

Kulenkampff, Gabriele; Ockenfels, Martin; Zoz, Konrad; Zuloaga, Gonzalo

Working Paper

Kosten von Breitband-Zugangsnetzen: Clusterbildung und Investitionsbedarf unter Berücksichtigung des bestehenden Ausbaus. bottom-up Modellierung und statistische Analyse

WIK Diskussionsbeitrag, No. 473

Provided in Cooperation with:

WIK Wissenschaftliches Institut für Infrastruktur und Kommunikationsdienste GmbH, Bad Honnef

Suggested Citation: Kulenkampff, Gabriele; Ockenfels, Martin; Zoz, Konrad; Zuloaga, Gonzalo (2020) : Kosten von Breitband-Zugangsnetzen: Clusterbildung und Investitionsbedarf unter Berücksichtigung des bestehenden Ausbaus. bottom-up Modellierung und statistische Analyse, WIK Diskussionsbeitrag, No. 473, WIK Wissenschaftliches Institut für Infrastruktur und Kommunikationsdienste, Bad Honnef

This Version is available at:

<https://hdl.handle.net/10419/228686>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Kosten von Breitband-Zugangsnetzen

Clusterbildung und Investitionsbedarf
unter Berücksichtigung
des bestehenden Ausbaus
– bottom-up Modellierung
und statistische Analyse –

Autoren:

Gabriele Kulenkampff

Martin Ockenfels

Konrad Zoz

Gonzalo Zuloaga

Bad Honnef, Dezember 2020

Impressum

WIK Wissenschaftliches Institut für
Infrastruktur und Kommunikationsdienste GmbH
Rhöndorfer Str. 68
53604 Bad Honnef
Deutschland
Tel.: +49 2224 9225-0
Fax: +49 2224 9225-63
E-Mail: info@wik.org
www.wik.org

Vertretungs- und zeichnungsberechtigte Personen

Geschäftsführerin und Direktorin	Dr. Cara Schwarz-Schilling
Direktor	Alex Kalevi Dieke
Direktor Abteilungsleiter Netze und Kosten	Dr. Thomas Plückebaum
Direktor Abteilungsleiter Regulierung und Wettbewerb	Dr. Bernd Sörries
Leiter der Verwaltung	Karl-Hubert Strüver
Vorsitzende des Aufsichtsrates	Dr. Daniela Brönstrup
Handelsregister	Amtsgericht Siegburg, HRB 7225
Steuer-Nr.	222/5751/0722
Umsatzsteueridentifikations-Nr.	DE 123 383 795

In den vom WIK herausgegebenen Diskussionsbeiträgen erscheinen in loser Folge Aufsätze und Vorträge von Mitarbeitern des Instituts sowie ausgewählte Zwischen- und Abschlussberichte von durchgeführten Forschungsprojekten. Mit der Herausgabe dieser Reihe bezweckt das WIK, über seine Tätigkeit zu informieren, Diskussionsanstöße zu geben, aber auch Anregungen von außen zu empfangen. Kritik und Kommentare sind deshalb jederzeit willkommen. Die in den verschiedenen Beiträgen zum Ausdruck kommenden Ansichten geben ausschließlich die Meinung der jeweiligen Autoren wieder. WIK behält sich alle Rechte vor. Ohne ausdrückliche schriftliche Genehmigung des WIK ist es auch nicht gestattet, das Werk oder Teile daraus in irgendeiner Form (Fotokopie, Mikrofilm oder einem anderen Verfahren) zu vervielfältigen oder unter Verwendung elektronischer Systeme zu verarbeiten oder zu verbreiten.
ISSN 1865-8997

Inhaltsverzeichnis

Zusammenfassung	VII
Summary	IX
1 Einleitung	1
1.1 Gegenstand	1
1.2 Aufbau der Studie	2
2 Methodische Vorgehensweise, Modellierungsrahmen und zentrale Annahmen	4
2.1 Methodische Vorgehensweise	4
2.1.1 Flächendeckende bottom-up Modellierung	4
2.1.2 Glasfaseranschlussnetze – Stand der Versorgung, Profitabilität und Förderbedarf eines flächendeckenden Breitbandausbaus	6
2.1.3 Regressionsanalyse – Geeignete Schätzer für die Kosten von Anschlussnetzen	6
2.1.4 Clusterbildung	7
3 Flächendeckende bottom-up Modellierung	8
3.1 Modified scorched node Ansatz	8
3.2 Gegenstand und zentrale Annahmen der Investitionswert- und Kostenberechnung	10
3.2.1 Betrachtete Anschlussnetztechnologien	10
3.2.2 Parametergesteuerte Netzauslegung	12
3.2.3 Technologiespezifische Eingangsparameter	14
3.2.4 Weitere anschlussbereichsbezogene Investitionsparameter	15
3.2.5 Sonstige Bestandteile der Kostenberechnung	15
4 Glasfaseranschlussnetze – Stand der Versorgung, Profitabilität und Förderbedarf eines flächendeckenden Breitbandausbaus	16
4.1 Greenfield-Analyse	17
4.1.1 Kosten des Netzausbaus nach Clustern	17
4.1.2 Kostenvergleich: FTTH P2P und hybride NGA- Technologien (FTTS/FTTC)	19
4.1.3 Profitabilität des Netzausbaus der verschiedenen Anslussttechnologien	20
4.2 Brownfield-Analyse: Berücksichtigung bestehender Breitbandversorgung und Abschätzung von Cu-GF Migrationskosten	26
4.2.1 Datenbasis und anschlussbereichsbezogene Aufbereitung	27
4.2.2 Bestehende Versorgungslage nach Anschlüssen	30

4.2.3	Bestehende Versorgungslage nach Netzausbau	36
4.2.4	FTTH Investitionen pro Anschluss und bisheriger Breitbandausbau	38
5	Regressionsanalyse – geeignete Schätzer für die Kosten von Anschlussnetzen	42
5.1	Datenbasis	43
5.2	Methodische Vorgehensweise	44
5.3	Benchmark-Modell: Multiple Regression auf Ebene der Anschlussbereiche	47
5.3.1	Multiple Regression – Auswahl der Variablen	47
5.3.2	Benchmark-Modell: Ergebnisse der multiplen Regression	50
5.4	Validierung der Regressionsergebnisse auf Ebene der Anschlussbereiche	52
5.5	Validierung der Regressionszusammenhänge für Gebietskörperschaften (Kreise und Bezirke)	56
5.6	Regressionsmuster mit vollständig öffentlich verfügbaren Daten	58
5.7	Zwischenfazit Regressionsanalyse	62
6	Clusterbildung	65
6.1	Investitionen pro Anschluss und Clusterbildung anhand von Dichtekriterien und gleicher Anschlusszahl pro Cluster	66
6.1.1	Regionale Verteilung der Anschlussbereiche über Cluster – ein Vergleich der WIK-NGA-Modell Cluster mit FTTH P2P Investitionen je Anschluss, gruppiert nach halben Standardabweichungen	66
6.1.2	Vergleich von Dichtekriterien und univariater k-means Clusterung nach Investitionen je Anschluss	70
6.2	Clusterbildung nach der k-Mittelwert Methode	75
6.2.1	Kriterien zur Bestimmung der optimalen Anzahl Cluster	75
6.2.2	Optimale Anzahl Cluster bei univariater k-Means Clusterung der FTTH-Investitionen je Anschluss	77
6.3	Multivariate Clusterbildung: Verwendung der Regressoren zur Investitionswertschätzung bei der Clusterung	79
6.4	Vergleich der Clusteransätze	82
6.5	Geografische Nachbarschaft: Hot Spot - Cold Spot Analyse	85
6.5.1	Nachbarschaftsanalyse auf Basis der Anschlussbereiche	85
6.5.2	Nachbarschaftsanalyse auf Basis von Kreisen	88
6.6	Zwischenfazit Clusterung	90
7	Zusammenfassung und Ausblick	93
8	Literaturverzeichnis	100

A1 Anhang 1: Datenquellen und Modellierungsannahmen zum Versorgungsgebiet	103
A1.1 Anschlussnachfrage	103
A1.1.1 Verortung der Nachfrage	103
A1.1.2 Höhe des Nachfragepotentials	104
A1.2 Erzeugung der Datenbasis für die anschlussbereichsbezogene, GIS-basierte bottom-up Modellierung	108
A1.2.1 Berechnung der ASB-Grenzen unter Verwendung eines öffentlichen HVT-Datensatzes	108
A1.2.2 Ermittlung der Lotpunkte und Abstand der Gebäude	110
A1.2.3 Verschnitt der Lotpunkte mit den ASB-Grenzen	111
A1.2.4 Datenbank: Erzeugung der Inputdaten für das TAL Netzplanungsmodul	112
A2 Anhang 2: Modellierungsrahmen	113
A2.1 Bestimmung des Mengengerüsts (WIK-TAL-Modell 3.0)	113
A2.2 Bestimmung der Investitionswerte und Kosten (WIK-NGA-Modell)	117
A2.2.1 Anschlussbereichsgenaues Mengengerüst	117
A2.2.2 Variable Kosten je Kunde	117
A2.2.3 Gebäudeinterne Verkabelung	118
A2.2.4 OPEX und indirekte Investitionen	118
A2.2.5 Upstream Kosten und Erlöse	118
A2.2.5.1 Kosten von Konzentrations- und Kernnetz	119
A2.2.5.2 Vertriebs- und Werbekosten (Retailkosten)	120
A2.2.5.3 Erlöse (ARPU)	120

Abbildungsverzeichnis

Abbildung 2-1:	Verwendete Tools der modellbasierten Kostenermittlung	5
Abbildung 3-1:	NGA Netzarchitekturen im Überblick	11
Abbildung 4-1:	Zeitpunktbezogener Vergleich von Kosten und ARPU für FTTH P2P nach Clusterstufen (pro Kunde/ Monat, Index 1,0 = Cluster 1, 2020)	18
Abbildung 4-2:	Greenfield – Vergleich der Kosten von FTTH P2P, FTTS und FTTC (pro Kunde/ Monat, Index 1,0 = Cluster 1 bei FTTC)	19
Abbildung 4-3:	Profitabilität eines flächendeckenden FTTH P2P Ausbaus	21
Abbildung 4-4:	Kosten, ARPU und Profitabilität FTTH P2P, FTTS und FTTC Greenfield (Index 1,0 = Cluster 1 bei FTTC)	23
Abbildung 4-5:	Vergleich der Kosten und ARPU pro Kunde/ Monat: FTTS und FTTC (Index 1,0 = Cluster 1, FTTC)	24
Abbildung 4-6:	Illustration der Datenverfügbarkeit aus dem Breitbandatlas: Breitbandanschlüsse eines Anschlussbereichs	28
Abbildung 4-7:	Ermittlung der Versorgungslage je Anschlussbereich	30
Abbildung 4-8:	Breitbandversorgung nach Clustern	32
Abbildung 4-9:	Ergebnis der Versorgung nach Clustern nach Bereinigung der Überschneidungen	34
Abbildung 4-10:	Ausbaustand von FTTH, CATV und FTTC – vergleichende Darstellung	41
Abbildung 5-1:	Struktur- und netzbezogene Daten als Einflussfaktoren des FTTH– Investitionsbedarfs pro Anschlussleitung	43
Abbildung 5-2:	Zusammenhang zwischen der Anzahl von Anschlüssen und den erforderlichen FTTH-Investitionen pro MPoP	44
Abbildung 5-3:	Zusammenhang zwischen Teilnehmerdichte und FTTH-Investition pro Anschluss (bezogen auf (a) ASB-Fläche bzw. (b) bebaute Fläche [km ²])	46
Abbildung 5-4:	Zusammenhang zwischen Teilnehmerdichte pro bebauter Fläche [km ²] und FTTH-Investition pro Anschluss, logarithmiert	47
Abbildung 5-5:	Robustheitsanalyse: Spanne der Schätzungen des Benchmark-Modells nach Eingang von zusätzlichen erklärenden Variablen	53
Abbildung 5-6:	NUTS-Abgrenzung: NUTS1 (Länder) und NUTS3 (Landkreise)	56
Abbildung 5-7:	Zusammenhang zwischen der Zielvariable (FTTH-Invest pro Anschluss) und den ausgewählten Proxy-Variablen aus öffentlich zugänglichen Daten (logarithmiert)	60

Abbildung 5-8:	Vergleich zwischen der Benchmark-Variable „HK-Trassendichte pro bebautem km ² “ (blau) und der entsprechenden Proxy-Variable „Straßendichte ((a) außerorts und (b) gesamt) pro bebautem km ² “ (rot)	62
Abbildung 6-1:	Räumliche Verteilung der FTTH Investitionen je Anschluss gruppiert nach halben Standardabweichungen	68
Abbildung 6-2:	Räumliche Verteilung der Clusterabgrenzung nach Teilnehmerdichte bezogen auf die ASB-Fläche (20 Cluster des WIK-NGA-Modells)	69
Abbildung 6-3:	Investitionen je Anschluss nach Cluster für unterschiedliche Clusterbildungen	72
Abbildung 6-4:	Räumliche Verteilung der Clusterabgrenzung für die Investitionen je Anschluss nach der k-Mittelwert Methode bei 20 Clustern	74
Abbildung 6-5:	Optimale Anzahl Cluster bei univariater k-Means Clusterung mit Investitionen je Anschluss als Clusterkriterium	77
Abbildung 6-6:	Clusterkriterium nach Maßgabe „Invest pro Anschluss“ – Vergleich univariater Clusterbildung für k = 4, 8 und 20	78
Abbildung 6-7:	Optimale Anzahl Cluster bei multivariater k-Means Clusterung der Regressorenmatrix zur Bestimmung der Investitionswerte pro Anschluss	80
Abbildung 6-8:	FTTH-Investitionswerte und multivariate Clusterbildung für k=10, 20	81
Abbildung 6-9:	Vergleich der Clusteransätze bezüglich der Anzahl Anschlussbereiche je Cluster	82
Abbildung 6-10:	Vergleich der Clusteransätze bezüglich der Anzahl Anschlüsse je Cluster	83
Abbildung 6-11:	Vergleich der Clusteransätze bezüglich der mittleren Investitionen pro Anschluss je Cluster	84
Abbildung 6-12:	Hot Spot / Cold Spot Analyse der Investitionen je Anschluss	87
Abbildung 6-13:	Vergleich der räumlichen Verteilung der Investitionswerte je Anschluss mit der Hot Spot /Cold Spot Clustern	88
Abbildung 6-14:	Vergleich der Siedlungstypologie nach Kreisen mit Hot Spot /Cold Spot Clustern	89
Abbildung 6-15:	Vergleich der Siedlungstypologie nach Kreisen mit k-Mittelwert Clusterbildung der Investitionen je Anschluss bei 4 Clustern (90% erklärte Varianz)	90

Tabellenverzeichnis

Tabelle 3-1:	Datenquellen	9
Tabelle 3-2:	Modellierte Glasfaserarchitekturen	12
Tabelle 4-1:	Clusterbildung und Anschlussdichte	16
Tabelle 4-2:	Klassifizierung des Breitbandatlas nach Geschwindigkeiten und Technologien	27
Tabelle 4-3:	Kategorialwerte zur Einschätzung der Versorgungslage im Raster	27
Tabelle 4-2:	Kategorialwerte des Breitbandatlas und deren Übersetzung für die Modellrechnungen	31
Tabelle 4-5:	Durchschnittlicher Versorgungsgrad nach Clustern	35
Tabelle 5-1:	Datenbasis aus dem WIK NGA-Modell zur Ableitung des „Benchmark“-Modells	48
Tabelle 5-2:	Testergebnisse und statistische Kennzahlen für die Variablenauswahl des „Benchmark“-Modells	49
Tabelle 5-3:	STATA-Regressionsoutput für das „Benchmark-Modell“	51
Tabelle 5-4:	STATA-Regressionsoutput des Benchmark-Modells und der Robustheitsmodelle (2 bis 10)	55
Tabelle 5-5:	STATA-Regressionsoutput des Benchmark-Modells auf Ebene der Anschlussbereiche, der Landkreise (NUTS3) und Regierungsbezirke (NUTS2)	57
Tabelle 5-6:	Korrelationsanalyse zwischen Proxy- und Benchmark-Variablen	59
Tabelle 5-7:	Ergebnisvergleich zwischen Proxy- und Benchmark-Regressionen	61

Zusammenfassung

Diesem Diskussionsbeitrag liegt ein Forschungsprojekt zu Grunde, das die Kosten von glasfaserbasierten Anschlussnetzen aus verschiedenen Perspektiven beleuchtet. Der Fokus liegt dabei darauf, verschiedene methodische Konzepte hinsichtlich ihrer Eignung für eine Abschätzung von Kosten und Förderbedarfen zu analysieren. Dazu wurde in einem ersten Schritt eine flächendeckende, GIS-basierte bottom-up Kostenmodellierung für die NGA-Technologien FTTH P2P, FTTS und FTTC unter Rückgriff auf öffentliche, allgemein verfügbare Daten durchgeführt, die die Grundlage für die weiteren Analysen bildet. Die durchgeführten Investitions- und Kostenberechnungen wurden in der Folge für drei Fragestellungen analysiert und ausgewertet:

Im ersten Analyseteil wurde unter Rückgriff auf das Analytische Kostenmodell für das Anschlussnetz 3.0 sowie das WIK-NGA-Modell die Profitabilität eines Glasfaserausbaus bzw. des benötigten Förderbedarfs anschlussbereichsgenau unter Berücksichtigung des bestehenden Ausbaustands ermittelt. Die Ergebnisse wurden für 20 Cluster abgeleitet und zeigen für Deutschland im Greenfield-Ansatz eine Spreizung der Kosten pro Anschluss zwischen den billigsten und den teuersten Clustern in Höhe von 3,2. Ein privatwirtschaftlicher Ausbau von FTTH P2P kann für 65% der Haushalte erwartet werden, unter der Annahme, dass die durchschnittlichen Erlöse nicht wettbewerbsfähig bestritten werden. Dabei wurde ein Marktanteil in Höhe von 90% mit der Absicht unterstellt, die die maximale Reichweite eines profitablen privatwirtschaftlichen Netzausbaus abzuschätzen.

Für den Brownfield-Ansatz wurde die Versorgungslage dem Breitbandatlas mit Datenstand 04/2018 entnommen und mit den Daten der Anschlussbereiche verschnitten. Erwartungsgemäß besteht für FTTH eine Versorgungslücke von über 90% der Haushalte. CATV und FTTC decken bereits gut 65% der Haushalte ab. Überraschend war das Ergebnis der visualisierten Aufbereitung des Ausbaustands, wonach die aktuelle Breitbandverfügbarkeit für alle Technologien in den Regionen vorzufinden ist, die unseren Berechnungen zu Folge tendenziell geringe Investitionen pro Anschluss aufweisen. Im Vergleich zu FTTH sind sowohl CATV als auch FTTC aufgrund ihres höheren Ausbaustands erwartungsgemäß stärker in der Fläche verbreitet. Eine alternative Brownfield-Analyse zielte darauf ab, die Einsparpotentiale abzuschätzen, die sich bei einem FTTH P2P Ausbau aufgrund des bereits realisierten Ausbaustands für FTTC bzw. CATV realisieren lassen. Ausgehend von FTTC konnten hier maximal 20%, für CATV (approximiert durch FTTS) maximal 30% unter der Annahme ableiten lassen, dass die bereits getätigte Auslegung der Hauptkabeltrassen bereits den Anforderungen einer FTTH P2P Architektur genügt und dort lediglich die Kabelinvestitionen noch zu tätigen sind.

Die im Zuge der Vollerhebung bottom-up ermittelten Investitionsbedarfe für ein flächendeckendes Glasfaseranschlussnetz wurden im zweiten Teil für eine statistische Regressionsanalyse herangezogen, um Bestimmungsfaktoren für die Kosten von An-

schlussnetzen zu identifizieren. Es konnte gezeigt werden, dass die Proxy-Variable „Anzahl Haushalte bezogen auf die *besiedelte* Fläche des Anschlussbereichs bessere Schätzergebnisse liefert als die Anzahl der Haushalte bezogen auf die *gesamte* Fläche eines Anschlussbereichs. Das angestrebte Ziel einer erwartungstreuen Kostenschätzung konnte lediglich mit einer Regressionsanalyse unter Rückgriff auf die vollständige Datenbasis der bottom-up Modellierung erreicht werden. Die angestrebte Überführung in ein Proxy-Modell unter Rückgriff auf allgemein verfügbare Daten ist noch nicht zufriedenstellend gelungen. Zwar erklären die aufgestellten Proxy-Modelle im Minimum über 90 Prozent der Datenvariationen der Zielvariablen, jedoch besteht keines der aufgestellten Proxy-Modelle eines der üblichen Testverfahren.

Der dritte Analyseteil dieses Forschungsprojekts befasste sich mit der Clusterbildung und den ihr zugrunde liegenden Kriterien. Der Clusterbildung des WIK-NGA-Modells nach Maßgabe der Anschlussdichte und konstanter Teilnehmerstärke der Cluster wurde mit univariaten und multivariaten Verfahren verglichen. Das Ziel, exogene Variablen für die Clusterbildung zu identifizieren, die eine möglichst homogene und trennscharfe Zuordnung zu den Clustern erlauben, konnte nicht erreicht werden. Hierzu sind weitere Analysen angestrebt, die auf den Erkenntnissen der Regressionsanalyse aufsetzen sollen. Darüber hinaus konnten erste Erkenntnisse mit Anwendung von Nachbarschafts-orientierten Methoden der Clusterbildung gewonnen werden (Hot Spot – Cold Spot Analysis). Aufgrund der Relevanz des Nachbarschaftskonzepts für den regionalen Netzausbau und damit in Verbindung stehende Fragen der Profitabilität und des Förderbedarfs streben wir hierzu weitere Untersuchungen an.

Summary

The assessment of NGA-network costs' is manifold. It is the objective of this paper to present three different methodological approaches towards the determination of NGA-access costs per subscriber and to give special emphasis on the determination of profitability of private investment and the requirement for state aid in order to achieve 100% coverage of broadband access by the means of FTTH P2P. A GIS-based bottom-up cost model calculation constitutes the basis for all three parts of the analysis (Greenfield approach). Besides FTTH P2P the modelling was also conducted for FTTS G.fast and FTTC Vectoring.

The cost analysis is based on 20 clusters, each cluster representing 5% of all German households, ranked according to density. The results show a relevant spread: the highest cluster represent costs per subscriber which – in comparison to the cheapest cluster – are scaled by the factor 3.2. The model results suggest profitability for 65% of all households under the assumption that the investor is able to contract 90% of the homes passed. This extreme assumption is taken in order assess a threshold with regard to the maximum reach of a profitable NGA rollout. Under these assumptions and in absence of price competition, but regional cross-subsidisation an investor at national scale could even profitably serve the whole country with FTTH P2P.

In order to interlink the analysis with the actual market situation and level of broadband coverage, we have analysed the data of the national broadband map as available by 04/2018 and matched it with the GIS data used for bottom-up modelling. As expected, we derived a lack of coverage with regard to FTTH which amounts to more than 90% of all households. Of course, CATV and FTTC show a much better coverage and require additional investment for approx. 35% of all households only. The graphical presentation of broadband availability and investment per subscriber on a map for Germany perfectly visualised that the existing access infrastructures apparently are located in the commercially most viable regions. Of course, CATV and FTTC cover more areas than FTTH due to their higher level of coverage.

In order to assess migration path related investment needs we established a further Brownfield approach, which focussed on potential savings due to the reusage of already existing infrastructure. In fact, these saving are very limited. Taking FTTC as a starting point for migrating towards FTTH P2P we have derived a maximum of 20% of savings in relation to 100% Greenfield investment in FTTH P2P. With FTTS as a starting point, a maximum of 30% was derived. In order to achieve these savings a 100% "re"-usage of feeder trenches is required.

In the second and third part of the study a statistical analysis was conducted. First, a regression model was established in order to overcome the heavy data requirements which a bottom-up modelling is facing. In a first step, we have derived a benchmark-model in order to estimate the regionally differentiated investment per subscriber on the

basis of the complete data set available from the bottom-up model. In the following, this benchmark-model served as a template in order to derive proxy-variables which allow estimating the investment per subscriber in absence of data base stemming from the bottom-up model. Our analysis revealed that “households per developed km²” is considered to be superior over the plain proxy-variable “households per km²”. Unfortunately, we haven’t been able to establish a proxy regression model for estimating regionally differentiated subscriber costs on the basis of publicly available data. The proxy-models are able to explain more than 90% of the data variation in invest per subscriber, but did not pass the statistical tests for robustness.

A second statistical analysis dealt with the methodology for clustering. Taking the 20 clusters of the WIK NGA model as a starting point, comparisons were made with univariate and multivariate approaches. Furthermore, the (optimal) number of clusters was subject to analysis, resulting in the finding that a relevant reduction of the 20 clusters can be implemented without noticeably endangering the quality of results (especially with regard to univariate methods). With the application of a neighbouring concept a further cluster approach was analysed (Hot Spot – Cold Spot Analysis). This concept allowed to derive connected areas of remarkable size (exceeding NUTS 2) and can be considered as promising with regard to assessment of private investments’ profitability and funding requirements. Because the location of profitable access areas and their adjacency do matter with regard to business plans, we expect to be able to achieve a level of analysis that is even more realistic with regard to private investment decisions and need for state aid.

1 Einleitung

1.1 Gegenstand

Der Ausbau von Glasfaseranschlussnetzen schreitet in Deutschland im internationalen Vergleich nur langsam voran. In der politischen Diskussion hingegen ist das Ziel einer glasfaserbasierten Breitbandversorgung unumstritten. Perspektivisch sollen in ganz Deutschland Gigabit-Bandbreiten über Glasfasernetze zur Verfügung gestellt werden. Von regulatorischer Seite wird erwartet, dass die Akteure im Markt den Glasfaserausbau vorantreiben. Die Politik will dort mit Fördergeldern helfen, wo ein privatwirtschaftlicher Ausbau nicht lohnend erscheint. Der Frage nach dem für einen Glasfaserausbau erforderlichen Investitions- und Subventionsbedarf ist das WIK bereits vor knapp 10 Jahren in einem Diskussionsbeitrag nachgegangen.¹ Zwar hat sich in Deutschland in den vergangenen zehn Jahren die Versorgung mit höheren Anschlussgeschwindigkeiten verbessert, ein Ausbau von Glasfaseranschlussleitungen für eine zukunftssichere Versorgung bis zum Endkunden ist bisher jedoch bei weitem nicht flächendeckend beobachtbar.

Diesem Diskussionsbeitrag liegt ein Forschungsprojekt zu Grunde, das die Kosten von glasfaserbasierten Anschlussnetzen aus verschiedenen Perspektiven beleuchtet. Der Fokus liegt dabei darauf, unterschiedliche methodische Konzepte hinsichtlich ihrer Eignung für eine Abschätzung von Kosten und Förderbedarfen zu analysieren. Dazu wurde in einem ersten Schritt auf Basis allgemein öffentlich verfügbarer Daten eine flächendeckende, GIS-basierte bottom-up Kostenmodellierung für Deutschland durchgeführt (Vollerhebung), die die Grundlage für die nachfolgenden Analysen bildet.

Die Kosten- und Profitabilitätsberechnungen wurden in der Folge für drei Fragestellungen analysiert und ausgewertet:

- (1) Ermittlung der Profitabilität eines Glasfaserausbaus bzw. des benötigten Förderbedarfs unter Berücksichtigung des bereits bestehenden Ausbaustands.
- (2) Um zukünftig in der Lage zu sein, regional differenzierte Kosten von Anschlussbereichen unter relativ geringen Datenanforderungen schätzen zu können, wurde eine Regressionsanalyse durchgeführt. Ziel ist es dabei, geeignete Regressoren zu identifizieren, die idealerweise in Form allgemeiner, öffentlich zugänglicher Daten verfügbar sind. Die im Zuge der Vollerhebung bottom-up ermittelten Kosten dienen dabei methodisch als wahres Modell (Benchmark-Modell), um in einem zweiten Schritt die Eignung von Proxy-Variablen für die Regressionsanalyse zu finden.

¹ Jay / Neumann / Plückebaum (2011)

- (3) Der dritte Analyseteil dieses Forschungsprojekts zielt auf eine Clusterbildung und ihren Beitrag für die Identifizierung und Quantifizierung des Förderbedarfs ab. Hier wird die Brücke zu der dem WIK-NGA-Modell zugrunde liegenden Clusterbildung geschlagen. Unter Rückgriff auf die Erkenntnisse sowohl der anschlussbereichsscharfen Kostenermittlung als auch der Regressionsanalyse werden Verfahren der Clusterbildung untersucht und angewendet und dem Clusteransatz des WIK-NGA-Modells vergleichend gegenüber gestellt.

Durch das Zusammenführen dieser drei Analyseschritte soll die Belastbarkeit und Qualität von Kostenschätzungen und clusterbasierten Analysen weiter verbessert und für Anwendungsfälle mit schlechteren Datenverfügbarkeiten nutzbar gemacht werden. Zudem sollen auf dem Nachbarschaftskonzept basierende Clusteransätze und Erkenntnisse hinsichtlich ihrer Eignung sowohl für Investitionsentscheidungen als auch für die Gewährung von Fördergeldern ausgewertet werden.

1.2 Aufbau der Studie

Wir stellen unsere Ergebnisse in 5 zentralen Kapiteln der Studie dar. Zunächst wird in Abschnitt 2 die methodische Vorgehensweise aller drei Analyseansätze skizziert. Die Abschnitte 3 und 4 stellen auf die flächendeckende bottom-up Modellierung zur Bestimmung der Kosten einer Breitbandversorgung ab. Dies legt die Grundlage zum einen für die Analyse der Profitabilität und des Förderbedarfs eines Breitbandausbaus, zum anderen für die statistischen Analysen. Diese gliedern sich in Analysen zur erwartungstreuen Schätzung der Investitionen pro Anschluss auf Basis von Regressionsanalysen und in Clusteranalysen.

Abschnitt 3 skizziert die methodische Vorgehensweise der Kostenmodellierung. Im Anhang ergänzen detaillierte Informationen über die zugrunde liegenden Daten und deren Aufbereitung. In Abschnitt 4 wird erläutert, welche NGA-Technologien in die Analyse einbezogen und wie diese in den Modellen operationalisiert werden, um die Kostenunterschiede zu erfassen. Die Ergebnisse werden anhand der 20 Cluster des WIK-NGA-Modells dargestellt und strukturell mit denen der Vorgängerstudie aus 2011 verglichen. Auf Basis eines Greenfield-Ansatzes werden die Kosten von drei NGA-Technologien (FTTH P2P, FTTS und FTTC) gegenüber gestellt und es wird ein profitabler Technologie-Mix ermittelt. Unter Einbeziehung des Breitbandausbaus, Stand 2018 (gemäß Breitbandatlas), wird im Rahmen eines Brownfield-Ansatzes der noch erforderliche Investitionsbedarf für ein zukunftssicheres, flächendeckendes Glasfasernetz (FTTH P2P) ermittelt.

Die mit Hilfe der bottom-up Modellierung ermittelten anschlussbereichsgenauen Ergebnisse werden in Abschnitt 4.2.4 dazu verwendet, um zunächst ein Benchmark-Modell im Rahmen einer Regressionsanalyse abzuleiten, das sich vollständig auf die aus der Modellierung verfügbaren Daten abstützt. Daran schließt sich die Ableitung von Proxy-

Variablen an, mit Hilfe derer eine Kostenschätzung ermöglicht werden soll, die sich von der Notwendigkeit aufwendiger Datenaufbereitung und bottom-up Modellierung befreit und dennoch in der Lage ist, erwartungstreue Ergebnisse für die durchschnittlichen Investitionen von NGA-Anschlussleitungen zu liefern.

In Abschnitt 1 werden die Erkenntnisse aus den vorangegangenen Abschnitten aufgegriffen, um mit Hilfe von Verfahren zur Clusterbildung einen vereinfachten Ansatz zur Bestimmung regional differenzierter Investitionen pro Anschlussleitung zu identifizieren. Darüber hinaus wird mit einem alternativen Clusteransatz versucht, dem Aspekt räumlich zusammenhängender Einheiten Rechnung zu tragen. Diesem wird sowohl im Kontext von privatwirtschaftlichen Investitionsentscheidungen als auch der Gewährung von Fördergeldern eine relevante Rolle zugesprochen.

Abschnitt 7 fasst die Erkenntnisse dieser Studie zusammen und gibt einen Ausblick auf noch offene Fragen, die im Rahmen dieses Forschungsprojektes nicht oder nicht abschließend analysiert werden konnten.

2 Methodische Vorgehensweise, Modellierungsrahmen und zentrale Annahmen

2.1 Methodische Vorgehensweise

2.1.1 Flächendeckende bottom-up Modellierung

Unter Rückgriff auf eine GIS-basierte bottom-up Modellierung sollen in Form eines Greenfield-Ansatzes die Kosten eines flächendeckenden Glasfaserausbaus bestimmt werden. Eine vergleichbare Studie war bereits 2011 veröffentlicht worden (nachfolgend auch als Vorgängerstudie bezeichnet).² Die Kostenberechnung basierte seinerzeit auf einer Schätzung, wobei die Anschlussbereiche in 20 Cluster annähernd gleicher Anschlusszahlen unterteilt worden waren und für jedes dieser Cluster eine repräsentative Kostenschätzung durchgerechnet wurde. Bei der jetzt angestrebten GIS-basierten, anschlussbereichsscharfen Berechnung ist eine höhere Genauigkeit angestrebt worden, mit der Erwartung, differenzierte Erkenntnisse für den Förderbedarf ermitteln zu können.

In Analogie zu der Vorgängerstudie findet auch hier das vom WIK entwickelte NGA-Modell Anwendung, welches aufbauend auf einem bestehenden Mengengerüst, clusterbezogenen Kosten und Profitabilität von Anschlussnetzen berechnet.³ Das Analytische Kostenmodell für das Anschlussnetz Version 3.0 des WIK (WIK-TAL-Modell AKM-AN 3.0)⁴ wird vorbereitend herangezogen, um das benötigte Mengengerüst sämtlicher Anschlussbereiche⁵ für die Befüllung des WIK-NGA-Modells zu generieren. Gegenstand der Modellierung waren dabei drei verschiedene NGA-Architekturen

- Fibre to the Home Punkt-zu-Punkt (FTTH P2P),
- Fibre to the Street (FTTS)
- Fibre to the Curb (FTTC)

Für die Modellberechnungen waren die folgenden Arbeitsschritte erforderlich:

² Jay / Neumann / Plückebaum (2011): Diese Studie hatte eine Abschätzung des Investitions- und Förderbedarfs für einen flächendeckenden Glasfaserausbau in Deutschland zum Gegenstand und kann insofern auch als Vorgängerstudie zu dem ersten Analyseteil dieses Diskussionsbeitrags bezeichnet werden.

³ Wir haben dieses Modell bereits in verschiedenen anderen Projekten verwendet und stetig weiterentwickelt. Relevante Vorgängerprojekte sind unter anderem: Elixmann / Ilic / Neumann / Plückebaum (2008), Ilic / Neumann / Plückebaum (2009), Ilic, D. / Neumann, K.-H. / Plückebaum, T. (2009a), Jay, S. / Plückebaum, T. / Ilic, D. (2010), Doose / Elixmann / Jay (2009), Jay, S., Neumann, K.-H., Vogelsang, I., et al. (2011), Neumann, K.-H. / Plückebaum, T. / Zoz, K. (2017), Braun, M., Wernick, C., Plückebaum, T., Ockenfels, M. (2020), Lucidi, S., Ockenfels, M., Sörries, B. (2020).

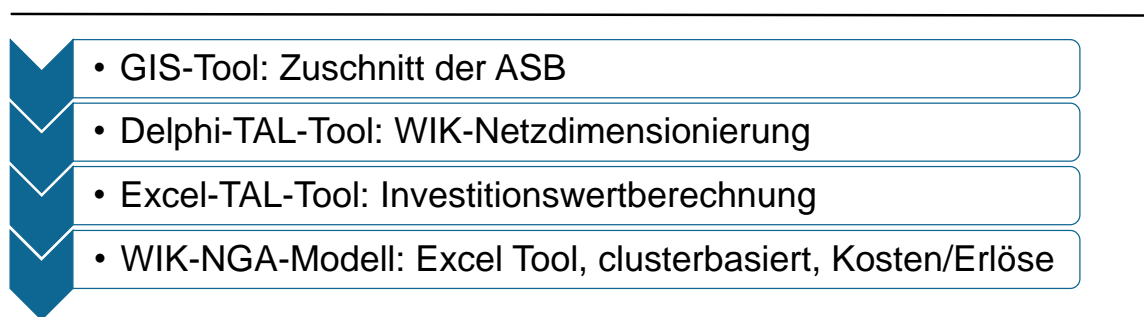
⁴ Kulenkampff, G. / Plückebaum, T. / Zoz, K. (2019)

⁵ In der Vorgängerstudie stand nur ein auf Cluster aggregiertes Mengengerüst im WIK-NGA-Modell zur Verfügung.

1. Aufbereitung der Anschlussbereichsdaten auf Basis von öffentlich verfügbaren Geodaten für die Basisbefüllung des WIK-TAL-Modells zur Erzeugung von Datensätzen je Anschlussbereich: MPoP (Metropolitan Points of Presence, abgeleitet aus den bisherigen Hauptverteilerstandorten (HVt)), Rasterdaten zu Gebäuden und Haushaltszahlen zur Ableitung der EVz-Standorte und zugehöriger Anschlussnachfrage, Festlegung von Anschlussbereichsgrenzen durch Zuordnung der Gebäude zu MPoP sowie Segmentierung von Straßen als Grundlage für eine bottom-up Netzauslegung und Bestimmung des anlagenbezogenen Mengengerüsts.
2. Bottom-up Modellierung unter Rückgriff auf das WIK-TAL-Modell (AKM-AN 3.0) zur Berechnung des anlagenbezogenen Mengengerüsts pro Anschlussbereich (ASB).
3. Transformation der Outputgrößen des WIK-TAL-Modells in die Eingangsgrößen des WIK-NGA-Modells auf Aggregationsebene der Anschlussbereiche.⁶
4. Festlegung der NGA-Modell-Eingangsparameter (Preisparameter) für die Berechnung des Investitionswertes und der Kosten.
5. Anschlussbereichsgenaue Modellberechnungen von Kosten und Profitabilität durch Anwendung des WIK NGA Modells für die Technologien FTTH P2P, FTTS und FTTC⁷.

Die nachfolgende Grafik illustriert das Ineinandergreifen der verschiedenen Analysewerkzeuge.

Abbildung 2-1: Verwendete Tools der modellbasierten Kostenermittlung



Quelle: WIK

⁶ Aufgrund einer Erweiterung des WIK-NGA-Modells steht diese Option zur Verfügung, so dass eine anschlussbereichsgenaue Investitionswert- und Kostenberechnung im NGA-Modell umgesetzt werden kann und eine Aggregation auf Cluster erst im Zuge der Aggregation der Ergebnisse auf Basis der Kosten pro Anschluss erfolgt.

⁷ In welchem Vorläuferprojekt wurden diese Technologien erläutert?

Mit den Modellberechnungen und der Ableitung anschlussbereichsgenauer Kosten wird eine Datenbasis für die drei Analyseaufgaben dieses Diskussionsbeitrags geschaffen, die in den nächsten drei Unterabschnitten skizziert werden.

2.1.2 Glasfaseranschlussnetze – Stand der Versorgung, Profitabilität und Förderbedarf eines flächendeckenden Breitbandausbaus

Unter Rückgriff auf die Ergebnisse der GIS basierten bottom-up Modellierung wird der Netzausbau auf Basis des NGA Modells bewertet.

Methodisch wird dabei in mehreren Schritten vorgegangen. Zunächst erfolgt eine Kostenbestimmung pro Anschlussbereich. Erst in der Folge wird eine Aggregation auf die 20 Dichtecluster des WIK-NGA-Modells vollzogen. Die Berücksichtigung der Erlösseite erfolgt für das gesamte Versorgungsgebiet einheitlich und differenziert lediglich nach Leistungsfähigkeit des Anschlussnetzes. Kosten und Erlöse werden bezogen auf die 20 Dichtecluster des NGA-Modells gegenübergestellt und hinsichtlich ihres Förderbedarfs ausgewertet.

Der bestehende Versorgungsstand des Breitbandausbaus wird erst im letzten Schritt der Analyse berücksichtigt. Dazu wird auf Daten des Breitbandatlas zurückgegriffen (Stand 04/2018). Um die Einsparungen im Vergleich zu einem flächendeckenden Neuausbau zu ermitteln, wird zum einen auf die durchschnittlichen Kosten pro Anschluss abgestellt. Zum anderen wird versucht, aus Perspektive des Netzaufbaus abzuschätzen, welche Investitionen bei dem gegebenen Versorgungsstand noch zu tätigen sind, um eine in Deutschland flächendeckende Bereitstellung von sog. Very High Capacity Networks (VHC) zu erreichen. Basis für diese Analyse ist die Auswertung und Überführung der Informationen des Breitbandatlas in die anschlussbereichs- und clusterbezogene Analyse.

Die Ergebnisse werden auch dahingehend ausgewertet, ob ein regionaler Verzicht auf die leistungsfähigeren und teuren Anschlussnetztechnologien den Umfang der profitablen Versorgung zu erhöhen vermag. Zu diesem Zweck wird der Frage nach einem optimalen Technologiemix für den flächendeckenden Ausbau nachgegangen.

2.1.3 Regressionsanalyse – Geeignete Schätzer für die Kosten von Anschlussnetzen

Die anschlussbereichsgenauen Ergebnisse der Kostenmodellierung bilden die Grundlage für die statistische Analyse der Kosten von Anschlussnetzen. Unter Rückgriff auf gängige statistische Verfahren werden die wesentlichen Einflussfaktoren des FTTH P2P-Glasfaserausbaus ausgewertet und identifiziert. Die Stärke des jeweiligen Zusammenhangs wird empirisch quantifiziert.

Ziel ist es, ein belastbares ökonometrisches Modell aufzustellen, das erwartungstreue Ergebnisse prognostiziert. Dazu wird in einem ersten Schritt ein Schätzmodell abgeleitet, welches sich der Daten bedient, die für die flächendeckende bottom-up Modellierung zur Verfügung standen. Mit dem für die Gesamtheit der Anschlussbereiche abgeleiteten und validierten „wahren Modell“ (im Weiteren auch als „Benchmark-Modell“ bezeichnet) wird in einem weiteren Schritt das Ziel verfolgt, das Modell auf unterschiedliche räumliche Abgrenzungen flexibel und dynamisch einzusetzen zu können. Dazu wird versucht, die im „wahren Modell“ identifizierten Einflussfaktoren über öffentlich verfügbaren Daten zu operationalisieren. Die Güte dieses zweiten, auf Basis von öffentlichen Daten formulierten Schätzmodells wird analysiert und der des wahren Modells vergleichend gegenüber gestellt. Die Datengrundlage, Vorgehensweise und Ergebnisse der Untersuchung werden im Abschnitt 4.2.4 dargestellt.

2.1.4 Clusterbildung

Das NGA-Modell ist originär auf eine clusterbasierte Kosten und Profitabilitätsanalyse ausgelegt.

Die Clusterbildung und damit auch Subventionsbedarfsermittlung erfolgt im WIK-NGA-Modell für 20 Cluster, wobei jedes Cluster 5% aller Anschlüsse umfasst und die Zuordnung der Anschlussbereiche nach Maßgabe der Besiedlungsdichte erfolgt.

Da das WIK NGA Modell infolge einer Weiterentwicklung nicht mehr auf eine maximale Anzahl von Clustern beschränkt ist, eröffnet sich die Möglichkeit zur Analyse und Bewertung verschiedener Clusteransätze.

Im Rahmen dieses Abschnitts wird die bisherige Clusterdefinition auf Basis der Besiedlungsdichte und einer Gleichverteilung der Haushalte über die Cluster aufgehoben und der Frage nachgegangen werden, welche Kriterien der Clusterbildung es erlauben, eine möglichst trennscharfe Differenzierung von Anschlussbereichen nach Maßgabe der Investitionen pro Anschluss zu leisten. Dazu werden sowohl alternative Dichtemaße herangezogen als auch statistische Verfahren, die auf eine möglichst homogene und trennscharfe Abgrenzung abzielen (k-means-Verfahren). Unter Rückgriff auf die aus der bottom-up Modellierung vorliegenden Investitionen pro Anschluss soll eine idealtypische Clusterbildung abgeleitet werden. Das k-means-Verfahren findet dabei auch für die Frage der optimalen Clusterzahl Anwendung. Multivariate k-means-Verfahren werden insbesondere herangezogen, um die aus der Regressionsanalyse abgeleiteten Regressoren für ihre Eignung zur Clusterbildung zu überprüfen.

Clusterbildung mit dem Fokus auf regional zusammenhängende Gebiete wird unter Rückgriff auf Methoden der Hot Spot / Cold Spot Analyse analysiert.

3 Flächendeckende bottom-up Modellierung

3.1 Modified scorched node Ansatz

Für die Auslegung neuer, rein glasfaserbasierter Anschlussnetze gibt es – anders als bei hybriden Architekturen, die zum Teil noch auf der kupferbasierten Anschlussleitung aufsetzen – a priori keine Beschränkungen in der Größe und Abgrenzung von Anschlussbereichen. Entsprechend erscheint es naheliegend für die Modellierung effizienter Investitionen und Kosten pfadabhängige Restriktionen, denen sich im Markt agierende Netzbetreiber ausgesetzt sehen, zu vernachlässigen und einen sogenannten scorched earth Ansatz⁸ zu verfolgen.

Im Rahmen dieser Studie wurde jedoch aus den folgenden Gründen von einem scorched earth Ansatz Abstand genommen:

- Die Bildung von Anschlussbereichen ist für die Größe der Fläche und Besiedlung Deutschlands ein (zeit-)aufwendiges Optimierungsproblem, insbesondere wenn weder Anzahl der Anschlussbereiche noch Standorte für die MPoP als bekannt angenommen werden können.
- Bestehende Anschlussbereiche reflektieren bereits Besiedlungsstrukturen. Gebäude und Liegenschaften in räumlicher Nähe sind so zu Anschlussbereichen zusammengefasst, dass es in vielen Fällen aufgrund der räumlichen Zuordnung keine sinnvollen Alternativen gibt.
- Selbst alternative Netzbetreiber, die bisher über kein eigenes Anschlussnetz verfügen, sehen sich einer gewissen Pfadabhängigkeit gegenüber als dass ihre eigenen Kern- und Aggregationsnetze auf eine Anschlussfähigkeit an bestehende Anschlussnetze ausgelegt wurden.

Stattdessen wird daher in dieser Studie ein sogenannter modified scorched node Ansatz verfolgt: Es wird auf die öffentlich bekannten, bisherigen Hauptverteilerstandorten (HVt) in Deutschland zurückgegriffen (scorched nodes)⁹. Um eine Investitionswert- und Kostenberechnung für Anschlussnetze durchführen zu können, ist es darüber hinaus erforderlich, für die Gesamtheit der Gebäude und Liegenschaften, die vom Breitbandausbau abgedeckt werden sollen, eine Zuordnung zu einem MPoP zu vollziehen. Erst über diese Zuordnung kann das Optimierungsproblem für eine bottom-up Modellierung effizienter Kosten hinreichend beschrieben werden. Die im Netz des flächendeckenden Anbieters bestehende räumliche Abgrenzung der Anschlussbereiche ist nicht öffentlich bekannt und kann daher für unsere Modellierungszwecke nicht angewendet werden.

⁸ Der scorched earth Ansatz stellt im regulatorischen Kontext das Gegenstück zum scorched node Ansatz dar, welcher in der regulatorischen Praxis die netzseitigen Abschlusspunkte des Anschlussnetzes („nodes“) und in der Regel auch die Zuordnung der Anschlüsse zu den Abschlusspunkten (auch als MPoP „Metropolitan Point of Presence“) bezeichnet.

⁹ <http://selke.de/privates/hvt-standorte/>, Veröffentlichungsdatum 13.4.2013.

Diesen Umstand erachten wir dabei nicht als problematisch, sondern sehen darin auch den Vorteil, dass für die bottom-up Modellierung – von bestehenden Ineffizienzen aufgrund pfadabhängiger Realisierung von Anschlusslinien losgelöst – eine eigene Zuordnung implementiert werden kann. Dabei wird wie folgt vorgegangen:

- HVt-Standorte werden als MPoP und damit Ausgangspunkt für die Abgrenzung der ASB herangezogen,
- Gebäude werden dem nächstgelegenen MPoP über ein Minimum-Distanz-Kriterium zugeordnet, wobei methodisch das kürzeste Routing über den GIS-Straßenlayer Anwendung findet.

Diesen methodischen Ansatz bezeichnen wir als „modified“ scorched node Ansatz. Aufgrund der bottom-up erzeugten Zuordnung der Gebäude (nachfolgend auch als Endverzweiger, EVz, bezeichnet) zu den MPoP weicht dieser Ansatz vom aus der regulatorischen Praxis bekannten scorched node Ansatz ab und wird daher als „modified“ bezeichnet.

Die verwendete Datenbasis ist in der nachfolgenden Tabelle 3-1 dokumentiert. Die für die Abbildung der Anschlussnachfrage maßgebliche Datenbasis betrifft die Gebäude und Haushalte und ist bezüglich letzterer mit einer Genauigkeit eines 250m x 250m Rasters berücksichtigt und repräsentiert den Stand von 2018.

Tabelle 3-1: Datenquellen

Daten	Quelle	Bemerkung
HVt-Standorte	Netzressource auf Grundlage der BMWI DSL-Hauptverteilerliste; http://selke.de/privates/hvt-standorte/	7971 Datensätze mit Adressdoppelungen nach Schnittstellen, Generierung von 7871 HVT-Standorten
Straßen	TeleAtlas 2018/04	Externe Beschaffung und Aufbereitung des Gesamtdeutschen Straßenlayers von TeleAtlas Release 2018/04
Gemeinde- und Verwaltungsgrenzen	Bundesamt für Kartographie und Geodäsie	VG250-EW (Kompakt) Verwaltungsgebiete der Bundesrepublik Deutschland mit Einwohnerzahlen (Stand 1.01.2018)
Gebäude	Adresspunkte TOM TOM	rd. 22.7 Mio. Gebäudepunkte
Haushalte	AZ-Direct Haushalts- und Gebäuderaster (250mx250m), basierend auf Daten der Bertelsmann-Gebäude-Datei, Stand 2018	41,5 Mio. Haushalte Deutschlands

In Anhang 1: Datenquellen und Modellierungsannahmen zum Versorgungsgebiet findet sich eine detaillierte Darstellung zu der durchgeführten Aufbereitung der GIS-Daten sowie der Ableitung der anschlussbereichsbezogenen Daten als Grundlage für die Modellberechnungen.

3.2 Gegenstand und zentrale Annahmen der Investitionswert- und Kostenberechnung

Für die Investitionswertberechnung wurde im Rahmen dieses Forschungsprojekts sowohl auf das analytischen Kostenmodells für das Anschlussnetz 3.0 (AKM-AN 3.0) als auch das WIK-NGA-Modell zurückgegriffen. Ersteres bildet die Grundlage für die GIS basierte Ableitung des Mengengerüsts. Letzteres findet Anwendung für die Kostenberechnung auf Anschlussbereichsebene sowie die clusterbasierte Bestimmung der Profitabilität und des Förderbedarfs. Im Anhang 2: Modellierungsrahmen) findet sich eine ausführliche Darstellung der Funktionsweise der angewendeten Modelle.

