

Leistungsbewertung im Air Traffic Management

-

Methodik, Analyse und Evaluation



Dissertation

zur Erlangung des akademischen Grades

Doktor rerum politicarum
(Dr. rer. pol.)

vorgelegt
der Fakultät Verkehrswissenschaften "Friedrich List"
der Technischen Universität Dresden

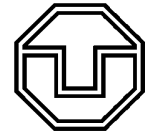
von

M.Sc. Thomas Standfuß

Institut für Luftfahrt und Logistik
Dresden, September 2020

Erstgutachter: Prof. Dr. rer. pol. habil. Georg Hirte
Zweitgutachter: Prof. Dr. Frank Fichert
Wissenschaftlicher Betreuer: Priv.-Doz. Dr.-Ing. habil. Michael Schultz

Eingereicht am: 17.09.2020
Verteidigt am: 23.04.2021

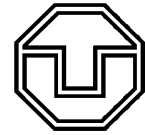


Inhaltsverzeichnis

Abbildungsverzeichnis	VII
Tabellenverzeichnis	IX
Abkürzungsverzeichnis	X
Verzeichnis der Flugsicherungsdienstleister.....	XIII
1 Einführung.....	1
1.1 Motivation	1
1.2 Zielsetzungen der Arbeit	3
1.3 Aufbau der Dissertation.....	4
2 Grundlagen Air Traffic Management	7
2.1 Stakeholder des Luftverkehrs	7
2.2 Luftverkehrsmanagement in Europa.....	10
2.3 Betriebliche Aspekte	11
2.4 Finanzierung europäischer Air Navigation Services	17
2.5 Herausforderungen des Luftverkehrsmanagements in Europa.....	19
3 Performance-Vergleich europäischer Flugsicherungen	23
3.1 Ökonomische Grundlagen	23
3.1.1 Produktivität und Effizienz	23
3.1.1.1 Leistungsbewertung von Unternehmen.....	23
3.1.1.2 Produktionsfunktionen	25
3.1.1.3 Effizienzmaße	27
3.1.2 Data Envelopment Analysis.....	31
3.1.2.1 Konstante Skalenerträge	31
3.1.2.2 Variable Skalenerträge	33
3.1.2.3 Sonderformen und Erweiterungen der DEA.....	35
3.1.3 Stochastic Frontier Analysis	36
3.1.4 Methodenauswahl für die Analyse im ATM	38
3.2 Status Quo der Leistungsbewertung im ATM	39
3.2.1 EUROCONTROL.....	39
3.2.1.1 Single European Sky	39
3.2.1.2 Performance Scheme.....	41
3.2.1.3 ATM Cost-Effectiveness Report	43
3.2.1.4 Performance Review Report	45
3.2.1.5 Weitere Berichte und Studien.....	46
3.2.2 Überblick zu akademischen Studien	47
3.2.3 Zusammenfassende kritische Würdigung	52



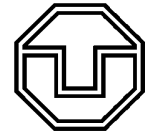
3.3	Datengrundlage.....	53
3.3.1	Vorgehen	53
3.3.2	Wertschöpfungsprozess	54
3.3.2.1	Grundmodell	54
3.3.2.2	Inputs	56
3.3.2.3	Outputs	58
3.3.3	Datenanalyse	60
3.4	Anwendung DEA.....	62
3.4.1	Modellierung	62
3.4.2	Ergebnisse	65
3.4.2.1	Konstante Skalenerträge	65
3.4.2.2	Variable Skalenerträge	67
3.4.2.3	Zeitliche Entwicklung, kompletter Datensatz.....	69
3.4.2.4	Zeitliche Entwicklung, homogener Datensatz	71
3.4.2.5	Räumliche Autokorrelation	74
3.4.3	Grenzen der Anwendung.....	77
3.5	Zwischenfazit	80
4	Ursachenanalyse	83
4.1	Ansatz	83
4.2	Regressionsanalyse auf ANSP Ebene	84
4.2.1	Methoden.....	84
4.2.1.1	Regressionsanalyse	84
4.2.1.2	Principle Component Analysis.....	85
4.2.1.3	Beschränkte und zensierte Modelle	86
4.2.2	Daten und Einflussgrößen	87
4.2.2.1	Auswahl und Gruppierung potentieller Faktoren.....	87
4.2.2.2	Endogene Faktoren	89
4.2.2.3	Teilexogene Faktoren.....	91
4.2.2.4	Exogene Faktoren	96
4.2.3	Einfluss auf die Produktivität.....	96
4.2.4	Einfluss auf die DEA-Scores.....	99
4.2.4.1	Vorgehen	99
4.2.4.2	Maximierung der Modellgüte	100
4.2.4.3	Stufenweiser Aufbau	102
4.3	Panel-Analyse	104
4.3.1	Daten und Modellierung.....	104
4.3.2	Methodik	105
4.3.3	Ergebnisse	106
4.3.4	Regression.....	109



4.3.4.1	Ansatz	109
4.3.4.2	Einfluss auf Produktivität	110
4.3.4.3	Einfluss auf TFP-Change	112
4.4	Disaggregierter Ansatz.....	114
4.4.1	Motivation und Vorgehen.....	114
4.4.2	Quantitative Analyse ACC	114
4.4.2.1	Datengrundlage	114
4.4.2.2	Leistungsbewertung	116
4.4.2.3	Cluster-Analyse	117
4.4.2.4	Regressionsanalyse	121
4.4.2.5	Panel Analyse.....	129
4.4.3	Qualitative Analyse ACC	132
4.4.3.1	Stakeholder Konsultationen.....	132
4.4.3.2	Strategischer Planungshorizont	133
4.4.3.3	Personalplanung und -einsatz.....	134
4.4.3.4	Kurzfristige Maßnahmen und Herausforderungen	135
4.4.3.5	Sektorkapazität.....	137
4.5	Größenvorteile im Air Traffic Management	141
4.5.1	Hintergrund	141
4.5.2	Analyse der Skalenerträge und Skaleneffizienz	142
4.5.3	Strukturelle Einflussfaktoren	146
4.6	Sonstige Einflussfaktoren.....	148
4.6.1	Prognosequalität.....	148
4.6.2	Anteil des General Aviation Verkehrs	153
4.6.3	Einfachbesetzung der Sektoren.....	158
4.6.4	Jährliche Arbeitszeit pro Lotsen.....	159
4.7	Zwischenfazit	160
5	Zusammenfassung und Handlungsempfehlungen	165
5.1	Ergebnisse der Dissertation	165
5.1.1	Überblick.....	165
5.1.2	Modellierung und Leistungsbewertung	166
5.1.3	Ursachenanalyse	167
5.1.3.1	Methodik	167
5.1.3.2	Betrieb	168
5.1.3.3	Luftraum.....	169
5.1.3.4	Nachfrage	171
5.1.3.5	Umweltfaktoren.....	174
5.2	Ausblick	175
5.2.1	Neue Marktteilnehmer	175



5.2.2	Dynamische Luftraumstrukturen.....	176
5.2.2.1	Hintergrund.....	176
5.2.2.2	Ansatz.....	176
5.2.2.3	Betriebliche und ökonomische Vorteile	179
5.2.2.4	Schlussfolgerungen	182
5.2.3	Weitere Forschungsaspekte.....	182
5.3	Handlungsempfehlungen	183
Anhang.....		XV
Literaturverzeichnis		XLV
Danksagung.....		LXI

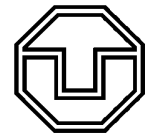


Abbildungsverzeichnis

Abbildung 1-1: Development Circle und Feedback Loop	4
Abbildung 2-1: Stakeholder des Luftverkehrsmarktes	7
Abbildung 2-2: Luftraumgröße europäischer Staaten (2017).....	10
Abbildung 2-3: Flight Information Regions, Unterer Luftraum, 2019.....	11
Abbildung 2-4: Tätigkeitsbereiche einer Flugsicherung und von EUROCONTROL	12
Abbildung 2-5: Phasen eines Fluges.....	12
Abbildung 2-6: Betriebliche Gliederung der Luftverkehrskontrolle	15
Abbildung 2-7: Betriebliche Ebenen und operative Ziele (Auswahl)	17
Abbildung 2-8: Density Plot der geplanten Trajektorien.....	20
Abbildung 2-9: Abweichung zwischen geplanter und tatsächlich geflogener Route (Beispiel).....	21
Abbildung 2-10: Umwege durch unterschiedlich hohe Gebührensätze der ANSPs (Beispiel).....	21
Abbildung 2-11: Functional Airspace Blocks (FABs).....	22
Abbildung 3-1: Effizienzmessverfahren.....	24
Abbildung 3-2: Produktionsfunktionen im Vergleich.....	26
Abbildung 3-3: Berechnung der Effizienz über Distanzfunktionen	27
Abbildung 3-4: Messung der Effizienz mittels radialer Distanz zu einer deterministischen Grenze	28
Abbildung 3-5: Skaleneffizienz	29
Abbildung 3-6: Bestimmung der allokativen Effizienz	30
Abbildung 3-7: Peers und Gewichte des dualen Modells.....	33
Abbildung 3-8: Nicht-sinkende und Nicht-steigende Produktionsfunktion	35
Abbildung 3-9: Stochastische Randproduktionsfunktionen.....	37
Abbildung 3-10: Gebühren- und Leistungssystem der europäischen Flugsicherungsunternehmen	41
Abbildung 3-11: Key Performance Areas und Indicators in RP2	43
Abbildung 3-12: Gate-to-Gate Kosteneffizienz, gemessen in Kosten pro CFH	44
Abbildung 3-13: Lotsenproduktivität 2017, gemessen in CFH pro Lotsenstunde	44
Abbildung 3-14: Prozesskette der Leistungsbewertung	53
Abbildung 3-15: Vereinfachtes Performance-Modell einer ANSP	55
Abbildung 3-16: Kosten pro Fluglotse pro Jahr (2017)	57
Abbildung 3-17: Individuelle Gewichte der ANSPs 2014	59
Abbildung 3-18: Entwicklung der Nachfrage 2003-2017	61
Abbildung 3-19: DEA-Scores 2016, Modell 2 und 2A, konstante Skalenerträge	66
Abbildung 3-20: DEA-Scores 2016, Modell 2 und 2A, variable Skalenerträge	69
Abbildung 3-21: Durchschnittlicher DEA-Score 2003-2016, kompletter Datensatz	70
Abbildung 3-22: Durchschnittlicher DEA-Score, Modell 2A, kompletter Datensatz	71
Abbildung 3-23: Durchschnittlicher DEA-Score 2003-2016, homogener Datensatz.....	72
Abbildung 3-24: Durchschnittlicher DEA-Score, Modell 2B, homogener Datensatz.....	73
Abbildung 3-25: Beispiele für räumliche Autokorrelation (Modell)	75
Abbildung 3-26: Bootstrapping DEA Frontier mit verschiedener Anzahl an Replika, Modell 2.....	79
Abbildung 3-27: Free Disposal Hull (Modell 2).....	80
Abbildung 4-1: Totale Faktorproduktivität.....	105
Abbildung 4-2: Modellvergleich TFP Wachstum	107
Abbildung 4-3: Durchschnittliche Wachstumsrate einzelner ANSPs (Modell 2A).....	108
Abbildung 4-4: ACCs mit der höchsten und niedrigsten ATCO-Produktivität 2017	116
Abbildung 4-5: Aufteilung der ACCs bei 14 Clustern	118
Abbildung 4-6: Unterschiede der Gruppenzuordnung bei k-means und PAM Algorithmus	119
Abbildung 4-7: Ergebnis Hierarchische Clusteranalyse	120
Abbildung 4-8: Zusammenhang zwischen Produktivität und Luftraum vor und nach Clustering.....	121
Abbildung 4-9: Anteil der ACCs an der Nachfrage innerhalb FABECs	138
Abbildung 4-10: Kapazität einzelner ACCs (Entry Counts).....	140
Abbildung 4-11: Interdependenz zwischen Streckenverkehr und Kapazität.....	140
Abbildung 4-12: Skalenindikator für die europäischen Flugsicherungen	144



