *Wettbewerb Energieeffiziente Stadt, Band 2: Energieversorgung, Energiebilanzierung und Monitoring*. Berlin: LIT-Verlag, 2014., S. 43–52. ISBN: 9783643127112
- [EU2011] Europäische Union: *Regionalpolitik, Städte von Morgen, Herausforderungen, Visionen, Wege nach Vorn*. Hg. v. Amt für Veröffentlichungen. Brüssel, 2011.
- [IBM2016] IBM: *IBM SPSS Statistics 22*, 2016. <http://www-01.ibm.com/software/de/stats22/>. Abruf: 13.01.2016.
- [MON2008] MONKA, M.; SCHÖNECK-VOß, N. M.; VOß, W. ET AL.: *Statistik am PC. Lösungen mit Excel*. 5., aktualisierte und erw. Aufl. München: Hanser, 2008. - ISBN: 978-3446415553.
- [Wal2014] WALL, W.; EICKELKAMP, T.; KOCH, M. K.: *Clusterung deutscher Städte anhand energetischer Kennwerte*. In: Marco K. Koch und Russell McKenna (Hg.): *Wettbewerb Energieeffiziente Stadt, Band 3: Methoden und Modelle*. Berlin: Lit-Verlag, 2014., S. 61–70. ISBN: 9783643128287.
- [WAL2016] WALL, W.: *Energetisch vergleichbare Städtegruppen – Eine gesamtheitliche Clusteranalyse und Clusterauswahl deutscher kreisfreier Städte auf Basis der typischen Verbrauchssektoren und sozio-energetischer Indikatoren*. Bochum: Selbstverlag des Lehrstuhls Energiesysteme und Energiewirtschaft, 2016. – In Vorbereitung