3.2.1 Betrachtete Anschlussnetztechnologien

Mit glasfaserbasierten Anschlussnetzen können die politisch formulierten Breitbandziele erreicht werden und den Anforderungen auch perspektivisch genügen. Mit einer dedizierten Glasfaseranbindung eines jeden Endkunden auf Basis einer FTTH P2P Architektur steht ein Potential zur Verfügung, welches sich durch den Einsatz aktiven Equipments für sehr hohe Datenübertragungsraten aufrüsten lässt, und den vorstellbaren Bedarf eines Haushaltes an Datenübertragungsrate mehr als überschreiten kann.¹⁰ – Aktuell aus der Praxis bekannte, glasfaserbasierte Angebote umfassen Datenübertragungsraten bis zu 500 Mbitps oder auch 1Gbitps.

Wir haben die FTTH P2P Architektur für die Bestimmung der Kosten- und Förderbedarfs zu Grunde gelegt, da diese uneingeschränkt den Anforderungen an Very High Capacity Networks (VHC) genügt und als die zukunftssicherste Architektur auch bei stark steigenden Bandbreitenanforderungen angesehen werden kann.

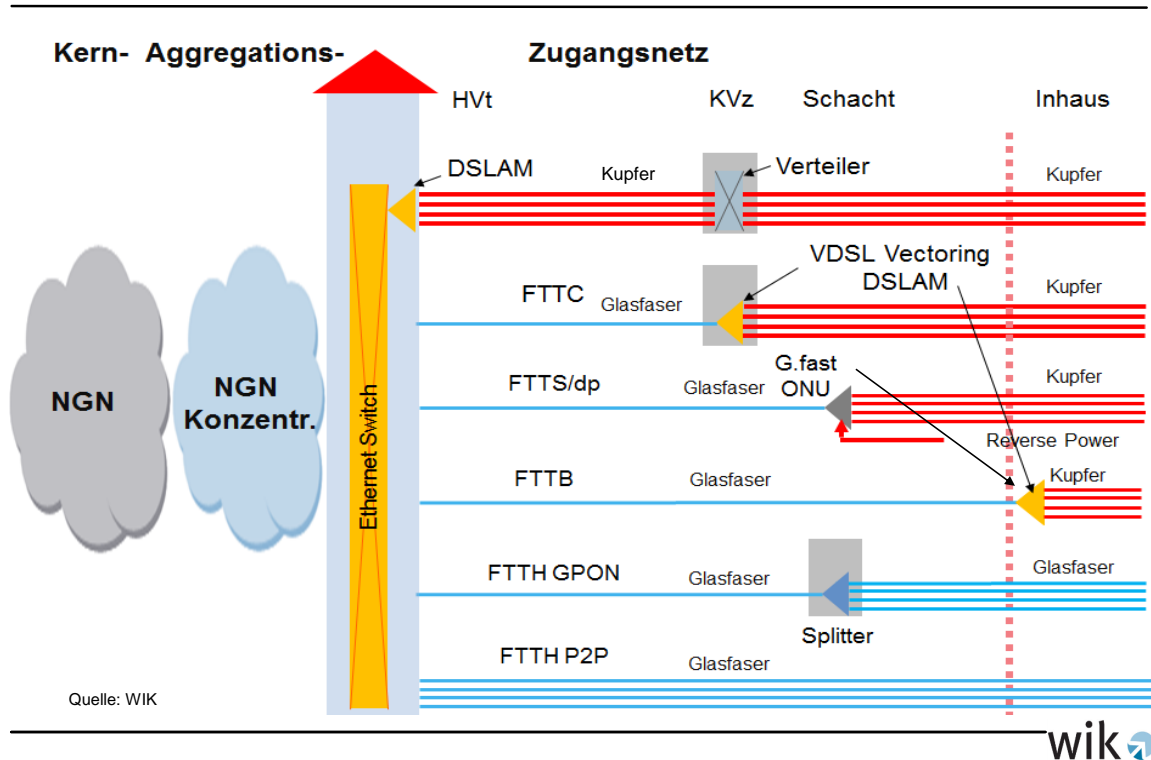
Aufgrund der empirisch beobachtbaren Migration der bestehender kupferbasierten Anschlussnetze zu hybriden und möglicherweise auch rein glasfaserbasierten Anschlussnetzen haben wir uns entschieden, ergänzend G.fast- sowie Vectoring als weitere NGA-Technologien mit in die Betrachtung einzubeziehen. Dabei wird G.fast durch eine FTTS-Topologie und Vectoring durch eine FTTC-Topologie beschrieben.

Die nachfolgende Abbildung 3-1 gibt einen Überblick zu den aktuell relevanten NGA Netzarchitekturen und illustriert auch für die in dieser Studie ausgewählten Varianten, in welchem Umfang das bestehende Kupfernetz noch zum Einsatz kommt (die rot dargestellten Linien repräsentieren die Kupferleitung). Bezüglich der technischen Leistungs-

¹⁰ Optische Verfahren erlauben Datenübertragungsraten von 400 Gbitps pro Faser. Mit Einsatz von Wellenlängenverfahren kann die Kapazität einer Glasfaser auf den Bereich von Terrabitps ausgeweitet werden. Entsprechendes Equipment ist bereits marktüblich, kommt aber natürlich nicht im Bereich der Anschlussnetze als vielmehr in den Kernnetzen auf Transportwegen mit hohen Nachfragen zur Anwendung.

merkmale der verschiedenen Anschlusstechnologien verweisen wir auf den WIK-Diskussionsbeitrag Nr. 457.¹¹

Abbildung 3-1: NGA Netzarchitekturen im Überblick



Quelle: WIK

Für die Auswahl von FTTS war neben dem Migrationsaspekt aus Perspektive der Kupfernetze auch die „Nähe“ zum bestehenden CATV-Netz, basierend auf Koaxialkabeln, Motivation. Die sog. Koaxialkabel-Netze sind gut geeignet für eine hochfrequente Übertragung, die jedoch herkömmlich nur für die Übertragung des TV-Signals im downstream ausgerichtet waren. Die bidirektionale Datenkommunikationsform für Koaxialkabel wurde mit dem DOCSIS Standard¹² festgelegt und für die Nutzung der CATV-Netze für den Breitbandzugang weiterentwickelt, wobei der Standard DOCSIS 3.2 (oder DOCSIS 3.1 full duplex) auch Datenübertragungsraten im upstream von bis zu 10 Gbps

¹¹ Plückerbaum, T., Ockenfels, M. (2020).

¹² Data over Cable Service Interface Specification fing bei Release 1.0 an; heute implementiert sind noch Reste von Release 2.1; ganz überwiegend verbreitet ist Release 3.0; mit der Implementierung von Release 3.1 wurde bereits begonnen. Das Kabel ist ein Shared Medium, ein zwischen allen angeschlossenen Kunden in seiner Nutzung geteiltes Übertragungsmedium. Die Zugriffsrechte auf diese gemeinsam genutzte Übertragungskapazität koordiniert ein sog. Cable Modem Termination System (CMTS) in Zusammenspiel mit den Cable Modem der Teilnehmer, in dem es jederzeit aktiv die Übertragungsrechte für die einzelnen Nutzer einräumt. Sofern nur ein Teilnehmer senden und empfangen will, steht die gesamte Kapazität dieser beiden gerichteten Kanäle (Down und Up) diesem Teilnehmer zur Verfügung. Andernfalls muss er sich die Kapazität mit den anderen kommunikationswilligen Teilnehmern teilen.

erlaubt. Allerdings sind hier die maximalen Teilnehmerzahlen je fibre node auf etwa 40 bis 50 begrenzt.¹³

Da das NGA Modell keine explizite Modellierung von CATV vorsieht, wurde CATV (DOCSIS 3.0) mit der Modellierung von FTTS approximiert. Sowohl die Topologie, als auch die jeweilige Anzahl Anschlüsse im Verzweigersegment sind vergleichbar, so dass wir diese methodische Vorgehensweise als angemessen erachten. Gleiches gilt für die Analogie zwischen CATV (DOCSIS 3.1DF) und FTTH.¹⁴ Im Zuge der weiteren Analyse wurde auf diese Vergleichbarkeit zurückgegriffen, um unter Rückgriff auf die Ergebnisse für FTTS die noch verbleibenden Ausbaukosten eines CATV Anschlussnetzes hin zu einer Punkt-zu-Punkt-Anbindung der Endkunden durch Glasfaser abzuschätzen.

In der nachfolgenden Tabelle sind die im Rahmen dieser Studie modellierten Glasfaserarchitekturen vergleichend gegenüber gestellt.

Tabelle 3-2: Modellierte Glasfaserarchitekturen

Kürzel	Glasfaser-Topologie zwischen Endkunde und Zentrale	Aktive Technik zur Beleuchtung der Fasern im MPoP	Kommentar
FTTH P2P	Punkt-zu-Punkt	Ethernet	
FTTS	Hybrid (Punkt-zu-"Multi"-Punkt)	G.fast	Dezentrale DSLAM zwischen MPoP und Kundenstandort Approx. von CATV DOCSIS 3.0
FTTC	Hybrid (Punkt-zu-"Multi"-Punkt)	Vectoring	Dezentrale MSAN am KVz zwischen MPoP und Kundenstandort

Quelle: WIK

3.2.2 Parametergesteuerte Netzauslegung

Um die topologischen als auch netztechnischen Besonderheiten der verschiedenen Anschlussnetztechnologien bei der Modellierung zu berücksichtigen, wurden an verschiedenen Stellen des Modells spezifische Annahmen getroffen. Wie in Abschnitt 3.2.1 ausgeführt, haben die eingesetzten Medien unterschiedliche Beschränkungen hinsichtlich der Reichweite (Länge der Anschlusslinie) oder auch Kapazitäten (maximale Zahl

¹³ Bis zum Standard DOCSIS 3.1 ist ein asymmetrisches Verhältnis von Up- und Downstream von 1:10 im Standard vordefiniert, was bei den heutigen Verschiebungen der Nachfrage in Richtung Symmetrie zu Engpässen führen kann. DOCSIS 3.0 begrenzt den Upstream auf maximal ca. 120 Mbps shared, bei DOCSIS 3.1 wird dies auf 1,2 Gbps erweitert.

Im Vergleich zu dem Vorgängerstandard DOCSIS 3.1 erfordert dies eine größere Anzahl Fibre Nodes. Das steuernde, aktive Equipment (cable modem termination system, CMTS) muss dazu kapazitiv aufgerüstet werden und auch die teilnehmerseitigen cable modem müssen hierfür erneuert werden. Bei DOCSIS 3.2 (oder DOCSIS 3.1 full duplex) werden die Upstream- und Downstream-Kanäle nicht mehr frequenztechnisch getrennt, sondern es wird nahezu der gesamte Frequenzbereich gemeinsam von beiden Kommunikationsrichtungen gleichzeitig genutzt.

¹⁴ Hier sind verschiedene Studie aus der Vergangenheit zu nennen, z.B. Ilic, D.; Jay, S.; Plückerbaum, T.; Stamm, P. (2013) oder auch Plückerbaum, T., Eltges, F., Ockenfels, M. (2019)

von Anschlüssen pro Splitter oder DSLAM). Zentrale Stellgrößen, die für die Abbildung der gewählten Anschlusstechnologien und ihrer Restriktionen herangezogen wurden, sind

- a) Erstreckung der Verzweigerbereiche, worüber die Länge des Kabels im Verzweigersegment gesteuert werden soll;
- b) maximale Anzahl Anschlüsse, die am Verzweigerknoten versorgt werden kann.

Bezüglich der Erstreckung der Verzweigerbereiche haben die folgenden Überlegungen die Annahmen motiviert:

FTTC

Bei der maximalen Anzahl von Anschlüssen pro Verzweigerbereich haben wir uns an den üblichen Kapazitäten des MSAN-Equipments orientiert und diese für FTTC mit 384 vorgegeben.

Bzgl. der Erstreckung des Verzweigerbereichs bei FTTC war zu berücksichtigen, dass die mit Hilfe der Vectoring-Technologie erreichbaren hohen Bandbreiten nur beschränkte Längen der Kupferdoppelader vom Endkunden bis zum MSAN tolerieren. Wir gehen hier von einer maximalen Länge in Höhe von 500 m aus. Modelltechnisch wird dazu der Standort des Kabelverzweigers (und damit auch MSAN) im Zentrum des Verzweigerbereichs gewählt und ein maximaler Radius von 360 m angesetzt. Dabei gehen wir von einem Umwegfaktor von knapp 1,4 aus.

FTTS (G.fast)

Bei FTTS gelten im Vergleich zu FTTC strengere Beschränkungen, da der (G.fast) DSLAM noch näher zum Kunden platziert werden muss und somit letztlich auch die Zahl der maximalen Anzahl von Anschlüssen pro DSALM und damit Verzweigerbereich geringer ausfallen. Wir haben Kapazitäten für den G.fast DSLAM mit 48 Anschlüssen vorgegeben und die Erstreckung des Verzweigerbereichs bei FTTS mit einem Radius von 180 m beschränkt, um maximale Längen des Kupferkabels von 250 m einhalten zu können.¹⁵

FTTH P2P

Für FTTH P2P haben wir die Kapazitätsbeschränkung bzgl. der maximalen Anzahl von Anschlüssen ggü. den anderen beiden Technologien mit 800 Anschlüssen pro Verzweigerbereich deutlich höher ausgelegt. Hier greifen Kapazitäten von Faserverteilern, an denen die teilnehmeranschließenden Glasfaserkabel (ein Kabel pro erschlossenes Gebäude) gespleißt und auf ein großes Glasfaserkabel aggregiert werden.

¹⁵ Auch hier gehen wir von einem Umwegfaktor von knapp 1,4 aus.

Da für die FTTH P2P Architektur bei der von uns zu Grunde gelegten Erstreckung der Anschlussbereiche das Glasfaserkabel keine relevanten Längenbeschränkungen aufweist, wurden auch hier die im Modell gewählten Beschränkungen im Vergleich zu FTTC und FTTS gelockert. Der maximale Radius der Verzweigerbereiche wurde auf 400 festgelegt und die Platzierung des Faserverzweigers so gewählt, dass die Entfernung zum MPoP minimiert wird. Somit werden auch Verzweigerkabelnängen von mehr als 600 m toleriert ($400 \text{ m} * \text{Umwegfaktor}$).

3.2.3 Technologiespezifische Eingangsparameter

Die betrachteten NGA-Technologien wurden auch mit Blick auf die relevanten Unterschiede in den Investitionsparametern analysiert. Zum einen finden mit Blick auf das aktive Equipment spezifische Preisparameter für G.fast DSLAM, MSAN oder Ethernet-Konzentratoren Anwendung. Technologie neutral sind demgegenüber Kabelpreise oder auch die Preise für die Installation von Glasfasern und Spleißen.

Ein relevanter Unterschied bei den Investitionen für FTTS und FTTC steht in Verbindung mit der Installation des aktiven Equipments „im Feld“ und der dafür erforderlichen Stromversorgung. Die Investitionen, die für die Realisierung des Stromanschlusses bei FTTC erforderlich sind, finden in den Investitionsparametern Berücksichtigung. Für FTTS wird hingegen darauf abgestellt, dass die Stromversorgung über das den Endkunden anschließende Kupferkabel realisiert werden kann (Reverse Powering). Aufgrund der geringen Leistungsaufnahme des DSLAM ist dies technisch möglich und im Standard vorgesehen, und wird in der Praxis auch entsprechend umgesetzt. Dieser Lösungsansatz reduziert die Standortkosten für das aktive Equipment bei FTTS relevant.

Bezüglich der Tiefbauinvestitionen unterstellen wir für die Verlegung der Glasfaserkabel bei den hybriden Anschlusstechnologien den Einsatz von erdverlegten Mikrorohrverbänden. Lediglich für die übrigen 10% wird eine reine Erdverlegung unterstellt. Demgegenüber sehen wir für FTTH P2P ausschließlich die Verwendung erdverlegter Mikrorohrverbände und aufgrund des Platzbedarfs im Hauptkabelsegment eine Mikrorohr-in-Rohr Verlegung vor.

Die Kosten der Trassenherstellung sind technologieunabhängig. Lediglich der Platzbedarf für die Kabel, getrieben durch die erforderlichen Fasern sowie Verlegetechnik, bestimmt die Grabengröße. Die zur Anwendung gelangenden Trassenpreise bewegen sich dabei zwischen gut 50€/m und knapp 200€/m, wobei letztere nur in sehr dicht besiedelten Gebieten, nahe dem MPoP Anwendung finden.

3.2.4 Weitere anschlussbereichsbezogene Investitionsparameter

Die Modellberechnungen erfordern eine Vielzahl weiterer Eingangsparametern. Wir beschränken uns darauf, diese nur punktuell zu beschreiben und verzichten weitestgehend auf eine Quantifizierung. Detailliertere Ausführungen finden sich im Anhang 2: Modellierungsrahmen.

3.2.5 Sonstige Bestandteile der Kostenberechnung

Um die Kostenmodellierung für eine Bewertung der Wirtschaftlichkeit des Glasfaserausbaus durchführen zu können, ist es erforderlich, netzunterstützende Investitionen und OPEX zu bestimmen, die ermittelten Investitionen in Kosten zu transformieren sowie die sonstigen upstream Kosten zu berücksichtigen. Diese Schritte werden innerhalb des NGA Modells vollzogen.

Methodisch wird dabei dergestalt vorgegangen, dass indirekte (netzunterstützende) Investitionen sowie Betriebskosten (OPEX) unter Rückgriff auf anlagebezogene Zuschlagsfaktoren auf die netzelementbezogenen Investitionen bestimmt werden. Die Festlegung der mark-up Faktoren wurde in dieser Studie technologieunabhängig vorgenommen. Netzbezogene upstream Kosten werden als durchschnittliche Kosten pro Kunde veranschlagt. Bei den Kosten des Konzentrations- und Kernnetzes wird dazu die Datenübertragungsrate in der Hauptverkehrsstunde (busy hour) als Kostentreiber berücksichtigt. Gleiches gilt für Vertriebs- und Angebotskosten (Retailkosten).

Für die beiden hybriden Anschlusstechnologien FTTC und FTTS wurde darauf verzichtet, die Investitionen in das kupferbasierte Verzweiger- und Hausanschlusssegment in Investitionen zu transformieren. Hier wurde stattdessen das preisregulierte Vorleistungsentgelt für die KVz-TAL zur Anwendung gebracht. Für FTTS wurde der Betrag dabei auf 2/3 des Vorleistungsentgelts der KVz-TAL reduziert. Lediglich für FTTH P2P mussten die bottom-up ermittelten Verzweigerbereichs- und Hauszuführungsinvestitionen in Kosten transformiert werden.

Auch wurden für FTTH P2P die Investitionen für gebäudeinterne Verkabelung sowie das Endkundenequipment ergänzt. Dafür wurde ein durchschnittlicher Investitionsbetrag von insgesamt knapp 250€ pro kontrahiertem Anschluss veranschlagt.

Den Modellrechnungen liegt die Annahme von 100% homes passed zu Grunde.

4 Glasfaseranschlussnetze – Stand der Versorgung, Profitabilität und Förderbedarf eines flächendeckenden Breitbandausbaus

Aus der bottom-up Modellierung und der Anwendung des WIK NGA-Modells konnten für die Gesamtheit von 7871 Anschlussbereichen die Kosten ermittelt werden.

Die Ergebnisdarstellung des NGA-Modells erfolgt auf Basis des ursprünglich implementierten Clusteransatzes, demzufolge Anschlussbereiche in Abhängigkeit der Teilnehmerdichte je km² zusammengefasst werden. Dem Ansatz liegen 20 Cluster zugrunde, wobei jedes Cluster die gleiche Teilnehmerstärke aufweist. Die nachfolgende Tabelle weist die Dichtegrenzen der Clusterstufen aus.

Tabelle 4-1: Clusterbildung und Anschlussdichte

Cluster	Min Anschl/km ²	Max Anschl/km ²
1	4.892	13.911
2	2.988	4.877
3	2.036	2.987
4	1.467	2.035
5	1.135	1.466
6	874	1.134
7	678	873
8	525	678
9	411	525
10	329	411
11	262	329
12	211	262
13	173	211
14	137	173
15	109	137
16	85	109
17	65	85
18	47	65
19	32	47
20	0	32

Quelle: WIK

4.1 Greenfield-Analyse

4.1.1 Kosten des Netzausbaus nach Clustern

Aus der Kostenmodellierung ergeben sich zunächst Kosten pro Anschlussbereich und Kunde. Im nächsten Schritt werden die Anschlussbereiche nach Dichte sortiert und den Clustern zugeordnet – nach Maßgabe, dass in jedem Cluster 5% der Anschlüsse repräsentiert sind (Durchschnittswertbildung).¹⁶

Die für diese Veröffentlichung gewählte Darstellung der berechneten Kosten eines FTTH P2P Anschlussnetzes stellt auf den Vergleich zu den Ergebnissen aus der WIK-Vorgängerstudie 2011 ab¹⁷ und verzichtet auf den Ausweis absoluter Kostenhöhen. Der damaligen Studie lagen ebenfalls 20 Cluster gleicher Teilnehmerstärke zu Grunde. Allerdings gibt es aufgrund der in dieser Studie durchgeführten Aktualisierung der Datenbasis Veränderungen in der Dichte der Anschlussbereiche.

Die Abbildung 4-1 zeigt einen clusterbasierten Vergleich der monatlichen Kosten pro Kunde für ein neues FTTH P2P-Netz zwischen der Vorgängerstudie und den im Rahmen dieser Studie durgeführten Berechnungen. Diesen Ergebnissen liegt der sog. Greenfield-Ansatz zugrunde, demzufolge der Netzbetreiber für das glasfaserbasierte Anschlussnetz sämtliche Investitionen neu tätigen muss und nicht auf bestehende, wiederverwendbare Infrastruktur zurückgreifen kann.

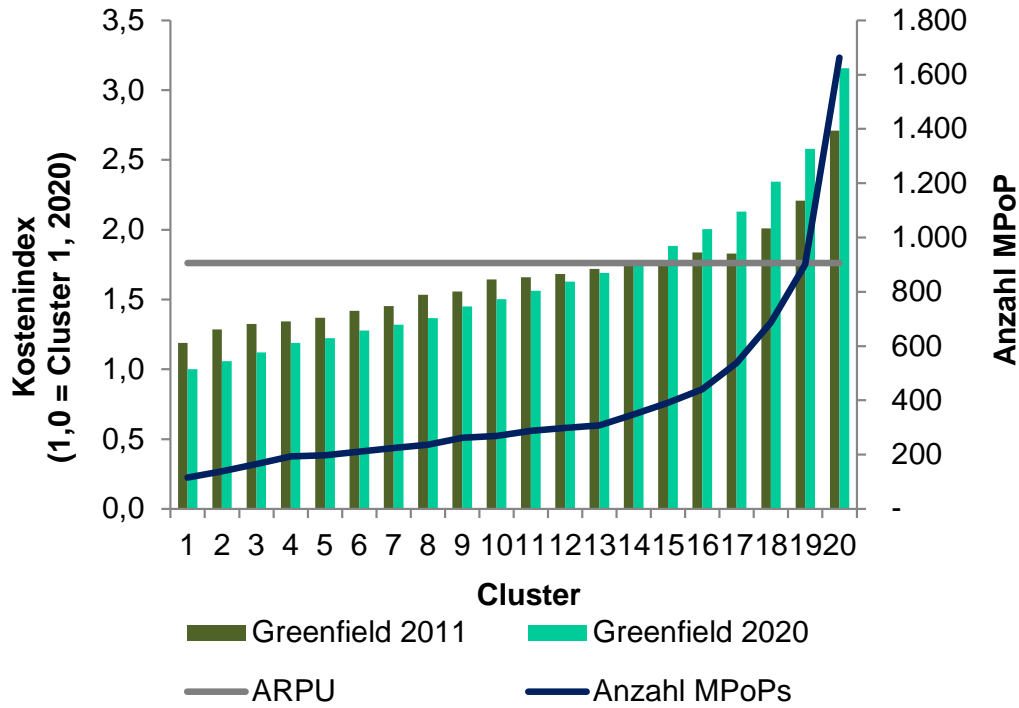
Die in dieser Studie berechneten Ergebnisse sind für die 20 Cluster mit den hellgrünen (Greenfield 2020), die von 2011 mit den dunkelgrünen Balken dargestellt. Die dunkelblaue Linie mit zugehöriger Skala rechts, beschreibt die Anzahl der MPoP, die auf die jeweiligen Cluster entfallen, wobei hier die Zuordnung der aktuellen Berechnung zu Grunde liegt. Sie verdeutlicht, dass die verwendete Clusterung Aufschluss über die Verteilung der durchschnittlichen Kosten pro Anschluss gibt und nicht auf die Anschlussbereiche abstellt. Aufgrund der vorherrschenden Besiedlungsstruktur fällt der Großteil der Anschlussbereiche nur in wenige Cluster.

Das aus der Vorgängerstudie bekannte Muster der Kosten pro Cluster stellt sich auch jetzt wieder dar. Mit Blick auf die Unterschiede ist aus der obenstehenden Abbildung ersichtlich, dass die monatlichen Kosten pro Kunde bis Cluster 13 bei der jetzigen Berechnung niedriger ausfallen als dies in der WIK-Vorgängerstudie (2011) der Fall war.

¹⁶ Dieser Methodik zufolge gibt es Anschlussbereiche, deren Anschlüsse in zwei Cluster fallen. Da die Ermittlung der clusterspezifischen Kosten auf Basis der durchschnittlichen Kosten pro Anschluss basiert, ist dies unproblematisch.

¹⁷ Jay, S., Neumann, K.-H., Plückebaum, T. (2011)

Abbildung 4-1: Zeitpunktbezogener Vergleich von Kosten und ARPU für FTTH P2P nach Clusterstufen (pro Kunde/ Monat, Index 1,0 = Cluster 1, 2020)



Quelle: WIK

Insgesamt weisen die Kosten pro Anschluss und Cluster in der Neustudie eine stärkere Spreizung auf als in der WIK-Vorgängerstudie (2011). In letzterer sind die Kosten der letzten Clusterstufe etwa 2,3 Mal teurer als beim ersten Cluster. In der Neustudie sind sie nun 3,2 Mal so teuer.

Auffällig ist darüber hinaus, dass sich nun – anders als in der Vorgängerstudie 2011 – die Kosten pro Kunde und Cluster monoton entwickeln. In 2011 waren zwischen Cluster 16 und 18 noch Unstetigkeit festgestellt worden – wenn auch in nur geringem Umfang.

Diese festgestellten strukturellen Veränderungen in den Ergebnissen lassen sich weitestgehend auf die beiden folgenden Aspekte zurückführen:

- neue, anschlussbereichsgenaue Datenbasis mit der zu Grunde liegenden GIS-basierten Anschlussnachfrage inkl. aktualisiertem GIS-Straßenlayer und Anschlussbereichsabgrenzung sowie deren Einfluss auf die Clusterbildung
- anschlussbereichs-individuelle Kostenermittlung und erst nachgelagerte Ableitung Cluster-spezifischer Durchschnittskosten pro Anschluss, im Unterschied zu einer Kostenermittlung pro Cluster auf Basis eines durchschnittlichen Mengenrüstes.

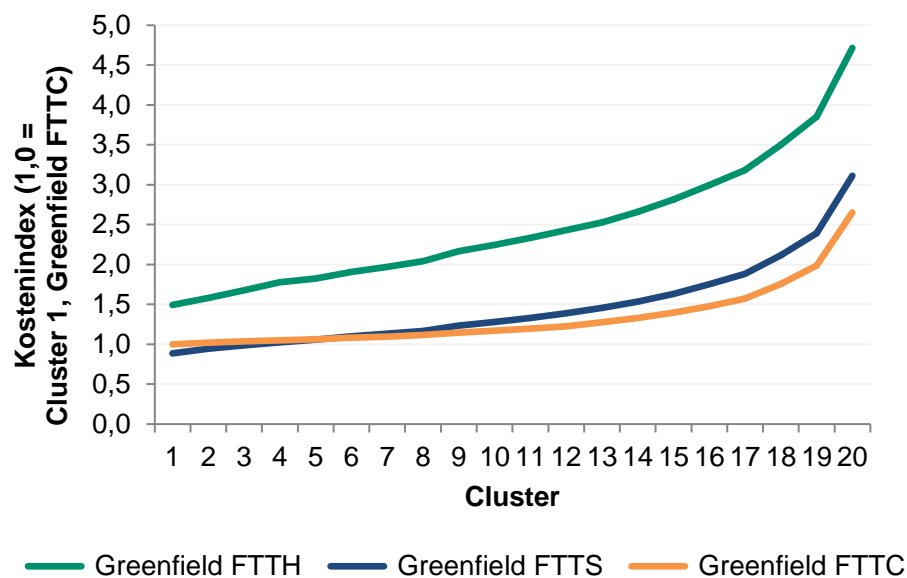
4.1.2 Kostenvergleich: FTTH P2P und hybride NGA- Technologien (FTTS/FTTC)

Analog zu FTTH P2P wurden unter Rückgriff der im Abschnitt 3.2.3 festgelegten technologieabhängigen Modellparameter der Investitionsbedarf, die Kosten und die Profitabilität für FTTS- und FTTC-Netzarchitekturen berechnet.

Vor dem Hintergrund, dass auf den erforderlichen Förderbedarf abgestellt werden soll, wurden für die beiden hybriden Anschlusstechnologien FTTS und FTTC die für den kupferbasierten Teil des Anschlussnetzes erforderlichen Investitionen nicht berechnet, sondern die Vorleistungsentgelte für die entbündelte KVz-TAL (im Fall von FTTS anteilig) herangezogen. Diese Kosten fallen für jeden kontrahierten Anschluss an.¹⁸

Die untenstehende Grafik stellt die Kosten pro Kunde/ Monat von FTTH P2P, FTTS und FTTC anhand der bisherigen Clusterung vergleichend gegenüber. Der Darstellung liegt eine Indexierung zu Grunde. Dazu wurden die Kosten für FTTC im Cluster 1 auf „1“ normiert und die übrigen Kostenwerte entsprechend indexiert.

Abbildung 4-2: Greenfield – Vergleich der Kosten von FTTH P2P, FTTS und FTTC (pro Kunde/ Monat, Index 1,0 = Cluster 1 bei FTTC)



¹⁸ In Abschnitt 3.2.5 wurde erläutert, dass für FTTC die Höhe des KVz-TAL Entgelts, für FTTS lediglich 2/3 des KVz-TAL Entgelts zum Ansatz kommen.

Die Abbildung 4-2 zeigt, dass auf Basis eines reinen FTTH P2P-Netztes (**grün**) die Kosten pro Kunde/ Monat je nach Clusterstufe zwischen 51 und 76 Prozentpunkte über einem reinen FTTS-Netz (**blau**) liegen. Die Kosten pro Kunde/ Monat eines FTTS-Netztes sind ab dem Clusterstufe 6 höher als bei FTTC (**orange**). Dieser Kostenunterschied steigt kontinuierlich mit zunehmenden Clusterstufen bis auf maximal 21 Prozentpunkte in Cluster 19.

Die Diskrepanz zwischen FTTH P2P und den beiden hybriden Anschlussnetzen liegt neben dem kostengünstigeren Ausbau des Hauptkabelsegments, welches einen relevant geringeren Faserbedarf aufweist, insbesondere in der Kostenstruktur von Hauszuführung und gebäudeinterner Verkabelung begründet. Fallen im Fall von FTTH P2P für 90% aller homes passed die Kosten für die Hauszuführung an und darüber hinaus die Kosten der gebäudeinternen Verkabelung für jeden kontrahierten Anschluss, so werden für die hybriden Anschlusstechnologien lediglich die (anteiligen) Vorleistungsentgelte pro kontrahiertem Kunden angesetzt.

Die Vorteilhaftigkeit von FTTS gegenüber FTTC in den ersten Clustern ist letztlich der Besiedlungsdichte zuzuschreiben. Die günstigere Realisierung der G.fast DSLAM, welche durch ein Reverse Powering keiner separaten Stromanbindung bedürfen, führt hier zu günstigeren Kosten. Mit abnehmender Dichte (ab Cluster 6) dominieren dann die Konzentrationsvorteile des MSAN (FTTC), der eine höhere Anschlusszahl, verbunden mit weniger strengen Längenbeschränkungen, ermöglicht.¹⁹

4.1.3 Profitabilität des Netzausbaus der verschiedenen Anschlusstechnologien

Für die Ermittlung der Profitabilität des NGA-Ausbaus haben wir Erlöse pro Kunde zu Grunde gelegt und diese den Kosten gegenüber gestellt.

Angesichts der Unterschiede in der technologieabhängigen Übertragungskapazität und des daraus entstehenden vermarktbareren Mehrwertes für die Kunden wird davon ausgegangen, dass die Höhe des durchschnittlichen Ertrags pro Kunde (engl. *ARPU*, *Average Revenue Per User*) von der eingesetzten Technologie abhängt. Im Einklang mit den Übertragungskapazitäten liegt der ARPU für FTTH P2P über dem ARPU für FTTS, und dieser wiederum über dem ARPU für FTTC. Die angenommenen ARPU wurden in Anlehnung an die marktüblichen Preise für die unterschiedlich möglichen Bandbreiten der drei Technologien gewählt.²⁰ Konkret gehen wir für FTTH P2P von einem ARPU in Höhe von 45,-€ aus. Aufgrund der geringeren Leistungsfähigkeit der

¹⁹ Mit kleineren G.fast DSLAM ließe sich der Auslastungsnachteil sicher verringern, allerdings bliebe die Notwendigkeit für eine größere Anzahl Standorte und die damit in Verbindung stehenden Kosten bestehen.

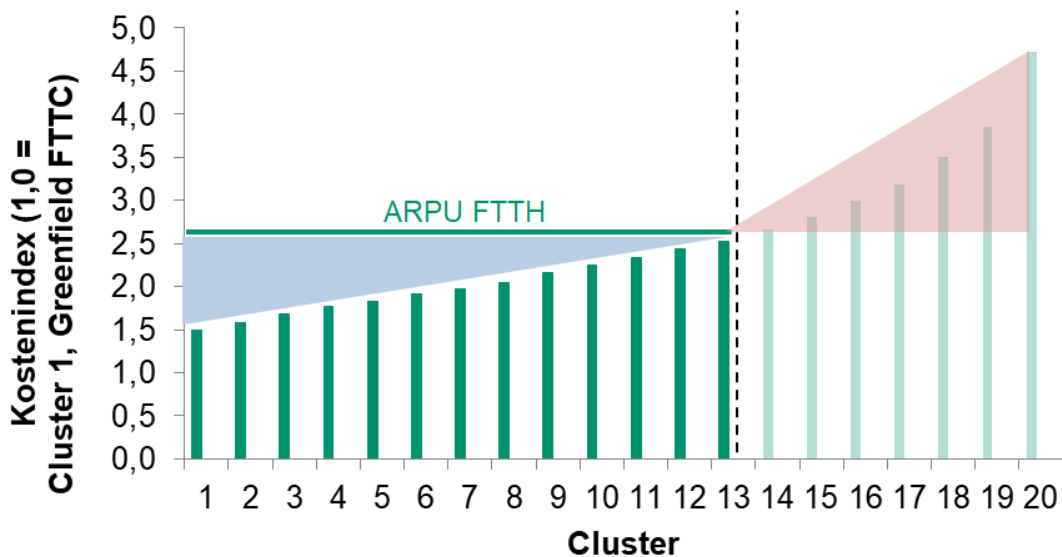
²⁰ Auf Basis von öffentlich verfügbaren Kundentarife der DTAG (siehe Webseite „Magenta Zuhause DSL-Tarife“). Der angenommene ARPU entspricht dem gewichteten monatlichen Durchschnittspreis (bei 24 Monaten Vertragslaufzeit) von 50 Mbit/s für FTTC, 250 Mbit/s für FTTS, und 500 Mbit/s für FTTH.

hybriden Anschlusstechnologien FTTS und FTTC haben wir den ARPU um jeweils 5,-€ reduziert, so dass für FTTS 40,-€ und für FTTC 35,-€ zum Ansatz kamen.

Auf Basis einer Gegenüberstellung von durchschnittlichen Kosten pro Anschluss und Cluster mit den ARPU lässt sich ermitteln, bis zu welcher Clusterstufe ein privatwirtschaftlicher Ausbau möglich ist.

Bereits in Abbildung 4-1 sind für FTTH P2P sowohl die Kosten (hellgrün) als auch die Erlöse dargestellt worden. Diese haben wir in der unten stehenden Abbildung 4-3 nochmals aufgegriffen. Aus der Abbildung 4-3 ist ersichtlich, dass bei dem angenommenen ARPU der Glasfaserausbau bis zur Clusterstufe 13 privatwirtschaftlich profitabel ist. Bezogen auf das gesamte Versorgungsgebiet können damit 65% der Haushalte bzw. Teilnehmer durch einen Glasfaseranschluss mit FTTH P2P-Technologie profitabel versorgt werden. Bezogen auf die Zahl der Anschlussbereiche ist damit jedoch weniger als die Hälfte aller Anschlussbereiche abgedeckt.

Abbildung 4-3: Profitabilität eines flächendeckenden FTTH P2P Ausbaus



Quelle: WIK.

Aus der grafischen Darstellung geht darüber hinaus hervor, dass in den profitablen Clustern ein relevanter Überschuss erzielt werden kann. Dieser wird durch das blaue Dreieck oberhalb der Kostensäulen und unterhalb der Erlös-Geraden illustriert. Für die nicht profitablen Cluster ist ein rotes Dreieck dargestellt, welches die Verluste repräsentiert. Beide Dreiecke haben annähernd denselben Flächeninhalt. Aufgrund des leicht exponentiellen Anstiegs der Kosten pro Anschluss in den Clustern 14-20 überschätzt das rote Dreieck die Verluste bei dem unterstellten uniformen ARPU. Es lässt sich da-

her schlussfolgern, dass ein flächendeckend ausbauender Netzbetreiber ein FTTH P2P Netz profitabel ausbauen und betreiben kann.

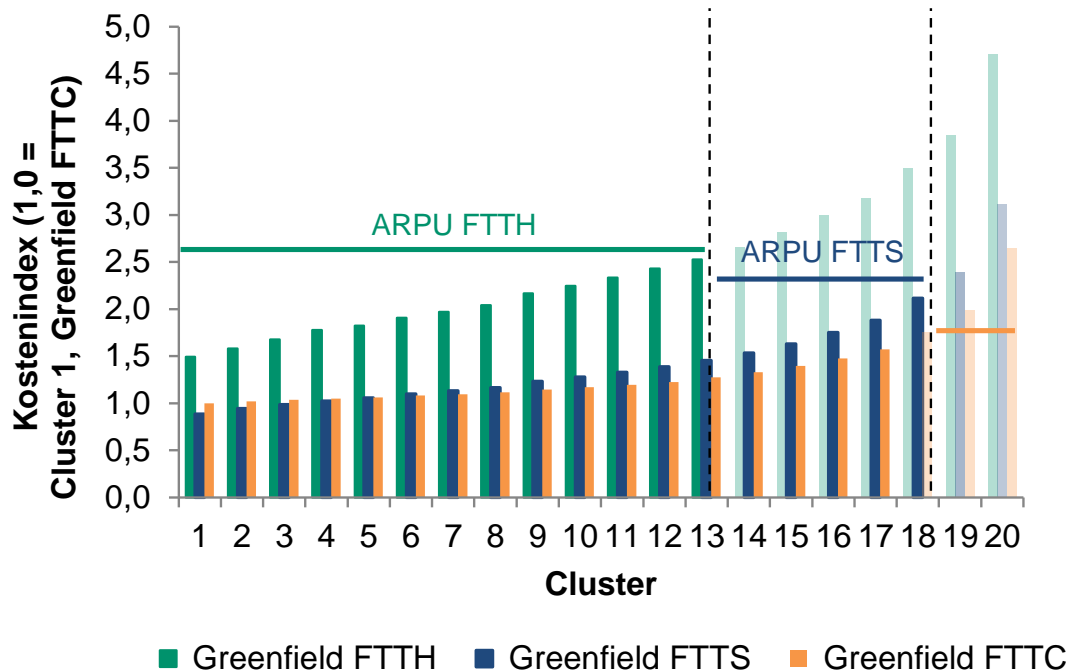
Profitabler Technologie-Mix

In die Betrachtung der Profitabilität wurden in einem weiteren Schritt die hybriden Anslusstechologien FTTS und FTTC P2P mit in die Betrachtung einbezogen. Ziel war es herauszufinden, wie durch eine Technologiewahl in Abhängigkeit der jeweiligen Profitabilität pro Cluster ein bestmöglicher Technologie-Mix aus FTTH P2P, FTTS und FTTC bestimmt werden kann.

Prämisse dieser Technologie-Mix-Untersuchung war dabei, dass die qualitativ höherwertige Anschlussnetztechnologie, die gerade noch profitabel ausgebaut werden kann, auch realisiert wird – losgelöst von strategischen Entscheidungen, beispielsweise mit weniger leistungsfähigen Technologien höhere Überschüsse erzielen zu können. Diesen Annahmen zufolge werden Glasfaseranschlüsse bis zum Endkunde ausgebaut, sofern die Kosten dieses FTTH P2P-Netzes deren Erträge nicht überschreiten. Ist FTTH P2P nicht mehr profitabel, wird dann der Glasfaserausbau mit „Zweitbeste“-Technologie – in unserem Fall FTTS – in denjenigen Clustern fortgesetzt, bei welchen er noch privatwirtschaftlich profitabel ist. Ist auch für diese „zweitbeste“ Technologie keine Profitabilität mehr erreichbar, wird die nächste, weniger leistungsfähige, aber günstigere Technologie herangezogen.

Die nachfolgende Abbildung 4-4 zeigt Kosten und ARPU pro Kunde/ Monat nach Clusterstufen (1 bis 20) und Technologie und den resultierenden Technologie-Mix.

Abbildung 4-4: Kosten, ARPU und Profitabilität FTTH P2P, FTTS und FTTC Greenfield
(Index 1,0 = Cluster 1 bei FTTC)

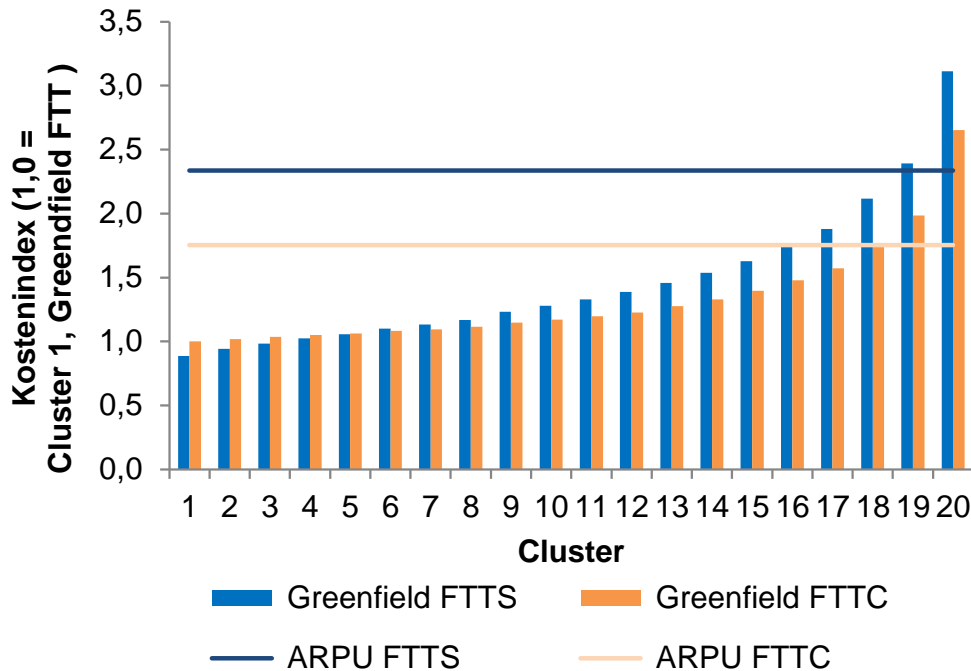


Quelle: WIK

Aufgrund der niedrigeren Kosten sind FTTS und FTTC bis zur Clusterstufe 18 bzw. 17 profitabel. Erwartungsgemäß schließt sich FTTS als profitable Anschlussnetztechnologie im Cluster 14 an den in den vorangehenden Clustern profitablen Ausbau von FTTH P2P an. Der Technologie-Mix ergibt von der Clusterstufe 14 bis zur Clusterstufe 18 den Ausbau der Glasfaser mit FTTS-Technologie. Die letzten zwei Clusterstufen 19 und 20 (ca. 10% der Haushalte) sind unter den angenommenen ARPU bei keiner der betrachteten Technologien (FTTH P2P, FTTS oder FTTC) privatwirtschaftlich profitabel.

Mit der nachfolgenden Abbildung 4-5 fokussieren wir auf die clusterbezogene Profitabilität für die hybriden Technologien FTTS und FTTC.

Abbildung 4-5: Vergleich der Kosten und ARPU pro Kunde/ Monat: FTTS und FTTC
(Index 1,0 = Cluster 1, FTTC)



Quelle: WIK.

wik 

Aus der grafischen Gegenüberstellung wird deutlich, dass die Kostenunterschiede zwischen FTTS und FTTC erst mit höherer Clusterstufe merklich werden.²¹ Trotz dieser Unterschiede in den Kosten unterscheiden sich die beiden Technologien kaum hinsichtlich des Umfangs für einen profitablen Ausbau, ausgedrückt durch die Profitabilität der Clusterstufen. Bei der höherwertigen Technologie FTTS und höherem ARPU ist FTTS profitabel bis Cluster 18, wohingegen FTTC eine Profitabilität bis Cluster 17 zeigt. Diese zunächst kontraintuitive Feststellung, dass der Ausbau für FTTC trotz geringerer Kosten weniger weitreichend profitabel sein soll, liegt in dem Verhältnis der Kosten zum erwarteten durchschnittlichen Erlös begründet. Letztlich zeigt sich hier die Sensitivität des Ergebnisses in Bezug auf den gewählten ARPU für FTTS und FTTC. Unter der Prämisse, dass die Haushalte in den Clustern 17 bis 20 keine alternative Breitbandversorgung zur Verfügung haben, ist zu vermuten, dass ihre Zahlungsbereitschaft die Höhe des marktüblichen Niveaus (hier repräsentiert durch den ARPU) übersteigt. In Abwesenheit von infrastrukturbasierendem Wettbewerb und der Hypothese, dass für FTTC ein ARPU, wie wir ihn für FTTS unterstellt haben, realisiert werden kann, wäre dann ein profitabler FTTC-Ausbau einschließlich Cluster 19 möglich.

²¹ Erläuterungen zu der anfänglichen Vorteilhaftigkeit von FTTS gegenüber FTTC wurden bereits in Abschnitt 4.1.2 gegeben.

Die Überlegungen, die wir für FTTH P2P anhand der Abbildung 4-3 dargestellt haben, lassen sich natürlich auch auf die hybriden Anschlussnetztechnologien übertragen. Für FTTS decken die Überschüsse aus den Clustern 14 bis 18 die Defizite der verbleibenden Cluster. Manifestierte sich ein solcher, wie hier „idealtypisch“ abgeleiteter Ausbau der Anschlussnetztechnologien, bliebe für ein FTTC Anschlussnetz kein Raum. Die verbleibenden Cluster 19 und 20 sind beide für sich genommen bei dem unterstellten ARPU nicht profitabel. Hier verbleiben als Lösungsansätze

- regional höhere Endkundenpreise (die ggf. bedarfsgerecht auf der Nachfrageseite subventioniert werden),
- regionale Förderung des Netzausbaus und damit Reduzierung der beim Netzbetreiber zu deckenden Kosten,
- verpflichtender Ausbau durch ausgewählte Betreiber, die durch das Mischungsverhältnis von profitablen und unprofitablen Regionen insgesamt profitabel wirtschaften können.

Die hier abgeleiteten Ergebnisse zielen auf die Profitabilität eines flächendeckenden Netzausbaus ab. Dabei haben wir anstelle von Sensitivitätsrechnungen für verschiedene Marktanteile eine Abschätzung aus der Perspektive eines dominanten Netzbetreibers vollzogen, der über einen Marktanteil von 90% verfügt.²² – Vor dem Hintergrund eines wettbewerblichen Marktes, dessen regulatorisches Umfeld auf einen infrastrukturasierten Wettbewerb abzielt, erscheint diese Annahme realitätsfern. Wir haben diesen Anteil in Höhe von 90% dennoch gewählt, um eine Einschätzung für die maximale Reichweite eines profitablen Netzausbaus zu gewinnen. Da Fixkosten in Anschlussnetzen ein relevantes Gewicht haben (insbesondere wenn von den Investoren ein flächendeckender Netzausbau erwartet wird),²³ schlägt diese Annahme auf die durchschnittlichen Kosten pro Anschluss und damit auf die hier festgestellte Reichweite eines profitablen Ausbaus der Anschlussbereiche durch: Mit sinkender Kundenzahl steigt der von einem einzelnen Kunden zu erbringende Kostendeckungsbeitrag zu den Fixkosten. Halbiert sich der Marktanteil, so verdoppeln sich die von jedem Kunden zu tragenden Fixkosten. – Anhand der voranstehenden Abbildungen zur Profitabilität des Netzausbaus ist ersichtlich,²⁴ wie hoch bzw. wie begrenzt der Spielraum für eine marktanteilsbedingte Erhöhung der durchschnittlichen Kosten pro Anschluss und Cluster für die Reichweite eines profitablen NGA-Ausbaus ist.

²² Die Größe von 90% ist auf die Gesamtheit der für die GIS-Modellierung erfassten Haushalte bezogen (gut 40 Mio. Haushalte).

²³ Zwar liegen den modellierten Ergebnissen auch mit der Teilnehmerzahl skalierende Kosten zu Grunde (bei FTTH P2P die gebäudeinterne Verkabelung sowie bei sämtlichen Anschlussnetztechnologien die upstream Kosten), diese fallen jedoch weniger ins Gewicht (insbesondere für die Technologien FTTS und FTTC). – Die Interpretation, dass sich mit der Halbierung des Marktanteils die durchschnittlichen Kosten pro Anschluss verdoppeln würden, ist sicher eine Überschätzung.

²⁴ Solange die durchschnittlichen Erlöse die durchschnittlichen Kosten pro Anschluss übersteigen, bleibt der Ausbau profitabel. In Abbildung 4-3 die Überschüsse auch über das blaue Dreieck zum Ausdruck gebracht.

Die von uns getroffene Annahme, dass in den Clustern 1 bis 13 FTTH P2P trotz relevanter Kostennachteile gegenüber den hybriden Technologien sich wird durchsetzen können, hängt nicht zuletzt von der Zahlungsbereitschaft und den verfügbaren Substituten ab.²⁵ Auf die Frage nach der Stabilität derartiger Marktanteile und die Fähigkeit des Investors, das eingesetzte Kapital zurückzugewinnen, gehen wir im Rahmen dieser Studie nicht ein.