Abbildung 4-13: Skaleneffekte nach Model 2A	145
Abbildung 4-14: Zusammenhang zwischen Komplexität und Skaleneffizienz (Modell 2B)	147
Abbildung 4-15: Zusammenhang zwischen Luftraum und Skaleneffizienz (Modell 2B)	148
Abbildung 4-16: Verkehrsprognose 2050	150
Abbildung 4-17: STATFOR Prognosefehler nach Jahr für die DFS	152
Abbildung 4-18: Die 25 meistfrequentierten GA Flughäfen in Europa	155
Abbildung 4-19: Die 25 meistfrequentierten GA Flughäfen in den USA	155
Abbildung 4-20: Density Plots für IFR und IFR GA, geplante Trajektorien	156
Abbildung 4-21: Fuzzy Clustering des europäischen GA Luftverkehrs	157
Abbildung 4-22: Jährliche Arbeitsstunden je ATCO	160
Abbildung 5-1: Analyse-Cluster und deren Elemente	165
Abbildung 5-2: Integration von Weltraumflügen in den europäischen Luftraum	175
Abbildung 5-3: Methodischer Ansatz der dynamischen Luftraumoptimierung	177
Abbildung 5-4: Von der Trajektorie zum Verkehrs-Cluster	178
Abbildung 5-5: Beispiel für ein Voronoi Diagramm	178
Abbildung 5-6: Kosteneinsparung durch dynamische Sektorisierung im unteren Luftraum	181
Abbildung A 1: Luftraumklassen	XV
Abbildung A 2: EUROCONTROL Staaten	XVI
Abbildung A 3: Arbeitskosten pro Lotsenstunde in Europa (2017)	XIX
Abbildung A 4: Schema Monte Carlo Simulation zur Berechnung der Composite Flight Hours	XIX
Abbildung A 5: Regressionsanalyse mit Hauptkomponenten	XXIV
Abbildung A 6: Logarithmische Transformation DEA Werte	XXIV
Abbildung A 7: Korrelationsmatrix der quantitativen Einflussfaktoren, ANSP Ebene	XXV
Abbildung A 8: Entwicklung Malmquist Indizes, Enroute Services	XXVIII
Abbildung A 9: Malmquist Indize europäischer Flugsicherungen, Enroute, Modell 3A	XXIX
Abbildung A 10: Entwicklung Malmquist Indizes, Terminal Services, Modell 5B	XXIX
Abbildung A 11: Malmquist Indizes europäischer Flugsicherungen, Terminal, Modell 5B	XXX
Abbildung A 12: Korrelationsmatrix Einflussfaktoren ANSP, Panel-Analyse	XXXI
Abbildung A 13: Korrelationsmatrix potentieller Einflussfaktoren auf ACC Ebene	XXXIII
Abbildung A 14: ATCO-Produktivität europäischer ACCs	XXXIV
Abbildung A 15: Grafische Bestimmung der Cluster-Anzahl	XXXV
Abbildung A 16: Silhouetten-Diagramm zur Überprüfung der Cluster-Anzahl	XXXVI
Abbildung A 17: Ergebnisse k-means Clustering	XXXVI
Abbildung A 18: Regressionsgüte ACC am Beispiel Cook-Graph	XXXVII
Abbildung A 19: Korrelationsmatrix Parameter der Regression auf ACC Ebene	XXXVII
Abbildung A 20: Anzahl der gleichzeitig in einem Sektor befindlichen Luftfahrzeuge	XXXIX
Abbildung A 21: Prognosequalität STATFOR	XLII
Abbildung A 22: Density Plots für IFR und IFR-GAT, tatsächliche Trajektorien	XLIII



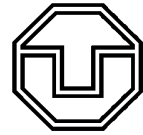
Tabellenverzeichnis

Tabelle 3-1: Berechnungsbeispiel für die Effizienz im Fall multipler Inputs	28
Tabelle 3-2: Ergebnisse RP3 Target Setting für ausgewählte ANSPs	50
Tabelle 3-3: Deskriptive Statistik PRU Daten 2016.....	61
Tabelle 3-4: DEA-Modelle	64
Tabelle 3-5: Modellvergleich konstante Skalenerträge	65
Tabelle 3-6: Modellvergleich variable Skalenerträge	68
Tabelle 3-7: Zeitübergreifend effiziente DMUs.....	74
Tabelle 3-8: Ergebnisse der räumlichen Autokorrelation mit Nachbarschaftsmatrix	76
Tabelle 3-9: Ergebnisse der räumlichen Autokorrelation mit Distanzmatrix	76
Tabelle 3-10: Supereffizienz ausgewählter ANSPs	78
Tabelle 4-1: Deskriptive Statistik der Faktoren (2016).....	88
Tabelle 4-2: Effizienzbeeinflussende Faktoren in der Regressionsanalyse (ANSP)	89
Tabelle 4-3: Volatilitäts-Scores ausgewählter ANSPs für das Jahr 2019	95
Tabelle 4-4: Regressionsmodell zur Prädiktion der abhängigen Variable PRU-Produktivität	97
Tabelle 4-5: OLS Regression ATCO-Produktivität.....	98
Tabelle 4-6: Regressionsanalyse für maximierte Modellgüte	101
Tabelle 4-7: Tobit und Truncated Regression für Modell 2A, 2016.....	103
Tabelle 4-8: Durchschnittliche Wachstumsraten, Gate-to-Gate.....	107
Tabelle 4-9: FEM und REM für Produktivität gemäß PRU.....	111
Tabelle 4-10: Fixed und Random Effects Model für Modell 2A.....	113
Tabelle 4-11: Deskriptive Statistik ACC Daten, 2017.....	115
Tabelle 4-12: Effizienzbeeinflussende Faktoren in der Regressionsanalyse (ACC).....	123
Tabelle 4-13: Regressionsanalyse ACC unter Maximierung der Modellgüte	126
Tabelle 4-14: Iterative Regressionsanalyse ACC Ebene, normierte Daten	127
Tabelle 4-15: Iterative Panelanalyse ACC Ebene, Fixed Effects	130
Tabelle 4-16: Iterative Panelanalyse ACC Ebene, Random Effects	131
Tabelle 4-17: Deskriptive Statistik Sektordaten	139
Tabelle 4-18: Anzahl der DMUs mit steigenden, konstanten oder fallenden SKE.....	143
Tabelle 4-19: Skalenerträge ausgewählter ANSPs.....	143
Tabelle 4-20: Prognosequalität der STATFOR Vorhersagen für die DFS, 2015-2017	151
Tabelle A 1: Luftraumklassen und deren Charakteristiken	XVI
Tabelle A 2: DEA-Ergebnisse nach CRS, 2016, Modellvergleich	XX
Tabelle A 3: DEA-Ergebnisse nach VRS, 2016, Modellvergleich	XXI
Tabelle A 4: DEA Ergebnisse 2007-2016, heterogener Datensatz, VRS, Modell 2.....	XXII
Tabelle A 5: DEA Ergebnisse 2007-2016, homogener Datensatz, VRS, Modell 2.....	XXIII
Tabelle A 6: Vergleich Ergebnisse Truncated Regression.....	XXVI
Tabelle A 7: Deskriptive Statistik Panel-Daten ANSP Ebene, Modell 2A	XXVII
Tabelle A 8: Panel-Analyse, Enroute-Modelle.....	XXVII
Tabelle A 9: Panel-Analyse, Terminal-Modelle	XXVII
Tabelle A 10: Panel-Analyse, Ergebnisse für Enroute	XXVIII
Tabelle A 11: Regression basierend auf Panel-Daten (Modell 1).....	XXXII
Tabelle A 12: Modellgüte Regression, ACC Ebene	XXXVI
Tabelle A 13: Iterative Regressionsanalyse ACC Ebene, originäre Daten	XXXVIII
Tabelle A 14: Bewertung von in FABEC genutzter ATFM-Tools	XXXIX
Tabelle A 15: Ansätze zur Berechnung der Komplexität auf Sektorebene	XL
Tabelle A 16: Skalenerträge, ANSP Ebene, 2016	XLI
Tabelle A 17: Prognosequalität STATFOR basierend auf DFS-Daten	XLII
Tabelle A 18: Potentielle Einsparungen durch Single Sector Operations	XLIII



Abkürzungsverzeichnis

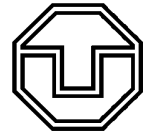
ACC	Area Control Center
ACDM	Airport collaborative decision-making
AFB	Absolute Forecast Bias
AIM	Aeronautical Information Management
AM	Airport Movements
ANS	Air Navigation Services
ANSP	Air Navigation Service Provider
ASM	Airspace Management
ASMA	Arrival Sequencing and Metering
ATC	Air Traffic Control
ATCO	Air Traffic Control Officer
ATFM	Air Traffic Flow Management
ATM	Air Traffic Management
ATS	Air Traffic Services
BADA	Base of Aircraft Data
BAF	Bundesaufsichtsamt für Flugsicherung
BIP	Bruttoinlandsprodukt
CANSO	Civil Air Navigation Services Organisation
CAPAN	ATC Capacity Analyser Tool
CDA	Continuous Descent Approach
CDV	Capacity Default Value
CFH	Composite Flight Hour
CHMI	Collaboration Human Machine Interface
CIU	Composite Infrastructure Unit
CNS	Communication, Navigation, Surveillance
CRCO	Central Route Charges Office
CRS	Constant Returns to Scale
DCB	Demand Capacity Balancing
DDR	Demand Data Repository
DEA	Data Envelopment Analysis
DLE	Dienstleistungseinheit
DMU	Decision Making Unit
DOS	Diseconomies of Scale
DPF	Durchschnittsproduktionsfunktion
DRS	Decreasing Returns to Scale
DSO	Dual ATCO Sector Operations
EBG	Einsatzberechtigungsguppe
ECAC	European Civil Aviation Conference
EOS	Economies of Scale
Exp	Expected
FAB	Functional Airspace Block
FABCE	Functional Airspace Block Central Europe
FABEC	Functional Airspace Block Europe Central