Auch an dieser Stelle möchten wir nochmal darauf hinweisen, dass der dieser Profitabilitätsanalyse zugrunde liegende Clusteransatz sich allein aus dem Dichtekriterium und einer Gleichverteilung der Anschlüsse ableitet. Dabei entfallen überproportional viele Anschlussbereiche auf die hohen, in der Abbildung 4-5 als unprofitable dargestellten Cluster. Die hinter einem Cluster liegenden Anschlussbereiche verteilen sich über die gesamte Republik und bilden kein zusammenhängendes Gebiet, welches realistisch nicht von einem einzelnen bestimmten Netzbetreiber versorgt wird (es sei denn, von einem flächendeckend agierenden, nationalen Netzbetreiber). An den Profitabilitätsgrenzen der hier betrachteten Technologien (FTTH zu FTTS sowie FTTS zu FTTC) werden Anschlussbereiche zum Teil mit der höherwertigen, zum Teil mit der minderwertigen Technologie berücksichtigt. – Mit Blick auf die Praxis ist hier Skepsis geboten, da ein Parallelbetrieb von Anschlussnetztechnologien – zumindest durch einen Betreiber – geringere Wirtschaftlichkeit aufweist, bzw. bei infrastrukturbasiertem Wettbewerb die preisliche Bedrohung der höherwertigen Technologie mit sich bringt.

Vor dem Hintergrund der hier genannten Beschränkungen wird der, diesen Ergebnissen zugrunde liegende methodische Ansatz der Clusterbildung in dem späteren Abschnitt 1 dieses Diskussionsbeitrags weitergehend analysiert.

4.2 Brownfield-Analyse: Berücksichtigung bestehender Breitbandversorgung und Abschätzung von Cu-GF Migrationskosten

Der in den letzten Jahren bereits erfolgte Ausbau von Breitbandanschlussnetzen lässt erwarten, dass die bottom-up ermittelten Investitionen des Greenfield-Ansatzes für einen neu in den Markt eintretenden Investor den sich aktuell darstellenden Investitionsbedarf übersteigen. Ziel dieses Abschnitts ist es daher die bestehende Breitbandversorgung im Sinne eines Brownfield-Ansatzes zu berücksichtigen und nur die inkrementell erforderlichen Investitionen abzuschätzen. Dazu musste

- eine Einschätzung zur bestehenden Versorgungslage je Anschlussbereich gewonnen und unter Rückgriff auf die Daten des Breitbandatlas des BMVI der Grad der noch fehlenden Versorgung ermittelt werden sowie

²⁵ Die Annahme eines Marktanteils in Höhe von 90% wurde für alle drei NGA Technologien unterstellt. Entsprechend diskriminiert sie nicht mit Blick auf die relative Vorteilhaftigkeit der einzelnen Technologien (ausgedrückt in Kosten pro Anschluss).

- darüber hinaus in einer zweiten Betrachtung für den Fall eines bereits bestehenden FTTC-Ausbaus eine Abschätzung der Migrationskosten von FTTC zu FTTH P2P sowie CATV/FTTS zu FTTH vollzogen werden.

4.2.1 Datenbasis und anschlussbereichsbezogene Aufbereitung

Die Datengrundlage zur Einschätzung der Breitbandverfügbarkeit bildet der Breitbandatlas des Bundesverkehrsministeriums mit einem Datenstand von 04/2018. Uns liegt dazu das gesamte deutschlandweite 250 m x 250 m Raster des Breitbandatlas vor, untergliedert in die unten angegebenen Versorgungsarten in Technologien und Übertragungsgeschwindigkeiten und inklusive der anschließend aufgeführten Breitbandverfügbarkeiten in Prozent der Haushalte nach 5 Kategorien:

Tabelle 4-2: Klassifizierung des Breitbandatlas nach Geschwindigkeiten und Technologien

	DSL	FTTH	CATV	Mobil
>= 1 Mbit/s	x			
>= 2 Mbit/s	x			x
>= 6 Mbit/s	x			
>= 16 Mbit/s	x			
>= 30 Mbit/s	x			
>= 50 Mbit/s	x	X	x	

Quelle: BMVI Breitbandatlas 04/2018

Pro Rasterzelle und Technologie und Geschwindigkeitskategorie liegt ein Kategoriewert zur Einschätzung der Versorgungslage bezogen auf die Haushalte im Raster vor:

Tabelle 4-3: Kategoriewerte zur Einschätzung der Versorgungslage im Raster

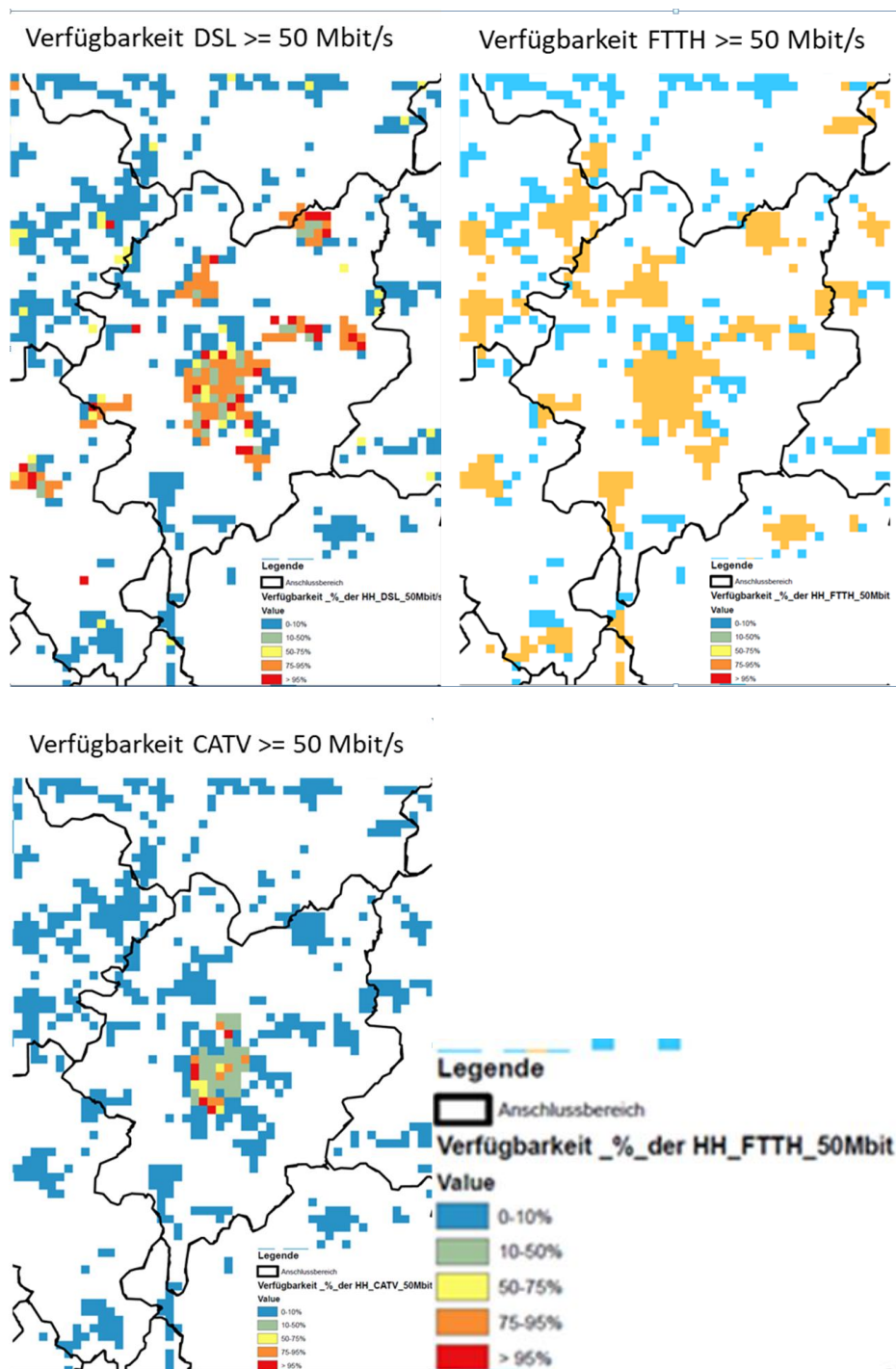
Kategorie	Bedeutung
1	0-10% der Haushalte sind versorgt
2	10-50% der Haushalte sind versorgt
3	50-75% der Haushalte sind versorgt
4	75-95% der Haushalte sind versorgt
5	> 95% der Haushalte sind versorgt

Quelle: BMVI Breitbandatlas 04/2018

Abbildung 4-6 zeigt die Versorgungslage laut Breitbandatlas für die Technologien DSL, FTTH und CATV in den relevanten Geschwindigkeitsklassen für einen Anschlussbereich. Die Farben kennzeichnen die 5 Kategorien der Versorgung in % der Haushalte

mit der jeweiligen Kombination aus Technologie und Übertragungsgeschwindigkeit. Somit liegen uns pro Rasterzelle 3 Informationen zum Versorgungsgrad vor.

Abbildung 4-6: Illustration der Datenverfügbarkeit aus dem Breitbandatlas: Breitbandanschlüsse eines Anschlussbereichs



Die zwischenzeitlich im Breitbandatlas verfügbaren Kategorien ≥ 100 Mbit/s, ≥ 200 Mbit/s ≥ 400 Mbit/s und > 1000 Mbit/s konnten zum Zeitpunkt unserer Datenbeschaffung und Datenaufbereitung für das Projekt nicht bereitgestellt werden. Ebenso wertvoll wäre diese Information für die Einschätzung hochbreitbandiger CATV-Versorgung gewesen. Weiterhin lag zum Zeitpunkt auch keine verwertbare Einschätzung hinsichtlich der Versorgungslage für Geschäftskunden vor.

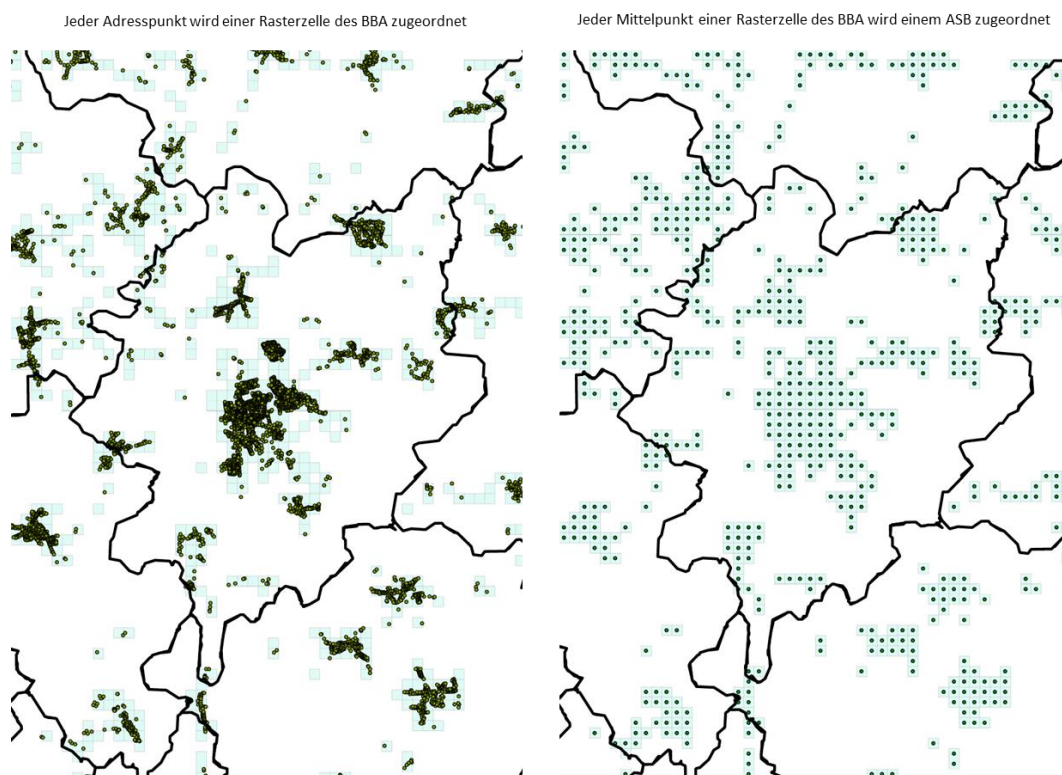
Um die Informationen über die Versorgungslage auf Rasterebene für unsere Studie zugänglich zu machen, musste eine Beziehung zwischen den Breitbandrasterflächen und den auf die jeweilige Rasterzelle entfallenden Haushalte hergestellt werden. Der direkte Weg über einen Abgleich zwischen dem Breitbandatlasraster und dem AZ-Haushalts- und Gebäuderaster war aufgrund von Lageverschiebungen der beiden Raster nicht gangbar. Deshalb wurde ein räumlicher Verschnitt der rd. 22,7 Mio. Adresspunkte mit den rd. 1,38 Mio. Rasterzellen des Breitbandatlas vorgenommen, so dass im Ergebnis für jeden Adresspunkt die Rasterzellen-ID derjenigen Zelle bekannt ist, in der der Adresspunkt zu liegen kommt. Eine Gruppierung der Daten auf Ebene der Rasterzellen liefert die Anzahl Adresspunkte und das Nachfragepotential je Rasterzelle des Breitbandatlas.

In einem weiteren Schritt war zu ermitteln, welche Rasterzelle in welchem Anschlussbereich liegt. Für die Auswertung der räumlichen Lage wurden die Mittelpunkte der Rasterzellen bestimmt und dem nächstliegenden Anschlussbereich zugeordnet. Ziel war hier für alle 1,38 Mio. Rasterzellen eine eindeutige Zuordnung zu einem der 7871 Anschlussbereich zu generieren.

Abbildung 4-7 illustriert das beschriebene Vorgehen zur Ermittlung der Versorgungslage in einem Anschlussbereich für die drei Technologien über die entsprechenden räumlichen Zuordnungen der vorliegenden Informationen. Im Ergebnis sind für jede Rasterzelle des Breitbandatlas folgende Informationen bekannt:

- Anschlussbereichszugehörigkeit
- Versorgungsgrad je Technologie (DSL, FTTH/B, CATV)
- Anzahl Adresspunkte im Raster
- Nachfragepotential im Raster

Abbildung 4-7: Ermittlung der Versorgungslage je Anschlussbereich



Quelle: WIK

wik 

4.2.2 Bestehende Versorgungslage nach Anschlüssen

Für die Operationalisierung der bestehenden Versorgungslage zum Zweck unserer Brownfield-Analyse wurde zunächst ermittelt, welche Zahl von Anschlüssen mit einer der verfügbaren Technologien bereits versorgt ist. Dabei haben wir uns aufgrund unseres Untersuchungsgegenstandes darauf beschränkt, aus den verfügbaren Daten lediglich die Geschwindigkeiten ≥ 50 Mbit/s je Technologie zu verwenden und die in Tabelle 4-2 dargestellten Differenzierungen unterhalb von 50 Mbit/s zu vernachlässigen. Die Daten haben wir dann in der Folge aufbereitet und mit der nachfolgenden Interpretation für unsere Analyse verwendet:

- „DSL_50Mbit“ > interpretiert als „FTTC“
- „CATV_50Mbit“ > interpretiert als „CATV“
- „FTTH_50Mbit“ > interpretiert als „FTTH“

Ideal für eine Einschätzung der bestehenden Abdeckung wäre es für jeden Hausanschluss die verfügbaren Technologien zu kennen. Leider waren diese Daten im Rahmen des Projektes nicht zu beschaffen.

Ein relevantes Defizit der verfügbaren Daten war die zugrunde liegende Kategorisierung. Für die Einschätzung der Versorgungslage wäre es sehr wünschenswert gewesen, wenn eine klare Trennung in Kategorie 1 zwischen 0% (Technologie nicht vorhanden) und 1-10% (es existiert eine Versorgung) vollzogen worden wäre. Somit war nicht eindeutig zu identifizieren, in welchen Regionen gar keine oder eben nur eine geringe Versorgung mit Breitbandanschlüssen vorzufinden ist.

Die verfügbaren Informationen in den vorliegenden Kategorien mussten deshalb zur Operationalisierung einer Interpretation unterzogen werden. Hierbei haben wir die im Breitbandatlas angegebenen Abdeckungen für DSL, CATV und FTTH wie folgt konkretisiert:

Tabelle 4-4: Kategoriewerte des Breitbandatlas und deren Übersetzung für die Modellrechnungen

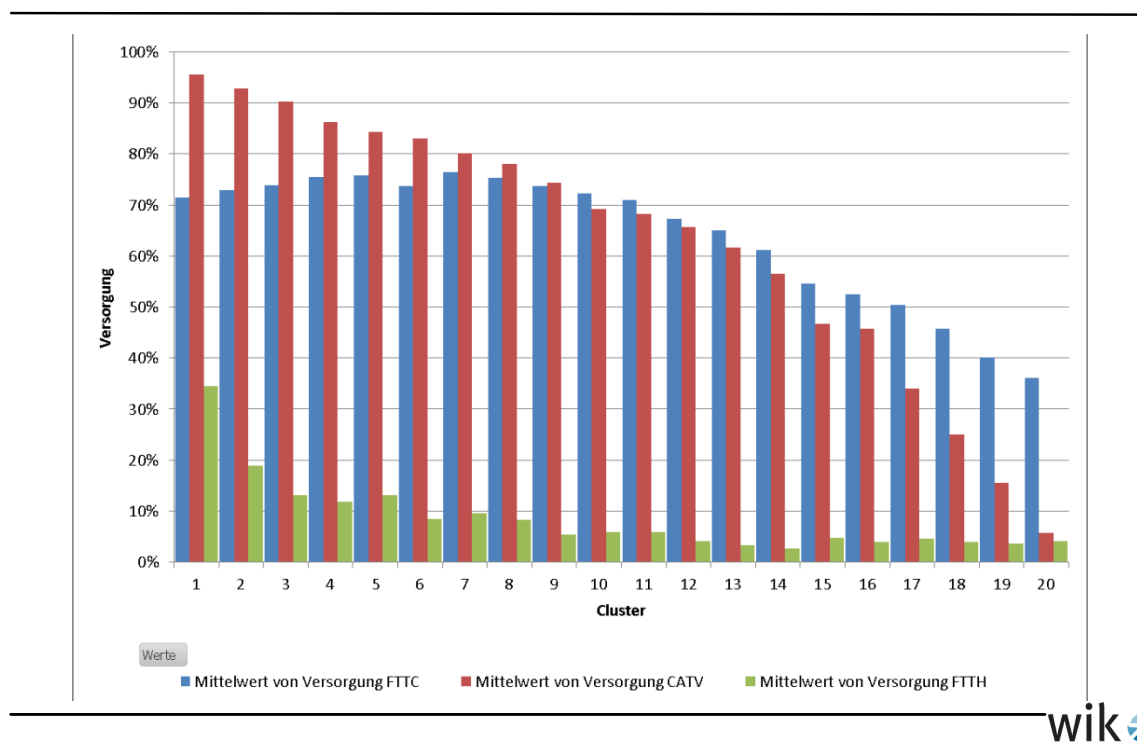
Kategorie	Bedeutung	Konkretisierung der Abdeckung
1	0-10% der Haushalte sind versorgt	0%
2	10-50% der Haushalte sind versorgt	30%
3	50-75% der Haushalte sind versorgt	60%
4	75-95% der Haushalte sind versorgt	85%
5	> 95% der Haushalte sind versorgt	100%

Quelle: WIK

Auf Ebene der Rasterdaten ermitteln wir die Anzahl Haushalte und Gewerbe (als „Anschlüsse“ bezeichnet). – Durch Anwendung des Prozentsatzes „Konkretisierung der Abdeckung“ aus Tabelle 4-4 berechnen wir dann die Anzahl „Anschlüsse“, die schon mit Breitband versorgt sind (pro Raster) und können diese Information auf Anschlussbereichsebene aggregieren.²⁶ Das Ergebnis der Aggregation und die Zusammenfassung in die 20 Cluster für den Modellrahmen sind in der nachfolgenden Abbildung 4-8: Breitbandversorgung nach Clustern dargestellt. In dieser Darstellung wurden die Versorgungsgrade unabhängig voneinander dargestellt, wie dies auch im Breitbandatlas erfolgt, es kann also auch eine Mehrfachversorgung mit zwei oder allen drei Technologien auf Gebäudeebenen geben. Für die Darstellung auf Cluster-Ebene wurden die Mittelwerte der zum jeweiligen Cluster gehörenden Anschlussbereiche verwendet.

²⁶ Leider enthalten die Daten des Breitbandatlas – anders als die Rasterscharfen Anschlussnachfragen – keine Informationen zur Abdeckung gewerblicher Nutzer. Diese Ungenauigkeit erscheint uns angesichts der fehlenden Informationen zur Abdeckung gewerblicher Anschlüsse und der groben Einschätzungen zur Haushaltsabdeckung vertretbar.

Abbildung 4-8: Breitbandversorgung nach Clustern



Quelle: WIK unter Rückgriff auf Breitbandatlas (04/2018)

CATV zeigt die höchste Abdeckung, die bis inklusive Cluster 8 oberhalb der Abdeckung von FTTC liegt. Mit anderen Worten, 40% der Haushalte in den dicht besiedelten Regionen können von einer umfangreicheren CATV Abdeckung Nutzen ziehen als von einer FTTC Abdeckung in diesen Regionen. Ausgehend von über 95% zeigt Abdeckung durch CATV einen konkaven Verlauf.

Überraschend ist, dass in den ersten 3 Clustern die Abdeckung durch FTTC zwischen 25% und 15% niedriger ausfällt als durch CATV. Erst ab Cluster 9 kehren sich die Verhältnisse um und die Abdeckung durch FTTC nimmt für die nachfolgenden Cluster langsamer ab als dies durch CATV der Fall ist.

Für FTTH stellt sich die Abdeckung über die Cluster weniger stetig dar als bei den beiden anderen Technologien. Zum einen ist das Niveau der Abdeckung lediglich im ersten, urbanen Cluster höher als 30% und ab Cluster 6 mit lediglich einer Ausnahme unterhalb von 10%. Überraschend erscheint auf den ersten Blick, dass die höchste Abdeckung im dichtesten Cluster 1 zu beobachten ist, da die vermeintlich größten Aktivitäten des FTTH-Ausbaus in ländlichen Regionen vermutet werden. Bei der Interpretation der hier vorliegenden Darstellung ist jedoch zu beachten, dass die vollzogene Clusterbildung auf eine Gleichverteilung der Haushalte pro Cluster abstellt und die Anzahl der dahinter liegenden Anschlussbereiche vernachlässigt. So verbirgt sich insbesondere in den letzten Clustern eine Vielzahl von Anschlussbereichen, die einen FTTH-Ausbau

erfahren – trotz oder auch wegen der dünnen Besiedlung und der damit in Verbindung stehenden Förderung des Breitbandausbaus.

Wie aus der bisherigen Darstellung ersichtlich (siehe auch Abbildung 4-8), sind die Angaben im Breitbandatlas nicht überschneidungsfrei, da ein Anschluss über mehrere Technologien versorgt sein kann. Die Verfügbarkeit je Anschlussbereich könnte sich hierbei auf bis zu 300% addieren, wenn sämtliche Anschlüsse mit allen 3 Technologien angebunden wären. Diese Feststellung ließe sich auch als Infrastrukturwettbewerb beschreiben. Tatsächlich liegen die Einzelbeobachtungen durchgängig unterhalb von 100% und die Daten lassen keine Rückschlüsse zu, ob der Versorgungsgrad der übrigen Technologien die Lücke der „Nicht-Abdeckung“ zu schließen vermag.

Bei der Identifizierung der Versorgungslücke sind wir dabei mit der Problematik der „Doppelversorgung“ folgendermaßen vorgegangen:

- In einem ersten Schritt wurde die nach Breitbandatlas Datenlage vorliegende Information „mit FTTH versorgt“ dergestalt interpretiert, dass kein weiterer Bedarf mehr für einen Breitbandausbau besteht. FTTH konstituiert insofern die höchstwertigste Anschlussnetztechnologie.
- CATV wurde als zweit höchstwertigste Technologie eingestuft, aber wegen der im Breitbandatlas erfolgten Gleichsetzung bzgl. der verfügbaren Bandbreite ($\geq 50\text{Mbit/s}$) als gleichwertig zu der FTTH angenommen. Entsprechend wurde für den noch bestehenden Investitionsbedarf unterstellt, dass sobald ein Haushalt mit CATV angebunden ist, kein weiterer Bedarf mehr für einen Breitbandausbau besteht (in den Brownfield-Szenarien, in denen CATV Berücksichtigung findet). Investitionsbedarf besteht lediglich für Anschlüsse, die noch keine Abdeckung durch FTTH oder CATV erfahren haben.
- FTTC wurde als dritt höchstwertigste Technologie bewertet. Sobald ein Haushalt mit FTTC angebunden werden könnte, besteht in den Berechnungen für einen flächendeckenden Ausbau mit FTTC kein weiterer Bedarf mehr. In den Szenarien für einen flächendeckenden Ausbau mit FTTS bzw. FTTH müsste noch ein Ausbau erfolgen.

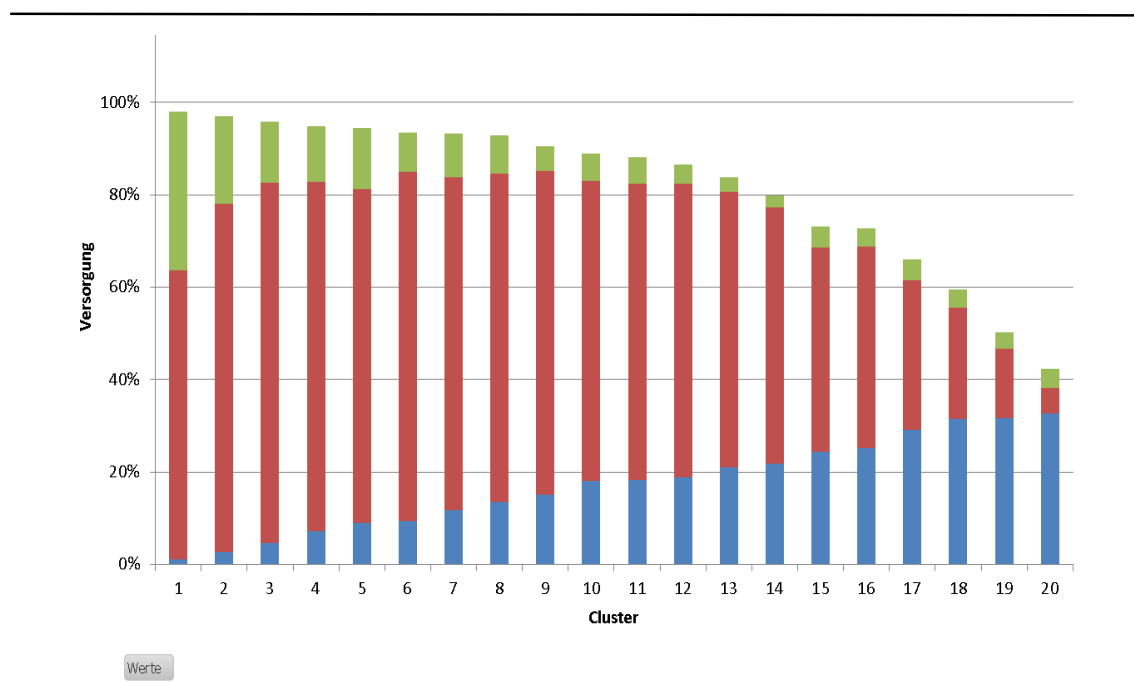
Gibt es weder eine Versorgung mit FTTC, CATV noch FTTH, so ist in jedem Falle ein Ausbau in einer höherwertigen Technologie erforderlich. Der Investitionsbedarf entspricht dann den in Abschnitt 4.1 und 4.1.2 abgeleiteten Größen des Greenfield-Ansatzes. Dabei stellen wir mit den berechneten Kosten für FTTH P2P der Investitionsbedarf auf die FTTH Versorgung und mit FTTS auf die CATV Versorgung ab. Die modellierten FTTC-Kosten sind stellvertretend für den xDSL-Ausbau mit bis zu 50 Mbit/s.

Folgendes Beispiel soll die Analyse illustrieren: Gesucht ist der Anteil an Haushalten, für die noch Investitionen in den FTTH Ausbau zu tätigen sind. Den Ausgangsdaten zufolge sind 20% mit FTTH und 20% mit CATV versorgt. 70% sind mit FTTC abge-

deckt. Für den flächendeckenden FTTH Ausbau interpretieren wir die Abdeckung durch die beiden höchstwertigen Anschlusstechnologien konservativ, in dem wir eine vollständige Überlappung unterstellen. Entsprechend gehen wir davon aus, dass in diesem Beispiel noch 80% ausgebaut werden müssen.

Die in der nachfolgenden Abbildung 4-9 dargestellten Balken zeigen in ihrer Höhe den Umfang der höchsten Abdeckung durch eine der drei Technologien an. Empirisch sind dies lediglich CATV (Cluster 1 bis 9) sowie FTTC (Cluster 10 bis 20). Die Anordnung der Anteile erfolgt dabei „von oben nach unten“, wobei die prozentuale Abdeckung durch FTTH die Anteile von CATV überdeckt. Gleiches gilt für die Darstellung von FTTC.

Abbildung 4-9: Ergebnis der Versorgung nach Clustern nach Bereinigung der Überschneidungen



Quelle: WIK unter Rückgriff auf Breitbandatlas (04/2018)

Für den weiterhin bestehenden Investitionsbedarf greifen wir auf die durchschnittlichen Kosten pro Anschluss zurück und kommen dabei zu folgenden clusterbezogenen Ergebnissen:

- Für FTTH ergibt sich, je nach Cluster, ein verbleibender Investitionsbedarf von mindestens 65% und maximal 97% bezogen auf den Greenfield-Investitionsbedarf.

- Für CATV bzw. FTTS ergibt sich, je nach Cluster, ein verbleibender Investitionsbedarf von mindestens 5% und maximal 97% bezogen auf den Greenfield-Investitionsbedarf.
- Für die Variante FTTC ergibt sich, ebenfalls je nach Cluster, ein verbleibender Investitionsbedarf zwischen 23% und 63% bezogen auf den Greenfield-Investitionsbedarf.

Mit Blick auf den gesamten, über alle Cluster hinweg bestehenden Investitionsbedarf lässt sich auf Basis der Daten (Stand 04/2018) feststellen, dass

- für FTTH 92% aller Haushalte
- für CATV/FTTS 37% aller Haushalte und
- für FTTC 36% aller Haushalte

noch zu versorgen sind.²⁷

Tabelle 4-5: Durchschnittlicher Versorgungsgrad nach Clustern²⁸

Cluster	FTTC	CATV	FTTH
1	71%	96%	34%
2	73%	93%	19%
3	74%	90%	13%
4	75%	86%	12%
5	76%	84%	13%
6	74%	83%	8%
7	76%	80%	10%
8	75%	78%	8%
9	74%	74%	5%
10	72%	69%	6%
11	71%	68%	6%
12	67%	66%	4%
13	65%	62%	3%
14	61%	57%	3%
15	55%	47%	5%
16	52%	46%	4%
17	50%	34%	5%
18	46%	25%	4%
19	40%	16%	4%
20	36%	6%	4%
Mittelwert	64%	63%	8%
fehlende Versorgung	36%	37%	92%

Quelle: WIK unter Rückgriff auf Breitbandatlas 04/2018

²⁷ Werden die Versorgungsanteile auf Ebene der ASB für eine Mittelwertbildung herangezogen, so stellt sich die durchschnittliche Versorgungslage etwas schlechter dar. Darin kommt zum Ausdruck, dass die Anschlusstechnologien in den dichter besiedelten Regionen, die mit einer geringeren Anzahl von Anschlussbereichen einhergehen, schon umfangreicher ausgebaut sind.

²⁸ Dies tabellarische Darstellung ist inhaltlich mit der Abbildung 4-8 deckungsgleich.

Sowohl für die abgeleiteten Investitionsbedarfe für CATV bzw. FTTS als auch FTTC ist zu beachten, dass die kundenseitige Anschlussrealisierung ab dem Kabelverzweiger bzw. Konzentrationspunkt als bereits verfügbar betrachtet wird. Wie bereits ausgeführt werden weder Investitionen noch Kosten für diesen Netzabschnitt betrachtet.

Die den Ergebnissen zu Grunde liegenden Daten veranlassen uns darüber hinaus, auf weitere Beschränkungen hinzuweisen. Der Umstand, dass im Breitbandatlas die Datenübertragungsrate für FTTH-Anschlüsse mit ≥ 50 Mbit/s angegeben wurde, lässt offen, um was für eine FTTH-Topologie es sich in den jeweiligen Gebieten handelt und welche Restriktionen hinsichtlich der maximal erreichbaren Datenübertragungsraten greifen. Hierbei ist an das Dämpfungsbudget von FTTH P2MP Topologien zu denken, ebenso wie an FTTH-Lösungen, hinter denen sich lediglich FTTB-Realisierungen verbergen. Die gebäudeinterne Auslegung von Glasfaserkabeln stellt sich in Deutschland als ein relevanter Hemmschuh dar.²⁹

4.2.3 Bestehende Versorgungslage nach Netzausbau

Im voranstehenden Abschnitt 4.2.2 wurde der Breitbandausbau mit dem Fokus auf den Anteil der bereits versorgten Bevölkerung adressiert und anhand der dem NGA Modell zu Grunde liegenden 20 Cluster, differenziert nach Besiedlungsdichte ausgewertet.

Die dabei vollzogene Linearisierung der Investitionsbedarfe nach Maßgabe der zu versorgenden Anschlüsse stellt eine vereinfachende methodische Vorgehensweise dar. Sie ist in der Regel mit einer Unterschätzung des noch bestehenden Investitionsbedarfs verbunden, da es sich bei der Netzauslegung zu einem relevanten Teil um gemeinsame Kosten handelt, die durch gemeinsame Verlegung reduziert werden können.

Um diesen Beschränkungen Rechnung zu tragen, wurde eine zweite methodische Betrachtung bezüglich der Migrationskosten angestellt, die nicht die durchschnittlichen Kosten pro Anschluss zum Ausgangspunkt wählt, sondern vielmehr die produktionstechnische Betrachtung. Methodisch wurde hierzu auf die Erkenntnisse des WIK-Diskussionsbeitrags Nr. 457 „Kosten und andere Hemmnisse der Migration von Kupfer- auf Glasfasernetze“ zurückgegriffen.³⁰ Dabei wird eine bestimmte Anschlussnetztechnologie als Startpunkt der Analyse gewählt und im Sinne von pfadabhängigen Investitionsentscheidungen überprüft, welche Netzbestandteile für den nächsten Ausbauschritt der Glasfaser in Richtung Endkunde investitionssparend wieder- bzw. weiterverwendet werden können.

Wie bereits ausgeführt, liegt den Ergebnissen für FTTC und FTTS bereits eine erste Pfadabhängigkeit zu Grunde als dass die Investitionen in das kupferbasierte teilneh-

²⁹ In diesem Kontext ist auch das bestehende Nebenkostenprivileg der Kabelnetzbetreiber zu sehen, welches diesen erlaubt, die gebäudeinterne Verkabelung über Nebenkosten den Mietern in Rechnung zu stellen.

³⁰ Plückerbaum, T.; Ockenfels, M. (2020)

meranschließende Netzsegment keine Berücksichtigung finden. Ausgehend von einem rein kupferbasierten Anschlussnetz konstituieren sie somit für FTTC und FTTS wiederverwendbare Anlagen.

FTTC – FTTH P2P Migration

Ausgehend von einer bestehenden FTTC-Infrastruktur wurde nun eine Abschätzung versucht, mit welchen Einsparungen pfadabhängig bei den Investitionen in eine FTTH P2P-Infrastruktur gerechnet werden kann.

Da sich die glasfaserbezogenen Investitionen des FTTC-Netzes auf das Hauptkabelsegment beschränken, stellt die Abschätzung des Einspareffekts auf bestehende Trassen und Kabel im Hauptkabelsegment ab. Linientechnisch ist das FTTH P2P Anschlussnetz jedoch mit einem deutlich höheren Faserbedarf verbunden, so dass – ohne Kenntnis über die tatsächlich getätigten Investitionen keine Aussage darüber möglich ist, ob ausreichende Leerrohr- oder Glasfaserkapazitäten im Hauptkabelsegment vorhanden sind.

Unter der Annahme, dass der Ausbau des FTTC-Netzes bereits vorausschauend auf eine spätere FTTH P2P Migration erfolgte, konnten wir Einsparungen von maximal 20% abschätzen. Dabei fand die Wiederverwertbarkeit sowohl von Trassen als auch Kabeln im HK-Segment Berücksichtigung.

Sofern aus historischen Gründen im Verzweigersegment ebenfalls noch Leerrohrkapazitäten aus dem kupferbasierten Anschlussnetz vorhanden wären, so könnten auch hier Einsparungen realisiert werden. Für Deutschland ist hier mit keinem nennenswerten Einsparpotential für den FTTH P2P Ausbau zu rechnen.

FTTS/CATV – FTTH P2P Migration

In Analogie zu der Migration von FTTC zu FTTH P2P haben wir die bestehende CATV Infrastruktur als Ausgangspunkt für die Abschätzung der Einsparpotentiale beim FTTH P2P Ausbau aufgrund bereits getätigter Investitionen untersucht.

Dabei unterstellen wir, dass die von uns ermittelten Kosten für die FTTS-Architektur zu DOCSIS 3.0 vergleichbar sind; und dass ferner die inkrementellen Kosten des Ausbaus von FTTS zu FTTH P2P mit denen des Ausbaus von DOCSIS 3.0 auf DOCSIS 3.1FD bzw. DOCSIS 4.0 angenähert werden können.

Da die Netztopologie von FTTS/ CATV im Vergleich zu FTTC ein kürzeres Verzweigerkabel- und entsprechend längeres Hauptkabelsegment aufweist, haben letztere bei den Investitionen auch ein größeres Gewicht. Wird auch hier unterstellt, dass der bestehende Ausbau des Hauptkabelsegments zu 100% wiederverwendbar sei, ergäben sich maximale Einsparungen in Höhe von 30%.

Realistisch ist weder bei FTTC noch bei DOCSIS 3.0 von einer vollständigen Wiederverwertbarkeit des Hauptkabelsegments auszugehen. Wir erachten das CATV-Netz als weitergehend wiederverwendbar. Dies lässt sich für das FTTC-Netz in dieser Form nicht bestätigen. Lediglich sofern – wie bereits oben angemerkt – beim FTTC rollout die Investitionen für FTTH P2P schon vorweggenommen wurden. Im Vergleich zum bisherigen Brownfield-Ansatz, der auf einer Linearisierung der Einsparungen nach Maßgabe der Anzahl versorgter Anschlüsse basiert (Abschnitt 4.2.2) lassen sich bei dieser Migrationspfad-Betrachtung geringere Investitionersparnisse ableiten.

4.2.4 FTTH Investitionen pro Anschluss und bisheriger Breitbandausbau

Die folgenden Abbildungen geben die räumliche Lage des Breitbandausbaus in Deutschland wieder, wie er sich aus dem Breitbandatlas des Bundes (Stand 2018) ergibt und stellen diesen in den Zusammenhang zu den von uns berechneten FTTH-Investitionen je Anschluss. Dabei wird auf die als Hintergrund dienende Karte der berechneten FTTH-Investitionen je Anschluss das nach Technologie und Ausbaugrad differenzierte Raster des Breitbandatlas gelegt. Der Hintergrund-Layer ist in der Hot Spot/Cold Spot Färbung dargestellt, bei der die blaue Färbung niedrige Investitionen je Anschluss bedeuten und die rot gefärbten Anschlussbereiche hohe FTTH Investitionen pro Anschluss signalisieren. Darübergelegt sind jeweils die Raster-Layer, um die je Rasterzelle verfügbaren Technologien, in Prozent der Haushalte je Rasterzelle anzuzeigen. Aus den drei Abbildungen geht jeweils die Verfügbarkeit von breitbandigen Anschlüssen in der Bandbreite 50 Mbit/s und mehr für FTTH, CATV und VDSL hervor.³¹ Dargestellt sind jeweils nur die Rasterzellen mit einer Verfügbarkeit für 50% der Haushalte und mehr (Kategorien $\geq 50\%$ -75%, $\geq 75\%$ -95% und $>95\%$ der Haushalte). Die Färbung der Zellen geht mit aufsteigendem Versorgungsgrad von hellblau nach dunkelblau.

Die erste Abbildung zeigt im Vordergrund die FTTH/B-Ausbaubereiche mit einem Versorgungsgrad $\geq 50\%$ der Haushalte. Die Abbildung macht den geringen Ausbaustand mit FTTH/B Leitungen in Deutschland augenfällig. Auf ganz Deutschland gesehen springen lediglich einige städtische Ausbaubereiche als gut erkennbare blaue Flecken ins Auge. Das sind größere Gebiete in und um Hamburg, Köln, Frankfurt und München. In der Vergrößerung werden noch weitere größere und kleinere Sprengsel erkennbar. Dem Anschein nach findet ein großer Teil des Ausbaus mit FTTH/B in Anschlussbereichen mit, gemäß unseren Berechnungen, niedrigen bis mittleren Investitionen je Anschluss statt. Selbst in Bayern, wo der Ausbau in unzähligen Sprengseln über das Land verteilt stattfindet, und wo überdies viele Anschlussbereiche hohe Ausbaukosten je Anschluss aufweisen, finden sich Ausbaucuster mit einer höheren Farbsättigung ebenfalls in den Anschlussbereichen mit niedrigeren Investitionen je Anschluss.

³¹ Die aktuell im Breitbandatlas dargestellten Bandbreitenkategorien von ≥ 100 , ≥ 200 , ≥ 400 und ≥ 1000 Mbit/s waren zum Zeitpunkt der Datenerhebung nicht verfügbar und sind in der Kategorie mindestens 50 Mbit/s enthalten.

Der Ausbaugrad der Kabel-TV Netze mit ≥ 50 Mbit/s ist im Vergleich dazu, deutlich präsenter innerhalb Deutschlands, wie man an der zweiten Abbildung erkennen kann. Der Ausbau findet gut über Deutschland verteilt statt, aber auch meistens in den Zentren und Zentrallagen kleinerer Städte und Ortschaften. Dabei deckt sich der sichtbare Ausbau im Kabel-TV Netz fast vorbildlich mit den Anschlussbereichen mit niedrigeren FTTH-Investitionen je Anschluss. Wie weit die Glasfaser in diesen Ausbaugebieten jeweils reicht, ist aufgrund unserer Datenbeschränkung nicht auszumachen (die Erstreckung der Ausbaugebiete höherer Bandbreiten sind hier nicht zu erkennen).

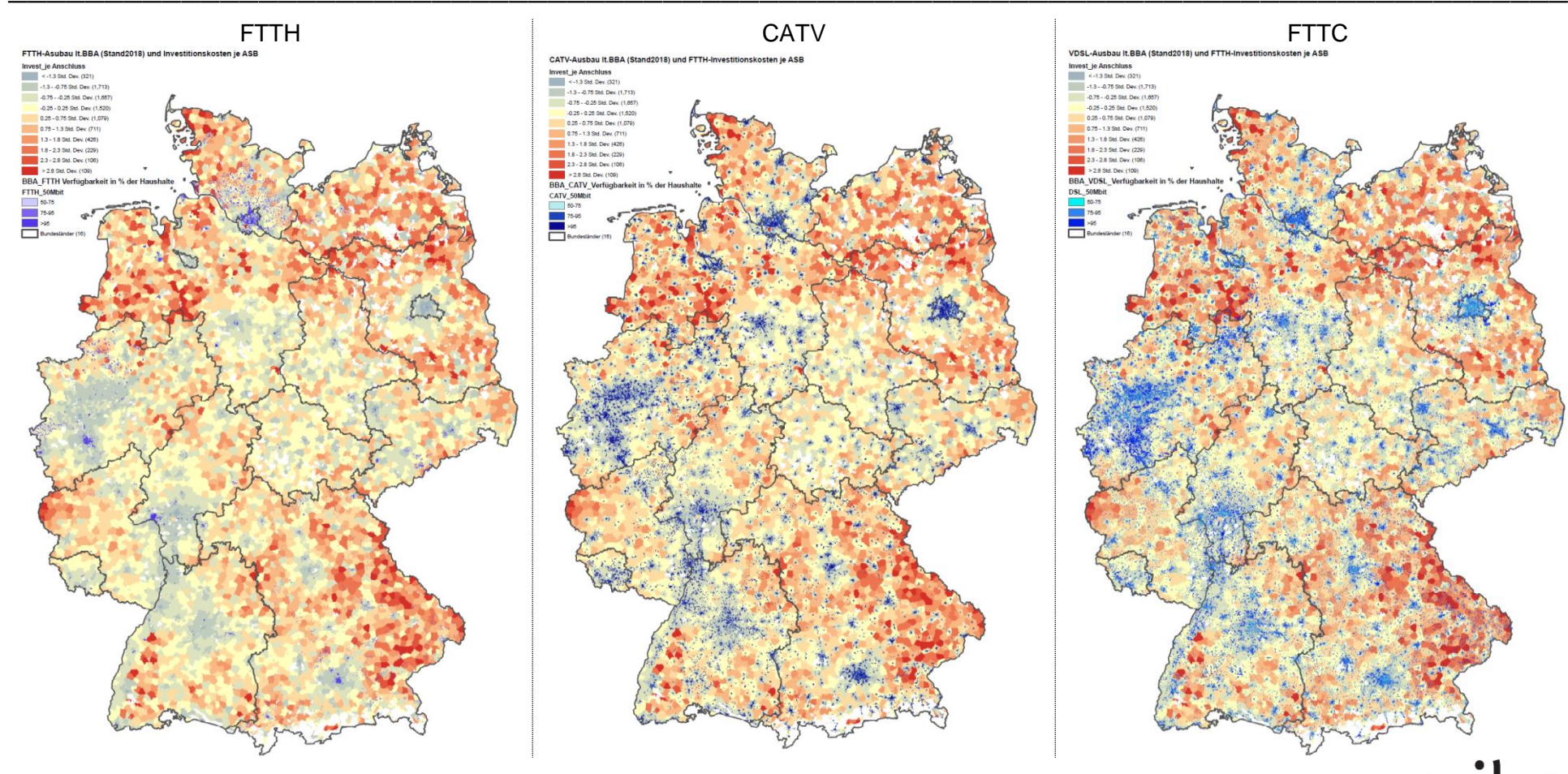
Die dritte Abbildung zeigt den VDSL-Ausbau für Deutschland. Dieser weist die stärkste Verbreitung in der Fläche und in den Zentren auf. Ebenso findet man in den dichten und für den FTTH Ausbau billigeren Gebieten natürlich eine stärkere Farbsättigung. Doch anders als bei den Technologien FTTH und CATV ist die Streuung in die Fläche deutlich höher.

Zusammenfassend lässt sich aus der räumlichen Lage des Breitbandausbaus erkennen, dass bei allen drei Technologien ähnliche Treiber des Ausbaus wirken, wie Skalenvorteile und Dichtevorteile. Auch scheinen die Kostenbestandteile des Netzes wie Trassenkosten, wenn auch nicht gleich, so doch hinreichend dominant für alle drei Technologien zu sein. Man erkennt, dass die in den Abbildungen rot gefärbten Gebiete in der überwiegenden Zahl auch nur wenig ausgebaut sind, während in den Ballungsräumen häufig die Wahl zwischen mindestens zwei breitbandigen Technologien besteht.

Aus der Verbreitung der drei Technologien wird auch klar, dass die Investitionskosten nicht die einzige Determinante des Ausbaus darstellen. Dazu zählen sicherlich die verfügbaren Technologien, die Nachfrage nach Bandbreite, die Nachfragedichte, die Wettbewerbssituation und natürlich historischen Gegebenheiten und Pfadabhängigkeiten.

Was in den drei Abbildungen nicht dargestellt ist, betrifft die räumliche Lage nicht versorgter Gebiete in allen drei Technologien. Diese finden sich sowohl in Gebieten niedriger Investitionen je Anschluss aber auch besonders in den teureren Gebieten. Das sind diejenigen Sprengel, die zu einer Flächendeckung fehlen. Aufgrund der hier gewählten grafischen Auflösung, ist die Darstellung dieser Gebiete nicht möglich.

Abbildung 4-10: Ausbaustand von FTTH, CATV und FTTC – vergleichende Darstellung



Quelle: WIK, Landesgrenzen und Kreisgrenzen BKG, Breitbandatlas 04/2018

5 Regressionsanalyse – geeignete Schätzer für die Kosten von Anschlussnetzen

In diesem Abschnitt werden die einzelnen Kostentreiber pro FTTH-Anschlussleitung anhand statistischer Methoden identifiziert, quantifiziert und bewertet. Zu diesem Zweck wird eine Regressionsanalyse durchgeführt, mit Hilfe derer valide statistische Zusammenhänge zu den Einflussvariablen auf den Investitionswert pro Anschluss identifiziert werden sollen.³²

Im vorliegenden Fall geht es darum, die Stärke des statistischen Zusammenhangs unterschiedlicher Kostentreiber (erklärende Variablen) auf die FTTH-Investitionshöhe pro Anschluss (erklärte Variable) zu untersuchen, um ein aussagekräftiges Modell mit möglichst kausalen Zusammenhängen aufzustellen. Damit streben wir an, die Ermittlung von Investitionswerten für Teilnehmeranschlussnetze bei veränderten räumlichen und zeitlichen Abgrenzungen zukünftig erwartungstreu schätzen zu können. Für diesen Zweck gehen wir wie folgt vor:

Im ersten Schritt erfolgt eine Darstellung der Datenbasis, gefolgt von einer Einführung in die methodische Vorgehensweise der Regressionsanalyse sowie die Illustration einer einfachen Regression anhand der Teilnehmerdichte. Daran anschließend (Abschnitt 5.3.2) wird ein multiples Regressionsmodell aufgestellt, welches auf der Datenbasis der bottom-up Modellierung beruht. Diese Datenbasis stellt den statistischen Analyserahmen mit flächendeckenden Informationen sowohl über die Kostentreiber als auch über die Zielgröße „Investitionen pro Anschlussleitung“ dar. Auf Basis vollständiger Datenverfügbarkeit wird somit zunächst ein „Benchmark-Modell“ erstellt. Diese Regressionsanalyse erfolgt sowohl auf Ebene der Anschlussbereiche als auch für größere, aggregierte Gebiete (Gebietskörperschaften). Die Ergebnisse dieser Regressionsanalysen sind in den Abschnitten 5.3.2 und 5.5 dargestellt.

In einem zweiten Schritt soll ein Modell aufgestellt werden, welches sich nicht mehr der vollständigen Datenbasis der bottom-up Modellierung bedienen kann, sondern auf allgemein verfügbare Daten beschränkt ist. Ziel ist es dabei, ein statistisches Modell aufzustellen, das zum Benchmark-Modell vergleichbare gute Ergebnisse erzeugt. Insofern dient das Benchmark-Modell als Richtschnur für diejenigen Modelle, die es in einem weiteren Schritt mit den üblichen Datenbeschränkungen aufzustellen sind.

³² Die Regressionsanalyse ist ein statistisches Verfahren, mit Hilfe dessen versucht wird, beispielsweise die statistische Abhängigkeit zwischen zwei Größen zu quantifizieren, die Entwicklung einer Kennzahl zu prognostizieren, oder auch die tatsächliche Wirkung einer Maßnahme auf eine Zielgröße empirisch zu bewerten.

5.1 Datenbasis

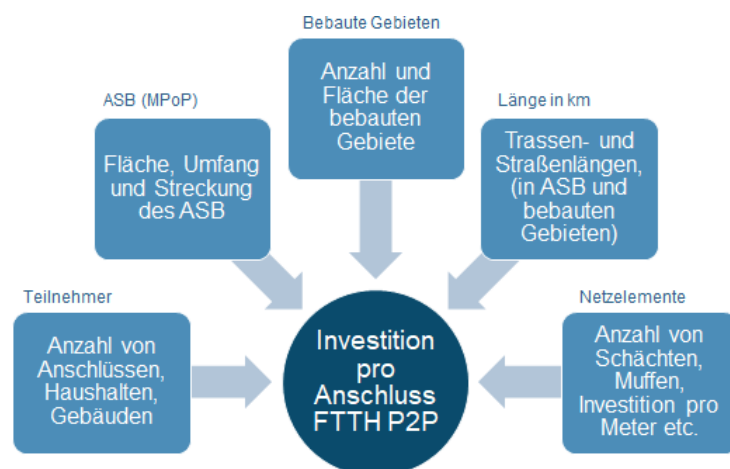
Als Datengrundlage für die oben beschriebene quantitative Analyse wird insbesondere auf die im bottom-up Modell verwendeten, georeferenzierten Daten abgestellt. Sie lassen sich als struktur- und netzbezogene Daten kennzeichnen.

Unter die Strukturdaten fallen beispielsweise sowohl die Anzahl von Anschlüssen, Haushalten oder Gebäuden, als auch Flächen, Umfang und Erstreckung des Anschlussbereichs (siehe Abbildung 5-1). Diese Größen werden auf Basis von GIS-Datenbeständen (AZ-Direct Bertelsmann, AZ-Haushalts und Gebäuderaster, TeleAtlas Straßenlayer, 2018) und der deutschen Verwaltungsgebiete des Bundesamtes für Kartographie und Geodäsie (VG250-EW (Kompakt) 2018) für jeden Anschlussbereich ermittelt.

Zu den netzbezogenen Daten zählen dagegen die durch Anwendung des bottom-up Modells ermittelten Größen, wie z.B. Verzweigerkabel- und Hauptkabeltrassenlängen, Anzahl der Kabelverzweiger, Anzahl der Schächte und Muffen (siehe z.B. „Netzelemente“ in Abbildung 5-1). Hierbei handelt es sich um anlagebezogene Mengengerüste, die Output der Modellrechnungen darstellen. Diese Informationen liegen mit ihrer Berechnung bereits in einer anschlussbereichsgenauen Abgrenzung vor (siehe Abschnitt 3 und 4).

Sowohl die georeferenzierten Strukturdaten als auch die über die Modellierung entstandenen netzbezogenen Daten werden in eine gemeinsame Datenbank eingepflegt und für die statistische Analyse verwendet.

Abbildung 5-1: Struktur- und netzbezogene Daten als Einflussfaktoren des FTTH-Investitionsbedarfs pro Anschlussleitung



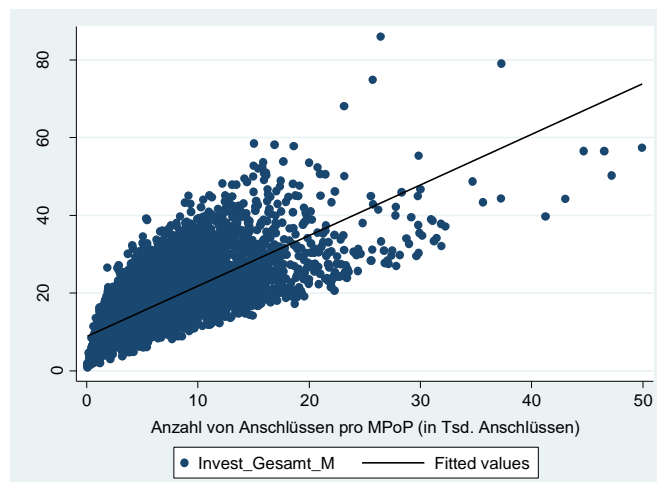
Quelle: WIK

5.2 Methodische Vorgehensweise

Im linearen Regressionsmodell ist der Zusammenhang zwischen zwei oder mehreren Merkmalen (Variablen) durch eine lineare Funktion beschrieben.³³ Diese Funktion (oder sogenannte „Regressionsgleichung“) kann, auf Basis von hinreichenden Beobachtungswerten, diejenige mathematische Gerade finden, die die Punktwolke möglichst „gut“ beschreibt. Dies wird dadurch erreicht, dass die Summe der Abweichungsquadrate zwischen den wahren Beobachtungswerten und dieser Geraden minimiert werden. Das entspricht der sogenannten KQ-Methode („Kleinste Quadrate“).

Im nachstehenden Punktwolkendiagramm werden exemplarisch die Anzahl der Anschlüsse pro Anschlussbereich (auf der X-Achse) mit der dazugehörigen Investitionshöhe bei flächendeckendem FTTH Ausbau (auf der Y-Achse) dargestellt:

Abbildung 5-2: Zusammenhang zwischen der Anzahl von Anschlüssen und den erforderlichen FTTH-Investitionen pro MPoP



Quelle: WIK

Aus Abbildung 5-2 ist ein positiver Zusammenhang zwischen der Anzahl der Anschlüsse und den erforderlichen FTTH-Investitionen pro MPoP ersichtlich. Wie erwartet steigt die erforderliche FTTH-Investition innerhalb eines Anschlussbereichs mit zunehmender Anzahl von Anschlüssen. Das vorliegende Beispiel, bei dem es sich um lediglich zwei Variablen handelt, entspricht einem sogenannten „einfachen“ linearen Regressionsmodell. Aus der Abbildung 5-2 ist ersichtlich, dass die Abstände zwischen der schwarzen Linie (Regressionsgerade) und den blauen Punkten (wahren Beobachtungswerten) sehr groß sein können. Wir haben untersucht, ob die Zielvariable „FTTH-Investition“ präziser erklärt werden kann, in dem weitere relevante Einflussfaktoren im Modell be-

³³ „Linear“ in den Parametern (d.h. hinsichtlich der Koeffizienten). Siehe Wooldridge (2012).

rücksichtigt werden. Methodisch kann dies mit einem sogenannten „multiplen Regressionsmodell“ analysiert werden.

In einem **multiplen linearen Regressionsmodell** sind mehrere erklärende Variablen aufzunehmen. Die Auswahl der Eingangsvariablen erfolgt nach Maßgabe der folgenden festgelegten Kriterien:

- **Theoretische Basis:** Der operationalisierte Einflussfaktor stützt sich auf eine ökonomische bzw. technische Überlegung, die in der wissenschaftlichen Theorie verankert ist. Er erklärt die Zielvariable durch nachvollziehbare Wirkungszusammenhänge, die mit den allgemeinen ökonomischen und technischen Grundsätzen konsistent sind.
- **Statistischer Beweis:** zwischen der erklärenden und der erklärten Variablen liegt ein robuster und stabiler statistisch signifikanter Zusammenhang vor. Die erklärende Variable trägt somit Erklärungsgehalt. Ihr Einbau ins Modell führt zu einer statistisch-signifikanten Erhöhung der Modellgüte, die über gängige ökonomische Kennzahlen, beispielsweise das (adjustierte) Bestimmtheitsmaß (adj. R^2), nachgewiesen werden kann.³⁴

In Erweiterungen eines einfachen Regressionsmodells sind daher nicht nur zusätzliche erklärende Variablen aufzunehmen, sondern darüber hinaus auch eine Transformation dieser Variablen dergestalt zu vollziehen, dass sie eine besser Erklärung der Investitionen pro Anschluss leisten können.

Dazu werden in der weiteren Analyse anhand der oben genannten Kriterien (theoretische Basis und statistischer Beweis), zwei grundlegende Anpassungen vorgenommen:³⁵

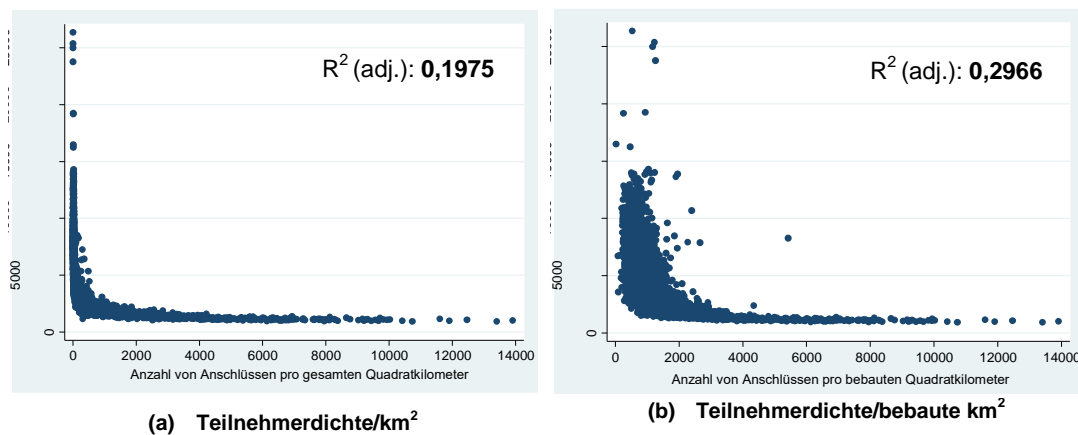
- Bei der ersten Anpassung wird die Variable „Zahl der Anschlüsse“ in eine relative Größe „Anschlüsse pro km^2 “ transformiert (bezogen auf die Gesamtfläche bzw. bebaute Fläche des Anschlussbereichs). Ebenso werden die Investitionen pro MPoP in Investitionen pro Anschluss überführt.
- Die sich daran anschließende zweite Transformation logarithmiert die verwendeten Variablen.

Die Zusammenhänge der ersten Transformation der Variablen werden in Abbildung 5-3 dargestellt.

³⁴ Exkurs: Das Bestimmtheitsmaß (R^2) beschreibt den Anteil der gesamten Variationen in den Daten, die durch das Modell erklärt werden können. Es liegt zwischen 0 und 1. Das Bestimmtheitsmaß erhöht sich automatisch mit zunehmender Anzahl von erklärenden Variablen, unabhängig davon, ob diese Erklärungsgehalt tragen. Um dies zu korrigieren, wird bei dem adjustierten Bestimmtheitsmaß die Aufnahme jeder neu hinzugenommenen erklärenden Variable „bestraft“. Diese Korrektur ist ein etabliertes ökonomisches Vorgehen und wird vom Statistik Programm STATA automatisch durchgeführt.

³⁵ Damit wir versucht, relevante Grundsätze der Produktionstheorie im Telekommunikationsbereich, wie Skaleneffekte und Dichtevorteile, zu berücksichtigen.

Abbildung 5-3: Zusammenhang zwischen Teilnehmerdichte und FTTH-Investition pro Anschluss (bezogen auf (a) ASB-Fläche bzw. (b) bebaute Fläche [km²])



Quelle: WIK

wik

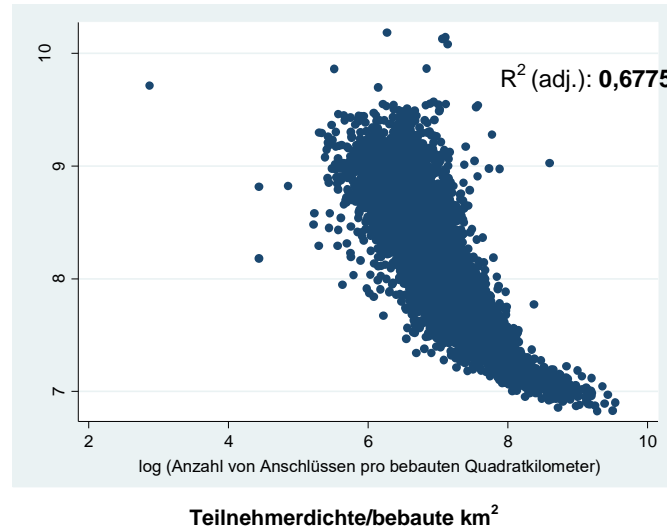
Wie erwartet sinkt die erforderliche FTTH-Investition pro Anschluss, je höher die Anschlussdichte des Anschlussbereichs ist. Da bei der Abgrenzung von Anschlussbereichen oft irrelevante Flächen für den FTTH-Ausbau miteinbezogen werden, beispielsweise land- oder forstwirtschaftliche Flächen, zeigt die oben dargestellte rechte Abbildung den Zusammenhang zwischen FTTH-Investition pro Anschluss und der Teilnehmerdichte, wenn diese sich ausschließlich auf die bebauten ASB-Flächen bezieht.³⁶ Dadurch ist eine Erhöhung des Erklärungsgehalts zu erwarten, da Kostenersparnisse aus Dichtevorteilen bei geringeren Messfehlern eindeutiger aufgedeckt werden können. Statistisch wird diese Hypothese bestätigt. Das adjustierte Bestimmtheitsmaß steigt dadurch von 0,1975 auf 0,2966.

Die zweite Anpassung betrifft die funktionale Form des vorliegenden Zusammenhangs: Hier geht es um die Frage, ob der Zusammenhang zwischen FTTH-Investition pro Anschluss und der Teilnehmerdichte pro bebauter Fläche linear oder nicht-linear ist. Ob tatsächlich eine lineare oder beispielsweise eine abnehmende oder exponentielle Beziehung vorliegt, wird dabei nicht a priori festgelegt. Vielmehr wird auf bereits vorliegende Beobachtungen zurückgegriffen, wie sie mit den Beobachtungswerten in Abbildung 5-3 vorliegen: Im vorliegenden Beispiel kann auf einen konvexen Zusammenhang geschlossen werden. In dünn besiedelten Gebieten (unter 2.000 Teilnehmer pro bebautem km²) wirkt sich die Änderung der Teilnehmerdichte auf die Höhe der FTTH-Investition pro Anschlussleitung wesentlich stärker aus als in den teilnehmerdichteren Gebieten. Um dieser Nichtlinearitäten entlang der Kurve Rechnung zu tragen, wird eine Log-Log-Transformation vollgezogen.³⁷

³⁶ Vgl. FTTH Council Europe (2017)

³⁷ Verwendet wurde hier der natürliche Logarithmus.

Abbildung 5-4: Zusammenhang zwischen Teilnehmerdichte pro bebauter Fläche [km²] und FTTH-Investition pro Anschluss, logarithmiert



Quelle: WIK

Im Gegensatz zu linear-linearen Modellen drücken log-log-Modelle die Zusammenhänge in relativen Änderungsraten (sogenannten „Elastizitäten“) aus.³⁸ Log-log transformierte Daten haben zusätzlich noch den Vorteil, die Spanne zwischen Beobachtungswerten zu verkürzen. Das verringert die Empfindlichkeit der KQ-Schätzung gegen extreme Werte. Potentiell unerwünschte Effekte aus einer heteroskedastischen³⁹ und schiefen Verteilung werden bei strikt positiven Variablen, wie es bei FTTH-Investitionen und Teilnehmerdichten der Fall ist, erheblich gemildert.⁴⁰ Das adjustierte Bestimmtheitsmaß springt somit von 0,2966 auf 0,6775.

5.3 Benchmark-Modell: Multiple Regression auf Ebene der Anschlussbereiche

5.3.1 Multiple Regression – Auswahl der Variablen

Im Rahmen einer multiplen Regressionsanalyse ist das Modell um maßgebliche Variablen zu erweitern. Für das „Benchmark“-Modell wird hauptsächlich auf verfügbare Netz-

³⁸ Das „linear-lineare“ Modell beschreibt eine Regressionsgleichung, bei der sowohl die erklärten als auch die erklärenden Variablen unter nominaler Skalierung verwendet werden. In „Log-Log“-Modellen setzt die Regressionsgleichung auf logarithmischen transformierten Variablen auf.

³⁹ D.h. die Varianz der Störterme ist nicht konstant.

⁴⁰ Vgl. Wooldridge (2012). Introductory Econometrics: a Modern Approach, S. 193

daten auf Anschlussbereichsebene aus dem NGA-Modell zurückgegriffen. Diese werden in der untenstehenden Tabelle dargestellt:

Tabelle 5-1: Datenbasis aus dem WIK NGA-Modell zur Ableitung des „Benchmark“-Modells

Netzelement	Variable aus dem NGA-Modell	Einheit
Teilnehmer	Anzahl von Anschlüssen	1
	Anzahl von Anschlüssen pro ASB-Fläche in qkm (gesamte qkm)	1/ km ²
	Anzahl von Anschlüssen pro bebauter ASB-Fläche in qkm (bebaute qkm)	1/ bkm ²
Hauszuführung	CPE	€
	gebäudeinterne Verkabelung	€
	Lotlänge	m
VZK & HK	Trassenlänge in VZK-Segment	km
	Trassenlänge in HK-Segment	km
	∅ Investition pro Trassenmeter	€/m
	Anzahl von Muffen pro km	1/km
	Investition pro Muffen & Schächte	€/1
Verteiler	Anzahl von KVZ pro MPoP	1
	Investition kundenseitiger MPoP	€
Investition	FTTH-Investition	€
	FTTH-Investition pro Anschluss	€/ 1

Quelle: WIK

Die oben gelb markierten Variablen stellen die bereits berücksichtigten Variablen dar. Die weiteren aufgelisteten erklärenden Variablen sind Outputgrößen des NGA-Modells. Diese verfügen bereits über eine „theoretische Basis“, weswegen die erste Anforderung an das statistische Modell, die wir in Abschnitt 5.2 eingeführt haben, somit erfüllt wird.

Entsprechend sind für die zweite Anforderung („statistischer Beweis“, siehe ebenfalls Abschnitt 5.2) bei der Variablenauswahl noch die statistischen Kriterien zu prüfen. Analog zur „Teilnehmerdichte“ werden dabei unterschiedliche Variablenvarianten getestet. Konkret werden sie nicht nur auf Basis ihrer absoluten Größe pro Anschlussbereich berücksichtigt, sondern auch transformiert, d.h. im Verhältnis zu ihrer Anzahl von Anschlüssen bzw. zur bebauten ASB-Fläche (bebaute Quadratkilometer) getestet. Ausgenommen werden diejenige Variablen, die aus dem NGA-Modell bereits in Relation gesetzt sind, beispielsweise „Anzahl von Muffen per km“ oder die „durchschnittliche Investition pro Trassenmeter“. Investitionen in CPE und gebäudeinterne Verkabelung sind pro Anschluss konstant und tragen nicht zum Erklärungsgehalt des Modells bei. Diese Variablen können ohne Verlust von der weiteren Analyse ausgeschlossen werden.

Den Ergebnissen des NGA-Modells zufolge machen die Trassenkosten im VZK- und HK-Segment **62%** der Gesamtinvestition für FTTH P2P aus, und werden daher vorrangig getestet. Die übrigen Netzelemente werden im Rahmen der darauf folgenden Robustheitstests berücksichtigt.

Die Daten erlauben es, die Trassenkosten über unterschiedliche Faktoren abzubilden: einmal über den mengenmäßigen Kostentreiber der Trassenlänge (in km) und ein anderes Mal über den Faktor Investitionen pro Trassenmeter [€/m], der die Kostenintensität widerspiegelt. Mit welchen operationalisierten Variablenvarianten und welcher funktionalen Form diese Kostenfaktoren letztendlich modelliert werden, wird empirisch anhand von Testverfahren bestimmt. Die nachstehende Tabelle zeigt die Ergebnisse der Testverfahren zur Modellwahl:

Tabelle 5-2: Testergebnisse und statistische Kennzahlen für die Variablenauswahl des „Benchmark“-Modells

		Erklärungs- gehalt	Funktionale Form	Modellüber- frachtung
		R ² adj. [0,1]	RESET [t-stat]	VIF>10, max.
	log (Teilnehmerdichte [1/ bkm ²])	0,6775	210,8	1,0
Trassen- längen	+ log (Trassenlänge VZK [km])	0,6775	210,6	1,0
	+ log (Trassenlänge HK [km])	0,8265	82,4	2,7
quadriert	+ log (Trassenlänge VZK [km]) ²	0,6791	221,25	1,1
	+ log (Trassenlänge HK [km]) ²	0,7603	99,09	1,4
Trassen- dichten	+ log (Trassenlänge VZK [bkm ²])	0,9618	1.690,6	1,0
	+ log (Trassenlänge HK [bkm ²])	0,9501	152,3	4,8
quadriert	+ log (Trassenlänge VZK [bkm ²]) ²	0,9619	1.717,9	1,0
	+ log (Trassenlänge HK [bkm ²]) ²	0,9563	8,91	1,4
Kosten/ Meter	+ log (ØTrassenkosten pro Meter [€/m])	0,8187	55,5	1,9
	+ log (ØTrassenkosten pro Meter [€/m]) ²	0,8380	231,5	1,9

Quelle: WIK

Die Erweiterungen des bestehenden Modells um jeweils „Trassenkosten pro Meter“ und „Trassenlängen“ erhöhen das Bestimmtheitsmaß von ursprünglich 0,6775 auf 0,8187 bzw. 0,8265. Dennoch sind die „Trassendichten auf Basis von bebauten Gebieten“, diejenige zusätzliche erklärende Variable, die den höchsten Anstieg am Erklärungsgehalt liefert. Dabei steigt das R² von 0,6775 auf 0,95 bzw. 0,96.

Die korrekte funktionelle Form der identifizierten Modelle wird statistisch durch den RESET-Test geprüft. Das Regressionsmodell mit der Variable „Trassendichte im VZK-Segment“ (sowohl nominal als auch quadriert) besteht diesen Test nicht. Die RESET Test-Statistik ist hier sehr hoch. Auch in den Regressionsresiduen spielt sich das hier stark wieder, in dem ein unerwünschtes Datenmuster in den Residuen sehr leicht erkennbar ist. Im Gegensatz zum Modell mit „Trassendichte im VZK-Segment“ liefert das Regressionsmodell mit dem Polynom der „Trassendichte im HK-Segment“ (gelb markiert) die niedrigste RESET-Teststatistik. Zugleich hat dieses Modell eine VIF (*Variance*

Inflation Factor) unter 10. Mit einem VIF-Wert in Höhe von 1,4 deutet dieser Test auf keine Modellüberfrachtung hin (sogenanntes „Multikollinearitätsproblem“).⁴¹

5.3.2 Benchmark-Modell: Ergebnisse der multiplen Regression

Auf Basis von Variablen, die mit den allgemeinen ökonomischen und technischen Zusammenhängen in Einklang stehen, zusammen mit empirischen Testergebnissen aus dem Abschnitt 5.2 wird folgende Regressionsgleichung für das „Benchmark-Modell“ spezifiziert:

$$\ln(\text{Invest } p/A)_i = \beta_0 + \beta_1 \ln(\text{Teilnehmerdichte})_i + \beta_2 \ln(\text{Trassendichte HK})_i + \beta_3 \ln(\text{Trassendichte HK})_i^2 + \beta_4 \ln(\text{Kostenintensität})_{NUTS1} + \beta_5 \ln(\text{Urbanisierungsgrad})_i + \varepsilon_i$$

wobei der Index i für den i -ten Anschlussbereich (MPoP) steht. *Invest p/A* ist die erforderliche FTTH-Investition pro Anschlussleitung, *Teilnehmerdichte* ist die Anzahl von Anschlüssen pro bebautem Quadratkilometer. *Trassendichte HK* entspricht der Trassenlänge in HK-Segment pro bebauten Quadratkilometer, zentriert. *Kostenintensität* bildet die durchschnittlichen Investitionskosten pro Trassenmeter (im VZK und im HK-Segment) auf Ebene des Bundeslandes (*NUTS1*)⁴² ab und *Urbanisierungsgrad* ist der Flächenanteil bebauter Gebiete pro Anschlussbereich. ε ist der Fehlerterm. *ln* steht für den natürlichen Logarithmus.

Unter Anwendung des KQ-Schätzers stellt die nachstehende Tabelle die Regressionsergebnisse für das „Benchmark-Modell“ zusammen:

⁴¹ Bei Modellen mit quadratischen Termen sind die quadrierten und nicht-quadrierten Variablen zentriert
⁴² Das entspricht einem höheren Aggregationsniveau als dem sonst bisher verwendeten Anschlussbereich. Da auf Ebene des Anschlussbereiches die netzbezogenen Elementen (wie Trassenlängen in VZK- und HK-Segment) mit der durchschnittlichen Investition pro Meter sehr stark korrelieren (und dadurch potentiell eine saubere Trennung beider Effekte behindert wird), und mit Hinblick auf die Untersuchung auf Basis von öffentlichen zugänglichen Daten bei hochaggregierten regionalen Abgrenzungen wurde die Variable der Kostenintensität (operationalisiert durch die „durchschnittliche Investition pro Meter“) auf die EU-Kommission standardisierten räumlichen Abgrenzungen NUTS1 abgestellt.

Tabelle 5-3: STATA-Regressionsoutput für das „Benchmark-Modell“

(1)	
Benchmark-Modell	
VARIABLEN	Ln (FTTH-Investition pro Anschluss)
Ln (Teilnehmerdichte in bebauten km ²)	-0,460*** (0,004)
Ln (HK-Trassendicht in bebauten km ²)	0,341*** (0,004)
Ln (HK-Trassendicht in bebauten km ²) ²	0,051*** (0,002)
Ln (Kosten pro Meter, NUTS1)	0,290*** (0,021)
Ln (Flächenanteil bebauter Gebiete)	-0,052*** (0,004)
Konstante	9,969*** (0,094)
Beobachtungen	7.871
Bestimmtheitsmaß (R ²)	0,959
F (5, 7865)	35.345
RESET (t-stat)	2,42
Prob > F	0,064
VIF (min./ max.)	1,07 - 8,57

*Robuste Standardfehler in Klammern. Statistische Signifikanz: *** p<0.01, ** p<0.05, * p<0.1*

Quelle: WIK

Auf Basis einer großen Anzahl von Beobachtungen (7.871) ergibt die Regressionsanalyse statistisch signifikante Zusammenhänge für die fünf identifizierten Kostenfaktoren. Die modellierten Kostenfaktoren können bereits ca. 96% der Gesamtvariationen in den Daten erklären. Die daraus resultierenden Vorzeichen sind im Einklang mit der ökonomischen Theorie.

Die dargestellten Koeffizienten quantifizieren die Stärke der vorliegenden Zusammenhänge. Diese können als „Elastizitäten“ interpretiert werden, d.h. eine x%-ige Änderung der erklärenden Variable führt zur einer y%-igen Änderung der Zielvariable. Für das Benchmark-Modell führt eine Verdoppelung der Teilnehmerdichte bzw. des Urbanisierungsgrads zu einer Reduktion der FTTH-Investition pro Anschluss um 46% bzw. 5%. Die Erhöhung der HK-Trassendichte erhöht den Investitionsbedarf in einem exponentiellen Verhältnis. Die Verdoppelung der Kostenintensität (operationalisiert als Trassenkosten pro Meter) erhöht den Investitionsbedarf um 29%.

5.4 Validierung der Regressionsergebnisse auf Ebene der Anschlussbereiche

Die grundlegenden Annahmen für die Validität einer KQ-Schätzung und die Robustheit des Benchmark-Modells werden unter Berücksichtigung folgender Anforderungen getestet:⁴³

- 1) Linearität in den Parametern
- 2) Zufallsstichprobe
- 3) Keine perfekte Multikollinearität
- 4) Erwartungswert der Störgrößen ist Null $E(u|x = 0)$
- 5) Homoskedastizität
- 6) Normalverteilung $\varepsilon \sim N(0, \sigma^2)$

Die Annahmen 1) und 2) sind per Konstruktion erfüllt. Durch die Anwendung von „robusten“ Standardfehlern ist die Annahme 5) unproblematisch.⁴⁴ Bei der Annahme 6) wird aufgrund der hinreichenden Anzahl von Beobachtungen von einer asymptotischen Normalverteilung ausgegangen.⁴⁵ Die Annahmen 3) und 4) müssen zwingend notwendig getestet werden. Das Nichterfüllen dieser Annahmen würde eine Regressionsanalyse mit verzerrten Koeffizienten erzeugen.

Die Annahme 3) *keine (perfekte) Multikollinearität*⁴⁶ wurde bereits im Benchmark-Modell über die VIF-Statistik evaluiert (siehe Tabelle 5-3). Da der ermittelte VIF-Wert den Grenzwert (i. H. v. 10) nicht übersteigt, liegt hier kein Indiz für Multikollinearität vor.

Eine weitere Stufe der Sicherheitsprüfung wird über die Robustheitsanalyse geschaffen. Anhand dieses Testverfahrens wird die Sensitivität der geschätzten Koeffizienten gemessen, indem potentiell ausgelassene Einflussfaktoren ins Modell aufgenommen werden. Dadurch wird sichergestellt, dass über die VIF ungedeckte Symptome von „Multikollinearität“ wie Vorzeichenwechsel, sprunghafte und instabile Koeffizienten bei Modelländerungen untersucht werden können. Zugleich kann die Robustheitsanalyse als Testverfahren für die Annahme 4) *Erwartungswert der Störgrößen ist Null $E(u|x = 0)$* verwendet werden.⁴⁷ Diese Annahme ist von zentraler Bedeutung. Sie unterstellt, dass die berücksichtigten Einflussfaktoren mit dem Fehlerterm nicht korreliert sind, d.h. dass

⁴³ Vgl. Wooldridge (2012). *Introductory Econometrics: a Modern Approach*, S.157.

⁴⁴ In der Statistiksoftware STATA kann die Regression mit robusten Standardfehlern durch den Befehl „reg [...] , robust“ durchgeführt werden.

⁴⁵ Vgl. Wooldridge (2012). *Introductory Econometrics: a Modern Approach*, S. 120.

⁴⁶ Das Multikollinearitätsproblem liegt bei multiplen Regressionsanalysen dann vor, wenn die Effekte von zwei oder mehreren erklärenden Variablen voneinander nicht eindeutig abgegrenzt werden können. Dies führt dazu, dass die Koeffizienten der betroffenen Variablen je nach Modellspezifikationen, selbst bei kleineren Änderungen, radikal unterschiedliche Ergebnisse aufweisen können.

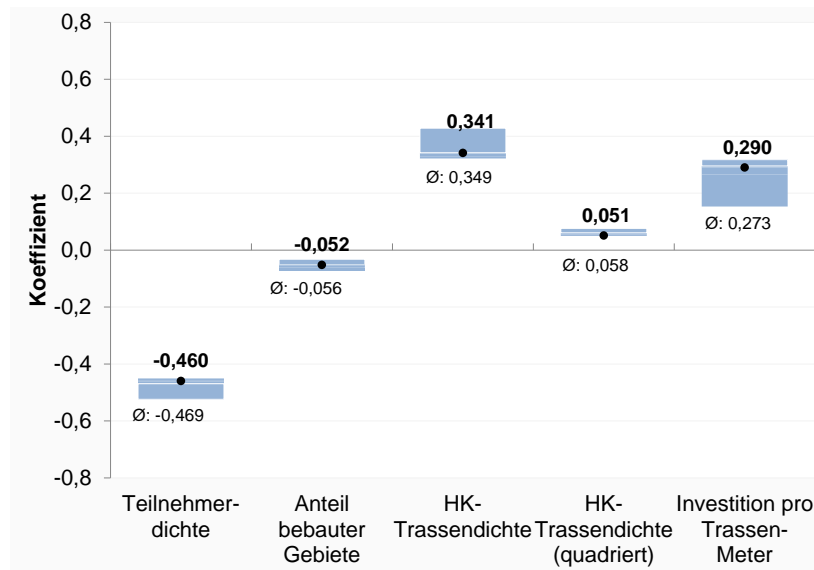
⁴⁷ Es liegt kein statistischer Diagnostik-Test (beispielsweise VIF, RESET-Test, etc.), mit dem man diese Annahme direkt testen kann. Daher wird in der Praxis diese Annahme über eine Robustheitsanalyse evaluiert.

ausgelassene Variablen keine Rolle spielen. Über die Robustheitsanalyse kann getestet werden, inwiefern ausgelassene Variable, sofern identifizierbar, unproblematisch sein können.

Für den Zweck der Robustheitsanalyse werden neun zusätzliche Regressionsmodelle aufgestellt. Jedes Modell schließt eine potentiell ausgelassene Variable ein. Berücksichtigte ausgelassenen Variablen umfassen die Lotlänge, Anzahl der KVZ pro Anschlussbereich, Anzahl von Muffen pro km, Trassenkosten pro Meter und ASB (im VZK-Segment), die Standardabweichung der Trassenlängen im HK-Segment pro ASB bezogen auf den Landkreisdurchschnitt („Dispersionsmaß“), die Gesamtfläche des Anschlussbereiches, die Trassenlängen in VZK-Segment und das HK/VZK-Verhältnis hinsichtlich der Trassenlängen. Für das erste Robustheitsmodell wird die Sensitivität des Benchmark-Modells gegenüber extremen Werten (sogenannten „Ausreißern“) getestet.

Die Ergebnisse der Robustheitsanalyse der identifizierten Kostenfaktoren im Benchmark-Modell werden grafisch in Abbildung 5-5 **Fehler! Verweisquelle konnte nicht gefunden werden.** und tabellarisch in Tabelle 5-4 Tabelle 5-4 dargestellt.

Abbildung 5-5: Robustheitsanalyse: Spanne der Schätzungen des Benchmark-Modells nach Eingang von zusätzlichen erklärenden Variablen



Quelle: WIK

In der voranstehenden Grafik sind für die Kostenfaktoren die geschätzten Koeffizienten dargestellt. Im Benchmark-Modell konnten wir eine Abweichung der Koeffizienten zwischen -12% und +6% vom Modelldurchschnitt feststellen. Vergleicht man diese mit den Koeffizienten der Robustheitsanalyse, so ist festzustellen, dass vor allem die Kostenfaktoren „Teilnehmerdichte in bebauten km²“ und „Trassendichte in bebauten km²“ robuste

Schätzgrößen darstellen. Für diese Variablen beträgt die Abweichung im Durchschnitt lediglich -2%. Keine Variable zeigt relevante Sprünge in den Werten der Koeffizienten, wechselt ihr Vorzeichen oder verliert ihre statistische Signifikanz. Somit liegen KEINE Symptome von „Multikollinearität“ vor. Das Benchmark-Modell zeichnet sich somit durch eine relative hohe Stabilität und Robustheit aus.

Tabelle 5-4: STATA-Regressionsoutput des Benchmark-Modells und der Robustheitsmodelle (2 bis 10)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SPEZIFIKATION:	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
VARIABLEN	Benchmark	ohne Ausreißer	inkl. Lotlänge	inkl. KVZ	inkl. Muffen	inkl. VZK Kosten/ m	inkl. HK Dispers.	inkl. Fläche ASB	inkl. VZK-TrassenL.	inkl. Verhältnis HK/ VZK
Ln (Teilnehmerdichte in bebauten km ²)	-0,460*** (0,004)	-0,467*** (0,004)	-0,470*** (0,004)	-0,460*** (0,004)	-0,451*** (0,004)	-0,474*** (0,004)	-0,455*** (0,004)	-0,464*** (0,004)	-0,459*** (0,004)	-0,524*** (0,005)
Ln (HK-Trassendicht in bebauten km ²)	0,341*** (0,004)	0,340*** (0,004)	0,342*** (0,004)	0,328*** (0,005)	0,332*** (0,004)	0,382*** (0,005)	0,329*** (0,004)	0,340*** (0,004)	0,322*** (0,004)	0,428*** (0,005)
Ln (HK-Trassendicht in bebauten km ²) ²	0,051*** (0,002)	0,053*** (0,002)	0,061*** (0,002)	0,053*** (0,002)	0,054*** (0,002)	0,060*** (0,002)	0,055*** (0,002)	0,050*** (0,002)	0,062*** (0,002)	0,076*** (0,002)
Ln (Kosten pro Meter, NUTS1)	0,290*** (0,021)	0,299*** (0,020)	0,291*** (0,021)	0,295*** (0,021)	0,318*** (0,021)	0,256*** (0,021)	0,296*** (0,021)	0,290*** (0,021)	0,257*** (0,021)	0,153*** (0,019)
Ln (Flächenanteil bebauter Gebiete)	-0,052*** (0,004)	-0,050*** (0,004)	-0,066*** (0,004)	-0,063*** (0,004)	-0,060*** (0,004)	-0,032*** (0,004)	-0,052*** (0,004)	-0,057*** (0,004)	-0,072*** (0,004)	-0,049*** (0,004)
Konstante	9,969*** (0,094)	9,982*** (0,089)	9,503*** (0,096)	9,853*** (0,096)	9,700*** (0,096)	9,276*** (0,100)	9,457*** (0,095)	10,010*** (0,094)	9,355*** (0,096)	11,197*** (0,094)
Beobachtungen	7.871	7.814	7.871	7.871	7.855	7.871	7.871	7.871	7.871	7.871
Bestimmtheitsmaß (R ²)	0,959	0,96	0,961	0,959	0,96	0,96	0,96	0,959	0,962	0,969

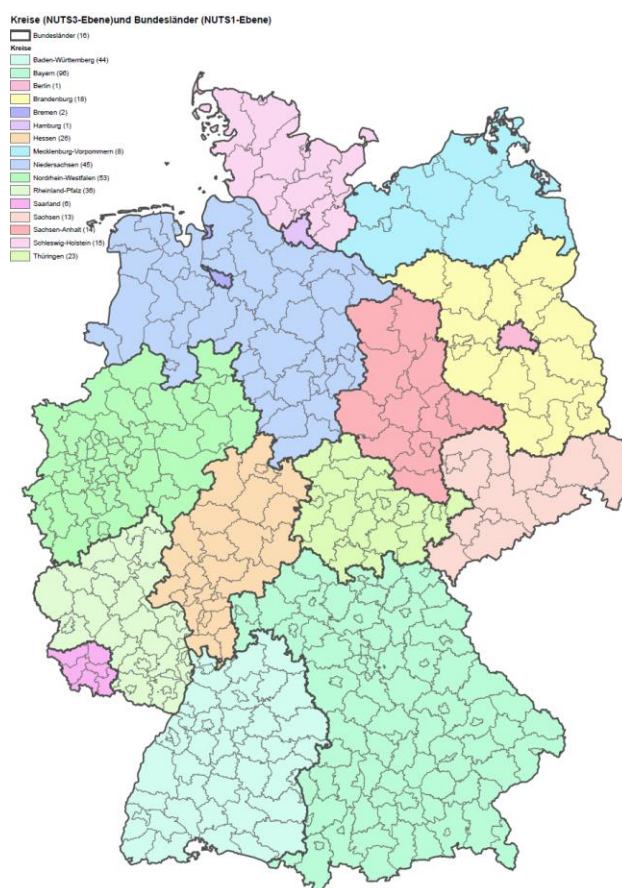
Robuste Standardfehler in Klammern. Statistische Signifikanz: *** p<0.01, ** p<0.05, * p<0.1

Quelle: WIK

5.5 Validierung der Regressionszusammenhänge für Gebietskörperschaften (Kreise und Bezirke)

Die Anwendbarkeit und Befunde des „Benchmark-Modells“ werden in dem nachfolgend dargestellten Analyseschritt auf einem höheren räumlichen Aggregationsniveau getestet. Dabei werden die MPoP-abhängigen Netzdaten auf Ebene der Landkreise und Regierungsbezirke aggregiert. Es wird weiterhin auf die Datenbasis der bottom-up Modellierung zurückgegriffen. Beide Aggregationsgrößen (Kreise und Bezirke) entsprechen den Abgrenzungen der von der EU-Kommission verwendeten NUTS-Klassifikation (*nomenclature of territorial units for statistics*) NUTS2 und NUTS3.

Abbildung 5-6: NUTS-Abgrenzung: NUTS1 (Länder) und NUTS3 (Landkreise)



Quelle: WIK; Verwaltungsgebietsgrenzen Bundesamt für Kartographie und Geodäsie

Auf Basis dieser aggregierten Datengrundlage wird die Regressionsgleichung im Abschnitt 5.3.2 erneut getestet. Die Ergebnisse der vollzogenen Regressionen werden in der nachstehenden Tabelle gezeigt.

Tabelle 5-5: STATA-Regressionsoutput des Benchmark-Modells auf Ebene der Anschlussbereiche, der Landkreise (NUTS3) und Regierungsbezirke (NUTS2)

VARIABLEN	Benchmark	Aggregation	
	ASB	NUTS3	NUTS2
Ln (Teilnehmerdichte in bebauten km ²)	-0,460*** (0,004)	-0,477*** (0,018)	-0,516*** (0,044)
Ln (HK-Trassendicht in bebauten km ²)	0,341*** (0,004)	0,445*** (0,018)	0,487*** (0,061)
Ln (HK-Trassendicht in bebauten km ²) ²	0,051*** (0,002)	0,034*** (0,012)	0,025 (0,031)
Ln (Kosten pro Meter, NUTS1)	0,290*** (0,021)	0,447*** (0,059)	0,853*** (0,195)
Ln (Flächenanteil bebauter Gebiete)	-0,052*** (0,004)	0,028 (0,018)	0,007 (0,061)
Konstante	9,969*** (0,094)	9,505*** (0,293)	7,887*** (0,941)
Beobachtungen	7.871	401	38
Bestimmtheitsmaß (R ²)	0,959	0,965	0,971

*Robuste Standardfehler in Klammern. Statistische Signifikanz: ***
p<0.01, ** p<0.05, * p<0.1*

Quelle: WIK

Auf Ebene der Landkreise (NUTS3-Ebene), zeigt sich ein zum Basismodell vergleichbares Muster der Regressionskoeffizienten. Diese variieren in ihrer Höhe etwas, aber Größenordnung, Richtung und Signifikanz bleiben im Wesentlichen erhalten. Auch die Modelltests bestätigen die statistische Validität der Regressionsergebnisse auf diesem Aggregationsniveau.

Auf der Aggregationsebene der Regierungsbezirke (NUTS2) kann das Modell insgesamt statistisch nicht abgesichert werden. Auf dieser Ebene liegen nur noch 38 Beobachtungen vor (gegenüber 401 Beobachtungen auf Kreisebene). Es ist trotzdem bemerkenswert, dass die Ergebnisse auf NUTS2-Ebene sich in das Gesamtbild gut einfügen. Diese Einschätzung lässt sich an den im Betrag nur geringfügig gestiegenen Koeffizienten für Teilnehmerdichte und HK-Trassendichte auf NUTS2-Ebene im Vergleich zu den NUTS3 Ergebnissen festmachen.

5.6 Regressionsmuster mit vollständig öffentlich verfügbaren Daten

Nach Validierung des Benchmark-Modells soll nun der Versuch unternommen werden, ein Regressionsmodell mit den in der Forschung üblichen Datenbeschränkungen aufzustellen. Diese Beschränkung bezieht sich auf die Daten, die aktuell durch die flächendeckende Anwendung des NGA-Modells zur Verfügung stehen. Auch wenn den Ergebnisse der bottom-up Modellierung nur öffentlich verfügbare, georeferenzierte Daten zugrunde liegen, bedarf es dennoch ihrer modellspezifischen Aufbereitung und Berechnung, um sie in der vorliegenden Form für die bisheriger Regressionsanalyse im Benchmark-Modell verwendbar zu machen. Das betrifft die bisherigen verwendeten Daten zu den Anschlusszahlen sowie den berechneten Trassenlängen im VZK- und HK-Segment und den durchschnittlichen Investitionen pro Trassenmeter auf ASB-Ebene.

Deshalb sollen in einem weiteren Schritt die bereits identifizierten Einflussfaktoren des mehrfach getesteten Benchmark-Modells mit reinen öffentlich zugänglichen Daten operationalisiert werden. In dieser Hinsicht dient das Benchmark-Modell als Richtschnur für die Auswahl der erklärenden Variablen, die in diesem Schritt nun lediglich auf Basis öffentliche verfügbarer Datenquellen approximiert werden sollen (sogenannte „Proxy-Variablen“).⁴⁸

Um die Eignung dieser Variablen als „Proxy“ zu bewerten, wird anhand einer Korrelationsanalyse die Stärke des Zusammenhangs zwischen Proxy- und Benchmark-Variablen quantifiziert. In Tabelle 5-6 werden die potentiellen Proxy-Variablen aus öffentlich zugänglichen Datenquellen, zusammen mit deren Korrelationskoeffizienten zum approximierenden Einflussfaktor, dargestellt.

⁴⁸ Basis der GIS-Datenbestände (AZ-Direct Bertelsmann, AZ-Haushalts und Gebäuderaster, TeleAtlas Straßenlayer, 2018) und Verwaltungsgebiete der BRD des Bundesamtes für Kartographie und Geodäsie (VG250-EW (Kompakt) 2018))

Tabelle 5-6: Korrelationsanalyse zwischen Proxy- und Benchmark-Variablen

	Korrelations- koeffizient
Ln (Teilnehmerdichte in bebauten km²)	
Ln (Haushaltsdichte in bebauten km ²)	0,9866
Ln (Haushaltsdichte pro km ²)	0,8577
Ln (Bevölkerungsdichte, NUTS3)	0,7059
Ln (Bevölkerungsdichte in bebauten km ² , NUTS3)	0,6551
Ln (HK-Trassendichte in bebauten km²)	
Ln (Straßendichte pro bebauten km ²)	0,8413
Ln (Straßendichte außerhalb bebauter Gebiete pro bebauten km ²)	0,4614
Ln (Straßendichte innerhalb bebauter Gebiete pro bebauten km ²)	0,3483
Ln (Straßendichte pro km ²)	- 0,5658
Ln (Straßendichte außerhalb bebauter Gebiete pro km ²)	0,2020
Ln (Straßendichte innerhalb bebauter Gebiete pro km ²)	- 0,8455
Ln (Kosten pro Meter, NUTS1)	
Ln (BIP pro Kopf, NUTS1)	0,4153

Quelle: WIK

Aus der obenstehenden Tabelle ist ersichtlich, dass die Proxy-Variable „Haushaltsdichte in bebauten km²“ den im Benchmark-Modell identifizierten Einflussfaktor „Teilnehmerdichte in bebauten km²“ sehr gut approximiert. Zwischen den beiden Variablen beträgt der Korrelationskoeffizient 0,9866.

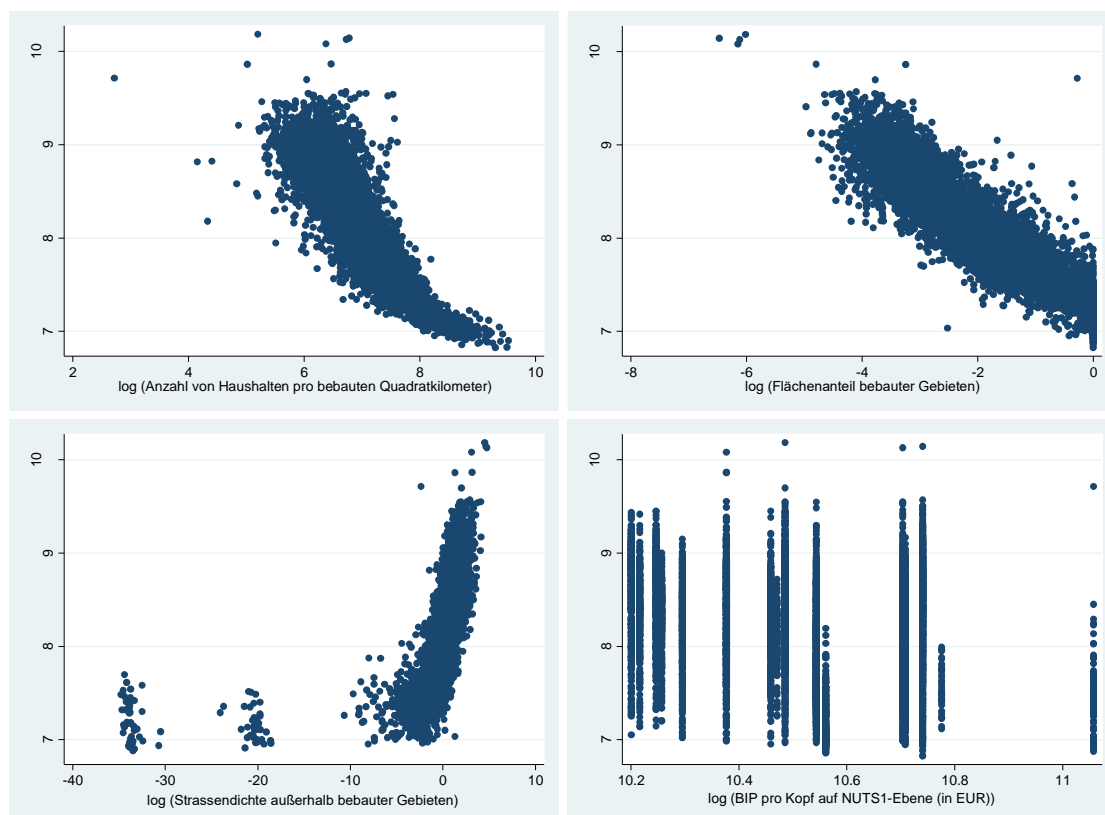
Für den Einflussfaktor „HK-Trassendichte in bebauten km²“ werden unterschiedliche Arten von Straßendichten⁴⁹ als potentielle Proxy-Variablen angenommen und ihre Korrelationskoeffizienten miteinander verglichen. Diese sind in der Tabelle 5-6 dargestellt. Sie beziehen sich auf den gesamten ASB, oder lediglich auf das Segment außerhalb oder innerhalb bebauter Gebiete des ASB. Konzeptionell ist es zu erwarten, dass Straßendichten außerhalb bebauter Gebiete, ähnliche wie die Trassendichten im HK-Segment, den Kosteneffekt aus der ASB-internen Dispersion der Teilnehmer am besten aufnehmen würden. Diese Erwartung stellt insbesondere auf den Aspekt der Streusiedlung ab, die einen höheren Trassenanteil außerhalb bebauter Gebiete erwarten lässt.

In der Tabelle 5-6 zeigt die Proxy-Variable „Straßendichte (im gesamten ASB) pro bebauten km²“ den höchsten positiven Korrelationskoeffizient, gefolgt von „Straßendichte außerhalb bebauter Gebiete pro bebauten km²“. Beide Variablen werden als Proxy-Variable für das Testverfahren aufgenommen (Proxy-Modell I bzw. II).

⁴⁹ Mit der Straßendichte wird auf das Verhältnis von Straßenlänge zur Anschlussbereichsfläche (bzw. Teilmengen der Fläche) abgestellt.

Schließlich werden die auf Ebene der Bundesländer (NUTS1) aggregierten Kosten pro Trassenmeter jeweils über das Bundeslands-Bruttoinlandsprodukt pro Kopf operationalisiert. Über die topografisch bedingten Unterschiede zwischen räumlichen Gebieten soll diese Proxy-Variable den Effekt des identifizierten Einflussfaktors widerspiegeln. Die nachstehenden Abbildungen stellen jeweils den Zusammenhang zwischen den ausgewählten Proxy-Variablen und der Zielvariable grafisch dar.

Abbildung 5-7: Zusammenhang zwischen der Zielvariable (FTTH-Invest pro Anschluss) und den ausgewählten Proxy-Variablen aus öffentlich zugänglichen Daten (logarithmiert)



Quelle: WIK

Nach der Ausgestaltung des validierten Benchmark-Modells und unter Rückgriff auf die oben identifizierten Proxy-Variablen werden zwei Regressionsmodelle auf Basis von lediglich öffentlich verfügbaren Daten aufgestellt. Die Tabelle 5-7 stellt die Regressionsergebnisse beider Proxy-Modelle dem Benchmark-Modell gegenüber.