FC-ATM	Flight Centered Air Traffic Management
FDH	Free Disposable Hull
FEM	Fixed Effects Model
FIR	Flight Information Region
FIS	Flight Information Services
FMP	Flow Management Position
FRA	Free Route Airspace
FSNC	Full Service Network Carrier
FTE	Full Time Equivalent
FUA	Flexible Use of Airspace
GAT	General Aviation
GDP	Gross Domestic Product
HFE	Horizontal Flight Efficiency
HHI	Herfindahl-Hirschman Index
ICAO	International Civil Aviation Organization
IFATCA	International Federation of Air Traffic Controllers' Associations
IFR	Instrument Flight Rules
IRS	Increasing Returns to Scale
KMO	Kaiser-Meyer-Olkin
KPA	Key Performance Area
KPI	Key Performance Indicator
LCC	Low Cost Carrier
LIZ	Leistungsinformationszentrum
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
MET	Meteorological Information Management
MI	Malmquist Index
MLE	Maximum Likelihood Estimation
MPSS	Most Productive Scale Size
MSA	Measure of Sampling Adequacy
MTOW	Maximum Take Off Weight
NEST	Network Strategic Modelling Tool
OAT	Operational Air Traffic
Obs	Observed
OLS	Ordinary Least Squares
OPS	Operations
OVB	Omitted Variable Bias
PAM	Partition Around Medoids
PBTF	Performance Benchmarking Taskforce
PCA	Principal Component Analysis
PE	Pure Efficiency
PMG	Performance Management Group
PPP	Purchasing Power Parities
PRB	Performance Review Body
PRC	Performance Review Commission



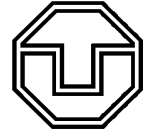
PRR	Performance Review Report
PRU	Performance Review Unit
REM	Random Effects Model
RFP	Relative Forecast Bias
RI	Runway Incursion
RP	Reference Period
RPF	Randproduktionsfunktion
SES	Single European Sky
SFA	Stochastic Frontier Analysis
SI	Scale Indicator
SKE	Skalenerträge
SMI	Separation Minima Infringement
SSO	Single ATCO Sector Operations
STU	Staffelungsunterschreitungen
SWOT	Strength Weaknesses Opportunities Threats
TCFH	Total Controlled Flight Hours
TF	Technischer Fortschritt
TFP	Totale Faktorproduktivität
TRA	Temporary Reserved Area
TSU	Total Service Units
VFR	Visual Flight Rules
VIF	Variance Inflation Factor
VRS	Variable Returns to Scale



Verzeichnis der Flugsicherungsdienstleister

Albcontrol	Albanien
ANA Lux (Administration de la Navigation Aérienne Luxembourg)	Luxemburg
ANS CR (Air Navigation Services of the Czech Republic)	Tschechische Republik
ANS Finland (Air Navigation Services Finland)	Finnland
ARMATS (Armenian Air Traffic Services)	Armenien
Austro Control	Österreich
Avinor	Norwegen
Belgocontrol	Belgien
BULATSA (Bulgarian Air Traffic Services Authority)	Bulgarien
Croatia Control	Kroatien
DCAC Cyprus (Department of Civil Aviation of Cyprus)	Zypern
DFS (Deutsche Flugsicherung)	Deutschland
DHMI (Devlet Hava Meydanları İşletmesi)	Türkei
DSNA (Direction des Services de la Navigation Aérienne)	Frankreich
EANS (Estonian Air Navigation Services)	Estland
ENAIRE	Spanien
ENAV (Ente Nazionale per l'Aviazione)	Italien
FAA (Federal Aviation Administration)	USA
HCAA (Hellenic Civil Aviation Authority)	Griechenland
HungaroControl	Ungarn
IAA (Irish Aviation Authority)	Irland
LFV (Luftfartsverket)	Schweden
LGS (Latvijas Gaisa Satiksme)	Lettland
LPS (Letové prevádzkové služby)	Slowakei
LVNL (Luchtverkeersleiding Nederland)	Niederlande
MATS (Malta Air Traffic Services)	Malta
M-NAV (Macedonian Air Navigations Services Provider)	Mazedonien
MOLDATSA (Moldovian Air Traffic Services Authority)	Moldawien
MUAC (Maastricht Upper Airspace Control Centre)	International
NATS (National Air Traffic Services)	Vereinigtes Königreich
NAV Portugal (Navegação Aérea de Portuga)	Portugal
NAVIAIR (Navigation Via Air)	Dänemark
Oro navigacija	Litauen
PANSA (Polish Air Navigation Services Agency)	Polen
ROMATSA (Romanian Air Traffic Services Administration)	Rumänien
Sakaeronavigatsia	Georgien
Skyguide	Schweiz
Slovenia Control	Slowenien
SMATSA (Serbia and Montenegro Air Traffic Services Agency)	Serbien/Montenegro
UkSATSE (Ukrainian State Air Traffic Service Enterprise)	Ukraine

Hinweis: Aus Gründen der Übersichtlichkeit werden die Abkürzungen der Flugsicherungsunternehmen im Text nicht eingeführt.



1 Einführung

1.1 Motivation

Zu Beginn des 20. Jahrhunderts etablierte sich weltweit, insbesondere in Europa und Amerika, der Luftverkehr als neue Reise- und Transportmöglichkeit. Der dadurch beschleunigten globalen Vernetzung stand ein uneinheitliches Luftrecht gegenüber. Erst das Chicagoer Abkommen von 1944 legte den Grundstein für einheitliche Normen und Prozesse und öffnete damit den Weg zur späteren Liberalisierung des Luftverkehrs (Fritzsche 2010; Pompl 2013; Schladebach 2007). Dies ermöglichte jene Strukturen, die heute vorherrschend sind, sowie die Entwicklung des Luftverkehrs zu einem einflussreichen Wirtschaftszweig.

2018 wurden weltweit 4,3 Mrd. Passagiere befördert und 58 Mio. t Fracht transportiert. Die Verkehrsleistung des Jahres betrug 8,3 Bio. Passagierkilometer, bzw. 231 Mrd. Tonnenkilometer. Berücksichtigt man sowohl die direkten als auch die indirekten, abgeleiteten und katalytischen Effekte des Luftverkehrsmarktes (siehe Abschnitt 2.1), so trugen die Unternehmen zu einer Wertschöpfung von ca. 2,7 Bio. \$ bei. Dies entspricht einem Anteil von rund 3,6% des weltweiten Bruttoinlandsproduktes. Die im Luftverkehr tätigen sowie daran beteiligten Firmen generierten 65,5 Mio. Arbeitsplätze. Dies umfasst u. a. 1303 Fluggesellschaften, 3759 kommerziell genutzte Flughäfen und 170 Flugsicherungen (ACI et al. 2019; ATAG 2018).

Die räumliche Verteilung des Luftverkehrs ist nicht gleichmäßig, sondern konzentriert sich insbesondere in Nordamerika, Europa und Südostasien. Im europäischen Netzwerk summierte sich im Jahr 2019 das Verkehrsaufkommen auf 11,11 Mio. Flüge - eine Steigerung von 0,8% im Vergleich zum Vorjahr. Das EU-weite Bruttoinlandsprodukt (BIP) stieg im gleichen Zeitraum um 1,5% (Statista 2020). Das verhältnismäßig niedrige Wachstum des Verkehrsaufkommens wird u. a. durch die Insolvenz von Fluggesellschaften (z. B. Germania) und dem Grounding der Boeing 737MAX erklärt. Dennoch ist eine zeitliche Konzentration des Verkehrs zu beobachten: Am 28. Juni 2019 wurden insgesamt 37.228 Flüge im europäischen Luftraum kontrolliert und damit der bis dahin bestehende Rekord aus dem Jahre 2018 übertroffen (EUROCONTROL 2019f, 2020e, 2020k).

Auch für das aktuelle Jahr (2020) prognostizierte STATFOR ein durchschnittliches Verkehrswachstum zwischen 0,5% (low scenario) und 3,9% (high scenario)¹. Mittelfristig (7 Jahre) wurde für Deutschland ein jährliches Verkehrswachstum von 1,3% prognostiziert (EUROCONTROL 2019i). In der langfristigen Perspektive (20 Jahre) werden Wachstumsraten zwischen 1,0% (Norwegen) und 4,7% (Armenien) vorhergesagt (EUROCONTROL 2019b). Jedoch kann aufgrund der Corona-Pandemie, analog zu anderen Krisen, ein vorübergehender Einbruch der Verkehrsnachfrage² beobachtet werden. Dies wird zu einer sinkenden Prognosequalität für die kommenden Jahre führen (siehe auch Abschnitt 4.6.1).

¹ Wachstumsraten der einzelnen Länder können abweichen, u. a. 8,1% in der Ukraine.

² Sofern nicht anders ausgewiesen ist mit „Nachfrage“ stets jene nach Flügen gemeint.



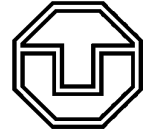
Eine sichere und effiziente Flugdurchführung erfordert eine übergeordnete Kontrollinstanz, welche den Luftverkehr überwacht und steuert. Diese Aufgaben übernehmen die Flugsicherungsdienstleister (Air Navigation Service Provider, ANSP). Sie sind u. a. für die Luftverkehrskontrolle (Air Traffic Control, ATC) und somit für eine konfliktfreie Durchführung kommerzieller und nicht-kommerzieller Flüge verantwortlich. Dadurch repräsentieren sie einen wesentlichen Bestandteil der Flugsicherungsdienste (Air Navigation Services, ANS). Diese Dienste umfassen neben der Steuerung des Luftverkehrs u. a. auch die Bereitstellung technischer Anlagen zur Ortung, Navigation und Kommunikation (ICAO 2006; Mensen 2014).

In den Leitlinien „Flightpath 2050“ werden die Ziele Europas zur Bewältigung des Luftverkehrsaufkommens der Zukunft formuliert (ACARE 2011). Neben ökologischen Forderungen werden auch Anforderungen an das europäische Luftverkehrsmanagement (Air Traffic Management - ATM) adressiert. EU-Projekte wie SESAR³ und CleanSky sollen die Umsetzung dieser Vision unterstützen. Langfristige Vorhersagen prognostizieren ein jährliches Wachstum der Anzahl der Flugbewegungen von bis zu 2,7% (EUROCONTROL 2013a). Zusätzlich erhöht der steigende Anteil der Low Cost Carrier (LCC) den Kostendruck der Airlines und, daraus folgend, der Flugsicherungsdienstleister (Buxbaum und Standfuss 2014).

Die Notwendigkeit eines effizienten Luftverkehrsmanagements führte in den letzten Jahren zu einem gesteigerten Interesse an einer Leistungsbewertung (Benchmarking, Performance Analysen) der europäischen Systeme. Dennoch, Performance-Analysen von ANSPs sind ein neues Feld in der Effizienzbewertung. Sowohl EUROCONTROL-Berichte (vgl. Abschnitt 3.2.1) als auch akademische Studien (vgl. Abschnitte 3.2.2) beschäftigen sich primär mit Teilaspekten auf einem hohen Aggregationslevel. Die Komplexität der Prozesse und Strukturen im Luftverkehr impliziert jedoch einen holistischen Ansatz um valide Maßnahmen abzuleiten.