Tabelle 5-7: Ergebnisvergleich zwischen Proxy- und Benchmark-Regressionen

VARIABLEN	Benchmark	Proxy I	Proxy II	PROXY-VARIABLEN
Ln (Teilnehmerdichte in bebauten km ²)	-0,460*** (0,004)	-0,442*** (0,006)	-0,480*** (0,006)	Ln (Haushaltsdichte in bebauten km ²)
Ln (HK-Trassendicht in bebauten km ²)	0,341*** (0,004)	0,029*** (0,004)	0,231*** (0,007)	Ln (Straßendichte (außerorts bei I) pro bebautem km ²)
Ln (HK-Trassendicht in bebauten km ²) ²	0,051*** (0,002)	0,001*** (0,000)	-0,005 (0,005)	Ln (Straßendichte (außerorts bei I) pro bebautem km ²) ²
Ln (Kosten pro Meter, NUTS1)	0,290*** (0,021)	0,298*** (0,012)	0,078*** (0,011)	Ln (BIP pro Kopf, NUTS1)
Ln (Flächenanteil bebauter Gebiete)	-0,052*** (0,004)	-0,252*** (0,006)	-0,132*** (0,006)	Ln (Flächenanteil bebauter Gebiete)
Konstante	9,969*** (0,094)	7,565*** (0,131)	10,423*** (0,131)	Konstante
Beobachtungen	7.871	7.618	7.871	Beobachtungen
Bestimmtheitsmaß (R ²)	0,959	0,907	0,928	Bestimmtheitsmaß (R ²)
F (5, 7865)	35.345	12.677	17.677	F(5, 7612 bzw. 7865)
RESET (t-stat)	2,42	245,65	159,61	RESET (t-stat)
Prob > F	0,064	0,000	0,000	Prob > F
VIF (min./ max.)	1,07 - 8,57	1,4 - 21,3	1,3 - 9,8	VIF (min./ max.)

Robuste Standardfehler in Klammern. Statistische Signifikanz:
 *** p<0.01, ** p<0.05, * p<0.1

Quelle: WIK

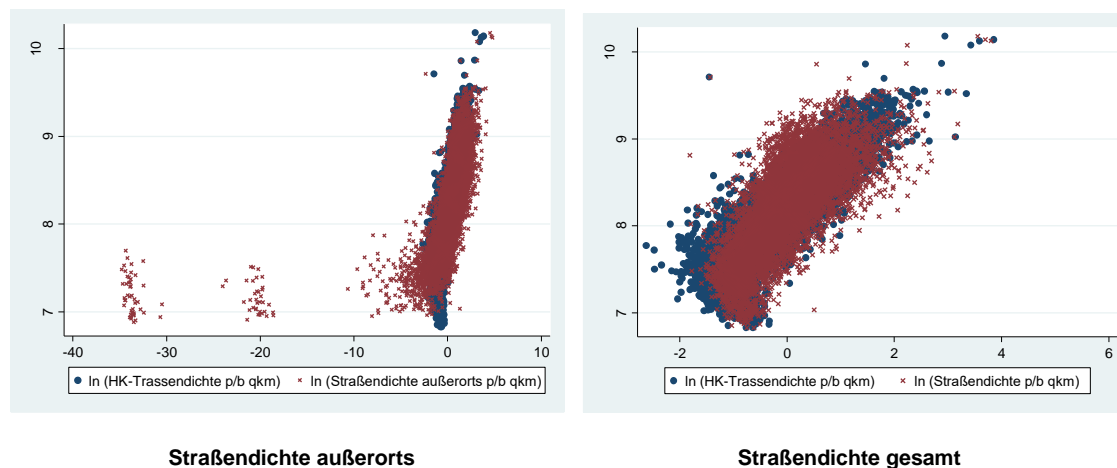
Aus der Tabelle 5-7 ist ersichtlich, dass über die aufgestellten Proxy-Modelle im Minimum über 90 Prozent der Datenvariationen der Zielvariable erklärt werden kann. Konsistent mit der Robustheitsanalyse zeigt sich die Proxy-Variable für die Teilnehmerdichte in bebauten km² „Haushaltsdichte in bebauten km²“ stabil. Bei anderen erklärenden Variablen liegen dagegen relative hohe Abweichungen zum Benchmark-Modell vor. Sowohl in der Höhe der Koeffizienten als auch in der statistischen Signifikanz.

Vor allem besteht keines der beiden aufgestellten Proxy-Modelle die üblichen Testverfahren, die im Abschnitt 5.4 für das Benchmark-Modell durchgeführt wurden. Das Bestehen dieser Tests ist für die Feststellung der Modellvalidität relevant (Feststellung ob die aufgestellten Modellannahmen als erfüllt angesehen werden können). Daher ist von einer Verzerrung in den Regressionskoeffizienten der Proxy-Variablen auszugehen. Somit ist die Validität der Proxy-Modelle nicht gegeben. Aufgrund der hohen t-Statistik wird der RESET-Test bei keinem Signifikanzniveau bestanden. Ein deutliches Datenmuster in den Residuen ist ebenfalls vorhanden. Zusätzlich deutet der VIF-Test für das Proxy-Modell I auf ein potentielles Multikollinearitätsproblem hin.

Diese Ergebnisse legen nahe, dass das Approximieren der Trassenlänge im Hauptkabel-Segment anhand öffentlicher Informationen über die Straßenlängen (gesamt oder außerorts) nicht hinreichend gut funktioniert. Hierzu haben wir die Daten nochmals grafisch aufbereitet: Die Abbildung 5-8 zeigt die Übereinstimmung der Approximation bei der öffentlich verfügbaren Datenbestände der Straßendichte (a) gesamt und (b) außerorts (**rot**) im Vergleich zur Trassendichte im Hauptkabelsegment bezogen auf die be-

bauten Flächen (**blau**). Es zeigen sich zwar relevante Überlappungen, jedoch ist diese offensichtlich nicht hinreichend um die Modelle statistisch absichern zu können.

Abbildung 5-8: Vergleich zwischen der Benchmark-Variable „HK-Trassendichte pro bebautem km²“ (**blau**) und der entsprechenden Proxy-Variable „Straßendichte ((a) außerorts und (b) gesamt) pro bebautem km²“ (**rot**)



Quelle: WIK

5.7 Zwischenfazit Regressionsanalyse

Auf Basis von ökonometrischen Analysen konnten die relevanten Einflussfaktoren des Investitionsbedarfs für den Glasfaserausbau bis zum Endkunde (FTTH P2P) ausgewertet, identifiziert und quantifiziert werden. Für die Investitionswertschätzung wurde die Relevanz der folgenden Einflussfaktoren empirisch bestätigt:

- a) **Teilnehmerdichte unter Berücksichtigung der Besiedlungsstruktur:** Anzahl der Anschlüsse pro bebautem Quadratkilometer
- b) **Urbanisierungsgrad:** Flächenanteil der bebauten Gebiete
- c) **Trassendichte:** pro bebautem Quadratkilometer im HK-Segment
- d) **Kostenintensität:** durchschnittliche Investition pro Trassenmeter

Die unter a) und c) gelisteten Einflussfaktoren weisen den stabilsten Zusammenhang auf. Unter Berücksichtigung der Besiedlungsstruktur führt eine Verdoppelung der Teilnehmerdichte zu einer Reduktion des FTTH P2P-Investitionsbedarfs pro Anschluss um **46%**. Der Investitionsbedarf steigt mit zunehmender HK-Trassenlänge pro bebauten Quadratkilometer in einem exponentiellen Verhältnis (Trassendichte). Mit einer Abwei-

chung von lediglich -2% zum Modelldurchschnitt haben die Testergebnisse die Robustheit der quantifizierten Schätzgröße für diese beiden Einflussfaktoren gezeigt.

Die Übertragung des Schätzansatzes auf höher aggregierte räumliche Abgrenzungen liefert, beispielsweise auf Kreisebene (NUTS3), vergleichbare Ergebnisse. Die statistische Signifikanz und Stärke der Zusammenhänge für Teilnehmerdichte und Trassendichte im HK-Segment bleiben nach der räumlichen Aggregation erhalten.

Die Nachbildung der Schätzung des Benchmark-Modells auf Basis der bisher identifizierten vollständig öffentlich verfügbaren Daten⁵⁰ liefert Schätzergebnisse, bei denen die grundlegenden Modellannahmen nicht erfüllt sind. Wir gehen davon aus, dass der Einflussfaktor HK-Trassenlänge anhand der öffentlich verfügbaren Straßenlängen nicht hinreichend approximiert wird. Wir vermuten, dass eine stärkere Ausdifferenzierung von Straßendaten helfen kann, die Schätzung zu verbessern. Dabei denken wir an eine Ausgrenzung von Straßen, die für die Auslegung des Anschlussnetzes keine oder nur untergeordnete Bedeutung haben. Beispielsweise sei hier auf Waldweg und Wege rein land- und forstwirtschaftlicher Nutzung verwiesen. So könnte zukunftsblickend die Straßenlänge so segmentiert bzw. transformiert werden, dass der Einflussfaktor „HK-Trassenlänge“ auf die Investition pro Anschluss präziser abbildet werden kann.

Für die Ableitung eines erwartungstreuen Proxy-Modells, das sich auch auf andere Länder und Gegebenheiten anwenden lässt, bedarf es noch weiterführender Analysen.

⁵⁰ Gemeint sind hier die Proxyvariablen für die im Benchmark Modell verwendeten Mengengerüstvariablen aus den Berechnungen je Anschlussbereich.

6 Clusterbildung

Nachdem im vorigen Kapitel die Frage nach den erklärenden Einflussgrößen auf die Höhe der Investitionen im Vordergrund stand, soll in diesem Kapitel der Frage nachgegangen werden, ob die Anschlussbereiche einer Gruppenbildung oder Clusterung zugänglich sind. Davon sind mehrere Aspekte betroffen:

- Lassen sich die Investitionen je Anschlussbereich in wenige homogene Gruppen einteilen?
- Gibt es beobachtbare Muster in den Eigenschaften der zugrundeliegenden Anschlussbereiche, die eine solche Gruppenbildung stützen?
- Bilden die Gruppen räumlich zusammenhängende Einheiten?

Die Motivlage für die drei Fragen ist unterschiedlich. In der ersten Frage geht es im Wesentlichen um eine Reduktion von Komplexität in der Darstellung von Ergebnissen und in der Kommunikation der Rechenergebnisse. Ebenso ergibt sich eine Vereinfachung des Rechenaufwandes in der Modellbildung, wenn die Ergebnisse es erlauben für nur eine geringe Zahl an Gruppen zu rechnen. Darüber hinaus besteht hier das Ziel, die in der Vergangenheit vorgenommene Gruppierung nach Teilnehmerdichte zu verbessern.

Die zweite Frage ist dem besonderen Erkenntnisinteresse geschuldet, ob bestimmte Siedlungsstrukturen oder andere Einflussgrößen sich mit den Investitionen eines FTTH-Ausbaus in Verbindung bringen lassen. Die Regressionsanalysen des Vorkapitels wurden unter anderem angestrengt, um diesen Punkt näher zu beleuchten und werden in Abschnitt 6.3 aufgegriffen. Hier bleibt zu untersuchen, ob eine Clusterbildung entlang der gefundenen Einflussgrößen eine brauchbare Gruppierung ergibt und insbesondere eine überzeugende räumliche Einteilung liefern kann.

Die dritte Frage thematisiert, ob die räumliche Korrelation der Einzelergebnisse für eine Ballung teurer oder billiger Ausbauregionen spricht. Auch wäre hier von Interesse herauszufinden, ob und inwiefern räumliche Korrelationen hoher und niedriger Ausbaukosten mit gängigen Einteilungen hinsichtlich Ballungszentren, städtischen Gebieten, ländlichen Gebieten und abgelegenen Regionen übereinstimmen.

In der folgenden Analyse erfolgt zunächst eine Gegenüberstellung der in Abschnitt 2 und 3 ermittelten anschlussbereichsgenauen Ergebnisse für die Investitionen pro Anschluss mit der dort verwendeten Clustereinteilung nach Teilnehmerdichte (WIK-NGA-Modell). Daran schließt sich mit der Frage nach der optimalen Anzahl von Clustern an. Dafür werden univariater k-means Clusterverfahren unter Rückgriff auf die Investitionen pro Anschluss als differenzierendes Kriterium für die Clusterbildung herangezogen (Abschnitt 6.2). Bei den multivariaten k-means Clusterbildungen werden die in der Regressionsanalyse abgeleiteten Größen verwendet und es wird ebenfalls der Frage einer

optimalen Anzahl von Clustern nachgegangen. Die vorgestellten Ansätze werden miteinander verglichen (Abschnitt 6.4), bevor abschließend eine Auseinandersetzung mit den räumlichen Korrelationen hoher und niedriger Investitionswerte sowie einer Gegenüberstellung alternativer räumlicher Einteilungen erfolgt (Abschnitt 6.5). Das Kapitel schließt mit einer Einordnung der Ergebnisse.

6.1 Investitionen pro Anschluss und Clusterbildung anhand von Dichtekriterien und gleicher Anschlusszahl pro Cluster

Das Grundanliegen einer Clusterbildung ist eine Zusammenfassung der heterogenen Ausgangslage in möglichst wenige gleichartige Gruppen. Die Gründe dafür liegen einmal in der Berechnung der Ergebnisse selbst, die für zunächst für die Analyse in Abschnitt 4 dieses Forschungsprojekt mittels bottom-up Modellierung aus einem mit hohem Aufwand erzeugten Mengengerüst in großer Granularität erzeugt wurden.⁵¹

Wünschenswert ist eine Reduktion der Komplexität mittels einer gut abgesicherten Gruppenbildung, Aussagen und Ergebnisse in hinreichender Auflösung zu erzeugen, und damit den Aufwand der Datenaufbereitung und Rechendauer deutlich zu reduzieren. Eine sinnvolle Gruppierung der Regionalbereiche nach deren Gleichartigkeit bezüglich objektiver Kriterien kann zur einfachen Abschätzung der Ausbaurkosten für unterschiedlich große Regionalbereiche genutzt werden. Gelingt eine solche Clusterbildung, kann zukünftig eine clusterbasierte bottom-up Modellierung die flächendeckende Berechnung ersetzen.⁵²

In der weiteren Analyse dieses Abschnitts werden wir die bottom-up ermittelten Investitionswerte pro Anschluss, differenziert nach Anschlussbereichen, heranziehen und diese im Sinne einer idealtypische Clusterung für die Bildung disjunkter Cluster verwenden. Dabei stellen wir zunächst auf die anschlussbereichsbezogenen Ergebnisse der FTTH P2P Investitionen je Anschluss, gruppiert nach halben Standardabweichungen ab (Abschnitt 6.1.1). Daran schließt sich ein Vergleich der Clusterbildung für 20 Cluster anhand von verschiedenen Dichtekriterien und univariater k-means Clusterung nach Investitionen je Anschluss an (Abschnitt 6.1.2).

6.1.1 Regionale Verteilung der Anschlussbereiche über Cluster – ein Vergleich der WIK-NGA-Modell Cluster mit FTTH P2P Investitionen je Anschluss, gruppiert nach halben Standardabweichungen

Abbildung 6-1 gibt die Ergebniswelt der Investitionswerte je Anschluss eines flächendeckenden FTTH-Ausbaus aller berechneter Anschlussbereiche (7871) in ihrer räumli-

⁵¹ Siehe hierzu die Abschnitte 3 und 4 sowie den Anhang.

⁵² Eine Reduzierung des Aufwands für Datenaufbereitung und Rechenzeiten erlaubt es im Gegenzuge die Ressourcen für relevante Sensitivitätsrechnungen zu verwenden.

chen Verteilung wieder.⁵³ Für die Gesamtheit der Anschlussbereiche wurden die durchschnittlichen Investitionen pro Anschluss herangezogen, der Mittelwert und die Standardabweichung ermittelt. Die Einfärbung erfolgt vom Minimum ausgehend in halben Standardabweichungen, von rot nach blau. Diese Darstellung dient in einem ersten Schritt als Referenz für unsere Clusteranalyse. Sie verdeutlicht die Streuung der Investitionswerte pro Anschluss und liefert eine Orientierungsgröße für die Bewertung des dem WIK-NGA-Modell zugrunde liegenden Clusteransatzes.

Dieser, auf Basis der Zielgröße gruppierten Darstellung der Anschlussbereiche soll nun die dem WIK-NGA-Modell zugrunde liegende Clusterung grafisch gegenübergestellt werden. Damit soll eine Einschätzung darüber gewonnen werden, wie gut das Merkmal der Besiedlungsdichte in der Lage ist, zielgenaue regionale Ergebnisse zu ermitteln.

Abbildung 6-2 zeigt die Anschlussbereiche nach ihrer Zugehörigkeit zu den 20 Clustern nach Teilnehmerdichte. Der Farbverlauf ist von hoher Teilnehmerdichte, dargestellt in rot, über gelb, grün hin zu geringer Teilnehmerdichte in blau dargestellt. Ein Vergleich mit Abbildung 6-1 zeigt, dass die Bereiche niedriger Investitionen gut mit den unteren Clustern der Abbildung 6-2 übereinstimmen (hohe Teilnehmerdichte), die Cluster niedriger Teilnehmerdichte allerdings das Bild dominieren, während in Abbildung 6-1 die blauen Farbwerte, die mit einer hohen Investition gekoppelt sind weniger stark hervortreten. Das bedeutet, dass der Zusammenhang zwischen niedriger Teilnehmerdichte und hoher Investitionen je Anschluss in Anschlussbereichen mit niedriger Teilnehmerdichte weniger stark greift, d.h. nicht immer Gültigkeit hat. Gerade in Gebieten niedriger Anschlussdichte hängen die Investitionen je Anschluss verstärkt auch von anderen Einflussgrößen ab.

53 Basis für die Darstellung der Investitionen pro Anschluss sind die anschlussbereichsgenauen Ergebnisse der Investitionswertmodellierung des WIK-NGA-Modells, welches auch die Grundlage für die Analyse in Abschnitt 4 bildet.

Abbildung 6-1: Räumliche Verteilung der FTTH Investitionen je Anschluss gruppiert nach halben Standardabweichungen

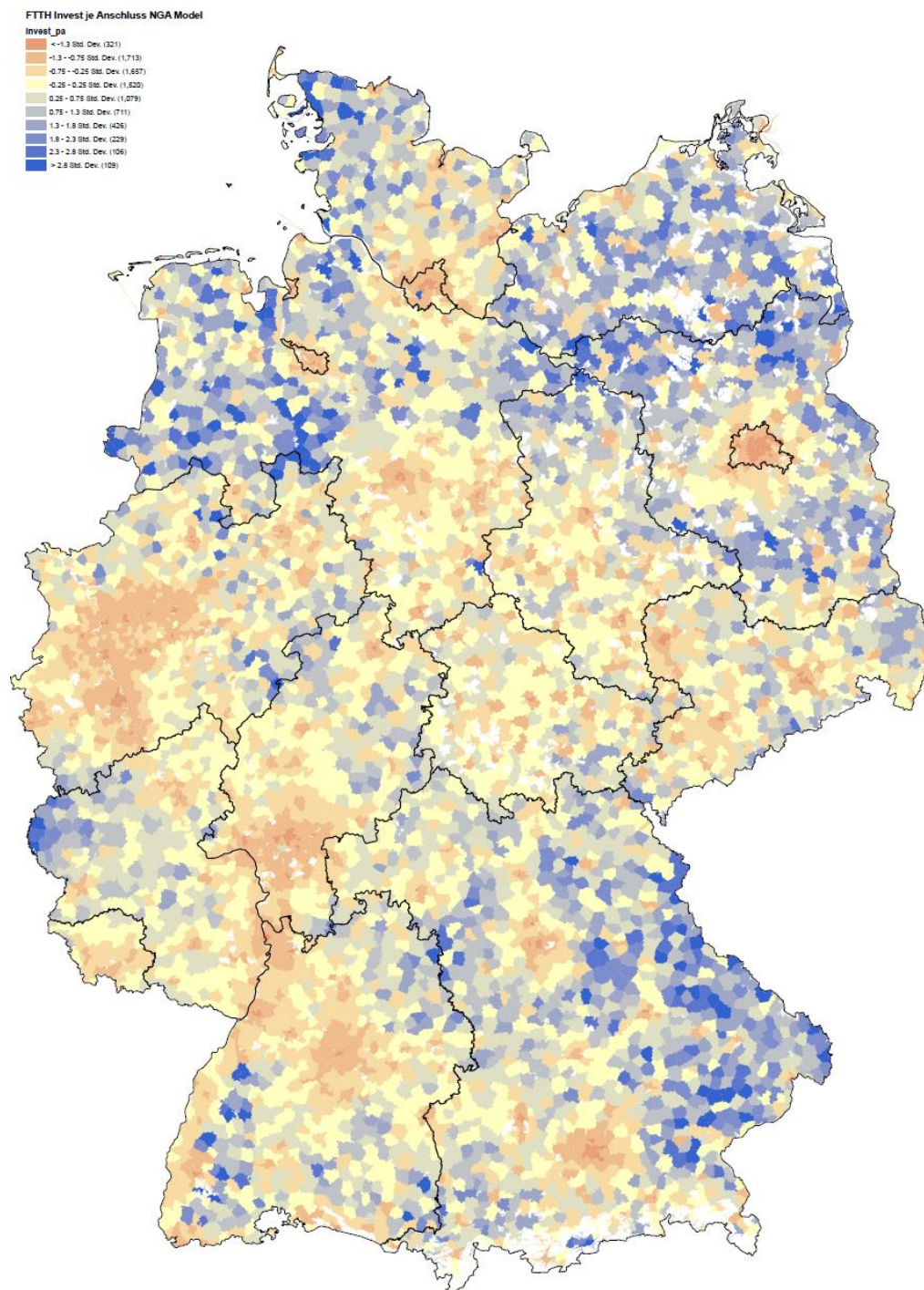
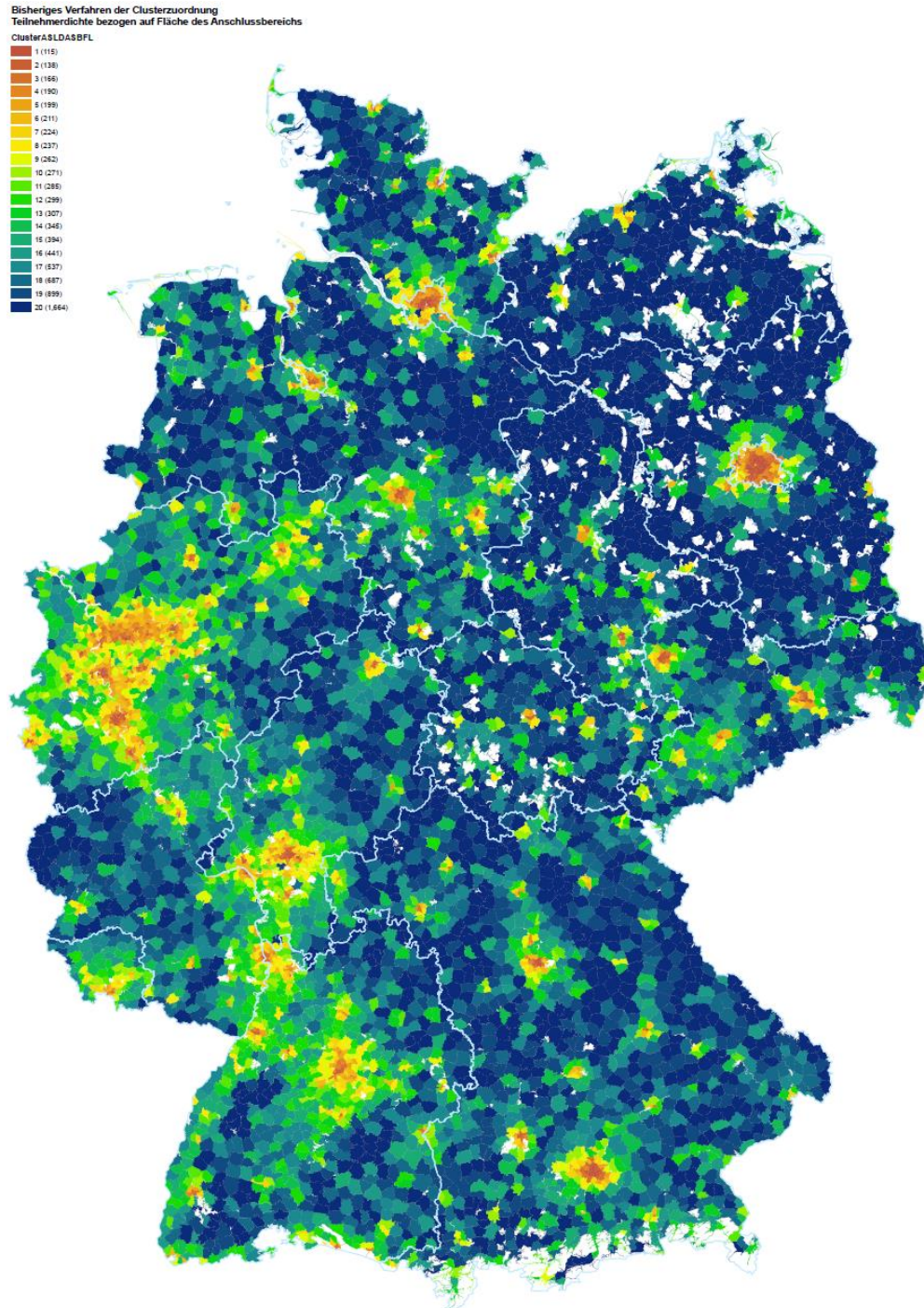


Abbildung 6-2: Räumliche Verteilung der Clusterabgrenzung nach Teilnehmerdichte bezogen auf die ASB-Fläche (20 Cluster des WIK-NGA-Modells)



Quelle: WIK, Landesgrenzen BKG

6.1.2 Vergleich von Dichtekriterien und univariater k-means Clusterung nach Investitionen je Anschluss

Um den Einfluss der Clusterbildung auf die Investitionsergebnisse näher zu beleuchten, wird auf die detaillierte Datenbasis abgestellt, die uns aus der bottom-up Modellierung zur Verfügung steht. Dabei soll nachfolgend zunächst nach alternativen Clusterkriterien gesucht werden, die das Dichtemaß „Anschlüsse pro km² des Anschlussbereichs“ ersetzen und zu einer stärker disjunkteren Clusterbildung führen können.

Dabei haben wir auf die in Abschnitt 5 gewonnenen Erkenntnisse der Regressionsanalyse⁵⁴ zurückgegriffen und die dort identifizierten Variablen für die hier betrachtete Clusterbildung herangezogen.

Aus der Regressionsanalyse lässt sich ein engerer erklärender Zusammenhang zwischen den Investitionen je Anschluss und der Teilnehmerdichte bezogen auf die bebauten Gebiete herleiten, im Vergleich zur Teilnehmerdichte bezogen auf die gesamte Anschlussbereichsfläche. Es liegt deshalb nahe, den Zusammenhang der Clusterbildung mit der Spanne der Investitionswerte je Anschluss je Cluster zu betrachten, um die Trennschärfe der Clusterbildung im Hinblick auf die Ausbaurkosten je Anschluss zu beurteilen.

Darüber hinaus liegen uns in dieser Studie weitere Maße für die Teilnehmerdichte je Anschlussbereich vor. Wir haben uns entschieden, die nachfolgenden drei Dichtemaße für eine vergleichende Gegenüberstellung der Clusterbildung zu verwenden, wobei das jeweilige Dichtemaß das Ordnungskriterium für die Cluster bildet und – in Konsistenz zu der Clusterbildung auf Basis des WIK-NGA-Modells – eine Gleichverteilung der Anschlüsse über 20 Cluster vollzogen wird. Die betrachteten Dichtemaße umfassen:

- die Teilnehmerzahl bezogen auf die ASB-Fläche (entsprechend dem WIK-NGA-Modell),⁵⁵
- die Teilnehmerzahl bezogen auf die bebauten Flächen je ASB (Wohngebietspolygonflächen)⁵⁶

54 So konnte mit der Regressionsanalyse gezeigt werden, dass ein enger Zusammenhang zwischen Investitionswert je Anschluss und Anschlussbereich und der Teilnehmerdichte besteht. Der Zusammenhang ist dergestalt, dass in Anschlussbereichen hoher Teilnehmerdichte in der Regel niedrige Investitionen je Anschluss resultieren.

55 ASB-Fläche: Aus der Anschlussbereichsabgrenzung liegen uns die Flächenpolygone der im GIS berechneten Anschlussbereiche vor. Diese Flächen werden mit der jeweiligen Teilnehmerzahl in Beziehung gesetzt um die Teilnehmerdichte zu bestimmen.

56 Wohngebietspolygonflächen: Polygone der Siedlungsflächen (aus WiGeoStreet Landnutzungslayer) liegen uns für ganz Deutschland vor und werden mit den Anschlussbereichen verschnitten, so dass eine Zuordnung der Siedlungsflächen zu den Anschlussbereichen ermöglicht wird. Die Summe der Siedlungsflächen innerhalb eines Anschlussbereichs dienen hier als Fläche zur Berechnung der Teilnehmerdichte des ASB. Die Siedlungsflächen sind in den meisten ASB kleiner als die ASB-Flächen, insbesondere in ländlichen Gebieten.

- und die Teilnehmerzahl bezogen auf die bebauten Flächen gemäß Gebäuderaster je ASB.⁵⁷

Die Clusterergebnisse für die drei verschiedenen Dichtemaße haben wir nachfolgend der Clusterung auf Basis eines univariaten k-means Verfahrens gegenübergestellt, wobei wir auch hier die Anzahl der Cluster auf 20 gesetzt haben („k“=20). Bei diesem univariaten k-means Clusteransatz greifen wir auf die Investitionswerte je Anschluss der einzelnen ASB als Clusterkriterium zurück und verwenden damit die Zielgröße der Clusterbildung (Cluster mit möglichst homogenen Investitionen pro Anschluss) als bestimmende Variable für die Clusterbildung.⁵⁸ Die k-means-Methode ersetzt die Gleichverteilung der Anschlüsse über die Cluster durch ein Zuordnungsverfahren, das die Abweichungen innerhalb eines Cluster minimiert und darüber eine trennscharfe Zuordnung der Anschlussbereiche zu den Clustern anstrebt.⁵⁹ Entsprechend repräsentieren die Cluster unterschiedliche Anschlusszahlen. Diese univariate k-means Clusterung nach Maßgabe der Investitionen je Anschluss stellt für unsere Analyse die idealtypische Clusterung dar. Es ist unser übergeordnetes Ziel, alternative Variablen zu finden, die eine vergleichbar gute, trennscharfe Clusterung mit Blick auf die Investitionswerte pro Anschluss erzeugen können.

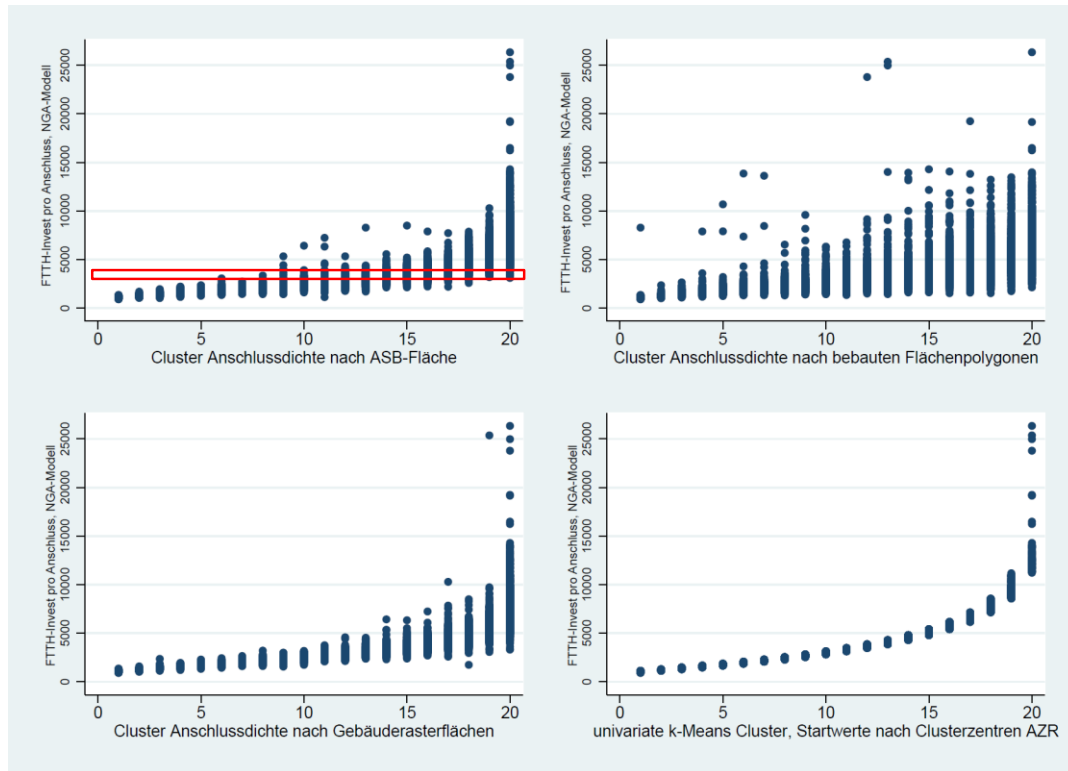
Abbildung 6-3 zeigt eine Zusammenstellung von vier Einzelgrafiken, in denen die Investitionswerte je Anschluss aller Anschlussbereiche gegen die Clusternummer abgetragen sind. Den Grafiken liegen die oben geschilderten Clusteransätze zugrunde. Die ersten drei Grafiken stellen den Zusammenhang für die Clusterbildung nach Teilnehmerdichte für die drei Dichtemaße dar (ASB-Fläche, Wohngebietspolygonflächen und Gebäuderasterflächen, s.o.). Allen drei Ansätzen liegt eine Gleichverteilung der Anschlüsse über die Cluster zugrunde. Die vierte Grafik repräsentiert die idealtypische Clusterbildung anhand der Investitionen je Anschluss unter Rückgriff auf die Methode der univariaten k-Means Clusterung, hier ebenfalls mit einer Clusterzahl von k=20.

57 Bebaute Flächen gemäß Gebäuderaster je ASB: Der verwendete Gebäuderasterdatensatz (Raster 250m x 250 m) weist nur dann eine Rasterfläche auf, wenn mindestens ein Gebäude in die Rasterzelle fällt. Die Zellen ohne Gebäude werden bei der Verarbeitung ausgeblendet (Zellen ohne Daten). Ein räumlicher Verschnitt mit den Anschlussbereichsflächen erlaubt eine Zuordnung der Rasterflächen mit mindestens einem Gebäude zu den Anschlussbereichen. Die Summe dieser Rasterflächen je Anschlussbereich dienen hier zur Bestimmung der Teilnehmerdichte des Anschlussbereichs. Die Gebäuderasterflächen beziehen im Gegensatz zu den Siedlungsflächen auch verstreute Gebäude und kleine Weiler mit in die Flächenberechnung ein und sind in der Regel kleiner als die ASB-Flächen.

58 Damit ergibt sich per definitione eine monotone Gruppierung der Anschlussbereiche nach Maßgabe der durchschnittlichen Investitionen pro Anschluss. Methodisch ist diese Vorgehensweise mit dem Benchmark-Modell für die Regressionsanalyse vergleichbar und dient dazu, eine Vorstellung über die „wahre“, kostenorientierte Clusterbildung zu generieren.

59 Siehe hierzu auch die ausführlichen Erläuterungen in Abschnitt 6.2.

Abbildung 6-3: Investitionen je Anschluss nach Cluster für unterschiedliche Clusterbildungen



Quelle: WIK

wik

Die Clusterbildung anhand der Teilnehmerdichte führt zu starken Überlappungen in den Investitionswerten je Anschluss der ASB, die aufgrund der Clusterbildung zu einer Gruppe zusammengefasst werden. Diese haben wir für den ersten Fall (Clustering nach Anschlussdichte) mit einem roten Balken illustriert. Er verdeutlicht, dass die durch den Balken beschriebene Investitionshöhe je Anschluss in sämtlichen Clustern 6 bis 20 vertreten ist. Nimmt man alternativ ein Cluster mit definierter Anschlussdichte als Ausgangspunkt der Betrachtung (vertikale Perspektive), so ist festzustellen, dass die zugehörigen Investitionen je Anschluss stark voneinander abweichen können: Verschiedene Anschlussbereiche, die eine nahezu gleiche Anschlussdichte aufweisen und daher in das gleiche Cluster fallen, weisen stark abweichende Investitionen je Anschluss auf (dies kommt durch die „vertikale“ Betrachtung der Balken pro Cluster zum Ausdruck).

Die Wertespanne in den unteren Clustern ist dabei in der ersten und dritten Grafik noch relativ eng und damit trennscharf, verliert diese Eigenschaft allerdings in den mittleren und spätestens in den oberen Clustern.

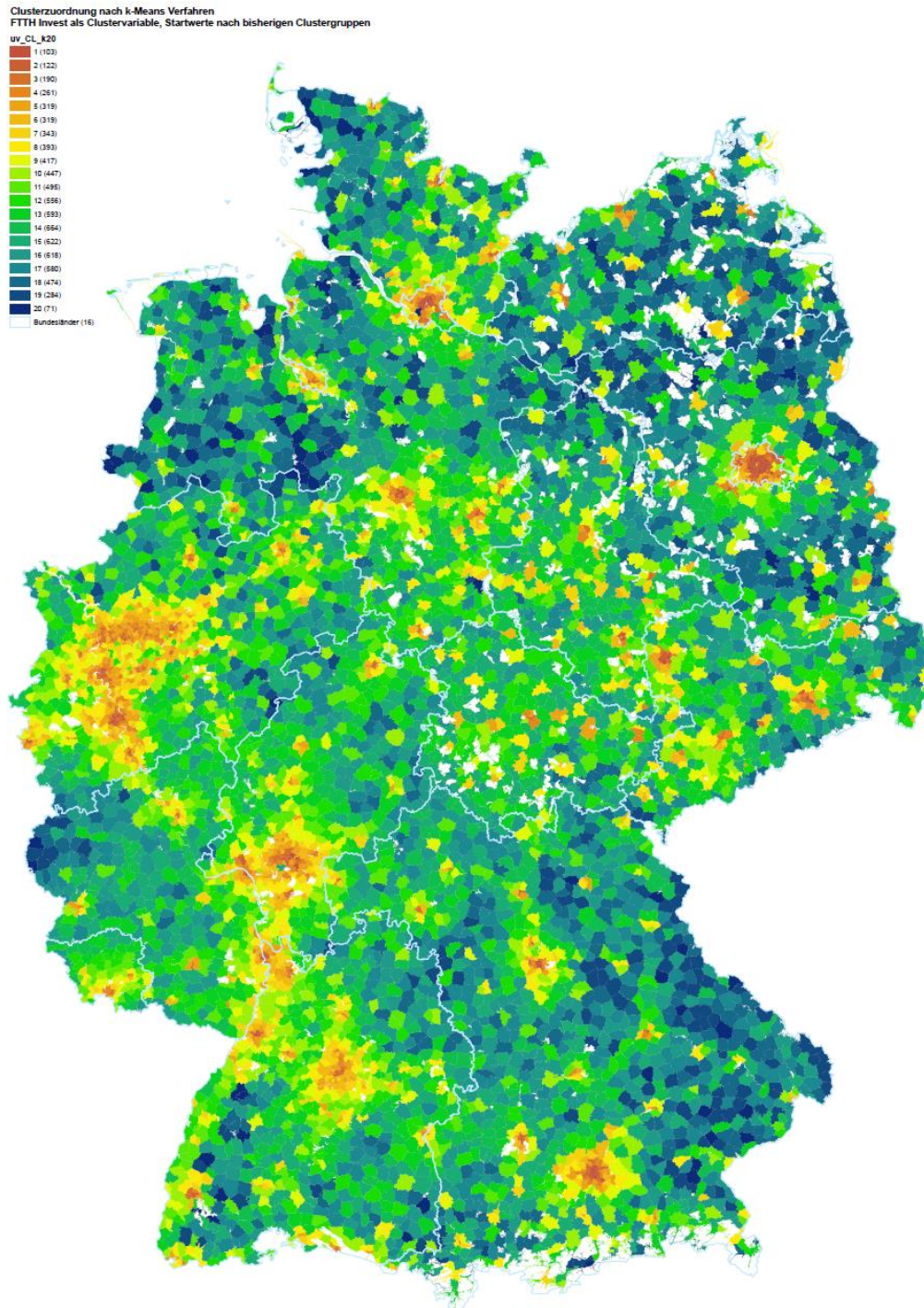
Am stärksten ausgeprägt ist der Überlappungseffekt bei der Clusterbildung unter Verwendung der Teilnehmerdichte nach bebauten Flächen der Siedlungspolygone. Hier ist in den oberen Clustern nahezu jeder Wert der Investitionen je Anschluss in einem Cluster vertreten. Nach der Spanne der Investitionen je Anschluss bemessen, liefert die Clusterung nach Teilnehmerdichte basierend auf den bebauten Flächenpolygone somit die Clusterung mit der geringsten Trennschärfe. Die Clusterung nach Teilnehmerdichte unter Verwendung der Gebäuderasterflächen schneidet hier am besten ab.

Allen drei Clusterungen ist gemeinsam, dass die Spanne der Investitionswerte je Anschluss mit zunehmender Clusterzahl, d.h. mit sinkender Teilnehmerdichte zunimmt. Überraschend ist das Auseinanderfallen der Beobachtung für die zweite und dritte Darstellung der Investitionen nach Anschluss. Letztlich ist für die Anschlussdichte nach bebauten Gebieten ein ähnliches Muster zu erwarten, wie es sich auch für die Cluster nach Anschlussdichte nach Gebäuderasterflächen darstellt.

Die idealtypische univariate k-Means Clusterung (k-Mittelwert Clusterung) anhand des Investitionswertes je Anschluss führt, wie aus der Grafik ersichtlich und wie erwünscht, zu homogenen Gruppen mit geringer Überlappung. Aber auch hier gilt für unsere Daten, dass die Spanne der Werte je Cluster in den oberen Clustern zunimmt.

Nachfolgend wollen wir die Erkenntnisse der Clusterbildung auch grafisch in Bezug auf die regionale Verteilung der Clusterzuordnung der Anschlussbereiche analysieren und einen Vergleich der aus dem WIK-NGA-Modell bekannten Zuordnung der ASB zu den Clustern mit der idealtypischen, univariaten Clusterung nach Maßgabe der Investitionen je Anschluss zu vollziehen. Die räumliche Verteilung der Clusterbildung nach Maßgabe der Investitionen je Anschluss (k-means-Methode) ist in Abbildung 6-4 wiedergegeben. Aus der Grafik wird ersichtlich, dass das Verfahren nun auch die mittleren Cluster stärker berücksichtigt und insgesamt die Verteilung der Investitionswerte je Anschluss sehr gut wiedergibt, wie sie aus Abbildung 6-1 ersichtlich ist. Ein Vergleich mit der Clusterbildung nach Teilnehmerdichte in Abbildung 6-2 zeigt, dass die Verteilung der Anzahl Anschlussbereiche über die Cluster sich stark verändert hat, und zwar von den oberen Clustern zu den mittleren Clustern

Abbildung 6-4: Räumliche Verteilung der Clusterabgrenzung für die Investitionen je Anschluss nach der k-Mittelwert Methode bei 20 Clustern



Quelle: WIK, Landesgrenzen BKG

6.2 Clusterbildung nach der k-Mittelwert Methode⁶⁰

Bei der k-Mittelwert Methode der Clusterbildung werden k Gruppen gebildet, wobei die Anzahl k vorzugeben ist. Ausgehend von k-Clusterzentren, die ebenfalls als Startwerte vorzugeben sind (oder als Zufallsziehung bereitgestellt werden), erfolgt eine Zuordnung aller Beobachtungen zu ihrem nächstgelegenen Startwert-Clusterzentrum. Danach wird auf Grundlage der gerade erfolgten Zuordnung für jede Gruppe das neue Zentrum ermittelt (Mittelwert über alle Beobachtungen des Clusters). Danach wiederholt sich die Zuordnung und Mittelwertbildung so lange, bis sich die k-Clusterzentren nicht mehr verändern oder das vorgegebene Iterationsmaximum überschritten wird. Aufgrund der gewählten Metrik der Euklidischen Distanz minimiert das Verfahren die Summe der quadrierten Abweichungen vom jeweiligen Clusterzentrum, was gleichbedeutend mit der Minimierung der Varianz innerhalb des Clusters ist. Es kommt zu einer Gruppenzuweisung nach größter Ähnlichkeit, wobei die Ähnlichkeit durch eine geringe euklidische Entfernung der Beobachtungsausprägung zum gemeinsamen Zentrum ausgedrückt ist.

Das Verfahren hat zwei Nachteile:

- Zum einen ist die Zahl der Clusterzentren k vorzugeben, die in der Regel unbekannt ist.
- Der Algorithmus berechnet ein lokales Minimum der Varianzsummen; das Verfahren führt bei unterschiedlichen Startwerten in der Regel zu unterschiedlichen Gruppierungen.

Die Lösung für das erste Problem besteht in der Berechnung vieler Clusterungen mit unterschiedlichen Werten für k und einem Vergleich der Varianzen aus den Clusterbildungen. Diesem Problem widmen wir uns im nächsten Abschnitt.

Eine gangbare Lösung zum zweiten Problem besteht darin, bei vorgegebenem k viele unterschiedliche Startwerte anzusetzen und ebenfalls die Varianzen der Clusterbildungen zu vergleichen. Diesem Problem sind wir in diesem Forschungsprojekt nicht nachgegangen.

6.2.1 Kriterien zur Bestimmung der optimalen Anzahl Cluster

Wenige, gut unterscheidbare Cluster wären für unsere Zwecke ein wünschenswertes Ergebnis. Die Reduktion der Anzahl Gruppen soll helfen, die Komplexität der Analyse und der Darstellung zu reduzieren.

⁶⁰ Die Berechnungen zur Clusterbildung mit der k-means Methode wurden mit STATA 14.2 durchgeführt.

Da bei der k-Mittelwert Methode unter Verwendung euklidischer Distanzen die Varianzsummen minimiert werden (within sum of squares WSS), kann die Varianzsumme als Maß für die Qualität der Clusterung herangezogen werden.

Zur Beurteilung welche Anzahl an Gruppen k unter sonst gleichen Bedingungen (gleicher Datensatz und gleiche Startwerte) gewählt werden sollte, wird folgendes Vorgehen gewählt⁶¹:

Berechnung einer Clusterung für eine fortlaufende Anzahl von Clustern k ($k = 1..n$)

Dazu werden für jede Clusterung (jedes k) folgende Größen berechnet:

Maß 1: "within sum of squares" bei k Gruppen WSS (k):

Abschätzung der Wirkung von k auf die Summe der Gruppenvarianzen unter sonst gleichen Bedingungen (bestimme WSS(k) für jedes k). Eine starke Reduktion von WSS bei einer Veränderung von k zeigt in die richtige Richtung. Ebenso weisen wir hier den Logarithmus der Varianzsumme $\log(\text{WSS})$ aus.

Maß 2: Anteil der „erklärten“ Varianz η^2

Ähnlich dem R^2 aus der Regressionsrechnung lässt sich die Varianzsumme bei k Gruppen WSS(k) zur maximalen Varianz des Datensatzes (WSS(1) bei einer Gruppe $k=1 = \text{TSS}$ ⁶²) in Beziehung setzen und als „erklärter Anteil der Gesamtvarianz“ η^2 darstellen:

$$\eta_k^2 = 1 - \frac{WSS(k)}{WSS(1)} = 1 - \frac{WSS(k)}{TSS}$$

Maß 3: Proportionaler Varianzreduktionsfaktor (PRE):

PRE gibt die proportionale Varianzänderung an, bei Erhöhung der Clusteranzahl um 1. Das Maß ist definiert als Veränderung der Summe der gruppenbezogenen Varianzen bei k Gruppen im Vergleich zu $k-1$ Gruppen, bezogen auf die Varianzsumme bei $k-1$ Gruppen:

$$PRE_k = \frac{WSS(k-1) - WSS(k)}{WSS(k-1)} \quad \forall k \geq 2$$

⁶¹ Die Clusteranalyse erfolgte in STATA und die Varianzindikatoren wurden analog zu folgendem STATA-Hinweis implementiert: Anna Makles, Stata tip 110: How to get the optimal k-means cluster solution, 2012, The Stata Journal (2012) 12, Number 2, pp. 347–351.

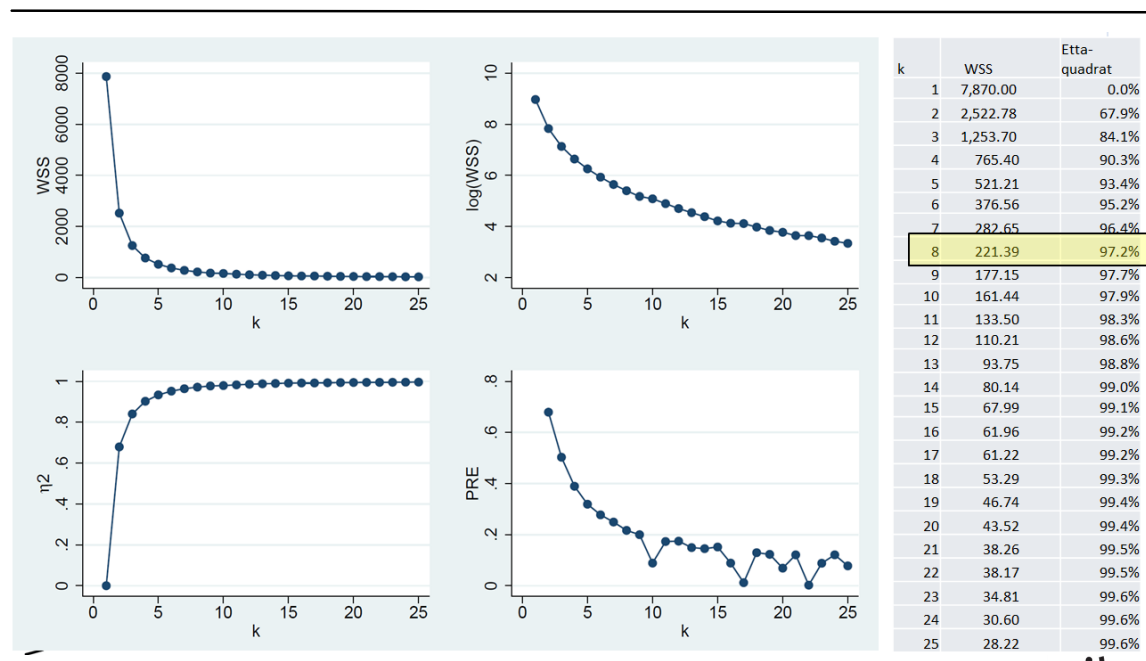
⁶² Total sum of squares (TSS)

6.2.2 Optimale Anzahl Cluster bei univariater k-Means Clustering der FTTH-Investitionen je Anschluss

Um zu einer Einschätzung zu kommen, welche Clusteranzahl k wir bei der univariaten Clustering der Investitionen je Anschluss unterstellen sollten, wurden 50 Clusterrechnungen durchgeführt und die oben dargestellten drei Maße zur Beurteilung der Güte berechnet.⁶³ Die Clustering greift in diesem ersten Schritt nicht auf eine exogene, erklärende Größe zurück,⁶⁴ sondern bedient sich – wie bereits in Abschnitt 6.1.2 – der durch die bottom-up Modellierung generierten durchschnittlichen Investitionen pro Anschlussleitung als Clusterkriterium.

Abbildung 6-5 zeigt die Ergebnisse der Berechnungen für $k = 1$ bis 25 Cluster. Die Plotmatrix gibt in der oberen Reihe die Varianzsummen ($WSS(k)$) und den natürlichen Logarithmus derselben ($\log(WSS(k))$), abgetragen gegen die Anzahl Cluster k , wieder. Darunter sind der Anteil der „erklärten“ Varianz η^2 und der proportionale Varianzreduktionsfaktor ($PRE(k)$ in Abhängigkeit von k abgetragen. Rechts der Plotmatrix sind die Ergebniswerte für die Varianzsummen ($WSS(k)$) und η_k^2 aufgelistet.

Abbildung 6-5: Optimale Anzahl Cluster bei univariater k-Means Clustering mit Investitionen je Anschluss als Clusterkriterium



Quelle: WIK

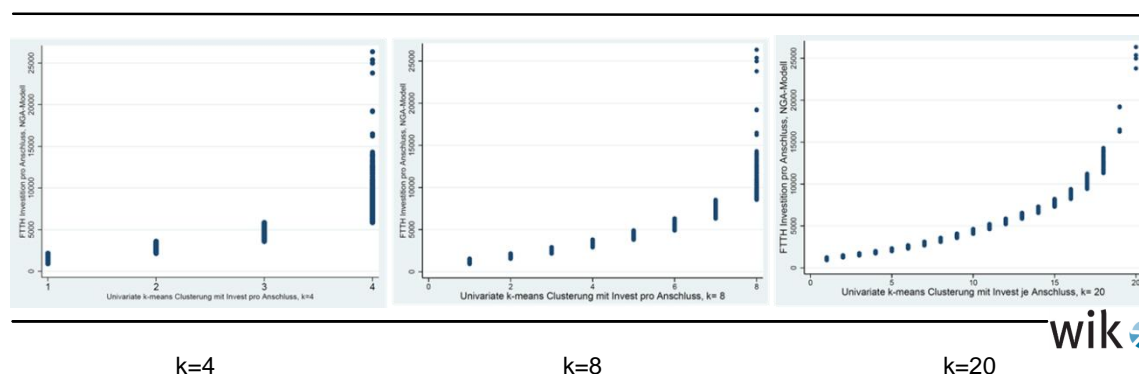
⁶³ Als Startwerte der Clusterberechnungen diente jeweils eine Zufallsauswahl von k Werten.
⁶⁴ Die Anwendung exogener Variablen findet erst in Abschnitt 6.3 Berücksichtigung.

Der Verlauf der Varianzkurven ($WSS(k)$ und $\log(WSS(k))$) zeigt eine bedeutende Reduktion der Varianz bei kleinen k mit einer starken Abnahme des Effekts bei Erhöhung der Clusteranzahl. Die Werte für η_k^2 zeigen uns, dass bei 4 Clustern bereits 90,3% der Varianz erklärt werden, bei 8 Clustern sind es 97,2% und bei 20 Clustern 99,4%. Der sinkende Verlauf der PRE-Kurve gibt diesen immer kleiner werdenden Beitrag jeder Zunahme der Anzahl Cluster um eine Einheit wieder.

Da wir einen stetig abnehmenden Verlauf der Varianzsummen vorliegen haben, empfiehlt sich hier keine der Anzahl Cluster als optimale Clusterzahl. Es wird deutlich, dass ein Trade-off zwischen geringer Clusterzahl und Anteil erklärter Varianz vorliegt. Liegt der Schwerpunkt auf einer Reduktion der Komplexität, können 4 Cluster bei 90% erklärter Varianz gegenüber 8 Clustern mit 97,3% oder gar 20 Clustern mit 99,4% überlegen sein. Hinzu kommt, dass 4 oder 8 Cluster besser zu interpretieren sind als 20 bei nur geringem Verlust an erfasster Varianz.

Abbildung 6-6 zeigt die Wertespanne für die Gruppierung der Investitionswerte in 4, 8 und 20 Clustern. Die Interpretation der Cluster fällt aufgrund der Eindimensionalität der untersuchten Variable „FTTH Investitionen je Anschluss“ und der Trennschärfe in den Werten leicht. Da die zur Clusterbildung herangezogene Variable identisch ist mit der Zielgröße der Clusterung „Investition je Anschluss“, kommt es zu einer klar disjunkten Clusterbildung. Zu beantworten bleibt lediglich die Frage nach der Granularität der Clusterbildung, die vom Verwendungszweck abhängen dürfte. Die Auswertungen zeigen, dass für die Grundgesamtheit der Anschlussbereiche in Deutschland der Informationsverlust bei 4 Clustern unter 10% liegt, bei 8 unterhalb von 3% und bei 20 Clustern weniger als ein Prozent der Information verloren geht.

Abbildung 6-6: Clusterkriterium nach Maßgabe „Invest pro Anschluss“ – Vergleich univariater Clusterbildung für $k = 4, 8$ und 20



Quelle: WIK

6.3 Multivariate Clusterbildung: Verwendung der Regressoren zur Investitionswertschätzung bei der Clusterung

Ausgangspunkt für die multivariate Clusterbildung ist die Frage, ob exogene Größen als (objektive) Kriterien für die Clusterbildung herangezogen werden können, die es erlauben, eine – gegenüber der bottom-up Modellierung - einfachere Abschätzung der NGA-Ausbaukosten zu leisten. Dazu soll geprüft werden, ob bestimmte Siedlungsstrukturen oder andere Einflussgrößen sich mit den Investitionen eines FTTH-Ausbaus in Verbindung bringen lassen und die Ausprägungen dieser Größen in gemeinsamen unterscheidbaren Konstellationen auftreten.

In der Regressionsanalyse (Abschnitt 5.3.2 und 5.4) wurden gerade solche Bestimmungsfaktoren für die FTTH-Investitionen pro Anschluss herausgearbeitet, weshalb es nahe liegt, diese für eine multivariate Gruppenbildung zu verwenden.

Hier wurde untersucht, ob eine Clusterbildung entlang der gefundenen Einflussgrößen eine brauchbare Gruppierung mit Blick auf die die Höhe der Investitionswerte ergibt und insbesondere eine überzeugende räumliche Einteilung liefern kann. Mit diesem Ansatz ist die Hoffnung verknüpft, mittels beobachtbarer Strukturmerkmale in einer ausreichenden Granularität auf die Kosten eines FTTH-Ausbaus schließen zu können.

Die in der Regressionsanalyse mit dem Benchmark-Modell abgeleiteten⁶⁵ und nachfolgend verwendeten Einflussgrößen⁶⁶ umfassen

1. die Teilnehmerdichte je Anschlussbereich (Anzahl Anschlüsse pro km² bebauter Gebiete),
2. die Trassendichte im Hauptkabelbereich (Trassenlänge im HK-Segment pro km² bebauter Gebiete, inkl. der quadrierten Variablen),⁶⁷
3. den Urbanisierungsgrad (Flächenanteil bebauter Gebieten pro ASB)
4. Streuung von Siedlungsflächen (Standardabweichung der Trassenlänge des Hauptkabels auf NUT3 Ebene)⁶⁸

Die Clusterbildung erfolgt ebenfalls auf Grundlage der k-Mittelwert Methode, jedoch für mehrere Variablen gleichzeitig. Dabei handelt es sich um eine mehrdimensionale Gruppenbildung, mit dem Ziel der Minimierung der Summe der gruppenspezifischen Varian-

⁶⁵ Siehe Abschnitt 5.6

⁶⁶ Alle verwendeten Variablen wurden vor Verwendung, wie in der Regression auch, mit dem natürlichen Logarithmus transformiert.

⁶⁷ Hier geht die Trassendichte im HK-Bereich sowohl als einfache logarithmierte Größe als auch als quadrierte logarithmierte Größe in das multivariate Verfahren als Variable ein.

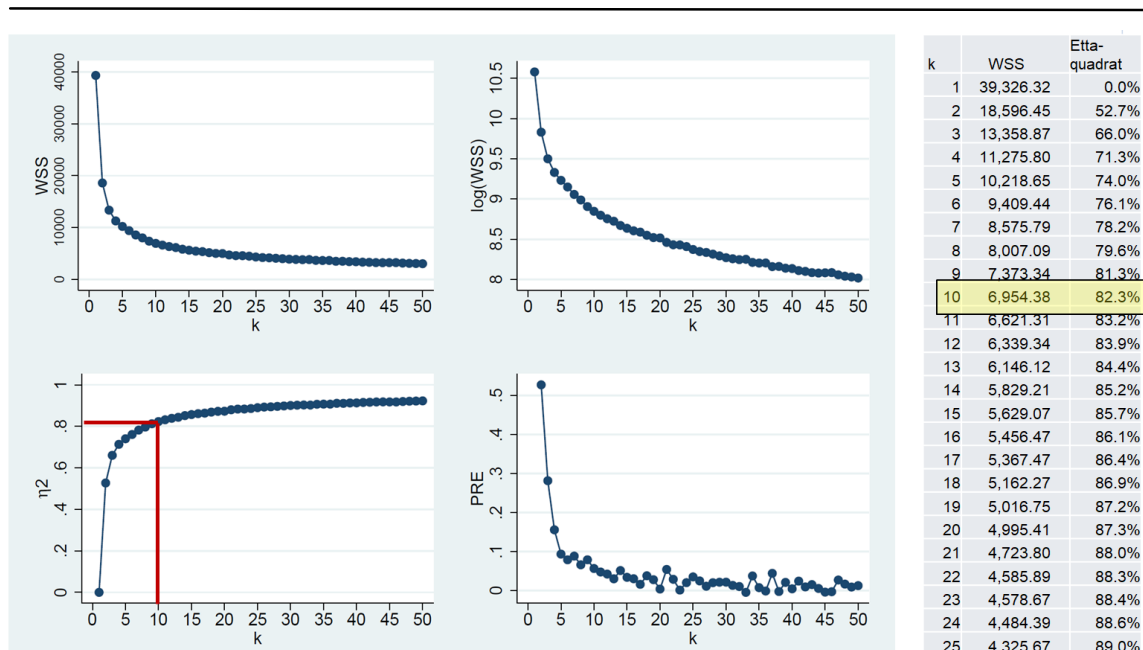
⁶⁸ Anstelle der Einflussgröße der Kostenintensität, die im Regressionsmodell verwendet wurde, wurde hier auf ein Streuungsmaß für die Besiedelung abgestellt. Beide Maße zeigten in der Regressionsmodellierung nahezu austauschbare Effekte, ein Maß für die Streuung der Besiedelung erscheint und im Kontext interessanter.

zen (within sum of squares (WSS)) über alle Variablen. Da das Verfahren von den Skalen der Variablen abhängt, wurden alle Größen vor Verarbeitung zentriert und standardisiert.

Wie bei der univariaten Clusterbildung des obigen Abschnitts, stellt sich bei der multivariaten Clusterung ebenfalls das Problem der Auswahl des Gruppenparameters k . Wir haben ein einheitliches Vorgehen gewählt und 50 Clusterungen durchgeführt. Zur Analyse und Bewertung der durchgeführten Clusterungen haben wir die im vorigen Abschnitt beschriebenen Maße zur Bestimmung einer optimalen Gruppenanzahl k generiert⁶⁹.

Die Prüfgrößen für die Wahl der Anzahl Cluster sind in Abbildung 6-7 wiedergegeben.

Abbildung 6-7: Optimale Anzahl Cluster bei multivariater k-Means Clusterung der Regressorenmatrix zur Bestimmung der Investitionswerte pro Anschluss



Quelle: WIK.

Wie bereits bei der univariaten Clusterung lässt sich aus den Prüfgrößen kein klares Signal für eine spezifische Clusteranzahl ableiten. Vielmehr besteht auch hier ein Trade-off zwischen der Reduktion von Komplexität bei kleiner Clusteranzahl und hinreichender Genauigkeit im Sinne der erklärten Varianz oder der genutzten Information. Nach bereits 4 Clustern verringert sich der Zugewinn an Information mit jedem weiteren Cluster deutlich. Dennoch kann mit 10 Clustern bereits 82,3% der Varianz in den

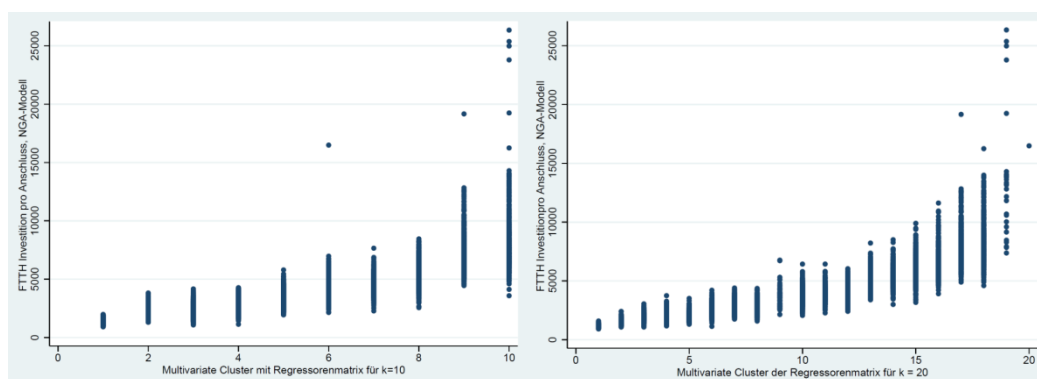
⁶⁹ Ebenso wurden die Starwerte für jede Berechnung zufällig ausgewählt.

5 Variablen erklärt werden. Die Erkenntnisgewinne durch zusätzliche Cluster sind beschränkt, weitere 6,7 Prozentpunkte können mit 10 weiteren Clustern erreicht werden (87,3 % bei 20 Clustern).

Abbildung 6-8 zeigt den Zusammenhang der Clusterbildung mit den FTTH-Investitionswerten je Anschluss für 10 und 20 Cluster. Demnach fängt die Clusterbildung zwar den Effekt steigender Investitionen in den oberen Clustern auf, allerdings weisen die Cluster starke Überlappungen in den Investitionswerten auf, können also keine trennscharfe Zuordnung hinsichtlich der Ausbaukosten leisten.⁷⁰ Die Hoffnung, mittels beobachtbarer Strukturmerkmale in einer ausreichenden Granularität auf die Kosten eines FTTH-Ausbaus schließen zu können, wird mit diesem Ansatz nicht erfüllt. Das Problem der Gruppenüberlappungen mit Blick auf die Investitionswerte je Anschluss ist markanter als wird es in Abschnitt 6.1.2 bei Verwendung allein der Teilnehmerdichte als Clusterkriterium hatten beobachten können (vgl. Abbildung 6-3, insbesondere die Clusterbildung nach Dichte, bezogen auf die Gebäuderasterflächen).

Für die Interpretation der gebildeten Cluster wären weitere, vertiefende Auseinandersetzungen nötig, ebenso für die Frage, ob sich exogene Clustergrößen finden lassen, die eine hinreichende Trennschärfe hinsichtlich der Investitionen je Anschluss aufweisen und einer Interpretation zugänglich sind.

Abbildung 6-8: FTTH-Investitionswerte und multivariate Clusterbildung für k=10, 20



Quelle: WIK

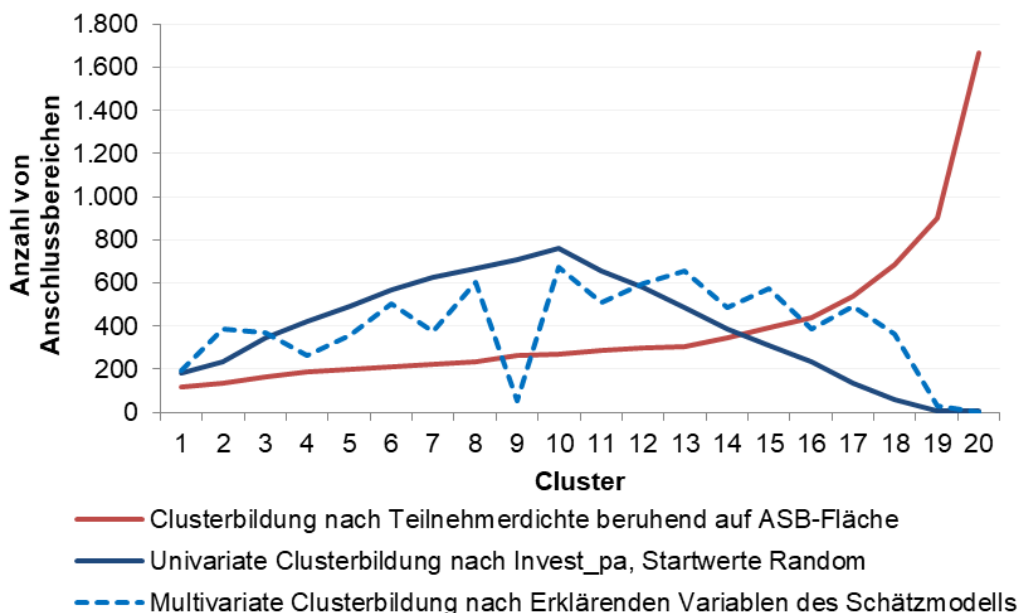
⁷⁰ Mit "Überlappung" wird die Feststellung bezeichnet, dass bei der gewählten Clusterbildung Anschlussbereiche, die dieselben durchschnittlichen Investitionen pro Anschluss aufweisen, in verschiedene Cluster fallen. Die Clusterbildung ist somit nicht überschneidungsfrei und führt vielmehr zu „Überlappungen“.

6.4 Vergleich der Clusteransätze

Die drei dargestellten Clusteransätze (nach Teilnehmerdichte, nach Investitionen je Anschluss und nach erklärenden Variablen des Schätzmodells) führen zu deutlichen Unterschieden in den Zuordnungen der Anschlussbereiche zu den Clustern. Dies soll für die Beispiele mit 20 Clustern für alle drei Ansätze dargestellt werden. Der Vergleich erfolgt anhand der Anzahl ASB je Cluster, der Anzahl Anschlüsse je Cluster und der mittleren Investitionen je Anschluss je Cluster.

Die Clusterbildung nach der Teilnehmerdichte beruhend auf der ASB Fläche ordnet den oberen (teuren) Clustern einen Großteil der ASB zu, wie aus Abbildung 6-9 (rote Linie) ersichtlich ist. Das hat seinen Grund in der Konstruktion der Clusterbildung, die jedem Cluster gleich viele Anschlüsse zuweist (ersichtlich aus Abbildung 6-10), bei gleichzeitig absteigender Sortierung nach Teilnehmerzahlen der Anschlussbereiche.⁷¹ Das führt dazu, dass die Anschlussbereiche mit niedriger Clusternummer die dichtesten Anschlussbereiche darstellen und mit wenigen Anschlussbereichen bereits 5% der Teilnehmer erreichen. Entsprechend sind immer mehr Anschlussbereiche notwendig, um in den weniger dichten Gebieten die 5% der Teilnehmer zu erreichen. In Abbildung 6-2 dominieren deshalb die Farben der hohen und damit weniger dichten Cluster die Abbildung der räumlichen Verteilung.

Abbildung 6-9: Vergleich der Clusteransätze bezüglich der Anzahl Anschlussbereiche je Cluster

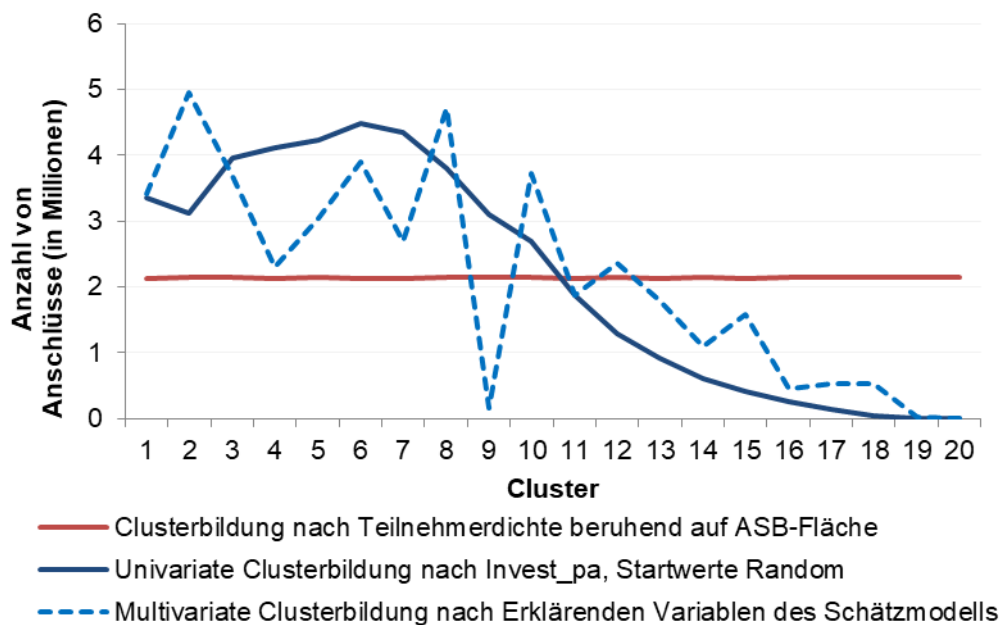


Quelle: WIK

⁷¹ Ein Vorteil der ursprünglichen Clusterbildung des WIK-NGA-Modells lag in der Aussagekraft der bzgl. der versorgten Haushalte, da jedes Cluster den gleichen Anteil an Haushalten repräsentiert.

Die univariate Clusterbildung nach den Investitionen pro Anschluss erfolgt als k-Mittelwert Clusterung. Die Anzahl der Anschlussbereiche ist eingipflig auf die Cluster verteilt mit einem Peak im mittleren Cluster (10) (dunkelblaue Linie in Abbildung 6-9), was direkt aus der Verteilung der Investitionen je Anschluss über die Anschlussbereiche folgt. Der Schwerpunkt der Teilnehmerzahlen je Anschlussbereich (Abbildung 6-10) liegt demgegenüber auf Cluster 6, weil eine ansteigende Zahl dichter Anschlussbereiche den Schwerpunkt automatisch nach vorne rückt. Nach der Logik der Clusterbildung sind in den unteren Clustern die Anschlussbereiche mit geringeren Ausbaurkosten vertreten, die zu einem Großteil auf die dicht besiedelten ASB entfallen.

Abbildung 6-10: Vergleich der Clusteransätze bezüglich der Anzahl Anschlüsse je Cluster



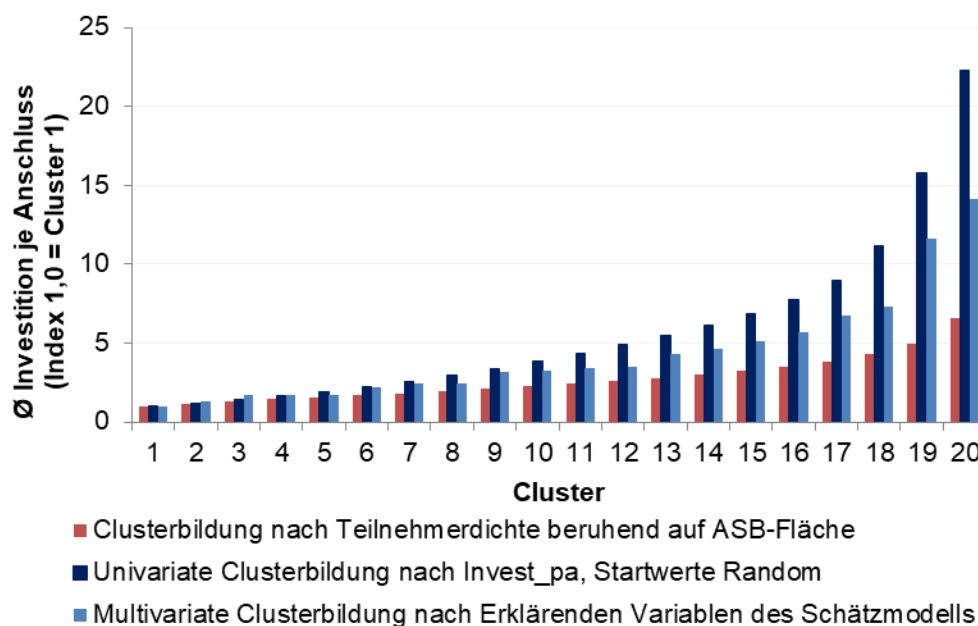
Quelle: WIK

Die multivariate Clusterung schwankt stark in ihrer Gruppengröße, weist aber ein Maximum der Anzahl Anschlussbereiche in Cluster 10 auf (Abbildung 6-9) hellblaue gestrichelte Linie). In den oberen Clustern sind mehr Anschlussbereiche vertreten als bei der univariaten Clusterung. Cluster 9 weist einen deutlichen Einbruch in der Gruppengröße auf. Dies ergibt sich letztlich aus k-means Clusteransatz und der empirischen Ausgangslage, die hier zu einer engen Grenzziehung für das Cluster 9 führt.⁷² Die Teilnehmerzahlen je Cluster sind maximal in Cluster 2 und 8 und zu den oberen Clustern hin sinkend.

⁷² Siehe auch das Minimum bei Cluster 9 in der Zahl Anschlussbereiche in Abbildung 6-9.

Der unstete Verlauf der Kurven im Fall der multivariaten Clusterung ist nur schwer zu interpretieren. Im Unterschied zu den beiden anderen Clusterverfahren findet bei der multivariaten Analyse nicht nur eine Variabel, sondern insgesamt fünf Einflussfaktoren Berücksichtigung. Die Vielzahl von Einflussgrößen erschwert die Interpretation.⁷³ Um die Ursachen des unsteten Verlaufs zu überprüfen, wären weitere umfangreichere Berechnungen mit alternativen Startwerten durchzuführen.

Abbildung 6-11: Vergleich der Clusteransätze bezüglich der mittleren Investitionen pro Anschluss je Cluster



Quelle: WIK

Abbildung 6-11 gibt für die drei Clusteransätze die Spreizung der mittleren Investitionen je Anschluss und Cluster bezogen auf Cluster 1 wieder. Die univariate Clusterung nach Investitionswert weist mit dem mehr als 20-fachen Wert in Cluster 20 gegenüber Cluster 1 die größte Spreizung auf. Die multivariate Clusterung mit dem etwa 14-fachen Wert in Cluster 20 bewegt sich in der Mitte und die geringste Spreizung weist die Clusterung nach Teilnehmerdichte auf. Dies passt zu den Eigenschaften der Clusteransätze: Bei der univariaten Clusterung sind die Clustergrenzen trennscharf: Anschlussbereiche mit gleichen Investitionen je Anschluss fallen ausnahmslos in dasselbe Cluster. Die Spreizung der Investitionen je Anschluss ist deswegen hier am deutlichsten. Bei der Clusterbildung nach Teilnehmerdichte hingegen geht diese trennscharfe Zuordnung verloren und es treten starke Überlappungen in den Investitionswerten je Cluster auf. Dies lässt die Spreizung der Mittelwerte schrumpfen. Beim multivariaten Ansatz besteht ebenfalls eine Überlappung der Investitionswerte je Anschluss, allerdings weniger deutlich als bei

⁷³ Siehe hierzu auch Abschnitt 6.3.

der Clusterung allein auf Basis der Teilnehmerdichte. Insbesondere in den hohen Clustern reicht er nicht an die Trennschärfe der idealtypischen univariaten Clusterbildung nach Maßgabe der Investitionen je Anschluss heran. Das beschriebene Muster ist über alle Cluster hinweg ab Cluster 4 sichtbar.

6.5 Geografische Nachbarschaft: Hot Spot - Cold Spot Analyse

In diesem Abschnitt wird der Frage nachgegangen, ob die FTTH-Investitionen je Anschluss erkennbar räumlich zusammenhängen. Bei der Clusteranalyse der vorherigen Abschnitte wurden Anschlussbereiche nach bestimmten Attributwerten einander zugeordnet. Die räumliche Nachbarschaft spielte dabei keine Rolle. Für einen privatwirtschaftlichen oder geförderten FTTH-Ausbau hingegen spielen räumliche und raumpolitische Erwägungen natürlich eine Rolle.

Wir gehen deshalb hier den räumlichen Zusammenhängen hoher und niedriger Investitionswerte je Anschluss nach. Dazu bedienen wir uns einer Hot Spot / Cold Spot Analyse⁷⁴. Die Hot Spot / Cold Spot Analyse identifiziert statistisch signifikante räumliche Cluster mit hohen Werten (Hot-Spots) und mit niedrigen Werten (Cold-Spots) der untersuchten Größe. Dabei wird geprüft, ob die beobachtete räumliche Cluster-Bildung von hohen oder niedrigen Werten auffälliger ist als bei einer zufälligen Verteilung der gleichen Werte. Sie bedient sich der Getis-Ord G_i^* -Statistik,⁷⁵ die einen z-Wert darstellt. Sie gibt eine Einschätzung darüber, ob sich hohe oder niedrige Werte einer Größe in einer vorgegebenen Nachbarschaft signifikant häufen. Das Rechenergebnis hängt dabei auch von der Art und Erstreckung der Nachbarschaft ab.

6.5.1 Nachbarschaftsanalyse auf Basis der Anschlussbereiche

In unserer Analyse wurden die FTTH-Investitionen je Anschluss für jeden Anschlussbereich als Untersuchungsgegenstand gewählt. Als Nachbarschaft wurde das Konzept eines fixen Distanzbandes übernommen. Die Erstreckung des Distanzbandes wurde so gewählt, dass sichergestellt ist, dass jeder Anschlussbereich mindestens einen Nachbarn aufweist. Diese Einstellung entspricht in unserem Datensatz rd. 43 km, was einer nahen regionalen Umgebung entspricht.⁷⁶

⁷⁴ Die Analyse wurde mit dem Geoverarbeitungswerkzeug Hot-Spot Analyse in ARCGIS 10.6 von ESRI durchgeführt.

⁷⁵ Getis, A. / Ord, J.K. (1992) sowie Ord, J.K. / Getis, A. (1995).

⁷⁶ Hier geht es darum, welche Nachbarschaft bei der Berechnung der G_i^* -Statistik unterstellt wird. D.h., es geht um die Auswahl von Werten der Anschlussbereiche, die an den Anschlussbereich, für den gerade die Statistik berechnet wird, in die Berechnung der Statistik einfließen sollen. Da die Statistik für alle ASB berechnet wird, wird, also bildlich gesprochen ein Rahmen einer bestimmten Größe über den zu berechnenden ASB gelegt und alle ASB, die in den Rahmen fallen, sind Nachbarn und werden bei der Berechnung der Statistik und damit bei der Entscheidung, ob der betrachtete Wert räumlich zufällig sein kann, berücksichtigt.

Die Statistik kann nur berechnet werden, wenn es auch Nachbarn gibt, deshalb wird vom Programm geprüft bei welcher Rahmengröße, um im Bild zu bleiben, dies für alle Anschlussbereiche gegeben

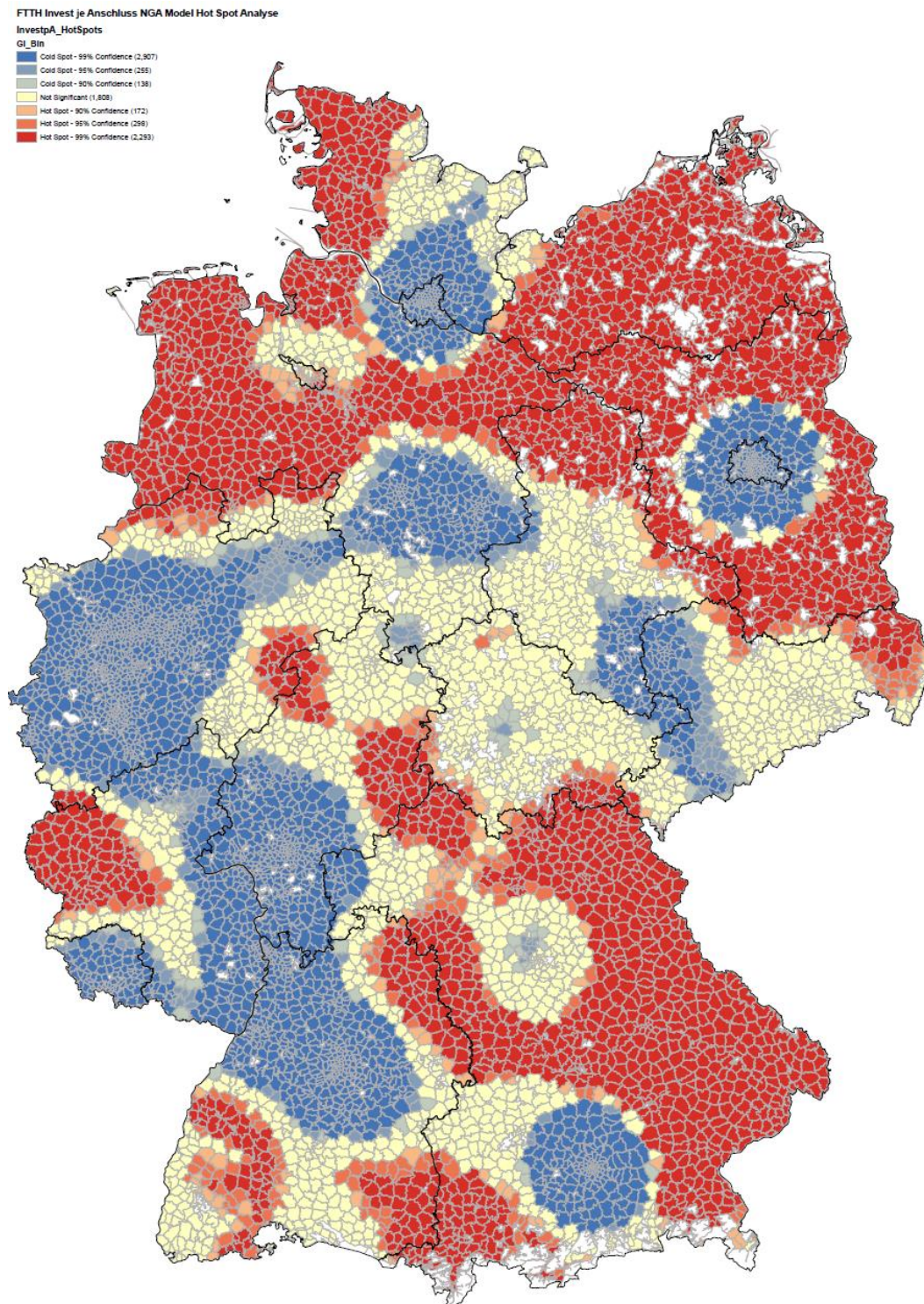
Das Ergebnis der Berechnung ist ein z-Wert (die G_i^* -Statistik) für jeden Anschlussbereich, der die Teststatistik für die Ablehnung der Nullhypothese wiedergibt, dass der Wert einer zufälligen räumlichen Verteilung entspricht. Auf dieser Basis wird die Kategorialvariable G_i -bin berechnet, die das statistische Signifikanzniveau für die Ablehnung des Tests statistischer räumlicher Unabhängigkeit auf den Niveaus 99%, 95% und 90% wiedergibt. Da die Teststatistik für hohe Werte (Hot Spots) positiv ausfällt und für niedrige Werte (Cold Spots) negativ, liefert die Kategorialvariable 7 Werte, wobei der mittlere Wert 0 die nicht signifikanten Bewertungen enthält (räumliche Zufälligkeit kann nicht abgelehnt werden).

Abbildung 6-12 gibt das Ergebnis der errechneten räumliche Muster wieder. Die Anschlussbereiche mit signifikant hohen Werten (Hot Spots) sind rot dargestellt, die signifikant niedrigen Werte (Cold Spots) sind blau dargestellt. Die Farbabstufungen folgen den Signifikanzniveaus, die nicht gefärbten Anschlussbereiche (Hintergrundfarbe) weisen insignifikante Werte der Teststatistik auf; die Nullhypothese der vollständigen zufälligen räumlichen Verteilung der Investitionen je Anschluss kann in diesen Fällen (weder rot noch blau) nicht verworfen werden.

Die Analyse zeigt klare, statistisch signifikante räumliche Cluster hoher und niedriger Investitionswerte je Anschluss. In Gebieten hoher Werte (Hot Spots, rot) ist der Ausbau teuer, in Gebieten niedriger Werte (Cold Spots, blau) sind die Ausbaukosten niedrig. Demnach bildet der gesamte Norden Deutschlands, mit Ausnahme der Großstadtregionen Hamburg und Berlin ein Cluster hoher Werte, ebenso die ländlichen Regionen in Bayern, die westliche Pfalz sowie Teile des Schwarzwalds und des Allgäus. Cluster niedriger Investitionen je Anschluss finden sich in den Großstadtregionen (Berlin, Hamburg, München, Hannover) und den Industriellen Regionen Deutschlands (Rhein-Ruhr, Rhein-Main, Stuttgarter Raum, Saarland, Leipzig-Zwickau).

ist. Das sind dann die 43 km Erstreckung des Distanzbandes (des Rahmens). Hier gibt es nun noch ganz andere Nachbarschaften die man unterstellen kann (z.B. nur angrenzende Anschlussbereiche, die k nächsten Nachbarn etc.) und natürlich kann man auch die Größe des Rahmens verändern. Die Wahrscheinlichkeit, dass die zufällige räumliche Verteilung abgelehnt wird, ist bei Vorliegen räumlicher Korrelation wahrscheinlicher. Den Peak räumlicher Korrelation erhalten wir bei rd. 64 km Erstreckung des Distanzbandes. Die Grafik dazu entspricht nahezu der mit 43 km Erstreckung. Eine umfassendere Analyse war im Rahmen dieser Forschungsarbeit nicht möglich, lässt aber noch weitere Erkenntnisgewinne erwarten.

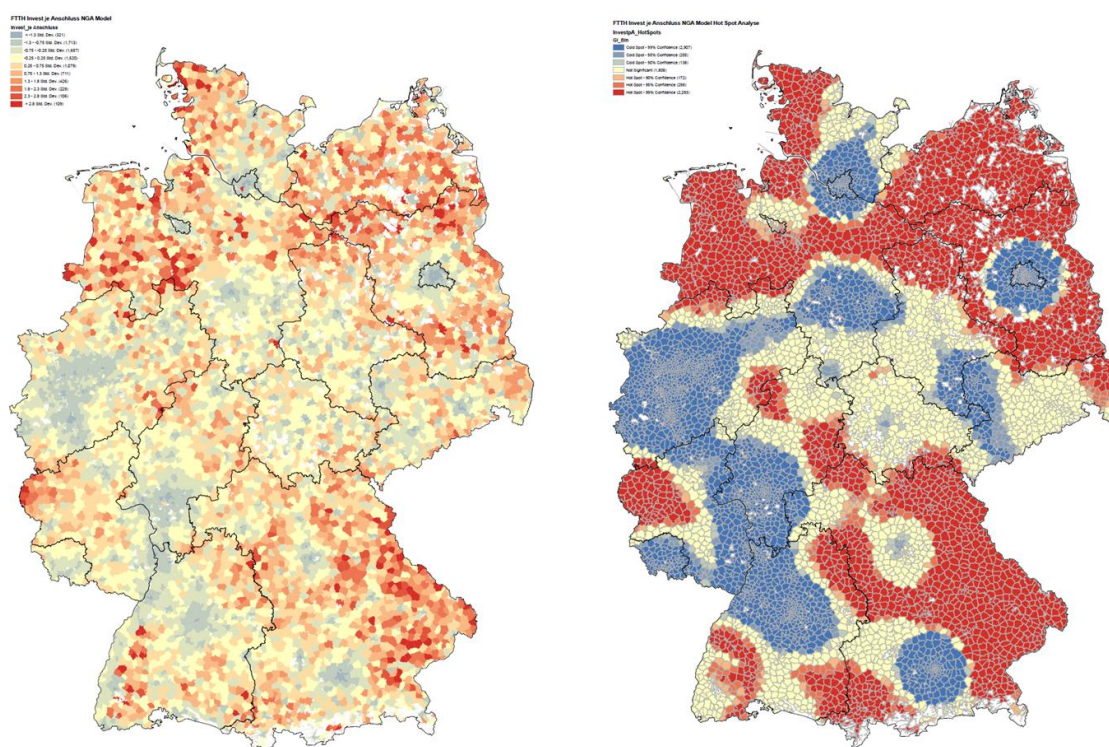
Abbildung 6-12: Hot Spot / Cold Spot Analyse der Investitionen je Anschluss



Quelle: WIK, Landesgrenzen BKG

Abbildung 6-13 stellt nochmals die Ausgangslage der räumlichen Verteilung der Investitionen je Anschluss (siehe Abbildung 6-1)⁷⁷ der räumlichen Clusterung auf Grundlage der Hot Spot/ Cold Spot Analyse gegenüber. Die Hot Spot Analyse untermauert den optischen Eindruck der Ergebnisdarstellung auf Anschlussbereichsebene und tritt sogar stärker hervor. Die Bereiche insignifikanter G_i^* -Statistik scheinen durch mittlere Werte der Investitionen je Anschluss geprägt und/oder durch wechselnde Werte in der näheren Umgebung bedingt zu sein.

Abbildung 6-13: Vergleich der räumlichen Verteilung der Investitionswerte je Anschluss mit der Hot Spot /Cold Spot Clustern



Quelle: WIK, Landesgrenzen BKG

wik 

6.5.2 Nachbarschaftsanalyse auf Basis von Kreisen

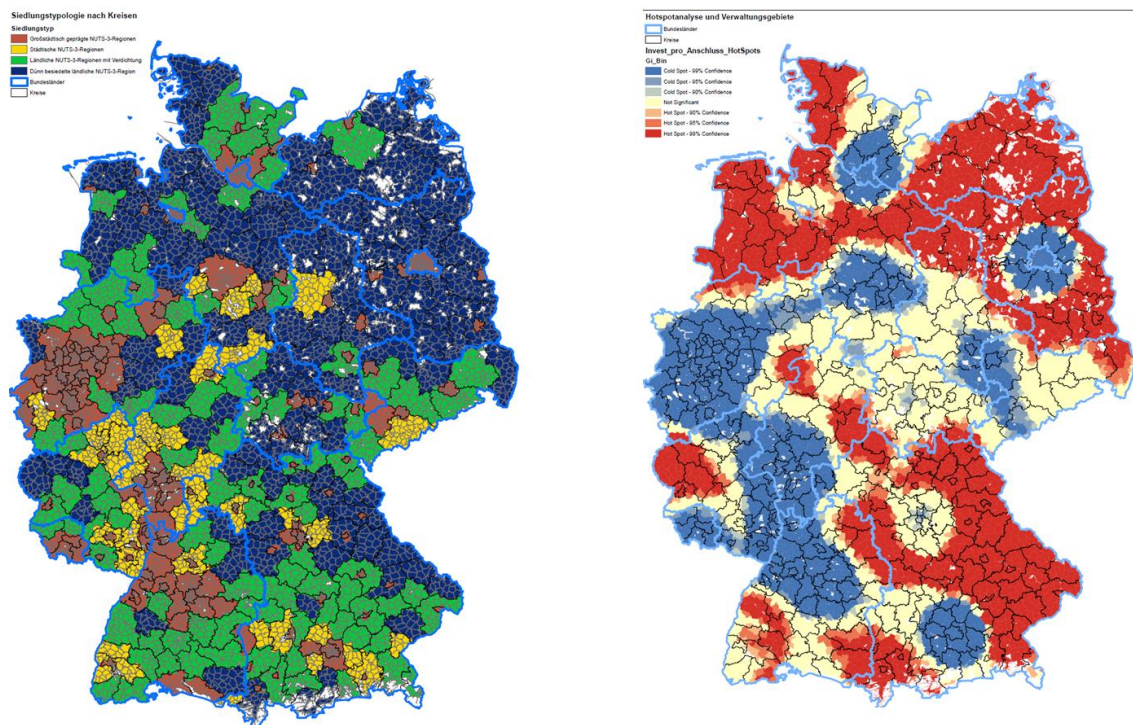
Abbildung 6-14 ist eine Gegenüberstellung einer Siedlungstypologie nach Kreisen des BBSR⁷⁸ und der berechneten räumlichen Clusterung der Investitionen je Anschluss. Da die Siedlungstypologie lediglich auf Kreisebene vorliegt, wurden der HotSpot/ Cold Spot Analyse für eine einfachere Vergleichbarkeit die Kreise hinzugefügt. Demnach besteht eine hohe Übereinstimmung der Cold Spot Cluster mit den großstädtisch geprägten

⁷⁷ Die Farbgebung wurde für den Vergleich der Hot Spot / Cold Spot Analyse angepasst (rot hoher Wert, blau niedriger Wert)

⁷⁸ Siedlungsstrukturelle NUTS-3-Regionen 2016 für Europa des Bundesamtes für Bauwesen und Raumordnung in Bonn.

Kreisen und den städtischen Kreisen (rot und gelbe Flächen der Siedlungstypologie). Die ländlichen Kreise mit Verdichtung (grün) stimmen häufig mit den insignifikanten Gebieten der Hot Spot / Cold Spot Analyse überein. Dünn besiedelte ländliche Regionen zeigen eine gute Übereinstimmung mit den Hot Spots unserer Analyse. Die gute Übereinstimmung ist eventuell auch der Erstreckung des Distanzbandes der Nachbarschaft in der Analyse geschuldet, die mit 43 km der Erstreckung von Kreisen gut entspricht. Eine statistische Absicherung dieses optischen Eindrucks ist weiteren Untersuchungen vorbehalten.

Abbildung 6-14: Vergleich der Siedlungstypologie nach Kreisen mit Hot Spot /Cold Spot Clustern

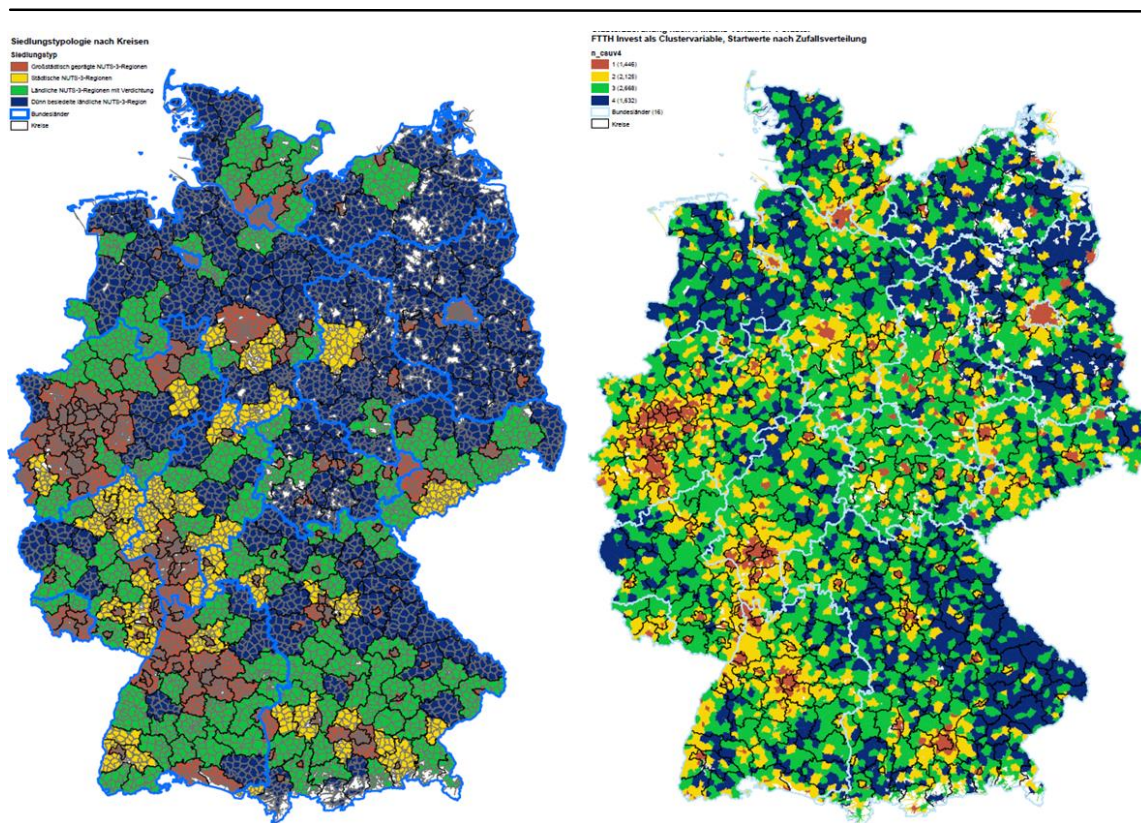


Quelle: WIK, Landesgrenzen und Kreisgrenzen BKG, Siedlungsstrukturelle NUTS-3-Regionen 2016 für Europa BBSR

Der Vergleich der Siedlungsstruktur mit der Clustereinordnung nach 4 Gruppen gemäß Investitionswert je Anschluss zeichnet ein etwas disparateres Bild. Bei der k-Mittelwert Clusterung bedeutet rot niedrige Ausbaurkosten, gelb niedrige bis mittlere Ausbaurkosten, grün mittlere bis hohe Kosten und blau sehr hohe Investitionen je Anschluss. Insbesondere scheinen die mittleren bis hohen Investitionen je Anschluss (grüne Flächen) ganz oder teilweise Flächen einzunehmen, die in der Siedlungstypologie als städtische, ländliche Gebiete mit Verdichtung und dünn besiedelte ländliche Gebiete gekennzeichnet sind. Cluster mit günstigen bis mittlere Ausbaurkosten (rote und gelbe Flächen)

scheinen gut mit großstädtisch geprägten Kreisen und nur zum Teil mit städtisch geprägten Kreisen übereinzustimmen. Auffällig ist ebenso, dass die Kreisflächen in der Darstellung der Cluster nach Investitionen je Anschluss häufig durch mehrere Farben geprägt sind, während die Siedlungstypologie in dieser Granularität nicht vorliegt.

Abbildung 6-15: Vergleich der Siedlungstypologie nach Kreisen mit k-Mittelwert Clusterbildung der Investitionen je Anschluss bei 4 Clustern (90% erklärte Varianz)



Quelle: WIK, Landesgrenzen und Kreisgrenzen BKG, Siedlungsstrukturelle NUTS-3-Regionen 2016 für Europa BBSR

6.6 Zwischenfazit Clusterung

Zusammenfassend lässt sich sagen, dass die bisherige, dem WIK-NGA-Modell zugrundeliegende Clusterbildung eine große Menge unterschiedlicher Investitionswerte zusammenfasst. Somit kann gerade für Gebiete niedriger Anschlussdichte keine klare Trennung nach Investitionshöhe pro Anschluss generiert werden. Hier hängen die Investitionen je Anschluss verstärkt auch von anderen Einflussgrößen ab. Alternative Dichtemaße, insbesondere die Anschlussdichte in Bezug auf die Gebäuderasterflächen, waren in der Lage trennschärfere Clusterbildungen zu generieren, ohne das Kriterium einer Gleichverteilung der Anschlüsse über die Cluster aufzugeben.