Potentielle Maßnahmen zur Verbesserung der Leistungsfähigkeit umfassen ein weites Spektrum. Dazu gehören sowohl neue Technologien im Zuge der fortschreitenden Digitalisierung (Lütjens et al. 2018), eine Restrukturierung des europäischen Luftraumes (Gedes et al. 2018; Korn et al. 2010) oder alternative Finanzierungskonzepte (Verbeek und Visser 2016). Die Auswirkungen dieser spezifischen „Stellschrauben“ sind jedoch noch nicht wissenschaftlich untersucht, mögliche Interdependenzen somit unbekannt. Dies führt dazu, dass proklamierte Konzepte hinsichtlich des ökonomischen und betrieblichen Nutzens überprüft werden müssen, um eine reale Steigerung der Leistungsfähigkeit zu erreichen.

³ Single European Sky ATM Research.



1.2 Zielsetzungen der Arbeit

Die vorliegende Dissertation beschäftigt sich damit, wie Produktivität und/oder Effizienz einer Flugsicherung bzw. des Luftverkehrsmanagements insgesamt erhöht werden können, um die prognostizierten Verkehrszuwächse sowie deren Unsicherheiten⁴ zu bewältigen. Der Fokus liegt dabei auf den europäischen ANSPs.

Ziel der vorliegenden Dissertation ist ein Beitrag zu folgenden grundlegenden Forschungsfragen:

1. Wie kann die Leistungsfähigkeit⁵ in diesem speziellen Marktsegment gemessen werden?
2. Welche endogenen und exogenen Faktoren haben einen Einfluss auf die Performance?
3. Welche Maßnahmen und Handlungsempfehlungen sind abzuleiten, um die Effizienz zu steigern?

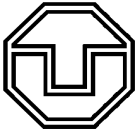
Gemäß der ersten Forschungsfrage ist es zunächst notwendig, eine geeignete Metrik für die Messung der Leistungsfähigkeit zu definieren. Effizienzanalysen im Luftverkehrsmanagement setzen eine Methodik voraus, welche multiple Inputs und Outputs berücksichtigen kann. In der Literatur werden u. a. deterministische und stochastische Verfahren beschrieben. Die verwendete Methode hängt maßgeblich von Analyseziel, Modellierung, betrieblicher Ebene und verfügbaren Daten sowie deren Qualität ab. Methodische Einschränkungen können zudem dazu führen, dass Inputs und/oder Outputs aggregiert werden müssen. Dabei ist zu klären, welche Parameter und Gewichtungsfaktoren zielführend sind. Schließlich wird untersucht, wie sich die Produktivität im Zeitablauf geändert hat. Diese multiperiodischen Betrachtungen ermöglichen die Analyse des Einflusses signifikanter globaler und lokaler Ereignisse, z. B. Vulkanausbrüche oder langfristige Luftraumsperrungen, auf die Performance innerhalb des europäischen ATMs sowie einzelner ANSPs.

Die zweite Forschungsfrage setzt sich mit den grundlegenden Einflussfaktoren auf die Performance auseinander. Hier sind zwei wesentliche Elemente zu unterscheiden. Zum einen endogene Effekte, die in der operativen Ausrichtung der ANSPs begründet sind. Hierbei ist sowohl die Zusammenwirkung der betrieblichen Ebenen innerhalb eines ANSP zu berücksichtigen als auch die ANSP-übergreifende Heterogenität. Zum anderen werden exogene Faktoren berücksichtigt. Diese umfassen u. a. Menge und Verteilung der Verkehrsnachfrage sowie politische und sozioökonomische Aspekte. Auch die Struktur des Luftraumes wird in dieser Forschungsfrage adressiert. Dabei wird bspw. untersucht, ob im europäischen ATM Größenvorteile existieren. Unter Annahme dieser Effekte ist seitens der Europäischen Kommission die Organisation der Lufträume in Functional Airspace Blocks (FABs) proklamiert und partiell bereits implementiert.

Auf Basis der Erkenntnisse aus den beiden ersten Forschungsschwerpunkten werden Handlungsempfehlungen, sowohl hinsichtlich der methodischen Umsetzung einer Leistungsbewertung im

⁴ Abweichung zwischen prognostiziertem und tatsächlichem Verkehrsaufkommen.

⁵ Der Begriff Leistungsfähigkeit wird in dieser Dissertation als Synonym für Performance bzw. Leistung verwendet.



Luftverkehrsmanagement als auch zur Steigerung der Leistungsfähigkeit und Kosteneffizienz einzelner Unternehmen abgeleitet.

Das Vorgehen ist schematisch in Abbildung 1-1 dargestellt. Basierend auf einer Leistungsbewertung der derzeitigen Technologie (Panelanalysen ermöglichen zudem die Betrachtung mehrerer vergangener Zeitabschnitte) können best practices identifiziert und die Performance der Flugsicherungen verglichen werden. Ausgehend davon kann mittels Ursachenanalyse bestimmt werden, welche Faktoren Einfluss auf die Performance haben. Basierend darauf können Maßnahmen und/oder Handlungsoptionen zur Performancesteigerung abgeleitet werden. Durch die Implementierung dieser Maßnahmen (z. B. Software, Systeme oder Verfahren) kann die Performance in der nächsten Periode gesteigert werden.

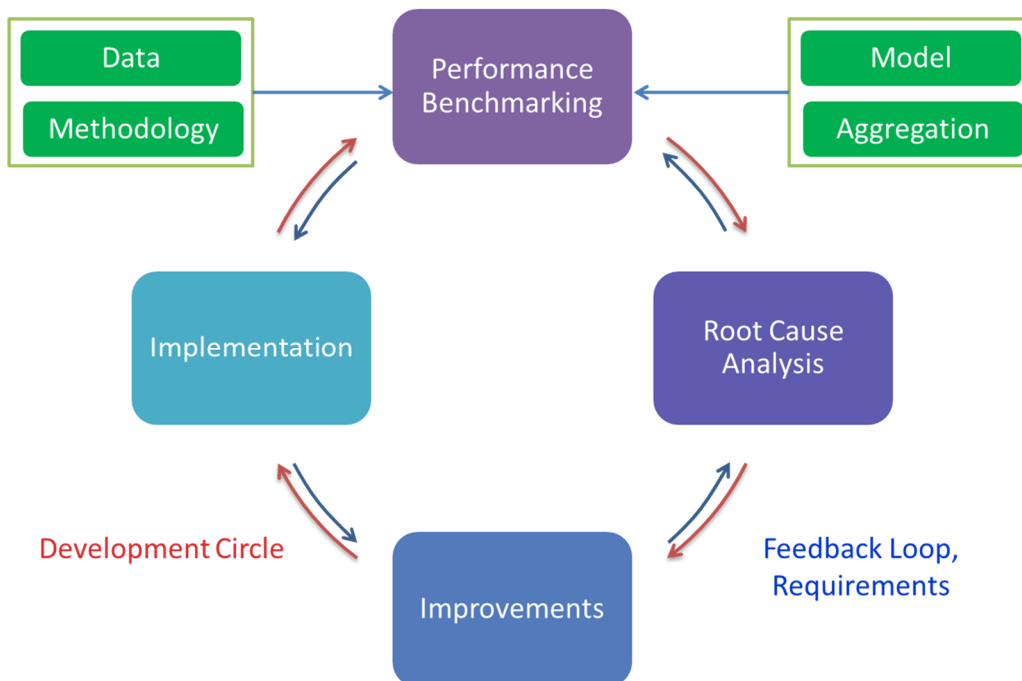
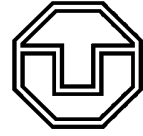


Abbildung 1-1: Development Circle und Feedback Loop

Die Implementierung valider Maßnahmen ist jedoch nur dann gewährleistet, wenn Performance Benchmarking und Ursachenanalyse alle wesentlichen Aspekte des betrieblichen Umfelds abdecken (Feedback Loop). Dies wiederum setzt die Nutzung adäquater Daten, eine geeignete Modellierung, differenzierter Aggregationsstufen sowie die spezifische Anpassung der verwendeten Methodik voraus.

1.3 Aufbau der Dissertation

Der holistische Ansatz der vorgelegten Dissertation berücksichtigt, wie im vorigen Abschnitt diskutiert, multiple Faktoren sowie deren Interdependenzen. Insbesondere die betriebliche Struktur eines ANSP erfordert die Anpassung von Methoden und Modellen. Aufbau und Funktionsweise sowie relevante Prozessketten haben ebenfalls einen direkten Einfluss auf die ökonomische Modellierung sowie die Methodenauswahl. Diese werden innerhalb der Dissertation kritisch diskutiert und verglichen.



Für die übergeordneten Ziele der Dissertation wird zunächst in Kapitel 2 der betriebliche und ökonomische Hintergrund einer Flugsicherung beleuchtet. Dadurch werden wesentliche Aspekte des ATM identifiziert, welche in Leistungsbewertung und Ursachenanalyse berücksichtigt werden müssen. Von einer makroperspektivischen Beschreibung aus wird die Betrachtungsebene sukzessiv auf den Beitrag der Flugsicherung zum System Luftverkehr fokussiert (Abschnitt 2.1). Dies beinhaltet auch die Finanzierung der Dienstleistungen sowie die innereuropäische Heterogenität (Abschnitte 2.2 bis 2.4). Zudem werden derzeitige und zukünftige Herausforderungen des ATM diskutiert (Abschnitt 2.5).

Zur Beantwortung der ersten Forschungsfrage werden in Kapitel 3 zunächst die ökonomischen Grundlagen der Leistungsbewertung herausgearbeitet (Abschnitt 3.1). Dies umfasst grundsätzliche Begriffe, Komponenten und Messmethoden der betrieblichen Leistungsfähigkeit. Darauf basierend werden zwei Methoden zur Berücksichtigung multipler Faktoren vorgestellt: Der nicht-parametrische, deterministische Ansatz der Data Envelopment Analysis (DEA) und die stochastisch-parametrische Stochastic Frontier Analysis (SFA). Abschließend werden die Methoden hinsichtlich Vor- und Nachteilen sowie Anwendbarkeit diskutiert.

Nach den ökonomischen Grundlagen wird in Abschnitt 3.2 der derzeitige Stand der Forschung aus betrieblicher und akademischer Sicht beschrieben. Darauf basierend werden die bisherigen Ansätze kritisch diskutiert und die Forschungslücke näher spezifiziert. Dies betrifft alle in Abbildung 1-1 illustrierten Bestandteile der Performance Analyse. Dabei werden insbesondere die verwendeten Daten sowie deren Qualität adressiert. Hier muss zunächst ermittelt werden, welche Daten auf welcher betrieblichen Ebene vorhanden sind. Diese werden schließlich in die ökonomische Modellierung überführt. Diese Datenauswertung wird in Abschnitt 3.3 adressiert und in Ressourcen und Dienstleistungen differenziert. Außerdem werden Aggregationsmöglichkeiten der einzelnen Faktoren diskutiert. Abschnitt 3.4 umfasst schließlich die Anwendung der Methode und Bewertung der Leistungsfähigkeit auf ANSP Ebene, inklusive Robustheitstests auf Basis verschiedener Modelle und Methoden. Die in Kapitel 3 gewonnenen Erkenntnisse werden in Abschnitt 3.5 zusammengefasst und darauf basierende Handlungsempfehlungen abgeleitet.

Die spezifischen Leistungsindikatoren werden durch verschiedene endogene, teil-exogene und exogene Ursachen beeinflusst. Die Gründe für Effizienz und Ineffizienz werden in Kapitel 4 analysiert. Zunächst wird der Ansatz dargestellt und mögliche Einflussgrößen diskutiert (Abschnitt 4.1). In Abschnitt 4.2 wird dann eine Ursachenanalyse auf ANSP Ebene durchgeführt. Diese umfasst die Nutzung verschiedener Methoden und Leistungsindikatoren. Dadurch wird sowohl der Einfluss von Faktoren als auch die Verwendbarkeit von Performance-Metriken untersucht.

Um die zeitliche Entwicklung zu berücksichtigen wird zudem eine Panel-Analyse durchgeführt (Abschnitt 4.3). Die verschiedenen Charakteristiken der betrieblichen Ebenen (siehe Abbildung 2-7) werden in einem disaggregierten Ansatz (Abschnitt 4.4) adressiert und auf Basis qualitativer und quantitativer Methoden verglichen. Abschnitt 4.5 untersucht die vermuteten Größenvorteile innerhalb des europäischen ATM. Weitere, nicht quantifizierbare Einflussgrößen werden in Abschnitt 4.6 diskutiert.



Abschnitt 4.7 beinhaltet die Erkenntnisse der Ursachenanalyse sowie die darauf basierenden Handlungsempfehlungen.

Kapitel 5 fasst die Erkenntnisse aus Leistungsbewertung und Ursachenanalyse zusammen und leitet potentielle Maßnahmen zur Steigerung der Leistungsfähigkeit ab. Dies umfasst sowohl betriebliche als auch politische Handlungsempfehlungen. In dem Kapitel werden zudem alternative Betriebskonzepte diskutiert. Dies ermöglicht einen Blick in die Zukunft sowie die Diskussion weiterer Forschungsschwerpunkte.

2 Grundlagen Air Traffic Management

2.1 Stakeholder des Luftverkehrs

Der heutige Luftverkehrsmarkt ist in Abbildung 2-1 dargestellt und umfasst neben den primär im Dienstleistungssektor tätigen Flughäfen, Fluggesellschaften und der Flugsicherung auch die Hersteller von Luftfahrzeugen, Triebwerken und anderem Equipment. Neben diesen direkt beteiligten Stakeholdern sind auch Unternehmen am Markt tätig, die lediglich mittelbar an der Leistungserstellung im Luftverkehr beteiligt sind. Diese indirekten Stakeholder sind aus Übersichtsgründen nur beispielhaft in der nachfolgenden Abbildung illustriert, profitieren vom Luftverkehr jedoch in umfangreichem Maße. Dies wird durch ökonomische Studien belegt, die dem Gesamtsystem Luftverkehr eine hohe Bedeutung im globalen Markt attestieren, u. a. (ATAG 2014).

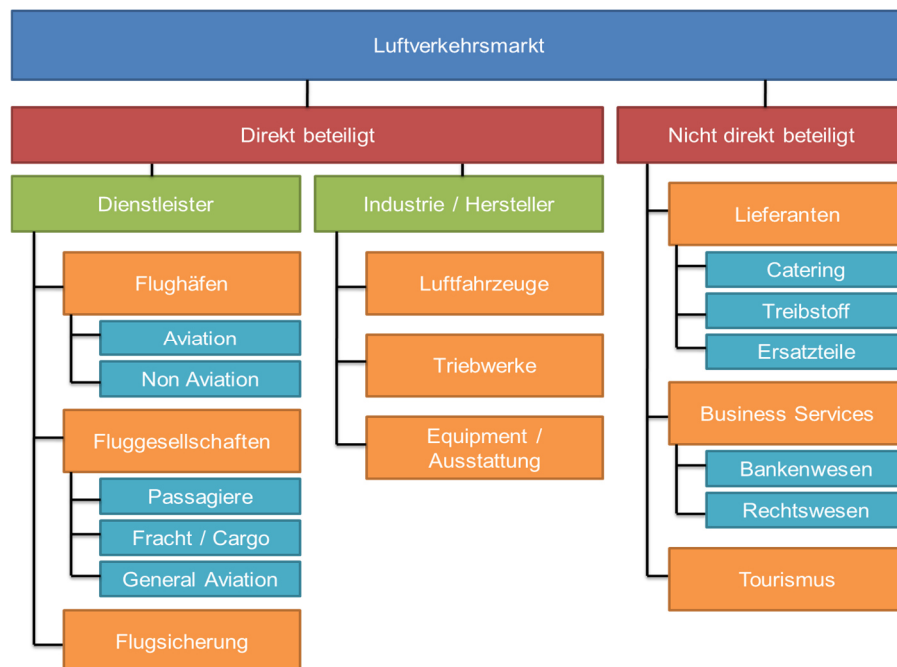
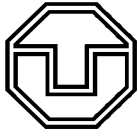


Abbildung 2-1: Stakeholder des Luftverkehrsmarktes
Angelehnt an ATAG (2005)

Insbesondere Studien von Industrievertretern unterscheiden die volkswirtschaftliche Bedeutung⁶ des Luftverkehrs in direkte, indirekte, induzierte und katalytische Effekte. Die direkten Einflüsse werden von allen unmittelbar am Luftverkehr beteiligten Stakeholdern repräsentiert. Vergeben diese Unternehmen Aufträge für die Erstellung einer Leistung, hat dieser Vorgang einen indirekten ökonomischen Einfluss. Beispiele hierfür sind Zulieferer von Einzel- und Ersatzteilen für Luftfahrzeuge, Catering- und Treibstofflieferanten sowie Dienstleister im Hard- und Softwarebereich. Zudem existieren indirekte Effekte für die Leistungserbringung im juristischen, kommunikativen, finanziellen oder betriebswirtschaftlichen Kontext⁷. Die induzierten Effekte repräsentieren Dienstleistungen und Güter, die aufgrund des

⁶ Damit ist bspw. die Verminderung der Kosten der Raumüberwindung und Arbeitskosten gemeint.

⁷ Hierunter zählen u. a. Versicherungen, Buchhalter, Callcenter, etc.

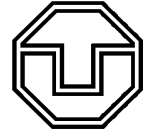


Flugbetriebes benötigt werden bzw. aus der Konsumnachfrage resultieren. Als Beispiel sei hier die Beförderung der Passagiere zum Flughafen mittels ÖPNV genannt. Während direkte, indirekte und induzierte Effekte infolge der Erstellung der Luftverkehrsdienstleistung entstehen, resultieren katalytische Effekte aus der Verwendung des Luftverkehrs und damit aus dem durch ihn entstehenden Standortvorteil. Diese Effekte betreffen insbesondere den Handel, den Tourismus oder regionale Investitionen und damit auch Wohlstand und Produktivität, aber auch Umweltfaktoren wie z.B. Emissionen (ATAG 2005; Hasche et al. 2008).

Diese Einteilung wird jedoch, insbesondere hinsichtlich der katalytischen Effekte, kritisch hinterfragt (Forsyth und Niemeier 2016; Niemeier 2001). Die Wertschöpfung von Volkswirtschaften profitiert von überregionaler bzw. internationaler Arbeitsteilung und die dadurch geringeren Produktionskosten. Die volkswirtschaftliche Bedeutung eines effizienten Verkehrssystems liegt u. a. in geringeren Transportkosten, welche unterstützend auf die Arbeitsteilung wirken (Aberle 2014; Biebig et al. 2017). Das heißt, je effizienter der Transport von Gütern organisiert werden kann, desto größer sind die Vorteile der Arbeitsteilung für die Volkswirtschaften. Ein wesentlicher Vorteil des Luftverkehrs kann daher in den geringeren Raumüberwindungskosten gesehen werden.

Der zivile Luftverkehrsmarkt kann als Vier-Säulen-Konstrukt bestehend aus Herstellern, Flughäfen, Flugsicherungen und Fluggesellschaften interpretiert werden. Flugzeughersteller, wie Boeing, Airbus oder Embraer, konstruieren und entwickeln Luftfahrzeuge. Diese werden an Kundenwünsche angepasst, z. B. hinsichtlich der Sitzplatzkonfiguration oder der Reichweite, sodass das Unternehmen für jedes angestrebte Marktsegment seine Produkte anbieten kann. Die Luftfahrzeuge werden von den Fluggesellschaften für den Betrieb von Verbindungen im Passagier- und Frachtbereich genutzt. In den vergangenen Jahren haben sich hier zwei Hauptgeschäftsmodelle herausgebildet: Low Cost Carrier und die Full Service Network Carrier (FSNC). Diese unterscheiden sich maßgeblich in der Kostenstruktur und den betrieblichen Konzepten (Forsyth et al. 2016a). Verschiedene Betriebskonzepte haben wiederum einen direkten Einfluss auf die Verkehrsströme. Durch das Hub-and-Spoke Netz werden Verkehre zwischen Großflughäfen gebündelt, während das Betriebskonzept der LCCs zu einer dezentralisierten Struktur führt. Durch die verschiedenen Antriebskonzepte⁸ der Flugzeuge unterscheiden sich die Geschwindigkeiten im Streckenflug sowie die Flughöhen. Die unterschiedliche Größe der Luftfahrzeuge führt zudem zu verschiedenen Staffelungsabständen beim Anflug auf den Flughafen. Dadurch haben diese Faktoren auch einen direkten Einfluss auf das Luftverkehrsmanagement und damit auf die Dienstleistungen der Flugsicherung. Durch die Etablierung der LCCs stieg auch der Kostendruck und damit die Notwendigkeit, eine Strecke bzw. Verbindung möglichst kostengünstig zu betreiben. Hierfür ist die Wahl eines geeigneten Fluggerätes ebenso wichtig wie eine wind- und gebührenoptimierte Flugroute (Rosenow et al. 2015). Darüber hinaus sind Fluggesellschaften bestrebt, den Turnaround-Prozess zu verbessern. Dies umfasst u. a. Boarding-Strategien, aber auch generell alle Vorgänge zwischen der Ankunft des Luftfahrzeuges am Gate und dem Push-Back (Rosenow und Schultz 2018; Schultz und Reitmann 2019). Auch hier können Überschneidungen zum ATM gesehen werden.

⁸ Bspw. kann man hier Turboprop-Triebwerke von Jet-Antrieben unterscheiden.