Unter Rückgriff auf die aus der bottom-up Modellierung flächendeckend vorliegenden Investitionen je Anschluss und Anwendung der univariaten k-means Clustermethode konnte eine idealtypische Clusterbildung mit entsprechender Trennschärfe zwischen den Clustern abgeleitet werden. Bei einer Anzahl von $k=20$ Clustern verbleibt auch hier in dem letzten Cluster eine sichtbare Abweichung, statistisch ist diese jedoch unbedeutend.

Der Frage nach der optimalen Anzahl von Clustern wurde ebenfalls unter Rückgriff auf die Investitionen je Anschluss als Clusterkriterium durch Anwendung der k-means Methode nachgegangen. Hierfür konnte eine deutliche Reduktion der Clusteranzahl mit nur geringem Informationsverlust (gemessen am Anteil der erklärten Varianz) aufgezeigt werden. Es bleibt weiteren, zukünftigen Analysen überlassen, für die Anwendung exogener Variablen im Rahmen univariater k-means-Clusterung die optimale Anzahl von Clustern zu bestimmen.

Multivariate k-means Clusterungen wurden unter Rückgriff auf die in der Regressionsanalyse abgeleiteten Einflussfaktoren durchgeführt. Die Ergebnisse zeigten einen erklärten Anteil der Varianz in Höhe von nahezu 89% bei einer Clusteranzahl von $k=25$. Dennoch waren die Ergebnisse nicht vollständig befriedigend und lieferten für die Grundgesamtheit der Anschlussbereiche in Deutschland schlechtere Trennschärfe als die Clusterbildung unter Anwendung des einfachen Dichtekriteriums nach Gebäuderasterflächen und Gleichverteilung der Haushalte über die Cluster (für $k=20$). Um dem Forschungsziel näher zu kommen, Kriterien für regional differenzierte Investitionen je Anschlussleitung auf Basis allgemein verfügbarer Informationen zu identifizieren, streben wir weitere Analysen an. Dabei wollen wir durch eine Zusammenführung der Erkenntnisse aus Regressionsanalyse und Clusterbildung eine Anwendung der Regressionskoeffizienten für die multivariate k-means-Clusterbildung heranziehen.

Räumlich korrelierte Cluster hoher und niedriger Investitionen je Anschluss konnten mit Hilfe einer Hot Spot/Cold Spot Analyse basierend auf der G_i^* -Statistik aufgezeigt werden. Die Konzeptualisierung der räumlichen Nachbarschaft hat einen deutlichen Einfluss auf die Clusterbildung, so dass auch hier ein Ansatzpunkt für weitere Arbeiten gesehen werden kann. Die Vorzüge dieses Konzepts bestehen in seiner Relevanz für unternehmensstrategische Investitionsentscheidungen oder breitbandpolitischen Ausbauförderung, welche in der Regel immer auf (größere) zusammenhängende Gebiete abzielen. Für zukünftige Arbeiten bietet sich daher an, die Erkenntnisse hinsichtlich der Profitabilität eines NGA Ausbaus, wie sie unter Rückgriff auf das WIK-NGA-Modell und die zugehörige Clusterung abgeleitet wurden (Abschnitt 4), auch für andere Clusteransätze zu untersuchen. Insbesondere mit Blick auf den erwarteten Umfang profitabler privatwirtschaftlicher Investitionstätigkeit und Förderbedarfe erwarten wir von einer Clusterbildung auf Basis des Nachbarschaftskonzepts weitergehende Erkenntnisse.

7 Zusammenfassung und Ausblick

In diesem Forschungsprojekt wurde durch das Zusammenführen von drei verschiedenen methodischen Ansätzen versucht, Erkenntnisse über die Kosten und Profitabilität eines Glasfaserausbaus zu gewinnen. Zu diesem Zweck wurde eine flächendeckende bottom-up Kostenmodellierung auf Basis öffentlich verfügbarer Daten und vom WIK entwickelter Kostenmodelle durchgeführt. Die Ergebnisse bildeten die Grundlage für alle drei Untersuchungsthemen dieses Diskussionsbeitrags.

Die Modellierung der durchschnittlichen Kosten von NGA-Anschlüssen je Anschlussbereich wurde im ersten Analyseteil (Abschnitt 4) dazu verwendet, clusterbasiert die Profitabilität eines Netzausbaus für die Technologien FTTH P2P, FTTN sowie FTTC zu bestimmen.

Der clusterbasierten Kostenanalyse zu Folge können mit einem FTTH P2P Ausbau 65% aller Haushalte in Deutschland profitabel erreicht werden. Dem liegt die konservative Annahme zugrunde, dass das ausbauende Unternehmen 90% der Haushalte versorgt. Damit soll der maximal erreichbare profitable Ausbau abgeschätzt werden. Bezogen auf die Anzahl der Anschlussbereiche ist jedoch weniger als die Hälfte profitabel, was die Konzentration von Haushalten auf einen Teil der Anschlussbereiche zum Ausdruck bringt. Für Berechnungen wurde ein durchschnittlicher Erlös in Höhe von 45 €, für FTTH P2P unterstellt.

Unter der Annahme von regionaler Quersubventionierung kann ein flächendeckend agierender Anschlussnetzbetreiber sogar insgesamt profitabel sein. Dazu muss er seine Überschüsse aus den profitablen Regionen für die Kostendeckung der unprofitablen Anschlussbereiche einsetzen. Dies gelingt jedoch nur, wenn die durchschnittlichen Erlöse (ARPU) und Marktanteile nicht relevant durch die Wettbewerber bestritten werden. Kann der Netzbetreiber nur geringere Marktanteile realisieren, führt der Fixkostencharakter seiner Investitionen in einen flächendeckenden Ausbau dazu, dass seine Margen spürbar sinken und Investitionen unprofitabel werden können. Antizipiert er dies, oder fürchtet, nur einen geringeren Marktanteil zu realisieren, wird er seinen NGA-Ausbau nur in Anschlussbereichen vorantreiben, deren Margen eine Profitabilität auch bei geringeren Marktanteilen erlauben. Diese Zusammenhänge mögen u.a. erklären, warum der in Deutschland beobachtbare FTTH Ausbau hinter den Erwartungen zurückbleibt.

Für die 20 Cluster des WIK-NGA-Modells wurde eine Kostenspreizung berechnet, der zufolge das teuerste Cluster um den Faktor 3,2 teurer ist als das billigste. Diese Spreizung fällt stärker aus als dies noch in der Vorgängerstudie 2011 mit dem Faktor 2,3 der Fall war. Diese Erkenntnis ist insbesondere auf die jetzt angewendete methodische Vorgehensweise einer jetzt anschlussbereichsgenauen Berechnung der Kosten pro Anschlussleitung zurückzuführen. Diese hat es ermöglicht, die im weiteren Zuge auf Cluster aggregierten Ergebnisse stärker ausdifferenzieren als dies noch in der Vorgängerstudie geleistet werden konnte. Seinerzeit basierten die Ergebnisse noch auf

Kostenberechnungen pro Cluster insgesamt (über alle zugehörigen Anschlussbereiche). Die dichten Cluster sind nun günstiger, die dünn besiedelten teurer. – In Abschnitt 6.4 wird auf die Spreizung nochmals Bezug genommen und herausgearbeitet, wie stark die Kriterien der Clusterbildung das Ergebnis beeinflussen.

Die identifizierten Kostenunterschiede zwischen der rein glasfaserbasierten FTTH P2P Technologie und den hybriden Technologien FTTS und FTTC sind nennenswert und liegen insbesondere im „letzten Segment“ der Hauszuführung und gebäudeinternen Verkabelung begründet (Abschnitt 4.1.2). Aufgrund der bereits vorhandenen kupferbasierten Infrastruktur zur Anbindung der Endkunden fallen Kosten lediglich für tatsächlich kontrahierte Kunden an. Demgegenüber sind für ein rein glasfaserbasiertes Anschlussnetz, wie es hier mit FTTH P2P betrachtet wird, die Investitionen in die Hauszuführung für die Gesamtheit der Haushalte (homes passed) und für die gebäudeinterne Verkabelung für sämtliche kontrahierten Haushalte noch zu berücksichtigen. Die bei hybriden Technologien zum Ansatz kommenden (anteiligen) Überlassungsentgelte der KVz-TAL beinhalten bekanntlich keine Kosten für die gebäudeinterne Verkabelung.

Die Auswertung der Daten des Breitbandatlas (Stand 04/2018) und deren Projizierung auf die Anschlussbereiche und Dichtecluster lieferte die Basis für einen Brownfield-Ansatz zur Einschätzung des noch ausstehenden Investitionsbedarfs für den flächendeckenden Glasfaserausbau (Abschnitt 4.2). Erwartungsgemäß wurde für FTTH ein noch bestehender Investitionsbedarf für mehr als 90% der Anschlüsse identifiziert. Demgegenüber weisen CATV und FTTC bereits eine weitaus größere Abdeckung auf und liegen mit 37% (CATV) und 36% (FTTC) noch zu realisierender Abdeckung deutlich hinter dem Investitionsbedarf von FTTH zurück. Leider mussten wir unsere Analyse auf Basis einer Datenqualität durchführen, die in mehrfacher Hinsicht unbefriedigend war: (1) Die Versorgungsangaben wurden nach Maßgabe von 6 Geschwindigkeitsklassen unterschieden, davon war jedoch die höchste, differenziert ausgewiesene Anschlussgeschwindigkeitsklasse „>= 50 Mbitps“ und daher mit Blick auf die heute relevanten Breitbandanschlussgeschwindigkeiten begrenzt aussagefähig. (2) Die Angaben wurden zusätzlich nach FTTH, CATV und xDSL differenziert, wobei hier die fehlende Angabe zur FTTH Topologie einerseits und die Abgrenzung gegenüber FTTB andererseits eine relevante Unsicherheit bezüglich des noch verbleibenden Investitionsbedarfs mit sich bringt (so erfordern Punkt-zu-Multipunkt Topologien weitere, nennenswerte Investitionen im Hauptkabelsegment). (3) Aus den Angaben des Breitbandatlas ließ sich nicht eindeutig identifizieren, welche Regionen noch über KEINE Versorgung mit einer der definierten Anschlussgeschwindigkeiten und Technologien verfügen. Die NICHT VERSORGTEN Regionen wurden mit denen bis einschließlich 10% versorgter Haushalte zusammengefasst. (4) Auch waren die Angaben nicht überschneidungsfrei, d.h. es war nicht erkennbar, ob sich die Verfügbarkeit verschiedener Infrastrukturen auf dieselben Haushalte erstreckt (Mehrfachversorgung) oder disjunkte Gruppen versorgt werden. Diese Defizite sind im aktuellen Release des Breitbandatlas zum Teil überwunden. Zum Zeitpunkt der Datenbeschaffung für diese Studie standen sie leider noch

nicht zur Verfügung und können von uns erst im Rahmen weiterer Forschungsarbeiten herangezogen werden.

Der Brownfield-Ansatz zur Bestimmung der noch zu leistenden Investitionen wurde um die Perspektive pfadabhängiger Investitionserfordernisse ergänzt (Abschnitt 4.2.3). Im Fokus der Betrachtung standen hier die „Erweiterungsinvestitionen“ aus der Perspektive eines FTTC bzw. FTTS/CATV-Netzes, um eine FTTH P2P Versorgung zu erzielen. Dabei wurde auf das Einsparpotential abgestellt, das durch den bereits realisierten Glasfaserausbau bis zum MSAN bzw. G.fast-DSLAM erreicht werden kann. Für FTTC wurde dieses auf maximal 20% des gesamten FTTH P2P Investitionsbedarfs abgeschätzt; für FTTS waren es mit 30% 10 Prozentpunkte mehr. Dabei ist Skepsis geboten, ob Einsparungen in diesem Umfang tatsächlich realisiert werden können, da (1) a priori nicht davon ausgegangen werden kann, dass die Trassenkapazitäten des Hauptkabelsegments bereits vollständig für die Implementierung einer FTTH P2P Topologie ausgelegt wurden. (2) Für unsere Analyse standen lediglich undifferenzierte Information zur Verfügung, aus denen nicht hervorging, wie vollständig der Ausbau der Verzweigerbereiche für CATV bzw. FTTC in den Anschlussbereichen bereits vollzogen wurde. Mit der Annahme, dass CATV bzw. FTTC in einem Anschlussbereich verfügbar ist, wurde für die Bestimmung des maximalen Einsparpotentials unterstellt, dass die Glasfaser bereits zu sämtlichen MSAN (FTTC) bzw. G.fast DSLAM (FTTS) ausgebaut wurde. – Letztlich wären Informationen über den konkreten Ausbaustand des flächendeckenden Netzbetreibers oder Gebäude-genaue Angaben zur Technologieverfügbarkeit notwendig, um hier die Aussagefähigkeit zu verbessern.

Mit einer regionalisierten grafischen Darstellung (Abschnitt 4.2.4) sowohl des Breitbandausbaustands als auch der anschlussbereichsgenaue FTTH P2P Investitionen konnte visualisiert werden, dass sich der bereits realisierte Breitbandausbau von FTTH, CATV und FTTC in diejenigen Regionen einbettet, die auch für den FTTH P2P Ausbau als kostengünstig und profitabel identifiziert werden konnten.

Die in dieser Studie angewendeten statistischen Analyseverfahren (Abschnitte 5 und 1) wurden zum einen durchgeführt, um die Belastbarkeit von Kostenschätzungen und clusterbasierten Analysen weiter zu verbessern. Damit sollte die Basis dafür geschaffen werden, auch unter Bedingungen schlechterer Datenverfügbarkeit und ohne eine aufwendige bottom-up Modellierung unter Anwendung statistischer Verfahren belastbare Ergebnisse über die Kosten von Anschlussnetzen abzuleiten. Zum anderen sollten die Beschränkungen des dem WIK-NGA-Modell zugrunde liegenden Clusteransatzes untersucht und die Clusterbildung hinsichtlich ihrer Aussagekraft bezüglich der Profitabilität privatwirtschaftlicher Investitionen verbessert werden.

Unter Rückgriff auf die Daten der flächendeckenden bottom-up Modellierung (Abschnitt 3 und 4) konnten in Abschnitt 5 wesentliche Einflussfaktoren auf die Investitionen je Anschluss mit Hilfe eines linearen Regressionsmodells ermittelt werden. Dieses erlaubt anhand weniger Strukturmerkmale eine erwartungstreue Schätzung der Investitionen je

Anschlussbereich mit hoher Qualität. Dieses als Benchmark-Modell bezeichnete Regressionsmodell dient in einem weiteren Schritt als Vorlage für die Ableitung und Überprüfung von Proxy-Variablen dieser Einflussgrößen mit allgemeinerer Verfügbarkeit.

Im Zuge der Analyse auf Basis des vollständigen Datensatzes der bottom-up Modellierung konnte gezeigt werden, dass anstelle des einfachen Dichte-Maßes (Anschlüsse pro km² des Anschlussbereichs) eine Referenzierung auf die bebaute Fläche (Anschlüsse pro Siedlungsfläche [km²] des Anschlussbereichs) bessere Ergebnisse liefert. Allerdings bleiben auch bei dieser Variablen in den teuersten Anschlussbereichen nennenswerte Varianzen bestehen.

Die Übertragung des Schätzansatzes auf höher aggregierte räumliche Abgrenzungen, beispielsweise auf Kreisebene (NUTS3), lieferte vergleichbare Ergebnisse. Die statistische Signifikanz und Stärke der Zusammenhänge für „Teilnehmerdichte“ und „Trassendichte im HK-Segment“ bleiben nach der räumlichen Aggregation weitestgehend erhalten.

Das angestrebte Ziel, bei der Kostenschätzung ausschließlich allgemein verfügbare Daten zu verwenden, konnte nicht erreicht werden. Zwar erklären die aufgestellten Proxy-Modelle im Minimum über 90 Prozent der Datenvariationen der Zielvariablen „Investitionen pro Anschluss“. Jedoch besteht keines der beiden aufgestellten Proxy-Modelle die üblichen Testverfahren. Eine Abschätzung der HK-Trassenlänge durch allgemein verfügbare Informationen zur Straßenlänge konnte keine befriedigenden Ergebnisse liefern. Hier gilt es daher weitere Untersuchungen durchzuführen, um adäquate Proxy-Variablen zu identifizieren, die für den hier zugrunde gelegten Analyserahmen der Investitionen in Teilnehmeranschlussnetze in Deutschland, perspektivisch aber auch für andere Länder und Regionen erwartungstreue Ergebnisse liefern können.

Die Analyse der Clusterbildung in Abschnitt 1 nimmt die dem WIK-NGA-Modell zugrunde liegenden Clusterkriterien als Ausgangspunkt und bedient sich der Erkenntnisse aus den vorherigen Abschnitten der bottom-up Modellierung (Abschnitte 3 und 4) und Regressionsanalyse (Abschnitt 5). Mit Hilfe der grafischen Darstellung für die Fläche Deutschlands wird visualisiert, dass die bisherige Clusterbildung (1) mit einer starken Versprenkelung der Anschlussbereiche eines Clusters einhergeht und (2) für einen Großteil der Anschlussbereiche keine differenzierten Erkenntnisse liefert (vgl. Abbildung 6-2 sowie Abbildung 6-9); gerade in Gebieten niedriger Anschlussdichte hängen die Investitionen je Anschluss verstärkt auch von anderen Einflussgrößen ab.

Ziel der Analyse der Clustermethoden war es, das bisherige Clusterkriterium der Gleichverteilung der Anschlüsse über die Cluster, geordnet nach Dichte (Anschlüsse pro km² der gesamten Anschlussbereichsfläche), zu Gunsten einer Clusterung aufzugeben, die eine belastbare Kostenermittlung nach Maßgabe regional identifizierbarer Unterschiede erlaubt. Die Anwendung alternativer Dichtemaße „Anschlüsse pro bebaute Fläche des Anschlussbereichs“, unter Rückgriff zum einen auf Polygondaten, zum

anderen auf Rasterdaten unter Beibehaltung der Gleichverteilung der Haushalte konnte für die 20 Cluster verbesserte Ergebnisse liefern: Die Überlappung von Investitionen pro Anschluss in den einzelnen Clustern konnte reduziert werden. Die besten Ergebnisse brachte dabei die Abgrenzung nach Maßgabe der „Anschlüsse pro bebauter Rasterfläche“ (Abschnitt 6.1.2).

Um eine Referenz im Sinne einer idealtypischen Clusterbildung heranzuziehen, wurde eine Clusterung unter Anwendung eines univariaten k-means-Verfahrens und Beibehaltung der 20 Cluster durchgeführt. Die Abgrenzung nach Maßgabe einer Gleichverteilung der Haushalte über die Cluster wird bei diesem Ansatz aufgegeben. Als Abgrenzungskriterium für die Cluster zieht diese Methode die Minimierung der summierten Varianzen heran. Methodisch wird damit eine möglichst homogene Gruppenbildung innerhalb der Cluster und klare Trennung zwischen den Clustern angestrebt. Durch die Verwendung der bottom-up ermittelten Investitionen pro Anschluss und Anschlussbereich als Clusterkriterium, wird eine überschneidungsfreie Clusterbildung erzielt. In Bezug auf die Zielsetzung einer regional differenzierten, möglichst überschneidungsfreien Clusterbildung hat dieser Ansatz idealtypischen Charakter, weist jedoch den zentralen Mangel auf, dass die hier für die Clusterbildung verwendete Größe „Investitionen pro Anschluss“ im allgemeinen nicht zur Verfügung steht. In unseren Analyserahmen dient diese Clusterbildung als Referenz, um beurteilen zu können, wie gut sich alternative, unter Rückgriff auf exogene Variablen abgeleitete Clusterbildungen eignen, regional differenzierte homogene, trennscharfe Cluster für die Investitionen je Anschluss zu bilden.

Der methodische Ansatz des k-means-Verfahrens wurde im weiteren Verlauf der Analyse dazu herangezogen, um der Frage der optimalen Anzahl von Clustern nachzugehen (Abschnitt 6.2). Dazu wurde bei der Anwendung univariater k-means-Verfahren zunächst an der Verwendung der Variablen „Investitionen je Anschluss“ festgehalten (Abschnitt 6.2.2). Auf Basis dieser idealtypischen Konstellation konnte für die in Deutschland betrachtete Grundgesamtheit von knapp 8000 Anschlussbereichen bereits eine Berücksichtigung von nur 8 Clustern mehr als 97% der Abweichungen erklären.

In einem weiteren Analyseschritt wurden unter Rückgriff auf die Erkenntnisse der Regressionsanalyse multivariate k-means Clusterverfahren durchgeführt und deren Ergebnisse denen der idealtypischen univariaten k-means Clusterung des durchschnittlichen Investitionswerts pro Anschluss gegenüber gestellt. Die für die multivariate Analyse verwendeten Einflussgrößen „Teilnehmerdichte (bebaute Gebiete)“, „Straßenlänge pro bebautes km²“, „Urbanisierungsgrad“ sowie „Streuung von Siedlungsflächen“ erlaubten jedoch keine befriedigenden Ergebnisse. Mit 10 Clustern konnten lediglich 82% der Abweichungen, mit 25 Clustern auch nur 89% der Abweichungen erklärt werden. Die Hoffnung, mittels beobachtbarer Strukturmerkmale in einer ausreichenden Granularität auf die Ähnlichkeit der Kosten eines FTTH-Ausbaus schließen zu können, konnte mit diesem Ansatz nicht erfüllt werden.

Ein Vergleich der Clusteransätze mit Blick auf die durchschnittlichen Investitionen pro Anschluss zeigt (Abschnitt 6.4), dass sowohl die Spreizung der Investition pro Anschluss zwischen Cluster 1 und Cluster 20 auf Basis der Clusterung des WIK-NGA-Modells als auch auf Basis der mit den multivariaten Verfahren abgeleiteten Clusterung hinter der Spreizung nach Maßgabe der Investitionen je Anschluss zurück bleibt. Darin kommt zum Ausdruck, dass das Problem der Gruppenüberlappungen (Anschlussbereiche mit denselben durchschnittlichen Investitionen je Anschluss fallen in verschiedene Cluster) – wie wir es bereits in der Ausgangs-Clusterung des WIK-NGA-Modells festgestellt haben – auch bei der multivariaten Clusterbildung Ursache für die stärkere Nivellierung der Investitionen pro Anschluss ist.

Um eine überschneidungsfreie Clusterbildung mit Blick auf die Zielgröße „durchschnittliche Investitionen pro Anschluss“ abzuleiten, sind letztlich weitere Untersuchungen erforderlich. Wir streben an, im Rahmen zukünftiger Forschungsprojekte unter Rückgriff auf die Erkenntnisse der Regressionsanalyse eine k-means Clusterung anhand der geschätzten Investitionen pro Anschluss durchzuführen. Dabei wollen wir uns der Koeffizienten des Regressionsmodells bedienen, um eine Gewichtung der Einflussfaktoren vornehmen und der Zielsetzung einer erwartungstreuen Investitionsverberechnung bei Reduzierung des Datenbedarfs und unter Berücksichtigung regionaler Unterschiede Rechnung tragen zu können.

Die Clusterbildung wurde im Rahmen dieses Forschungsprojekts auch unter einem weiteren Fokus analysiert, der auf die Nachbarschaftsbeziehung von Anschlussbereichen abstellt (Abschnitt 6.5) und somit im Kontext der Analyse und Bewertung von unternehmensstrategischen Investitionsentscheidungen oder breitbandpolitischen Ausbauförderung von Interesse ist. Letztlich zielt der Ausbau von Anschlussnetzen in der Regel immer auf größere, zusammenhängende Gebiete ab. Die im ersten Analyseteil (Abschnitte 6.1 bis 6.3) adressierten Aspekte der Clusterung fokussierten auf die möglichst überschneidungsfreie Bestimmung der Investitionen pro Anschluss anhand regional beobachtbarer Merkmale. Demgegenüber wurde bei diesem Untersuchungsansatz der Fokus darauf gelegt, herauszufinden, in welchem Maße NGA-Investitionen pro Anschluss erkennbar räumlich zusammenhängen.

Mit Hilfe der Nachbarschaftsanalyse (Hot Spot/Cold Spot Analyse), basierend auf der G_i^* -Statistik, konnten in einem ersten Analyseschritt räumlich korrelierte Cluster hoher und niedriger Investitionen je Anschluss aufgezeigt und statistisch abgesichert werden. Auch konnten die abgeleiteten Hot Spot/Cold Spot Flächen der Siedlungstypologie nach Kreisen (NUTS 3) des Bundesamtes für Bauwesen und Raumordnung gegenübergestellt werden, denen eine Untergliederung in 4 Kategorien zugrunde lag. Die visuelle Aufbereitung zeigte eine hohe Übereinstimmung der billigen Cold Spot Cluster mit den großstädtisch geprägten Kreisen und den städtischen Kreisen. Die ländlichen Kreise mit Verdichtung fallen häufig mit den insignifikanten Gebieten der Hot Spot / Cold Spot Analyse zusammen. Dünn besiedelte ländliche Regionen zeigten eine gute

Übereinstimmung mit den teuren Hot Spots unserer Analyse. Eine statistische Absicherung dieses optischen Eindrucks ist im Rahmen weiterer Untersuchungen angestrebt.

Ausblick

Mit dem hier vorgelegten Diskussionsbeitrag wurden verschiedene Facetten des NGA-Ausbaus in Deutschland betrachtet und die Heterogenität der durchschnittlichen Investitionen und Kosten je Anschluss herausgearbeitet. Der Versuch, mit statistischen Methoden den Erkenntnisgewinn einer aufwendigeren bottom-up Modellierung unter Rückgriff allein auf allgemein verfügbare, öffentliche Daten nachzubilden, blieb für die Gesamtheit der Anschlussbereiche in Deutschland hinter den Erwartungen zurück und konnte den statistischen Robustheitsmaßen nicht standhalten. Mit weiteren Forschungsarbeiten und einer Zusammenführung der Erkenntnisse aus der hier durchgeführten Regressions- und Clusteranalyse soll zukünftig versucht werden, die Schätzung regional heterogener Kosten von Anschlussnetzen weiter zu verbessern.

Die Analysen zur flächendeckenden Versorgung mit Breitbandanschlüssen haben wegen der unbefriedigenden Datenqualität nur begrenzt verwertbare Schlussfolgerungen zugelassen. Hier steht eine Aktualisierung der Datenbasis und der sich daraus ableitenden Erkenntnisse für nächste Forschungsarbeiten im Raum. Dabei streben wir an, dem Nachbarschaftskonzept im Kontext der Profitabilität eines privatwirtschaftlichen NGA-Ausbaus und ggf. bestehenden Förderbedarfs nachzugehen und den hier präsentierten Ergebnissen eines profitablen NGA-Ausbaus gegenüber zu stellen.

8 Literaturverzeichnis

- AZ-Direct Bertelsmann, AZ-Raster, TeleAtlas Straßenlayer (2018) GIS-Datenbeständen und der Verwaltungszuordnung des Bundesamtes für Kartographie und Geodäsie
- Braun, M., Wernick, C., Plückebaum, T., Ockenfels, M. (2020): Parallele Glasfaserausbauten auf Basis von Mitverlegung und Mitnutzung gemäß DigiNetzG als Möglichkeiten zur Schaffung von Infrastrukturwettbewerb, WIK-Diskussionsbeitrag Nr. 456, Januar 2020
- Bundesamt für Kartographie und Geodäsie Dokumentation Georeferenzierte Adressdaten GA, S.5; https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/ga.pdf
- Bundesministerium für Wirtschaft und Technologie (BMWi) (2013) (<http://selke.de/privates/hvt-standorte/>).
- Bundesministerium für Wirtschaft und Technologie DSL-Hauptverteilerliste; <http://selke.de/privates/hvt-standorte/>.
- Deutsche Energie-Agentur Gebäude Report Kompakt (2019), https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2019/dena-GEBAEUDEREPORT_KOMPAKT_2019.pdf
- Doose / Elixmann / Jay (2009): "'Breitband/Bandbreite für alle': Kosten und Finanzierung einer nationalen Infrastruktur". In diesem WIK-Diskussionsbeitrag wurden die Investitionen eines Ausbaus von FTTC und FTTH in Deutschland modelliert.
- Elixmann, D., Ilic, D., Neumann, K.-H., Plückebaum, T. (2008): The Economics of Next Generation Access, Study for the European Competitive Telecommunication Association (ECTA), 10 September 2008
- FTTH Council Europe (2017)
- Getis, A. / Ord, J.K. (1992): The Analysis of Spatial Association by Use of Distance Statistics, in Geographical Analysis, 24(3)
- IHS (2017): Broadband Coverage in Europe 2016: Coverage in Switzerland, 22 May 2017
- Ilic, D., Neumann, K.-H., Plückebaum, T. (2009): Co-investment and risk sharing in the context of multi fibre strategies; study for ECTA
- Ilic, D., Neumann, K.-H., Plückebaum, T. (2009a): Szenarien einer nationalen Glasfaserausbaustrategie in der Schweiz, Studie für das BAKOM, Dezember 2009
- Ilic, D.; Jay, S.; Plückebaum, T.; Stamm, P. (2013): Migrationsoptionen für Breitbandkabelnetze und ihr Investitionsbedarf, WIK-Diskussionsbeitrag Nr. 380
- Jay, S. / Plückebaum, T. / Ilic, D. (2010): "Der Einfluss von Next Generation Access auf die Kosten der Sprachterminierung". In diesem WIK-Diskussionsbeitrag wurden die Kosten der Terminierung in verschiedenen FTTx Architekturen untersucht.
- Jay, S., Neumann, K.-H., Plückebaum, T. (2011): Implikationen eines flächendeckenden Glasfaserausbaus und sein Subventionsbedarf, WIK Diskussionsbeitrag Nr. 359, Oktober 2011
- Jay, S., Neumann, K.-H., Vogelsang, I., et al. (2011): Architectures and competitive models in fibre networks, in: Vodafone, The Policy Paper Series, Nummer 13, June 2011, .

- Jay, S.; Plückebaum, T.; Ilic, D.; Stamm, P. (2013): Migrationsoptionen für Breitbandkabelnetze und ihr Investitionsbedarf, WIK Diskussionsbeitrag Nr. 380, Bad Honnef, September 2013
- Kroon, P.; Plückebaum, T.; Sanchez Gracia, J.; Sabeva, D.; Zoz, K. (2017): Study into current and future technological access options to all fixed telecommunications infrastructures in the Netherlands, den Haag, 21. June 2017, <https://www.acm.nl/nl/publicaties/publicatie/17463/Onderzoek-toegang-tot-vaste--telecommunicatienetwerken/>
- Kulenkampff, G., Plückebaum, T., Zoz, K. (2019): Analytisches Kostenmodell für das Anschlussnetz AKM-AN Version 3.0, WIK-Consult, Bad Honnef, 14.01.2019, Referenzdokument im Auftrag der Bundesnetzagentur https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Marktregulierung/Massstaebe_Methoden/Kostenmodelle/Anschlussnetz/20190114_AKM_AN_RefDokpdf.pdf?__blob=publicationFile&v=2
- Lucidi, S., Ockenfels, M., Sörries, B. (2020): Anhaltspunkte für die Replizierbarkeit von NGA-Anschlüssen im Rahmen des Art. 61 Abs. 3 EKEK, WIK Diskussionsbeitrag Nr. 460, März 2020
- Nett, L., Jay, S. (2014): Entwicklung dynamischer Marktszenarien und Wettbewerbskonstellationen zwischen Glasfasernetzen, Kupfernetzen und Kabelnetzen in Deutschland, WIK Diskussionsbeitrag Nr. 388, September 2014
- Neumann, K.-H. (2013): Der dynamische Investitionswettbewerb als Leitbild der Entwicklung des Telekommunikationsmarktes, Studie für den BREKO, November 2013
- Neumann, K.-H. / Plückebaum, T. / Zoz, K. (2017): Modellierung der Kosten eines flächendeckenden Hochbreitbandnetzes in der Schweiz, Studie für das Bundesamt für Kommunikation der Schweiz
- Ord, J.K. / Getis, A. (1995): Local Spatial Autocorrelation Statistics: Distributional Issues and an Application, in Geographical Analysis 27(4)
- Plückebaum, T., Eltges, F., Ockenfels, M.; BNetzA Studie (2019) Potentiell anzunehmende Vorleistungsprodukte in Kabelnetzen auf der Basis von DOCSIS, Studie im Auftrag der BNetzA, https://www.bundesnetzagentur.de/DE/Service-Funktionen/Beschlusskammern/1_GZ/BK1-GZ/2019/2019_0001bis0999/2019-0001bis0099/BK1-19-0001_WIK-Kabelgutachten_BA.pdf?__blob=publicationFile&v=2
- Plückebaum, T.; Jay, S.; Neumann, K.-H. (2014): Benefits and regulatory challenges of VDSL vectoring (and VULA), Florence School of Regulation, Communications Media 2014 Scientific Seminar, March 28 – 29, 2014, Florence, EUI Working Papers RSCAS 2014/69, <http://fsr.eui.eu/Publications/WORKINGPAPERS/ComsnMedia/2014/WP201469.aspx>
- Plückebaum, T., Eltges, F., Ockenfels, M. (2019): Potentiell anzunehmende Vorleistungsprodukte in Kabelnetzen auf der Basis von DOCSIS, Studie im Auftrag der BNetzA, Bad Honnef, 4. Februar 2019, https://www.bundesnetzagentur.de/DE/Service-Funktionen/Beschlusskammern/1_GZ/BK1-GZ/2019/2019_0001bis0999/2019-0001bis0099/BK1-19-0001_WIK-Kabelgutachten_BA.pdf?__blob=publicationFile&v=2
- Plückebaum, T.; Ockenfels, M. (2020): Kosten und andere Hemmnisse der Migration von Kupfer- auf Glasfasernetze, WIK-Diskussionsbeitrag Nr. 457, Bad Honnef, Februar 2020,

elektronisch verfügbar unter:

https://www.wik.org/uploads/media/WIK_Diskussionsbeitrag_Nr_457_01.pdf

Statistisches Bundesamt, Anzahl Betriebe (Unternehmensregister) nach Beschäftigtengrößenklassen und Wirtschaftszweigen für 2017 Stand 7.10.2020

Statistisches Bundesamt, Fachserie 5 Reihe 3 (2018), S.9.

Strube Martins, Wernick, Plückebaum, Henseler-Unger (2017): Die Privatkundennachfrage nach hochbitratigem Breitbandinternet im Jahr 2025, Bad Honnef, März 2017, <http://www.wik.org/index.php?id=871>

Swisscom (2017): FY 2016 analyst presentation, Analyst meeting in Zurich – 8 February 2017

Wooldridge (2012) Introductory Econometrics: a Modern Approach.

A1 Anhang 1: Datenquellen und Modellierungsannahmen zum Versorgungsgebiet

A1.1 Anschlussnachfrage

Für die Modellierung des flächendeckenden Ausbaus wurde eine geografische Verortung der Nachfrage in Form von Festnetzanschlüssen für ganz Deutschland vorgenommen. Dazu wurde auf umfangreiche GIS-Datensätze für Deutschland zurückgegriffen. Für die räumliche Lage der Nachfrage wurden hausnummerngenau vermessene Adresspunkte des Produktes „Tom Tom Multinet“ des GIS Anbieters Tom Tom verwendet. Weiterhin wurde die Anzahl der Haushalte in einer räumlichen Genauigkeit von 250 m x 250 m mithilfe des AZ Haushalts- und Gebäuderasters für Deutschland ermittelt. Durch die Kombination der beiden Datensätze wurde eine adressgenaue Verteilung der Haushalte für ganz Deutschland möglich. In einem ersten Schritt wurden dafür die Adressen ermittelt, die in eine Rasterzelle des AZ-Rasters fallen, um dann in einem weiteren Schritt das Haushaltspotential der betreffenden Rasterzelle auf diese Adresspunkte zu verteilen. Zur Erfassung von Nichtwohngebäuden und Unternehmensstandorten wurde jedem Adresspunkt mindestens eine Nachfrageeinheit zugewiesen. Die Verteilung der Haushalte auf die Adresspunkte erfolgte mit einer dafür entwickelten Verteilungsroutine.

A1.1.1 Verortung der Nachfrage

Bei den verwendeten Adresspunkten handelt es sich um einen Auszug der im Tom Tom Multinet Datenbestand enthaltenen hausnummerngenauen Verortung tatsächlicher Straßenadressen zu physischen Gebäuden oder Grundstücken für ganz Deutschland. Verwendet wurde ein Auszug mit Releasestand für das Jahr 2018 mit 22.726.396 Adresspunkten (Koordinaten).

Diese Adresspunkte dienen als Grundlage zur Verortung der Nachfrage. Wichtig ist es, die richtige Lage der Nachfrage zu erfassen, weil die Lage des Gebäudes die Netzlänge maßgeblich bestimmt. Der Adresspunktdatensatz ist hierfür besonders geeignet, weil er alle Adressen in ihrer räumlichen Lage erfasst, insbesondere auch mit Blick auf Unternehmensstandorte außerhalb von Wohngebieten.

Der Vergleich mit den uns vorliegenden statistischen Informationen zum Gebäudebestand stellt sich wie folgt dar:

Das Statistische Bundesamt weist einen Wohngebäudebestand für Deutschland in Höhe von 19.053.216 für 31.12.2018 aus⁷⁹. Die Deutsche Energie Agentur DENA berich-

⁷⁹ Statistisches Bundesamt, Fachserie 5 Reihe 3 (2018), S.9.

tet von 21,7 Mio. Wohn- und Nichtwohngebäuden⁸⁰ und beziffert die Nichtwohngebäude auf ca. 2,7 Mio.⁸¹ Das in der Datenaufbereitung ebenfalls verwendete AZ Gebäude-raster beruht nach Produktbeschreibung auf den Daten der Bertelsmann-Gebäude-Datei (BGD) mit fast allen Gebäuden in Deutschland. Darin werden Informationen sowohl zu den Gebäuden als auch zusätzlich über Personen und Haushalte gespeichert. Im Raster sind nach unseren Auswertungen 21.674.571 Gebäude erfasst, die sich auf 15.395.962 reine Wohngebäude, 5.561.665 gemischt genutzte Gebäude und 689.944 reine Firmengebäude aufteilen. Diese Aufteilung wird über einen Abgleich der bei AZ-Direct vorhandenen Haushaltsadressdatenbank mit der Unternehmensdatenbank erreicht, indem für eine Gebäudeadresse ermittelt wird, ob diese in jeweils nur einer der beiden Datenbanken oder eben in beiden Datenbanken vorhanden ist.

Die georeferenzierten Adressdaten des Bundesamtes für Kartographie und Geodäsie umfassen laut Angaben in deren Dokumentation etwa 22,6 Mio. Datensätze⁸². Dieser basiert auf dem Datensatz der Amtlichen Hauskoordinaten Deutschlands (HK-DE), der Landesvermessungsämter und berücksichtigt nur Hauptgebäude mit Hausnummer. Der TOMTOM Adressdatensatz liegt mit 22,7 Mio. Adresskoordinaten somit sehr nahe am amtlichen Bestand an Gebäudeadressen.

Die Anzahl Gebäude des AZ-Haushalts- und Gebäuderasters stimmt sehr gut mit den von DENA gemeldeten 21,7 Mio. Wohn- und Nichtwohngebäuden überein. Weist aber eine Untererfassung von rund 1 Mio. Adressen gegenüber dem TOMTOM Datensatz und den Adressdaten des Bundes auf.

Unterschiede in den Datenbeständen können auf Unterschiede in der zeitlichen Erfassung der Daten (Updatestand und Nachführungszeiträume) zurückgehen oder in unterschiedlichen Erfassungsstrategien (Zweck des Datenbestands, Herkunft der Daten) begründet sein. Wichtig für unser Anliegen war es Haushalts- und Firmenadressen im Adressbestand in ihrer Lage zu berücksichtigen. Dazu sind die Adressdaten des Bundes und der TOMTOM Adresspunktdatensatz hervorragend geeignet. Wir haben die AdressPoints von TOMTOM verwendet.

A1.1.2 Höhe des Nachfragepotentials

Die Höhe des Nachfragepotentials wurde über eine Anreicherung dieser Punktinformationen mit Informationen des AZ-Haushalts- und Gebäuderasters erzielt. Das verwendete Raster weist ebenfalls den Release-Stand von 2018 auf. Das verwendete Raster mit Gitterweite von 250 m X 250 m enthält Informationen zur Anzahl Gebäude, der Anzahl Haushalte und der Anzahl Personen je Rasterzelle, sowie Angaben zur Gebäudetypologie. Für ganz Deutschland liegen 1.214.491 Rasterzellen vor, die mindestens ein Ge-

⁸⁰ DENA Gebäude Report Kompakt (2019), S.6;

⁸¹ DENA Gebäude Report Kompakt (2019), S.14

⁸² Dokumentation Georeferenzierte Adressdaten GA, S.5

bäude beinhalten. Für ganz Deutschland sind in dem Rasterdatensatz folgende Potentiale erfasst:

Anzahl Gebäude	Anzahl Haushalte	Anzahl Personen
21.647.571	41.684.651	81.209.177

Das Statistisch Bundesamt beziffert die Anzahl Haushalte für 2018 mit 41,378 Mio. was rund 300Tsd weniger Haushalte sind als im AZ-Raster erfasst werden. Eine Alternativgröße für die Abschätzung der Privatkundennachfrage für Festnetzanschlüsse besteht in der Zahl der Wohnungen in Wohn- und Nichtwohngebäuden, die das Statistische Bundesamt mit 42.235.402 für das Jahr 2018 beziffert⁸³. Diese Zahl übersteigt die Zahl der Haushalte um knapp eine Mio. Da es unwahrscheinlich ist, dass jede Wohnung auch einen Festnetzanschluss besitzt, und die Haushaltsinformationen in einer räumlich sehr viel disaggregierteren Form vorliegt, benutzen wir die Anzahl Haushalte je Raster bei der Anschlussallokation.

Zur Abschätzung der Geschäftskundennachfrage käme die Anzahl der Betriebe in Deutschland in Betracht, für die das Statistische Bundesamt eine Anzahl von 3.761.534⁸⁴ für 2017 angibt (Anzahl Unternehmen: 3.481.860). Die Zahl der Betriebe aus offiziellen Quellen liegt allerdings lediglich hoch aggregiert auf Kreis- oder Regierungsbezirksebene vor, so dass allenfalls eine Zufallsverteilung der Anzahl Betriebe auf die Adresspunkte für solche Aggregate in Betracht käme. Wie aus den AZ-Gebäude und Haushaltsregisterdaten hervorgeht, werden die betrieblich genutzten Gebäude zum Großteil gemischt auch als Wohngebäude genutzt (rd. 5,5 Mio. gemischte Nutzung zu rd. 0,7 Mio. reine Firmennutzung). Wie sich diese rd. 6.2 Mio. Gebäude mit Unternehmensadressen zu den 3,76 Mio. Betrieben des Unternehmensregisters verhalten ist ebenso unklar. Darüber hinaus handelt es sich bei den meisten Betrieben (3,296 Mio.) um Betriebe mit 0-9 Beschäftigten. Eine Vielzahl dieser Betriebe wird in Gebäuden mit gemischter Nutzung verortet sein. Diese sind aber über die Haushaltsnachfrage bereits erfasst. Die Gebäude in reiner Firmennutzung sind über die Adressen dem Standort nach erfasst. Eine Zufallsverteilung über alle Adressen würde hier die Genauigkeit der Verortung des Potentials nicht erhöhen, da die Adresspunkte der Mischnutzung und der reinen Firmennutzung bereits berücksichtigt sind. Das Problem der genaueren Verortung von Unternehmen könnte allenfalls durch den Zukauf von Unternehmensadressdaten verbessert werden. Diese Mittel standen im Rahmen dieser Studie jedoch nicht zur Verfügung.

Verwendet man die Anzahl Haushalte, um das Nachfragepotential für private Festnetzanschlüsse abzuschätzen und die Anzahl Betriebe für die Geschäftskundenanschlüsse,

⁸³ Statistisches Bundesamt, Fachserie 5 Reihe 3 (2018), S.6

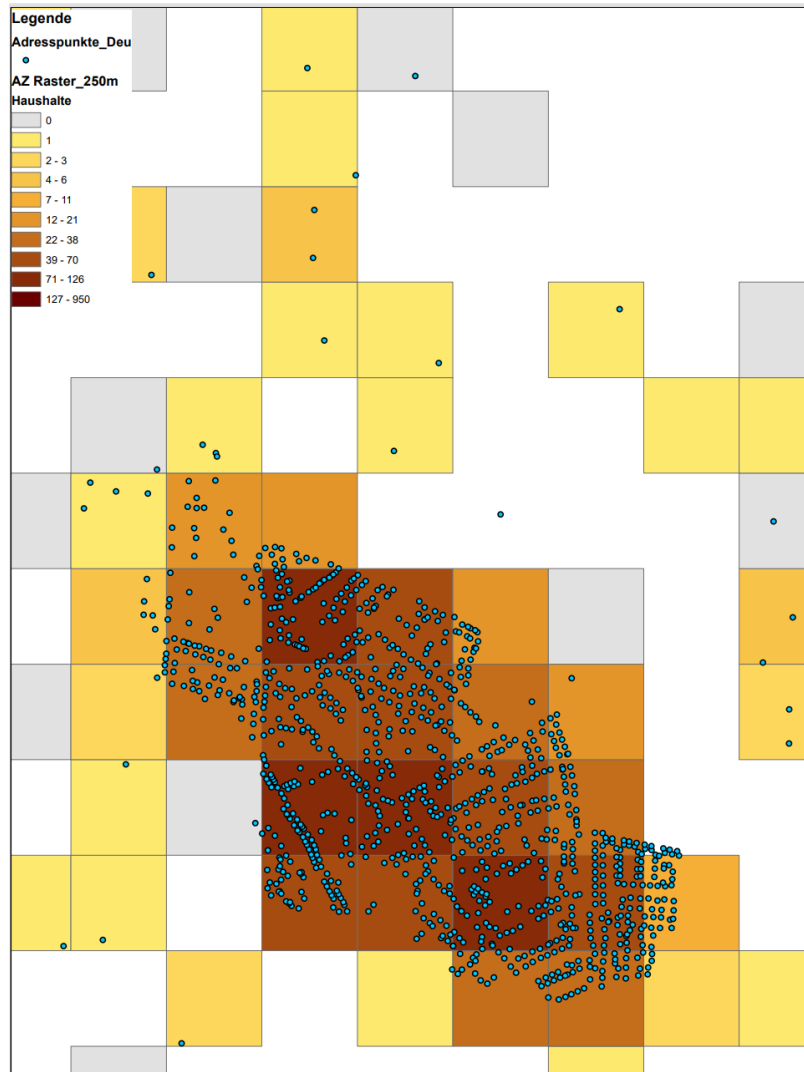
⁸⁴ Statistisches Bundesamt(2020)

dann ergibt sich ein Nachfragepotential von 45,24 Mio. Dieses erscheint uns jedoch vor dem Hintergrund von aktuell gut 35 Mio. Festnetzanschlüssen als zu hoch. Unser Vorgehen ist es deshalb, allen Adresspunkten eine Nachfrage zuzuordnen und in einem zweiten Schritt die Anzahl Haushalte eines Rasters auf die Adresspunkte innerhalb eines Rasters zu verteilen. Dazu wird vor der Zufallsverteilung die Anzahl Adresspunkte je Raster abgezogen und nur der verbleibende Rest zufallsverteilt.

Um die Potentialallokation auf die Adresspunkte durchzuführen war es notwendig den Adresspunktdatensatz mit dem AZ-Raster zu verschneiden und anschließend die Zufallsverteilung durchzuführen. Abbildung A 1-1 zeigt einen Ausschnitt mit der Überlagerung beider Datensätze. Die Kacheln des Rasterdatensatzes sind nach der Anzahl Haushalte eingefärbt. Überlagernd ist die Lage der Adresspunkte als hellblaue Punkte dargestellt.

Wie ersichtlich, gibt es Kacheln des Rasters, die Haushalte anzeigen (Färbung gelb bis braun) aber keine Adresspunkte aufweisen (Haushaltspotential von 36719), dieses Nachfragepotential geht bei der Nachfrageverteilung verloren. Umgekehrt gibt es Adresspunkte, denen keine Kacheln zugeordnet werden können, weil die Adresspunkte außerhalb der Kacheln liegen (die räumliche Schnittmenge ist leer). Diese Punkte werden bei der Nachfrageverteilung nicht berücksichtigt (das betrifft 80139 Adresspunkte). Weiterhin gibt es Adresspunkte auf grauen Kacheln, was den Fall markiert, dass die AZ-Rasterzelle mindestens ein Gebäude aufweist aber keine Haushalte und Personen. Diese Adresspunkte werden bei der Nachfrageverteilung mit einer Nachfrage versehen.

Abbildung A 1-1: Ausschnitt des AZ Haushalts- und Gebäuderaster und der Adresspunkte



Quelle: WIK

Das Vorgehen der Verteilung ist nun so, dass allen Adresspunkten die in eine Rasterzelle fallen eine Nachfrage zugeordnet wird und falls die Anzahl Haushalte in der Zelle die Anzahl Adresspunkte übersteigt wird dieses Restpotential auf die Adressen der Rasterzelle zufallsverteilt.

Das Ergebnis der Nachfrageverteilung stellt sich wie folgt dar:

Anzahl Adresspunkte mit Nachfrage	Allozierte Nachfrage
22.646.257	42.818.329

Im weiteren Verlauf der Datenaufbereitung werden die Adresspunkte mit Nachfragepotential über eine Nearest-Neighbour-Suche der räumlich nächstliegenden Straße zugeordnet und der Lotpunkt auf der Straße ermittelt sowie die Luftlinienentfernung zwischen Lotpunkt und Adresspunkt. Die Straßensegmente und damit implizit die Gebädepunkte werden zuvor berechneten Anschlussbereichen zugeordnet, damit eine Netzoptimierung je Anschlussbereich möglich wird.

Das Ergebnis der Gesamtallokation der Nachfrage auf die Straßensegmente und die Zuordnung der Straßensegmente zu den Anschlussbereichen stellt sich wie folgt dar:

Anzahl Anschlussbereiche	Anzahl Adresspunkte mit Nachfragepotential	Alloziiertes Nachfragepotential
7871	22.642.265	42.926.863

A1.2 Erzeugung der Datenbasis für die anschlussbereichsbezogene, GIS-basierte bottom-up Modellierung

A1.2.1 Berechnung der ASB-Grenzen unter Verwendung eines öffentlichen HVT-Datensatzes

Die Abgrenzung der Anschlussbereiche beruht auf einem modified scorched node Ansatz, bei dem die Gebietsabgrenzungen um vorzugebende Netzaggregationspunkte herum gebildet werden. Die vorgegebenen Standorte dienen als Zentren der Versorgungsgebiete, die nach Routinglängen auf Basis des Straßennetzes gebildet werden. Die dabei entstehenden Grenzen der Versorgungsgebiete haben die Eigenschaft gleich weit von den nächstgelegenen Zentren entfernt zu sein, oder eine vorgegebene Maximalentfernung nicht zu überschreiten, für die wir 30 km angesetzt haben.