Für den Verkehr am Boden, d. h. zwischen der Landung eines Luftfahrzeuges und dem erneuten Start, greifen Airlines auf die Infrastruktur des Flughafens zurück. Der Flughafen stellt alle notwendigen Dienstleistungen und Geräte sowie Infrastruktur bereit, die zur Abfertigung eines Luftfahrzeugs am Boden, zur Betreuung der Passagiere oder zur Abwicklung der Fracht notwendig sind. Flughäfen weisen meist monopolistische Strukturen auf, weshalb sie staatlich reguliert werden (Forsyth et al. 2016b, 2017). Hauptkunden des Flughafens können FSNCs und/oder LCCs sein. Dies hat einen direkten Einfluss auf zeitliche Bündelungen des Verkehrs („Morning-“ und „Evening-Peak“). Außerdem unterscheiden sich Flughäfen hinsichtlich der architektonischen Charakteristik, wie z. B. die Anzahl und Konfiguration⁹ der Start- und Landebahnen. Dementsprechend werden standardisierte An- und Abflugrouten konzipiert. Sowohl die Architektur des Flughafens als auch die zeitliche Konzentration von Flügen haben wiederum einen direkten Einfluss auf die Flugsicherungsdienstleistungen.

Die primäre Aufgabe der Flugsicherung (Säule 4) ist die kollisionsfreie Flugführung. Dafür steht den zuständigen Fluglotsen (Air Traffic Control Officer, ATCO) ein Portfolio an technischen Hilfsmitteln zur Verfügung, um potentielle Konflikte frühzeitig zu erkennen. Die Interaktion mit anderen Stakeholdern erfolgt primär zu den Airlines und zu den Flughäfen. Obwohl die Flugsicherungen weltweit grundsätzlich nach ähnlichen Prinzipien arbeiten, gibt es doch signifikante Unterschiede hinsichtlich Prozessabläufen, Verfahren, Unternehmensstruktur und Betriebskonzepten (FABEC 2019; Mensen 2014). Im beschriebenen Marktumfeld bildet die Flugsicherung einen Sonderfall. Während Hersteller, Fluggesellschaften und teilweise auch Flughäfen innerhalb eines Wettbewerbsmarktes agieren, sind ANSPs ökonomisch gesehen ein Beispiel für natürliche Monopole. Der Staat ist in einem solchen Fall bestrebt, regulatorische Maßnahmen mit dem Ziel zu ergreifen, die aus dem Marktversagen resultierenden negativen Effekte zu minimieren. Jeder ANSP ist gleichzeitig eine sogenannte „Essential Facility“ (zu den wirtschaftstheoretischen Hintergründen dieser Doktrin siehe Hagenfeldt (2011), Keitel (2015) sowie Pfähler et al. (1999)). Ein Wettbewerb im Markt ist in diesem Sektor entweder nicht möglich oder nicht sinnvoll. Der volkswirtschaftliche Nutzen eines Wettbewerbs innerhalb des europäischen Luftverkehrsmanagements wird jedoch in einer Studie von Ewers und Tegner (2002) diskutiert.

Die Zusammenarbeit der einzelnen Stakeholder hat einen entscheidenden Einfluss auf die effiziente, zuverlässige und sichere Durchführung von Flügen. Ein ANSP kann durch seine Entscheidungskompetenz sowie durch die begrenzte, physisch nicht erweiterbare Kapazität des Luftraumes als ein Bottleneck des derzeitigen und zukünftigen Luftverkehrssystems gesehen werden, sodass bei Prozessverbesserungen im Flugsicherungsbereich auch positive Auswirkungen für andere Stakeholder, primär die Airlines, zu erwarten sind. In der vorliegenden Arbeit wird daher die Flugsicherung fokussiert. Für einen ausführlichen Einblick in die Finanzierung, Organisation und Politik des Luftverkehrsmanagements siehe Oster und Strong (2017).

⁹ Z. B. parallele oder kreuzende Bahnen.



2.2 Luftverkehrsmanagement in Europa

Aufgrund der staatlichen Lufthoheit betreibt jeder europäische Staat eine eigene Flugsicherung. Eine wesentliche Ausnahme bildet der obere Luftraum über den Beneluxstaaten sowie dem Nordwesten von Deutschland. Dieser wird durch die internationale Organisation Maastricht Upper Area Control Centre betreut. Die Grenzen der Lufträume orientieren sich dabei an jenen des zugehörigen Staatsgebietes. Dementsprechend unterscheiden sich auch die Lufträume signifikant in ihrer Größe (Abbildung 2-2). Den größten europäischen Luftraum verwaltet die spanische ENAIRE¹⁰. Dieser ist etwa 107-mal größer als der slowenische Luftraum (EUROCONTROL 2019a, 2020h).

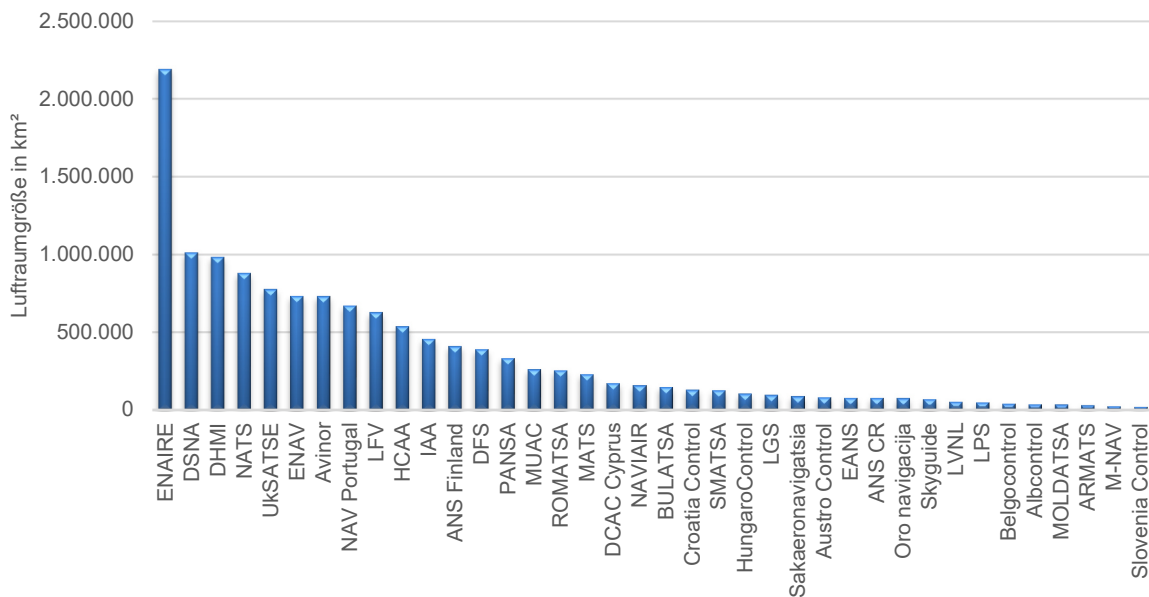


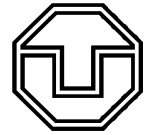
Abbildung 2-2: Luftraumgröße europäischer Staaten (2017)
Quelle: PRU

Der europäische Luftraum kann demnach je nach zugehörigem Staat bzw. der korrespondierenden Flugsicherung untergliedert werden. Die International Civil Aviation Organization (ICAO) teilt den Luftraum in sogenannte Flight Information Regions (FIRs), wie in Abbildung 2-3 dargestellt ist. Ein Staat bzw. ein ANSP kann aus einer oder mehreren FIRs bestehen. Diese sind horizontal oder, wie im Beispiel MUAC, vertikal¹¹ angeordnet (EUROCONTROL 2019c, 2019d). Eine weitere Möglichkeit besteht in der betrieblichen Aufgliederung hinsichtlich Kontrollzentralen oder Sektoren (siehe dazu Abschnitt 2.3).

Die in Abbildung 2-3 eingefärbten Staaten repräsentieren die 41 Mitglieder der internationalen Organisation EUROCONTROL. Diese ist verantwortlich für die zentrale Koordinierung des europäischen Luftverkehrs. Die Leistungen umfassen: „*Betrieb, Leistungserbringung, Konzeptentwicklung, For-*

¹⁰ Zu beachten ist, dass im Gegensatz zu Spanien andere Lufträume keine „Oceanic Areas“ beinhalten, u. a. Großbritannien (NATS) oder Portugal (NAV Portugal).

¹¹ In diesem Fall eine „Upper Information Area“, UIR.



schung, europaweite Projektumsetzung, Leistungsverbesserungen, Koordination mit wichtigen Luftfahrtakteuren auf verschiedenen Ebenen sowie Unterstützung für die zukünftige Entwicklung und strategische Ausrichtung der Luftfahrt“ (EUROCONTROL 2020a)¹².



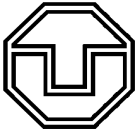
Abbildung 2-3: Flight Information Regions, Unterer Luftraum, 2019
(EUROCONTROL 2019c)

EUROCONTROL publiziert u. a. Berichte zu prognostizierten Verkehrszahlen (EUROCONTROL 2019i) sowie Performance-Vergleiche der Flugsicherungen (EUROCONTROL 2019a, 2019h). Zudem stellt sie umfangreiche Daten (z. B. Flüge, Verspätungen, etc.) zur Verfügung, die auch in der vorliegenden Dissertation genutzt werden. Die Aufgaben einiger EUROCONTROL-Funktionseinheiten sind in den folgenden Abschnitten näher erläutert.

2.3 Betriebliche Aspekte

Flugsicherungsunternehmen sind im Wesentlichen in drei Tätigkeitsfeldern aktiv, welche in Abbildung 2-4 dargestellt sind. Die zentrale Aufgabe wird durch das Air Traffic Management repräsentiert und umfasst neben den Air Traffic Services (ATS) das Airspace Management (ASM) und das Air Traffic Flow Management (ATFM).

¹² Übersetzt durch den Autor.



Das operative Geschäft einer Flugsicherung befindet sich primär im Feld der Air Traffic Services. Ziel der sich darin befindenden Elemente ist eine kollisionsfreie und effiziente Flugdurchführung. Dies wird durch Kontrolle und Führung des Fluges durch Fluglotsen (Bereich ATC) gewährleistet, die primär für kommerzielle Flüge nach Instrumentenflugregeln (Instrument Flight Rules, IFR) verantwortlich sind. Des Weiteren kommunizieren Lotsen des Fluginformationsdienstes (Flight Information Service, FIS) auf gesonderten Frequenzen und sind bspw. für Verkehrsinformationen, speziell bei Flügen nach Sichtflugregeln (Visual Flight Rules, VFR), oder für navigatorische Unterstützung verantwortlich.

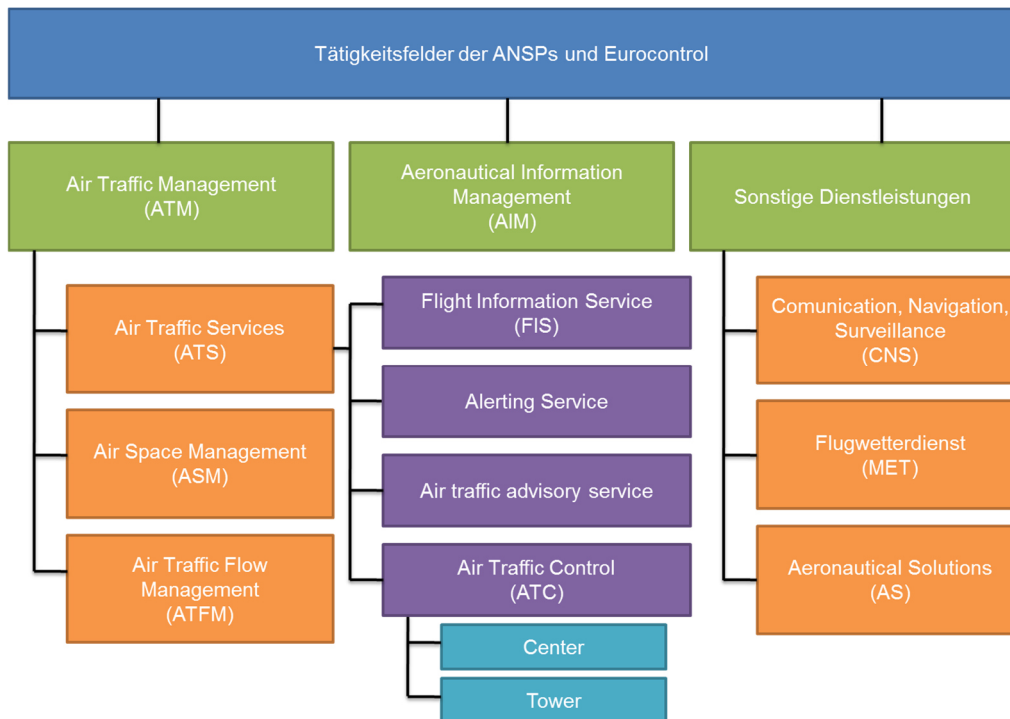


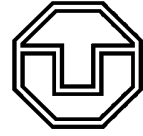
Abbildung 2-4: Tätigkeitsbereiche einer Flugsicherung und von EUROCONTROL
Angelehnt an Kreuz (2015)

Im Bereich ATC ist zudem eine Differenzierung zwischen Tower und Center notwendig. Die Luftraumüberwachung wird in drei Kategorien unterteilt: Die Streckenkontrolle (Center), die An- und Abflugkontrolle (Center oder Tower¹³) sowie die Überwachung für den Bereich des Flughafens selbst (Tower). Je nach Flugphase (Abbildung 2-5) sind die jeweiligen Lotsen zuständig.



Abbildung 2-5: Phasen eines Fluges
(Albright 2015)

¹³ Hier existiert keine homogene Zuordnung innerhalb Europas.



In der Kontrollzentrale wird der Streckenverkehr sowie in den meisten europäischen ANSPs der An- und Abflugverkehr kontrolliert. Die Center-Lotsen sind verantwortlich für:

- die Zuweisung von Luftstraßen und/oder „Tactical Directs“¹⁴
- die Zuweisung von Flughöhen bzw. Flugflächen¹⁵ sowie Steig- und Sinkraten
- die Zuweisung von Geschwindigkeiten
- Anweisung für Warteschleifen (Holdings)
- die vertikale und horizontale Staffelung
- Informationen über aktive Landebahnen am Zielflughafen sowie die Zuweisung des zu nutzenden Standard-Anflug-Verfahrens

Die Kollisionsprävention in der Luft wird durch vertikale und horizontale Staffelung der Flugzeuge realisiert. Je nach Flughöhe beträgt die vertikale Staffelung 1.000 bis 2.000 Fuß (ICAO 2016). Die zugewiesene Flughöhe ist dabei maßgeblich abhängig von der Flugrichtung des Flugzeuges. Somit wird verhindert, dass sich zwei entgegenkommende Luftfahrzeuge auf derselben Flugfläche befinden. Die horizontale Staffelung sorgt für den Abstand zwischen Maschinen auf der gleichen Flugfläche, in gleicher Flugrichtung, sowie beim Start oder im Landeanflug. Eine Staffelung ist insbesondere beim Anflug auf den Flughafen wichtig, da Luftfahrzeuge mit höheren Anstellwinkeln stärkere Wirbelschleppen verursachen. Diese Art der Staffelung wird auch als Wirbelschleppenstaffelung bezeichnet (ICAO 2016).

Die Lotsen im Tower sind verantwortlich für den Flughafen-Bereich¹⁶ sowie den umliegenden Luftraum. Im Gegensatz zum Center haben Towerlotsen in der Regel direkten Sichtkontakt zu den startenden und landenden Luftfahrzeugen. Sie nutzen dennoch Radar-Technik für zusätzliche Informationen über den Verkehr. Im Verantwortungsbereich der Tower-Lotsen liegen:

- Die Freigabe des aufgegebenen Flugplanes sowie Zuweisung einer ersten Flughöhe bzw. Flugfläche
- Die Freigabe zum Start der Triebwerke
- Die Zuweisung der Startbahn und (beim IFR-Verkehr) der zu nutzenden Abflugroute
- Die Freigabe für das Line-up und/oder den Start
- Die Landefreigabe
- Ggf. Anweisung zum Durchstarten eines anfliegenden Luftfahrzeuges, sofern eine sichere Landung nicht gewährleistet ist
- Ggf. Anweisung des zu nutzenden Abrollweges

Die Towerlotsen sind zudem verantwortlich für das Einsprechen eines Wetterberichtes¹⁷, welcher durch die Piloten im IFR-Verkehr abgehört werden muss. Dieser Bericht umfasst u. a. Informationen über Wind, Sichtweite, Taupunkt und aktive Bahnen.

¹⁴ Unter Umständen kann die im Flugplan aufgeführte Route abgekürzt werden. Man spricht dann von Directs, da das Luftfahrzeug direkt zu einem bestimmten Wegpunkt fliegen darf.

¹⁵ Flugflächen entsprechen einer Flughöhe in Fuß, bei der die letzten beiden Nullen weggelassen werden. Flugfläche 300 entspricht 30.000 Fuß.

¹⁶ Primär für die Start- und Landebahnen sowie einen Teil der Rollbahnen.

¹⁷ Automatic Terminal Information Service (ATIS).



Die Unterscheidung zwischen beiden Dienstleistungen ist auch im ökonomischen Sinne bedeutsam. So werden u. a. in EUROCONTROL-Berichten neben der Gate-to-Gate Perspektive (gesamter Flug) auch Enroute Services (Center) und Terminal Services (Tower) getrennt betrachtet. Analog werden auch unterschiedliche Maßeinheiten für den erzeugten Output verwendet (siehe Abschnitt 3.3.2).

EUROCONTROL übernimmt im Bereich des Air Space Managements die Planung des Luftraumes sowie dessen Aufteilung in unterschiedliche Regionen oder Klassen. An die Luftraumstruktur wird vor allem der Anspruch auf Flexibilität gestellt, sodass auf kurzfristig eintretende Situationen reagiert werden kann. Dies umfasst u. a. die Konzepte des Flexible Use of Airspace (FUA) und des Free Route Airspace (FRA) (EUROCONTROL 2015; Rosenow et al. 2018b). Diese Flexibilität trägt zu einer Effizienzsteigerung bei, setzt aber die Mitwirkung einer möglichst hohen Anzahl an Flight Information Regions voraus, um das Design an die europäischen Verkehrsströme anzupassen (Schwenk und Schwenk 1998).

Verkehrsströme orientieren sich an Luftstraßen und konzentrieren sich vor allem zwischen Gebieten mit Großflughäfen wie Frankfurt, London oder Paris. Das ATFM sorgt dabei für eine effiziente und sichere Nutzung der spezifischen Luftraumkapazität. Dadurch soll vermieden werden, dass das Verkehrsaufkommen die Kapazität übersteigt. Die Aufgaben werden nach einem zeitlichen Horizont differenziert. Dieser wird in taktische, prätaktische und strategische Prozesse unterschieden.

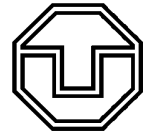
Auf strategischer Ebene (sechs Monate bis zwei Tage vor dem Flug) umfassen die Maßnahmen u. a. die Bestimmung der Verkehrsnachfrage sowie die Prognose über deren Verteilung (Traffic Flow). Ihr schließt sich die prätaktische Ebene an. Sie umfasst die beiden Tage vor dem Flug. Basierend auf der strategischen Planung und den bis dahin verfügbaren Informationen werden regulatorische Maßnahmen¹⁸ für den Verkehrsfluss eingeleitet und mit den Flow Managern der einzelnen Standorte, respektive Kontrollzentralen, abgestimmt. Zudem werden diese Informationen an die Luftfahrtnutzer, bspw. die Airlines, weitergeleitet und ggf. auf Basis von Slot¹⁹-Berechnungen Alternativrouten vorgeschlagen (EUROCONTROL 2020n).

Die taktische Phase umfasst den Tag des Fluges selbst. Sie beinhaltet insbesondere die Slot-Allokation, die jedoch aufgrund der hohen Dynamik im Luftverkehr (Wetterereignisse, Kapazitätsengpässe, Flugstreichungen etc.) stetig angepasst werden muss. Nach der errechneten Startzeit des Fluges gilt die Kapazität als gebunden und der Flug wird an die Luftverkehrskontrolle übergeben (Daduna und Voß 2000).

Zur sicheren und effizienten Flugdurchführung ist neben den Verkehrsplanungs- und -steuerungsmaßnahmen auch ein Informations- und Datenmanagement notwendig. Diese Aufgabe übernimmt das Aeronautical Information Management (AIM). In dessen Rahmen werden die luffahrtechnischen Daten (z. B. Kartenmaterial) für alle Flugphasen zur Verfügung gestellt. Dies ist vor allem aufgrund der fortschreitenden Digitalisierung notwendig, welche neue Verfahren (z. B. Raumnavigation) ermöglicht

¹⁸ Z. B. Änderungen der Flugroute oder die Neuzuweisung eines Slots.

¹⁹ Gemeint sind hier ATM-Slots, keine Airport-Slots.



und dadurch zu einem erhöhten Bedarf an Datenaustausch führt (EUROCONTROL 2016a; Lütjens et al. 2018). Die ATM- und ATS-Dienstleistungen repräsentieren sowohl den Kern einer Flugsicherung als auch den Schwerpunkt der vorliegenden Dissertation. Die weiteren Tätigkeitsfelder können gemäß Abbildung 2-4 als ergänzende Geschäftsfelder verstanden werden.

Der Luftraum ist eine gemeinschaftlich genutzte Ressource, u. a. für Airlines, Militär und Freizeitverkehr. Um eine Überwachung des Luftraumes zu gewährleisten, bedarf es einer Strukturierung nicht nur hinsichtlich von Luftraumklassen (siehe Anhang A1), sondern auch im Rahmen von betrieblichen Prozessen. Im Bereich Streckenflug sind die Kontrollzentralen (Center, vgl. Abbildung 2-4) verantwortlich. Die Streckenkontrolle ist nach fünf operativen Ebenen zu differenzieren. Diese Ebenen sind beispielhaft in Abbildung 2-6 dargestellt.