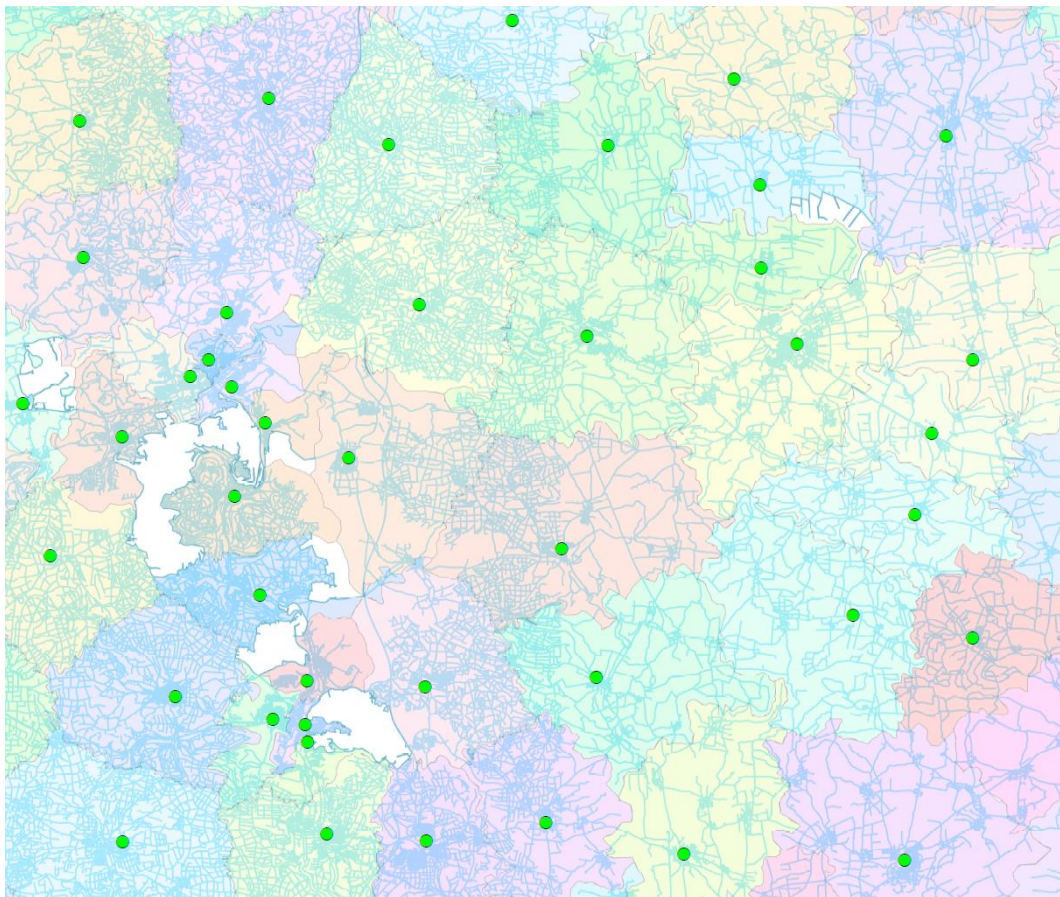
Als Zentren für die Abgrenzung dienen uns die Standorte einer im Internet veröffentlichten Liste der DSL-Hauptverteiler der Telekom. Die Liste enthält Angaben zu 7971 Hauptverteilern, die mit Adressen und geografischen Koordinaten aufgeführt sind.⁸⁵ Aufgrund von Adressdopplungen enthält die Liste lediglich 7871 verwertbare Standorte, für die eine Gebietsabgrenzung auf Grundlage des deutschen Straßennetzes vorgenommen wurde. Diese HVt Standorte sind als "Scorched Nodes" in die Modellierung eingegangen. Grundsätzlich sind wir davon ausgegangen, dass die existierenden HVt-

⁸⁵ Bundesministerium für Wirtschaft und Technologie (BMWi) (2013). Den Angaben der Fundstelle (Selke) entsprechend stammen die Daten aus einem „pdf-Dokument von der Webseite des Bundesministerium für Wirtschaft und Technologie (BMWi) und sind von vielen fleißigen Helfern korrigiert worden“ (<http://selke.de/privates/hvt-standorte/>). Auf der Website ist ein Veröffentlichungsdatum vom 13.4.2013 angegeben.

Standorte in Deutschland erhalten bleiben und als MPoP (Metropolitan Point of Presence) fungieren. Dies entspricht durchaus auch den Glasfaserausbaustrategien großer europäischer Telefongesellschaften, weil an den bisherigen HVt-Standorten die bereits bestehenden Backbone Netze enden. Je mehr ein MPoP auch Funktionen höherer Netzebenen ausübt, desto weniger ist er „verschiebbar“.

In Abbildung A 1-2 ist das Ergebnis der Anschlussbereichsabgrenzung für einen Ausschnitt dargestellt. Die grünen Punkte stellen die vorgegebenen Zentren der Versorgungsgebiete dar (MPoP), blau eingezeichnet ist das Straßennetz, auf dem die Abgrenzung nach Routinglänge zum Zentrum hin basiert. Die entstehenden Versorgungsgebiete sind farblich voneinander abgegrenzt, mit schwarzen Umrisslinien. Wie man an den weißen Gebieten erkennen kann sind Gebiete ohne Straßen nicht in der Gebietsabgrenzung enthalten, weil hier aufgrund fehlender Straßen kein Routing möglich ist.

Abbildung A 1-2: Anschlussbereichsabgrenzung Ausschnitt



Für die Erstellung der Anschlussbereichsabgrenzung haben wir das Tool „Einzugsgebiete erstellen“ aus der Toolbox Network Analyst der GIS Software ArcGIS verwendet.

A1.2.2 Ermittlung der Lotpunkte und Abstand der Gebäude

Ebenfalls in ArcGIS erfolgt eine Zuordnung der Adresspunkte zu den Straßenabschnitten des Straßenlayers. Diese Zuordnung basiert darauf, für jeden Adresspunkt über eine lokal eingegrenzte Suche das nächstgelegene Straßensegment zu ermitteln, den Lotfußpunkt auf dem Straßensegment zu erzeugen und den Abstand zwischen Lotfußpunkt und Adresskoordinate für jeden Adresspunkt zu speichern. Auf diese Weise werden die Längen der sogenannten Hausstiche oder Stichleitungen (Anschlussleitung zum Gebäude, i.d.R. auf privatem Grund) bestimmt.

Abbildung A 1-3: Zuordnung der Adresspunkte zu den Straßensegmenten



Quelle: WIK

Abbildung A 1-3 illustriert das Vorgehen zur Allokation von Adresspunkten und Straßensegmenten an einem Ausschnitt. Wiedergegeben sind die Straßensegmente (grüne Linien), die Adresspunkte (hellblaue Punkte), die bei der Zuordnung erstellten Lotfußpunkte auf den zugeordneten Straßensegmenten (rote Punkte) und die Abstandslinie

zwischen Adresspunkt und zugehörigem Lotfußpunkt (blaue Linien). Es wird ersichtlich, dass eine Zuordnung zum nach Luftlinie nächstgelegenen Straßensegment erfolgt, auch wenn diese Zuordnung nicht in jedem Fall rechtwinklig erfolgen kann.

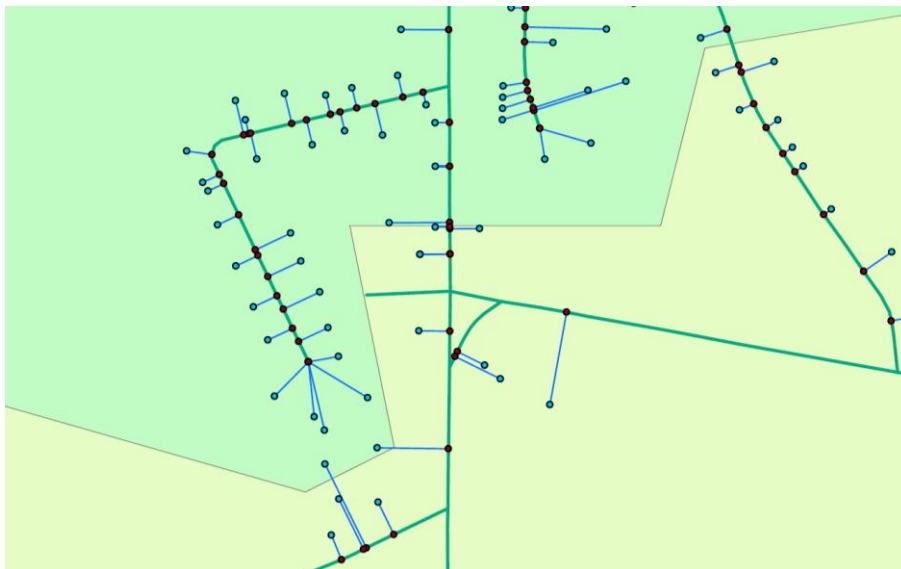
Wichtig für die Weiterverarbeitung sind die Allokationsbeziehungen zwischen Straßensegment-ID und Adresspunkt-ID, die Lotfußpunkte und die Abstände zwischen Lotfußpunkt und Adresspunkt, die die Basis für die Modellierung der Hauszuführung bilden. Die Zuordnung erfolgte mit Hilfe von Werkzeugen der Toolbox „Linear Referencing“ der GIS Software ArcGis.

A1.2.3 Verschnitt der Lotpunkte mit den ASB-Grenzen

Die Zuordnung der Adresspunkte zu den Anschlussbereichen erfolgt über die Zuordnung der Lotfußpunkte zu den Anschlussbereichen.

Abbildung A 1-4 zeigt die Lage der Adresspunkte, Straßensegmente und Lotpunkte an der Grenze zweier Anschlussbereiche. Da die Anschlussbereiche wie oben ausgeführt anhand der Straßensegmente gebildet werden, ist die Zugehörigkeit eines Straßenabschnitts zu einem ASB in der Regel eindeutig. Das gleiche gilt für die Lotfußpunkte die eine Projektion der Adresspunkte auf das nächstgelegene Straßensegment darstellen. Ein räumlicher Verschnitt der Lotfußpunkte mit den Anschlussbereichsflächen stellt deshalb eine Zuordnung von Adresspunkten zum Anschlussbereich des zugehörigen Straßensegments sicher, über den später das Netzrouting erfolgen soll. Eine Zuordnung über den Verschnitt der Polygonflächen mit den Adresspunkten würde in einigen Fällen dazu führen, dass die Anbindung der Adresspunkte beim Routing nicht sichergestellt wäre. In der Abbildung betrifft dies die beiden Adresspunkte links unten, die bei einem Verschnitt dem oberen Polygon zugeordnet würden, deren Straßenanbindung aber über die Straßensegmente des unteren Anschlussbereichs erfolgen.

Abbildung A 1-4: Zuordnung der Adresspunkte zu den Anschlussbereichen



Quelle: WIK

wik 

A1.2.4 Datenbank: Erzeugung der Inputdaten für das TAL Netzplanungsmodul

Die beschriebenen Aufbereitungsschritte dienen zur Erstellung der geografischen Inputdaten, die dem Netzplanungstool in Form einer SQL-Datenbank bereitgestellt werden. Mit dem Netzplanungstool wird das Mengengerüst für die unterschiedlichen Netzausbaustrategien berechnet. Der Output des Netzplanungstools besteht in den pro Anschlussbereich und nach Netzsegment aggregierten Mengengrößen aller Netzelemente wie Kabel, Trassen und Schnittstellen sowie den aggregierten Aufteilungsfaktoren für die dienstespezifischen Zurechnungen.

Die Berechnungen erfolgen pro Anschlussbereich, wofür die räumlichen und mengenmäßigen Beziehungen in der Datenbank hinterlegt werden und einer Netzberechnung zugänglich abgelegt sind.

Die in Abschnitt 3.1 dargestellte Tabelle 3-1 fasst die von uns für die Geodatenaufbereitung verwendeten Daten und deren Herkunft zusammen.

A2 Anhang 2: Modellierungsrahmen

Für die Investitionswertberechnung wurde im Rahmen dieses Forschungsprojekts sowohl auf das analytischen Kostenmodells für das Anschlussnetz 3.0 (AKM-AN 3.0) als auch das WIK-NGA-Modell zurückgegriffen.

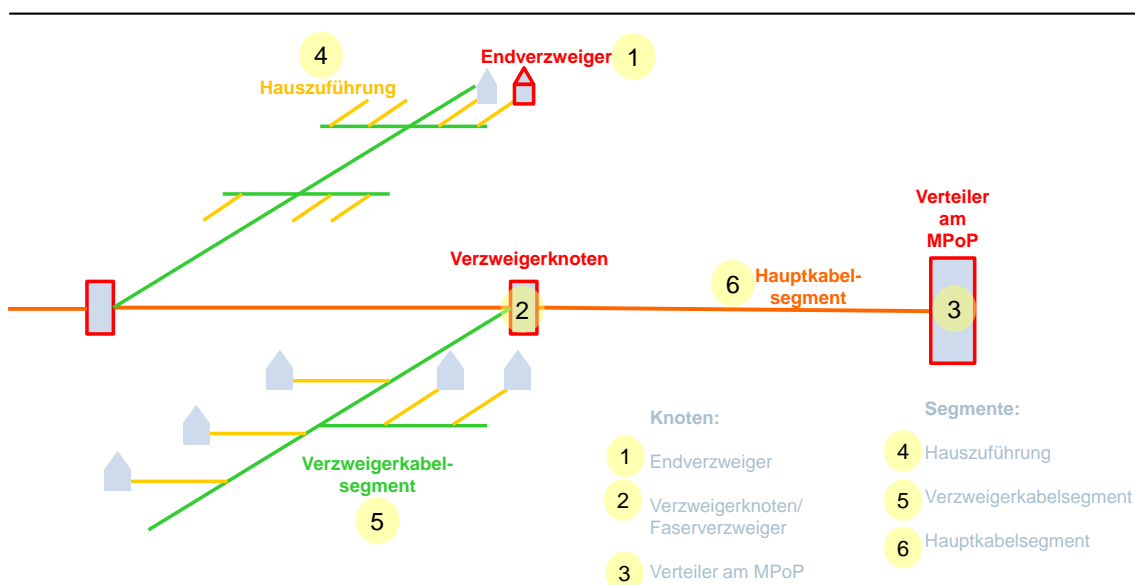
Ersteres bildet die Grundlage für die GIS basierte Ableitung des Mengengerüsts, Letzteres findet Anwendung für die Kosten- und Profitabilitätsberechnung auf Anschlussbereichsebene. Die methodische Vorgehensweise für alle drei Berechnungsschritte ist in den nachfolgenden Unterkapiteln skizziert.

A2.1 Bestimmung des Mengengerüsts (WIK-TAL-Modell 3.0)

Die abgeleiteten und georeferenzierten Nachfragedaten, ihre Verortung auf dem Straßenlayer und die beschriebene Methodik des modified scorched node Ansatzes bilden die Ausgangsbasis für die Berechnungen mit dem vom WIK erstellten analytischen Kostenmodell für das Anschlussnetz 3.0 (AKM-AN 3.0).

Mit einem zweistufigen Verfahren werden minimale Investitionen für das Anschlussnetz ermittelt. Im ersten Schritt wird eine Verzweigerbereichsbildung vollzogen, die den Anschlussbereich in mehrere kleine Versorgungsgebiete zergliedert und eine erste Bündelung der Anschlussnachfrage vornimmt, bevor diese gebündelt – über das Hauptkabel – zum MPoP geführt wird. Die Verzweigerbereichsbildung erfolgt dergestalt, dass ausgehend vom MPoP ein erster Verzweigerbereich, das sogenannte „starre Netz“, unter Rückgriff auf einen Radius (Modelleingangsparameter) gebildet wird, so dass diese Adresspunkte einem Verzweigerbereich zugeordnet werden, dessen Aggregationspunkt sich im MPoP befindet. Ausgehend vom zum MPoP nächstgelegenen Straßensegment vollzieht sich die weitere Verzweigerbereichsbildung unter Einhaltung von Erstreckungs- und Anschlusszahlrestriktionen.

Abbildung A 2-1: Generische Struktur eines Anschlussnetzes



Quelle: Kulenkampff, G., Plückebaum, T., Zoz, K. (2019)

Sind die Verzweigerbereiche eines Anschlussbereichs abschließend gebildet und der Standort des Kabelverzweigers (KVz) bzw. Faserverzweigers (FVz) festgelegt, beginnt mit der Auslegung der Verzweigerbereichstrassen die zweite Stufe der algorithmischen Netzbestimmung. Unter Anwendung eines shortest-path-Algorithmus werden ausgehend vom KVz/FVz die Endverzweiger mit ihrer zugehörigen Anschlussnachfrage an das Netz angebunden. Ist die Auslegung des Verzweigerkabels abgeschlossen, wird im Modell die Anbindung der KVz/FVz an den MPoP vollzogen. Dies geschieht ebenfalls unter Anwendung eines shortest-path Algorithmus.

Als kostentreibende Netzelemente werden dabei die folgenden berücksichtigt:

- Trassen und Kabel unter Berücksichtigung verschiedener Verlegearten (mit zugehörigen Kabelverbindungen (Muffen mit Gruben und Schächten) sowie
- Schalt- und Abschlusseinrichtungen (EVz, HVt/ODF am MPoP sowie Verzweigerknoten (KVz/FVz).

Die Erfassung von Größen- und Verbundvorteilen bei Netzdimensionierung und Kostenzurechnung erfolgt anhand von Beilauffaktoren (Modellinput).

Die Modellversion 3.0 stellt auf eine Kostenberechnung glasfaserbasierter Anschlussnetze ab und ist in der Lage, verschiedene Architekturen und auch Technologien abzubilden. Dem Modell liegt daher ein generischer Ansatz zu Grunde. Die abbildbaren

NGA-Technologien und –Architekturen umfassen sowohl hybride Kupfer/Glasfaser- als auch reine Glasfaseranschlussnetze:

- (1) FTTH P2P Ethernet
- (2) FTTB (G.fast)
- (3) FTTH PON P2MP
- (4) FTTC
- (5) FTTE_x

Der Technologiebezug entsteht durch

- die Berücksichtigung technologiespezifischen Equipments, welches wiederum sowohl eine individuelle Konfektionierung als auch eine entsprechende Preisparametrisierung mit sich bringt⁸⁶ sowie
- die technologieabhängige Aggregation von Anschlusslinien auf eine reduzierte Ader-/Fasernachfrage.

Spezifika der Verlegung von glasfaserbasierten Anschlussnetzen kommen in der Auslegung der Glasfaserkabel unter weitestgehender Vermeidung von Spleißstellen durch Verlegung individueller Kabel pro EVz zum Ausdruck (Verzweigerkabelsegment). Somit kommt dem Faserverzweiger die Rolle einer Abzweigmuffe zu.

Damit in Verbindung stehen auch die für Glasfaserkabel relevanten Verlegearten. Im Unterschied zu Kupferkabeln, die entweder als Erd- oder Röhrenkabel verlegt werden, finden bei der Verlegung von Glasfaserkabel Mikrorohre Anwendung. Diese werden entweder als erdverlegbare Mikrorohrverbände in den Gräben eingebracht oder in der Variante Mikrorohr-in-Rohr verlegt, beide sind Bestandteil des Modells.

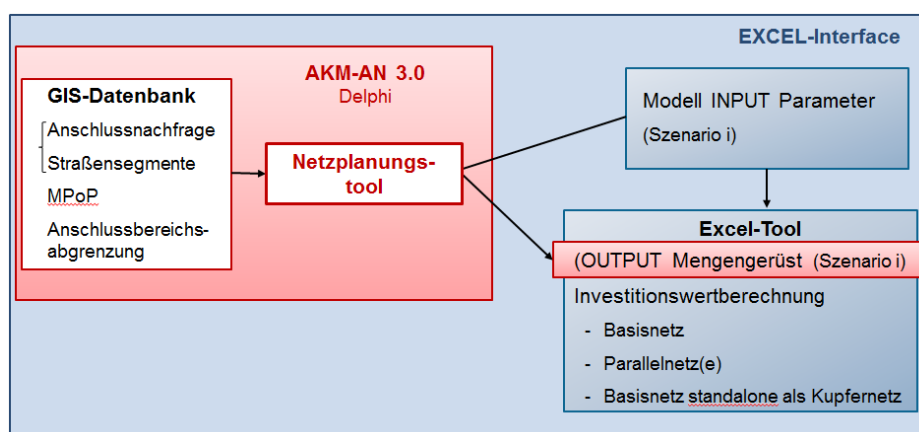
Als intermediäres Ergebnis produziert das Modell ein Mengengerüst, welches für eine Bewertung in Excel und die Anwendung von Preis- und Strukturparametern zugänglich gemacht wird. In der vorliegenden Modellversion 3.0 wird dabei die Netzplanung in ein Delphi programmiertes Planungs- und Optimierungswerkzeug und ein in Excel-VBA programmiertes Tool getrennt, das als Nutzerschnittstelle und Bedienoberfläche zur Eingabe und Definition grundlegender Parameter und zur Berechnung und Ausgabe der dienstebezogenen Ergebnisse des Modells dient.

Abbildung A 2-2 zeigt schematisch die Trennung der Modellierungsaufgaben in Netzplanung und Investitionswertberechnung sowie das Zusammenspiel der Modellkompo-

⁸⁶ Technologiespezifisches Equipment ergänzt an verschiedenen Knotenpunkten des Netzes die oben aufgelisteten Netzelemente. Dazu zählen: Splitter (PON-Technologien), DSLAM/ONU (G.fast) und MSAN inkl. Multifunktionsgehäuse (MFG; VDSL Vectoring)
Für den MSAN und MFG erfolgt keine Investitionswertberechnung, sondern lediglich eine Betrachtung des Mengengerüsts. Diese Komponenten werden entweder im NGN-Modell für den MSAN (WIK, Analytisches Kostenmodell für das Breitbandnetz) oder im Rahmen von Entgeltregulierungsverfahren von der Beschlusskammer für das MFG festgelegt.

nenten. Das Excel-Tool dient als grafische Benutzerschnittstelle zur Verwaltung der Inputparameter, zur Szenario-Erzeugung, zur Erzeugung der Datenübergabe für das Delphi-Netzplanungs-Werkzeug und zur Investitionsberechnung.⁸⁷

Abbildung A 2-2: Modellschema: Separierung Netzplanung und Bewertung



Quelle: Kulenkampff, G., Plückebaum, T., Zoz, K. (2019)

Der Output des Netzplanungstools besteht in den pro Anschlussbereich und nach Netzsegment aggregierten Mengengrößen aller Netzelemente wie Kabel, Trassen und Schnittstellen sowie den aggregierten Aufteilungsfaktoren für die dienstespezifischen Zurechnungen. Diese Mengenaggregate werden in ein Excel Arbeitsblatt ausgegeben und bilden den Ausgangspunkt für die Investitionsrechnung, die Mengengerüstaussgabe und die dienstespezifischen Ausgaben des Excelmodells.

Das TAL-Excel-Tool erlaubt eine Investitionswertberechnung. Auf diese Funktionalität wurde im Rahmen dieser Studie jedoch nicht zurückgegriffen. Vielmehr wurden die Zwischenergebnisse des Netzplanungstools inklusive gewichteter Investitionsparameter für eine Übergabe an das WIK-NGA-Modell aufbereitet und dort die weitere Investitionswert- und Kostenberechnung vollzogen.

⁸⁷ In der Modellsoftware Version 3.0 erfolgt die Erfassung der Inputs sowie Änderungen einzelner Parameter direkt in Excel. Dazu ruft der Nutzer das Excel Tool auf (z.B.: Szenario 1.xlsm) und gelangt zur programmierten Benutzerschnittstelle (GUI) des Szenarios (bzw. Excel-Tools). Das Netzplanungstool wird mit geografische Daten und netztechnischen sowie ökonomischen Inputparametern gespeist. Die geografischen Daten werden über eine SQL-Datenbank dem Modell zur Verfügung gestellt. Die netztechnischen und ökonomischen Inputparameter werden vom Nutzer über die programmierte Nutzerschnittstelle des Exceltools erfasst und dem Netzplanungstool als Excel Arbeitsblatt bereitgestellt.

A2.2 Bestimmung der Investitionswerte und Kosten (WIK-NGA-Modell)

Das WIK-NGA-Modell ermöglicht die Modellierung landesweiter Netze, wobei der Fokus auf den Kosten der Anschlussbereiche unter Verwendung von z.B. Anschlusszahlen, Trassenlängen, Investitions- und Kostenparametern usw. vollzogen wird. Komplettiert wird das Modell um eine vereinfachte Kostenbestimmung für die upstream Kosten. Es werden nicht alleine die benötigten Investitionen ermittelt, sondern auch die daraus resultierenden monatlichen Kosten, welche den monatlichen Erlösen gegenübergestellt eine Aussage über die Profitabilität von Anschlussbereichen zulassen. Die Modellergebnisse werden als Kosten pro Cluster ausgegeben. Darüber erfolgt im WIK-NGA-Modell der Ausweis von nicht profitablen Clustern sowie des benötigten Subventionsbedarfs.

A2.2.1 Anschlussbereichsgenaues Mengengerüst

Die Kostenmodellierung für sämtliche Anschlussbereiche basiert auf einem Mengengerüst, welches anschlussbereichsgenau mit dem TAL-Modell generiert wurde. Dazu wurden Trassen-, Kabel- und Verzweiger-bezogenen Daten aus dem TAL Modell für jeden Anschlussbereich verdichtet und an das NGA-Modell übergeben, welches die weitere Kostenberechnung vollzieht. Die Daten werden dabei dergestalt transformiert, dass im WIK-NGA-Modell mit einem einheitlichen Trassenpreis (pro Meter) je Anschlussbereich vorgesehen gearbeitet wird, der die Trassenstruktur des jeweiligen Anschlussbereichs reflektiert (verdichtete Ergebnisse des TAL-Modells). Gleiches gilt für die Kostenberechnung für Kabel- und Verzweigerkosten.

Das Verzweigerkabel besteht bei den Architekturen FTTC und FTTS aus den noch bestehenden Kupferdoppeladern und wird im Modell durch die entsprechende regulierte Sublooptmiete (FTTC) bzw. eines angemessenen Bruchteils davon (2/3) bei FTTS berücksichtigt.

Bei der Netzauslegung wird davon ausgegangen, dass ein Netzbetreiber einen Anschlussbereich zu 100% der adressierbaren Kunden ausbaut, denn jeder von ihnen könnte im Prinzip als Kunde gewonnen werden und sein Anschluss soll nicht durch lang anhaltende Baumaßnahmen verzögert werden (100% homes passed).

A2.2.2 Variable Kosten je Kunde

Im WIK-NGA-Modell werden auch variable Kosten berücksichtigt, die für die Anschaltung der einzelnen Kunden anfallen und für die Profitabilitätsanalyse zu berücksichtigen sind. Darunter fallen folgende Bestandteile: Der Netzbetreiber stellt ausschließlich für realisierte und angeschlossene Kunden aktives Equipment zur Verfügung (z.B. den Teilnehmerport im Ethernetswitch des MPoP (FTTH) bzw. im DSLAM (FTTC) oder der ONU (FTTS) und das CPE (Netzabschlussgerät beim Kunden)).

Die variablen Kosten je Kunde unterscheiden sich je nach der Architektur, sind aber im Vergleich mit den Kosten, die der Grundausbau (homes passed) im Ausbaugbiet erfordert, gering.

A2.2.3 Gebäudeinterne Verkabelung

Die Installation von Glasfasern innerhalb des Gebäudes kann das NGA-Modell als Kosten des Glasfaserausbaus berücksichtigen. Sie fallen erst an, wenn der erste Kunde im Gebäude akquiriert ist, und nicht bereits beim flächendeckenden Erschließen eines Ausbaugbietes. Im Rahmen dieser Studie wird die gebäudeinterne Verkabelung für FTTH P2P wertmäßig berücksichtigt. Für FTTS und FTTC wird von einer bestehenden Kupferkabelinfrastruktur ausgegangen, die keine weiteren Investitionen mehr benötigt.

A2.2.4 OPEX und indirekte Investitionen

Für die verschiedenen Anschlussnetztechnologien sind im Modell anlagebezogenen Zuschlagsfaktoren auf die netzelementbezogenen Investitionen für Betriebskosten und indirekte Investitionen vorgehalten (mark-up). Dabei werden die folgenden Positionen differenziert:

OPEX	
<ul style="list-style-type: none"> • Passives Equipment 	<ul style="list-style-type: none"> • Aktives Equipment
Indirekte Investitionen:	
<ul style="list-style-type: none"> • Investitionen für Fahrzeuge 	<ul style="list-style-type: none"> • Investitionen Büros und Gebäude
<ul style="list-style-type: none"> • Investitionen Werkstatt und Ausrüstung 	<ul style="list-style-type: none"> • Investitionen für IT Netzunterstützung und Ausrüstung
<ul style="list-style-type: none"> • Investitionen für das Netzmanagement des IP Netzes 	<ul style="list-style-type: none"> • Investitionen für Grundstücke und Gebäude

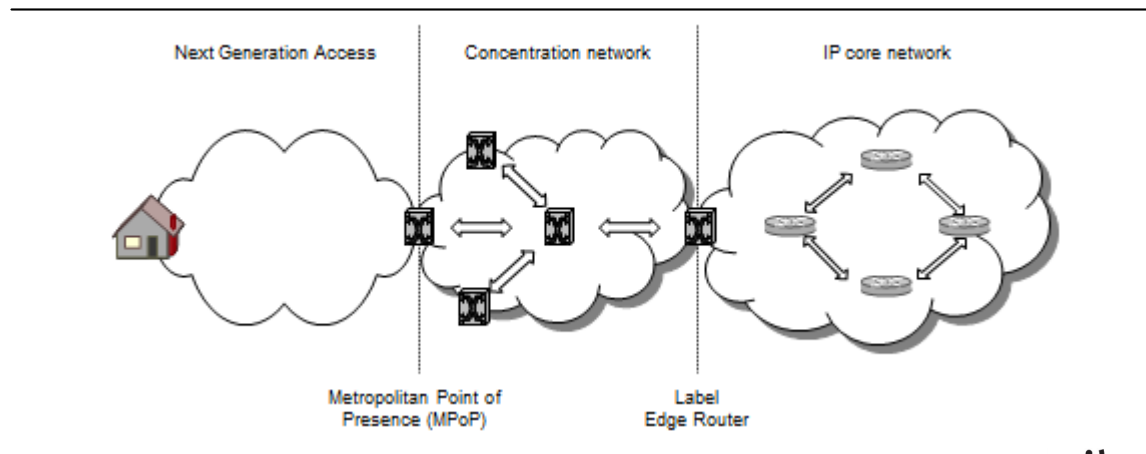
A2.2.5 Upstream Kosten und Erlöse

Die Profitabilität von Investitionsvorhaben hängt vom Verhältnis der Kosten zu den Erlösen ab. Neben den Kosten des Anschlussnetzes sind die sonstigen Upstream Kosten des Netzbetreibers sowie seine Erlöse zu betrachten. Beides erfolgte für dieses Forschungsprojekt unter Rückgriff auf das WIK-NGA Modell.

A2.2.5.1 Kosten von Konzentrations- und Kernnetz

Die Kostenbestimmung für Konzentrations- und Kernnetz sowie allgemeine Konnektivität erfolgt im WIK-NGA-Modell unter Rückgriff auf einen vereinfachten Modellierungsansatz. Gleiches gilt für die Berücksichtigung der Vertriebskosten.

Abbildung A 2-3: NGA, Breitbandzugang und zu Grunde liegende Wertschöpfung



Quelle: WIK

Abbildung A 2-3 beschreibt die Wertschöpfungsstufen des Breitbandzugangs, bestehend aus dem Teilnehmeranschlussnetz (NGA), Konzentrationsnetz sowie IP Kernnetz (Backbone) und seiner Zusammenschaltung mit anderen Netzen zur Herstellung einer vollständigen Konnektivität öffentlicher Telekommunikationsnetze. Das Teilnehmeranschlussnetz, dessen Kosten unter Rückgriff auf das AKM-AN 3.0 bestimmt werden, stellt für den Breitbandzugang das eigentliche Bottleneck mit den noch zu tätigen Investitionen dar. Die darüber liegenden Netzsegmente basieren bereits auf einer Glasfaserinfrastruktur und die erforderlichen technischen Einrichtungen der Übertragungstechnik lassen sich weitestgehend bedarfsgerecht skalieren.

Kosten des Konzentrationsnetzes: Das Konzentrationsnetz verbindet das glasfaserbasierte Teilnehmeranschlussnetz und das Kernnetz und ist konzeptionell zwischen MPoP und dem IP-PoP angesiedelt. Kosten für die Verkehrsübertragung sind als volumenbezogener, variabler Kostensatz je Mbps im Modell dargestellt. Insgesamt ergeben sich monatliche Kosten pro Kunde als Produkt aus eben diesem Kostensatz und dem kundenbezogenen Verkehr in der Busy Hour. Der Busy Hour Verkehr ist der statistische Verkehr, den ein durchschnittlicher Kunde in der auslastungsbezogenen Hauptverkehrszeit verursacht und anhand dessen die konzeptionelle Netzdimensionierung erfolgt. Hinzu kommen ein angenommener Fixkostenanteil für das nationale Konzentrationsnetz sowie ein über die pro IP-TV-Kanal benötigte Bandbreite berechneter Fixkostenanteil.

Kosten des Kernnetzes: Netzseitig vom IP-PoP befindet sich das Backbone-Netz, dessen volumengetriebene Übertragungskosten im Modell konzeptionell identisch sind mit denen im Konzentrationsnetz. Auch hier ergeben sich kundenbezogene Transportkosten nach Maßgabe der Busy Hour Verkehrsnachfrage (Mbps) eines angenommenen Durchschnittskunden. In Analogie zu den Transportkosten für das Konzentrationsnetz gehen auch Übertragungskosten für das Backbone-Netz als monatliche, kostenseitige Position in die Modellberechnung ein. Hinzu kommen hier ebenfalls ein angenommener Fixkostenanteil für das nationale Kernnetz sowie ein über die pro IP-TV-Kanal benötigte Bandbreite berechneter Fixkostenanteil.

Dienste im Konzentrations- und Kern-Netz: Konzeptionell gleich sind die variablen Anteile beider Kostenpositionen auch darin, dass das zugrunde liegende Verkehrsaufkommen nur Voice- und Internet-Dienste (inkl. Video on Demand), jedoch kein IP-TV berücksichtigt. Der Grund verbirgt sich hinter den unterschiedlichen technischen Konzeptionen und Aussendungsarten des IP-TV Dienstes und der beiden anderen Dienste. IP-TV (inkl. PayTV) ist ein Broadcast-Dienst, das heißt er verteilt sich unabhängig von TV nachfragenden Kundenzahlen im NGN-Netz. Voice- und Internet-Dienste sind dagegen kundengetrieben. Weiter unterscheidet sich das Nutzungsverhalten von IP-TV gegenüber anderen Diensten darin, dass sich IP-TV Backbone-seitig fächerartig ausbreitet bis zur kundenedizierten gebäudeinterne Verkabelung, während sich Voice- oder Internet-Dienste nicht fächerartig, sondern mittels technischer Punkt-zu-Punkt Verbindungen ausbreiten. Voice- und Internet-Dienste erzeugen die gleiche Verkehrsbelastung je Kunde unabhängig von der betrachteten Netzebene. Ungleich ist sie dagegen beim IP-basierten Fernsehdienst, weil dessen fächerartige Ausbreitung zu unterschiedlichen IP-TV-Verkehrsvolumina je Kunde pro betrachteter Netzebene führt.

A2.2.5.2 Vertriebs- und Werbekosten (Retailkosten)

Kosten für Customer Care, Order Management, Clearing und Kundensupport sind kostenmäßig im Modell erfasst. Unter dem Oberbegriff Retailkosten sind diese Positionen zusammengefasst. Der hierfür angesetzte Wert in Höhe von 5€ ist als monatlich wiederkehrender Aufwand je Kunde im Modell implementiert und greift unabhängig von der betrachteten Anschlussnetztechnologie.

A2.2.5.3 Erlöse (ARPU)

Profitabilität und Reichweite des Glasfaserausbaus werden ermittelt unter Rückgriff auf mit dem Ausbau verbundene Kosten einerseits und zu erwartende Erträge aus realisierten Diensten andererseits. Der hier angenommene Ertrag (Average Revenue per User, ARPU) ist ein Mischwert, der sich anteilig aus Erträgen für Single Play (VoIP), Double Play (VoIP und Internetdatendienst) und Triple Play (VoIP, Internet, IP-TV) ergibt und letztlich als Netto-Größe zu verstehen ist, die um etwaige Vergünstigungen bereinigt wurde.

Als "Diskussionsbeiträge" des Wissenschaftlichen Instituts für Infrastruktur und Kommunikationsdienste sind zuletzt erschienen:

- Nr. 398: Annette Hillebrand, Martin Zauner:
Qualitätsindikatoren im Brief- und Paketmarkt, Mai 2015
- Nr. 399: Stephan Schmitt, Marcus Stronzik:
Die Rolle des generellen X-Faktors in verschiedenen Regulierungsregimen, Juli 2015
- Nr. 400: Franz Büllingen, Solveig Börnsen:
Marktorganisation und Marktrealität von Machine-to-Machine-Kommunikation mit Blick auf Industrie 4.0 und die Vergabe von IPv6-Nummern, August 2015
- Nr. 401: Lorenz Nett, Stefano Lucidi, Ulrich Stumpf:
Ein Benchmark neuer Ansätze für eine innovative Ausgestaltung von Frequenzgebühren und Implikationen für Deutschland, November 2015
- Nr. 402: Christian M. Bender, Alex Kalevi Dieke, Petra Junk:
Zur Marktabgrenzung bei Kurier-, Paket- und Expressdiensten, November 2015
- Nr. 403: J. Scott Marcus, Christin Gries, Christian Wernick, Imme Philbeck:
Entwicklungen im internationalen Mobile Roaming unter besonderer Berücksichtigung struktureller Lösungen, Januar 2016
- Nr. 404: Karl-Heinz Neumann, Stephan Schmitt, Rolf Schwab unter Mitarbeit von Marcus Stronzik:
Die Bedeutung von TAL-Preisen für den Aufbau von NGA, März 2016
- Nr. 405: Caroline Held, Gabriele Kulenkampff, Thomas Plückebaum:
Entgelte für den Netzzugang zu staatlich geförderter Breitband-Infrastruktur, März 2016
- Nr. 406: Stephan Schmitt, Matthias Wissner:
Kapazitätsmechanismen – Internationale Erfahrungen, April 2016
- Nr. 407: Annette Hillebrand, Petra Junk:
Paketshops im Wettbewerb, April 2016
- Nr. 408: Tseveen Gantumur, Iris Henseler-Unger, Karl-Heinz Neumann:
Wohlfahrtsökonomische Effekte einer Pure LRIC - Regulierung von Terminierungsentgelten, Mai 2016
- Nr. 409: René Arnold, Christian Hildebrandt, Martin Waldburger:
Der Markt für Over-The-Top Dienste in Deutschland, Juni 2016
- Nr. 410: Christian Hildebrandt, Lorenz Nett:
Die Marktanalyse im Kontext von mehrseitigen Online-Plattformen, Juni 2016
- Nr. 411: Tseveen Gantumur, Ulrich Stumpf:
NGA-Infrastrukturen, Märkte und Regulierungsregime in ausgewählten Ländern, Juni 2016
- Nr. 412: Alex Dieke, Antonia Niederprüm, Sonja Thiele:
UPU-Endvergütungen und internationaler E-Commerce, September 2016 (in deutscher und englischer Sprache verfügbar)
- Nr. 413: Sebastian Tenbrock, René Arnold:
Die Bedeutung von Telekommunikation in intelligent vernetzten PKW, Oktober 2016
- Nr. 414: Christian Hildebrandt, René Arnold:
Big Data und OTT-Geschäftsmodelle sowie daraus resultierende Wettbewerbsprobleme und Herausforderungen bei Datenschutz und Verbraucherschutz, November 2016
- Nr. 415: J. Scott Marcus, Christian Wernick:
Ansätze zur Messung der Performance im Best-Effort-Internet, November 2016
- Nr. 416: Lorenz Nett, Christian Hildebrandt:
Marktabgrenzung und Marktmacht bei OTT-0 und OTT-1-Diensten, Eine Projektskizze am Beispiel von Instant-Messenger-Diensten, Januar 2017
- Nr. 417: Peter Kroon:
Maßnahmen zur Verhinderung von Preis-Kosten-Scheren für NGA-basierte Dienste, Juni 2017

- Nr. 419: Stefano Lucidi:
Analyse marktstruktureller Kriterien und Diskussion regulatorischer Handlungsoptionen bei engen Oligopolen, April 2017
- Nr. 420: J. Scott Marcus, Christian Wernick, Tseveen Gantumur, Christin Gries:
Ökonomische Chancen und Risiken einer weitreichenden Harmonisierung und Zentralisierung der TK-Regulierung in Europa, Juni 2017
- Nr. 421: Lorenz Nett:
Incentive Auctions als ein neues Instrument des Frequenzmanagements, Juli 2017
- Nr. 422: Christin Gries, Christian Wernick:
Bedeutung der embedded SIM (eSIM) für Wettbewerb und Verbraucher im Mobilfunkmarkt, August 2017
- Nr. 423: Fabian Queder, Nicole Angenendt, Christian Wernick:
Bedeutung und Entwicklungsperspektiven von öffentlichen WLAN-Netzen in Deutschland, Dezember 2017
- Nr. 424: Stefano Lucidi, Bernd Sörries, Sonja Thiele:
Wirksamkeit sektorspezifischer Verbraucherschutzregelungen in Deutschland, Januar 2018
- Nr. 425: Bernd Sörries, Lorenz Nett:
Frequenzpolitische Herausforderungen durch das Internet der Dinge - künftiger Frequenzbedarf durch M2M-Kommunikation und frequenzpolitische Handlungsempfehlungen, März 2018
- Nr. 426: Saskja Schäfer, Gabriele Kulenkampff, Thomas Plückebaum unter Mitarbeit von Stephan Schmitt:
Zugang zu gebäudeinterner Infrastruktur und adäquate Bepreisung, April 2018
- Nr. 427: Christian Hildebrandt, René Arnold:
Marktbeobachtung in der digitalen Wirtschaft – Ein Modell zur Analyse von Online-Plattformen, Mai 2018
- Nr. 428: Christin Gries, Christian Wernick:
Treiber und Hemmnisse für kommerziell verhandelten Zugang zu alternativen FTTB/H-Netzinfrastrukturen, Juli 2018
- Nr. 429: Serpil Taş, René Arnold:
Breitbandinfrastrukturen und die künftige Nutzung von audiovisuellen Inhalten in Deutschland: Herausforderungen für Kapazitätsmanagement und Netzneutralität, August 2018
- Nr. 430: Sebastian Tenbrock, Sonia Strube Martins, Christian Wernick, Fabian Queder, Iris Henseler-Unger:
Co-Invest Modelle zum Aufbau von neuen FTTB/H-Netzinfrastrukturen, August 2018
- Nr. 431: Johanna Bott, Christian Hildebrandt, René Arnold:
Die Nutzung von Daten durch OTT-Dienste zur Abschöpfung von Aufmerksamkeit und Zahlungsbereitschaft: Implikationen für Daten- und Verbraucherschutz, Oktober 2018
- Nr. 432: Petra Junk, Antonia Niederprüm:
Warenversand im Briefnetz, Oktober 2018
- Nr. 433: Christian M. Bender, Annette Hillebrandt:
Auswirkungen der Digitalisierung auf die Zustelloogistik, Oktober 2018
- Nr. 434: Antonia Niederprüm:
Hybridpost in Deutschland, Oktober 2018
- Nr. 436: Petra Junk:
Digitalisierung und Briefsubstitution: Erfahrungen in Europa und Schlussfolgerungen für Deutschland, Oktober 2018
- Nr. 437: Peter Kroon, René Arnold:
Die Bedeutung von Interoperabilität in der digitalen Welt – Neue Herausforderungen in der interpersonellen Kommunikation, Dezember 2018
- Nr. 438: Stefano Lucidi, Bernd Sörries:
Auswirkung von Bündelprodukten auf den Wettbewerb, März 2019
- Nr. 439: Christian M. Bender, Sonja Thiele:
Der deutsche Postmarkt als Infrastruktur für europäischen E-Commerce, April 2019

- Nr. 440: Serpil Taş, René Arnold:
Auswirkungen von OTT-1-Diensten auf das Kommunikationsverhalten – Eine nachfrageseitige Betrachtung, Juni 2019
- Nr. 441: Serpil Taş, Christian Hildebrandt, René Arnold:
Sprachassistenten in Deutschland, Juni 2019
- Nr. 442: Fabian Queder, Marcus Stronzik, Christian Wernick:
Auswirkungen des Infrastrukturwettbewerbs durch HFC-Netze auf Investitionen in FTTP-Infrastrukturen in Europa, Juni 2019
- Nr. 443: Lorenz Nett, Bernd Sörries:
Infrastruktur-Sharing und 5G: Anforderungen an Regulierung, neue wettbewerbliche Konstellationen, Juli 2019
- Nr. 444: Pirmin Puhl, Martin Lundborg:
Breitbandzugang über Satellit in Deutschland – Stand der Marktentwicklung und Entwicklungsperspektiven, Juli 2019
- Nr. 445: Bernd Sörries, Marcus Stronzik, Sebastian Tenbrock, Christian Wernick, Matthias Wissner:
Die ökonomische Relevanz und Entwicklungsperspektiven von Blockchain: Analysen für den Telekommunikations- und Energiemarkt, August 2019
- Nr. 446: Petra Junk, Julia Wielgosch:
City-Logistik für den Paketmarkt, August 2019
- Nr. 447: Marcus Stronzik, Matthias Wissner:
Entwicklung des Effizienzvergleichs in Richtung Smart Grids, September 2019
- Nr. 448: Christian M. Bender, Antonia Niederprüm:
Berichts- und Anzeigepflichten der Unternehmen und mögliche Weiterentwicklungen der zugrundeliegenden Rechtsnormen im Postbereich, September 2019
- Nr. 449: Ahmed Elbanna unter Mitwirkung von Fabian Eltges:
5G Status Studie: Herausforderungen, Standardisierung, Netzarchitektur und geplante Netzentwicklung, Oktober 2019
- Nr. 450: Stefano Lucidi, Bernd Sörries:
Internationale Vergleichsstudie bezüglich der Anwendung und Umsetzung des Nachbildbarkeitsansatzes, Dezember 2019
- Nr. 451: Matthias Franken, Matthias Wissner, Bernd Sörries:
Entwicklung der funkbasierten Digitalisierung in der Industrie, Energiewirtschaft und Landwirtschaft und spezifische Frequenzbedarfe, Dezember 2019
- Nr. 452: Bernd Sörries, Lorenz Nett:
Frequenzmanagement: Lokale/regionale Anwendungsfälle bei 5G für bundesweite Mobilfunknetzbetreiber sowie für regionale und lokale Betreiber unter besonderer Betrachtung der europäischen Länder sowie von China, Südkorea und den Vereinigten Staaten von Amerika, Dezember 2019
- Nr. 453: Martin Lundborg, Christian Märkel, Lisa Schrade-Grytsenko, Peter Stamm:
Künstliche Intelligenz im Telekommunikationssektor – Bedeutung, Entwicklungsperspektiven und regulatorische Implikationen, Dezember 2019
- Nr. 454: Fabian Eltges, Petra Junk:
Entwicklungstrends im Markt für Zeitungen und Zeitschriften, Dezember 2019
- Nr. 455: Christin Gries, Julian Knips, Christian Wernick:
Mobilfunkgestützte M2M-Kommunikation in Deutschland – zukünftige Marktentwicklung und Nummerierungsbedarf, Dezember 2019
- Nr. 456: Menessa Ricarda Braun, Christian Wernick, Thomas Plückerbaum, Martin Ockenfels:
Parallele Glasfaserausbauten auf Basis von Mitverlegung und Mitnutzung gemäß DigiNetzG als Möglichkeiten zur Schaffung von Infrastrukturwettbewerb, Dezember 2019

- Nr. 457: Thomas Plückebaum, Martin Ockenfels:
Kosten und andere Hemmnisse der Migration von Kupfer- auf Glasfasernetze, Februar 2020
- Nr. 458: Andrea Liebe, Jonathan Lennartz, René Arnold:
Strategische Ausrichtung bedeutender Anbieter von Internetplattformen, Februar 2020
- Nr. 459: Sebastian Tenbrock, Julian Knips, Christian Wernick:
Status quo der Abschaltung der Kupfernetzinfrastruktur in der EU, März 2020
- Nr. 460: Stefano Lucidi, Martin Ockenfels, Bernd Sörries:
Anhaltspunkte für die Replizierbarkeit von NGA-Anschlüssen im Rahmen des Art. 61 Abs. 3 EKEK, März 2020
- Nr. 461: Fabian Eltges, Gabriele Kulenkampff, Thomas Plückebaum, Desislava Sabeva:
SDN/NFV und ihre Auswirkungen auf die Kosten von Mobilfunk und Festnetz im regulatorischen Kontext, März 2020
- Nr. 462: Lukas Wiewiorra, Andrea Liebe, Serpil Taş
Die wettbewerbliche Bedeutung von Single-Sign-On- bzw. Login-Diensten und ihre Relevanz für datenbasierte Geschäftsmodelle sowie den Datenschutz, Juni 2020
- Nr. 463: Bernd Sörries, Lorenz Nett, Matthias Wissner
Die Negativauktion als ein Instrument zur Versorgung weißer Flecken mit Mobilfunkdiensten, Dezember 2020
- Nr. 464: Sebastian Tenbrock, Christian Wernick:
Incumbents als Nachfrager von Vorleistungen auf FTTB/H-Netzen, Dezember 2020
- Nr. 465: Marcus Stronzik, Gonzalo Zuloaga:
Empirische Untersuchung der FTTB/H-Ausbauaktivität im europäischen Vergleich, Dezember 2020
- Nr. 466: Antonia Niederprüm mit Unterstützung von Gonzalo Zuloaga und Willem van Lienden:
Verbundproduktion im Zustellmarkt: Briefnetze mit Paketen oder Paketnetze mit Briefen?, Dezember 2020
- Nr. 467: Serpil Taş, Lukas Wiewiorra (in Zusammenarbeit mit dem Weizenbaum-Institut):
Multihoming bei Plattformdiensten – Eine nachfrageseitige Betrachtung, Dezember 2020
- Nr. 468: Menessa Ricarda Braun, Julian Knips, Christian Wernick:
Die Angebotsentwicklung auf dem deutschen Mobilfunkmarkt 2017-2020, Dezember 2020
- Nr. 469: Isabel Gull, Lisa Schrade-Grytsenko, Martin Lundborg:
Cloud-Lösungen und KI-as-a-Service – Aktuelle und potenzielle Anwendungsszenarien und Marktentwicklungen, Dezember 2020
- Nr. 470: Bernd Sörries, Matthias Franken, Dajan Baischew, Stefano Lucidi:
Einfluss von Versorgungsaufgaben auf die Mobilfunkabdeckung in der EU, Dezember 2020
- Nr. 471: Julian Knips, Christin Gries, Christian Wernick:
Consumer-IoT in Deutschland – Anwendungsbereiche und möglicher Regelungsbedarf, Dezember 2020
- Nr. 472: Saskja Schäfer, Ahmed Elbanna, Werner Neu, Thomas Plückebaum:
Mögliche Einsparungspotentiale beim Ausbau von 5G durch Infrastructure Sharing, Dezember 2020
- Nr. 473: Gabriele Kulenkampff, Martin Ockenfels, Konrad Zoz, Gonzalo Zuloaga:
Kosten von Breitband-Zugangsnetzen, Clusterbildung und Investitionsbedarf unter Berücksichtigung des bestehenden Ausbaus – bottom-up Modellierung und statistische Analyse –, Dezember 2020

ISSN 1865-8997