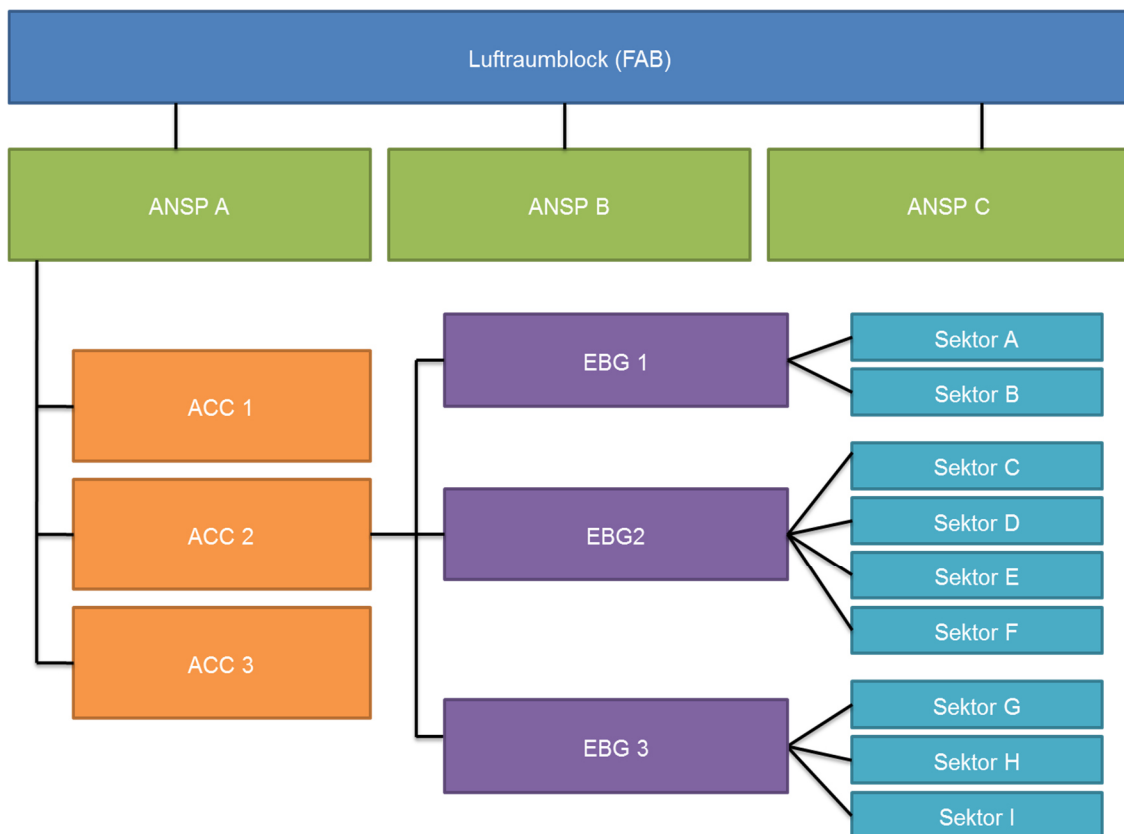
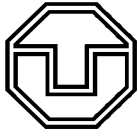


Abbildung 2-6: Betriebliche Gliederung der Luftverkehrskontrolle

Die unterste betriebliche Ebene wird durch die Sektoren repräsentiert. Geographisch sind sie sowohl horizontal als auch vertikal begrenzt. Pro Sektor sind zwei Fluglotsen aktiv, unterstützt durch technisches Equipment zur Ortung (z. B. Radarschirm), Organisation (z. B. Kontrollstreifen) und Kommunikation. Während ein Lotse die Kommunikation mit den Luftfahrzeugen übernimmt (Executive-ATCO), koordiniert der zweite Lotse den Verkehr mit den umliegenden Sektoren (Planner-ATCO)²⁰. Form und Lage der Sektoren sind statisch. Diese Grundstruktur ist daher nur mittel- bis langfristig an-

²⁰ Die Aufteilung der Aufgaben und auch die Anzahl der Lotsen pro Arbeitsplatz können abweichen.



passbar. Kurzfristig können jedoch, je nach Verkehrsaufkommen, Sektoren kombiniert oder geteilt werden und so die angebotene Kapazität an die Nachfrage der Luftraumnutzer angepasst werden. Gemäß Abbildung 2-6 könnten bspw. Sektoren C und D kombiniert werden. Die möglichen Kombinationen werden in Sektor-Konfigurationen zusammengefasst und sind u. a. über das Network Strategic Modelling Tool (NEST) der EUROCONTROL abrufbar. Kann ein Sektor nicht weiter geteilt werden, so wird er als Elementarsektor bezeichnet.

Sektoren weisen sehr spezifische Charakteristiken auf. Daraus resultiert eine hohe Spezialisierung des Lotsen, da dieser gezielt z. B. auf geographische Besonderheiten oder operative Verfahren geschult werden muss. Dies wiederum bedeutet, dass Lotsen nur in bestimmten Regionen, den Sektoren einer spezifischen Einsatzberechtigungsgruppe (EBG), eingesetzt werden können. In Bezug auf (Abbildung 2-6) kann ein Lotse des Sektor C auch in den Sektoren D, E und F eingesetzt werden, jedoch bspw. nicht in den Sektoren A und B. Außerdem können nur Sektoren kombiniert werden, welche der gleichen EBG zugeordnet sind.

Auf der dritten Ebene umfasst ein Area Control Center (ACC) meist mehrere Einsatzberechtigungsgruppen. Jeder ANSP kann die Anzahl und Standorte dieser Kontrollzentralen selbst bestimmen. Auch hinsichtlich der vertikalen Lufträume bzw. Flugflächen können die ACCs unterschieden werden. Teilweise sind sie nur für den oberen (z. B. MUAC) oder unteren (z. B. LVNL) Luftraum zuständig, teilweise decken sie beide Lufträume ab (Säulenstruktur, z. B. Skyguide). Die Deutsche Flugsicherung betreibt derzeit vier Kontrollzentralen in Langen, Karlsruhe, Bremen und München.

Der Air Navigation Service Provider bildet unternehmensintern die höchste betriebliche Ebene. Zum ANSP gehören sowohl die Tower als auch die Center mit dem jeweiligen Equipment, dem Personal und allen technischen Anlagen. Auch die für die Dienstleistung erhobenen Gebührensätze sind, bis auf wenige Ausnahmen, ANSP-weit gleich. Die Konsolidierung mehrerer ANSPs zu sogenannten Luftraumblocken wurde im Zuge des Single European Sky (SES) konzipiert. Sie können als höchste betriebliche Ebene bezeichnet werden (vgl. Abschnitt 2.5).

Jede betriebliche Ebene verfolgt eigene Zielfunktionen und berücksichtigt individuelle Determinanten. Abbildung 2-7 stellt den einzelnen Ebenen Beispiele ihrer individuellen Aspekte und Zielfunktionen gegenüber. Während bei den Sektoren (Mikroebene) primär Kapazitätsaspekte eine Rolle spielen, entscheiden ACCs u. a. über verwendete Tools oder die Personaleinsatzplanung. In der Makroebene wiederum wird die Aufteilung (ANSP) oder Konsolidierung von Lufträumen (FAB) adressiert, z. B. um Größenvorteile zu realisieren. Der holistische Ansatz der Effizienzmessung setzt voraus, dass nicht nur die Heterogenität zwischen einzelnen Einheiten einer betrieblichen Ebene (bspw. zwischen der deutschen und der polnischen Flugsicherung) berücksichtigt wird, sondern auch die Unterschiede zwischen den betrieblichen Ebenen (z. B. ANSP versus Sektor) in das ökonomische Modell einfließen.

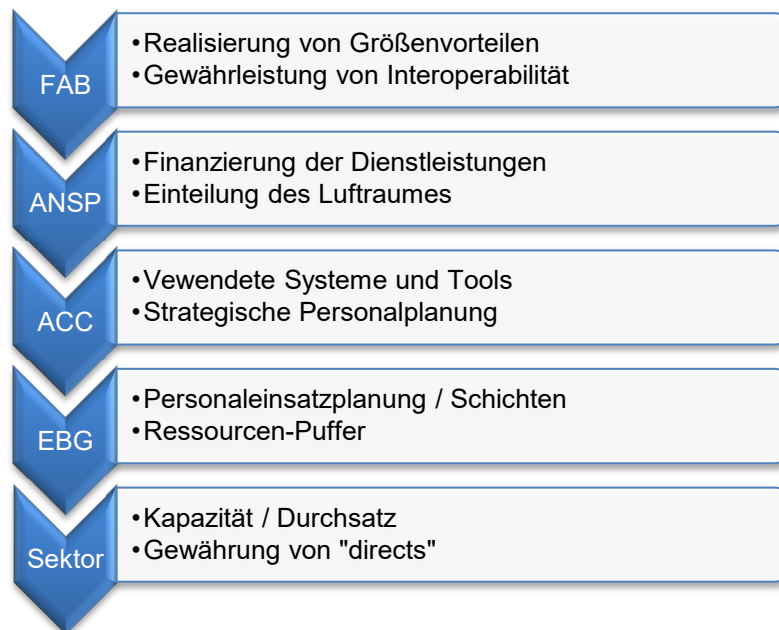
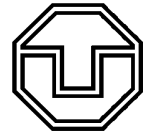


Abbildung 2-7: Betriebliche Ebenen und operative Ziele (Auswahl)

2.4 Finanzierung europäischer Air Navigation Services

Während die Flugsicherung in den USA größtenteils steuerfinanziert ist (Whittome und Standfuss 2018), müssen in Europa Luftraumnutzer für Flüge nach Instrumentenflugregeln Gebühren für die Nutzung der Lufträume zahlen. Die Grundlage der Finanzierung bildet European Commission (2010a). Die Gebührensätze sind dabei von Land zu Land, bzw. von ANSP zu ANSP unterschiedlich. Teilweise gibt es auch innerhalb eines ANSP unterschiedliche Gebührensätze²¹. Die Höhe der Gebühr R für einen Flug berechnet sich aus der Summe an Einzelgebühren r_i der zuständigen ANSPs. Die spezifische ANSP-Gebühr besteht aus drei Einzelkomponenten, wobei:

- d_i die Großkreisdistanz zwischen Startflughafen oder Einflugpunkt und dem Zielflughafen oder Ausflugpunkt innerhalb des Luftraumes des ANSP (angegeben in 100 km),
- $MTOW$ ²² das maximale Startgewicht des verwendeten Luftfahrzeuges in (angegeben in t),
- t_i dem derzeitigen Gebührensatz des ANSP pro Dienstleistungseinheit²³ (DLE) in € entspricht.

$$R = \sum_{i=1}^n r_i \quad (2-1)$$

$$r_i = d_i \cdot \sqrt{\frac{MTOW}{50}} \cdot t_i \quad (2-2)$$

²¹ Z. B. Portugal.

²² Maximum Take Off Weight.

²³ Eine Dienstleistungseinheit entspricht den ersten beiden Faktoren der Formel (2-2).



Beispielsweise hätte die Streckengebühr bei der Nutzung eines Airbus A320-200 ($MTOW = 70\text{ t}$) auf einem Flug von Hamburg nach München ($d_i = 600\text{ km}$) im Juni 2016 ($t_i = 82,68\text{ €}$) $586,97\text{ €}$ betragen. Die DFS wies im genannten Jahr nach der Schweiz und Großbritannien den dritthöchsten Gebührensatz²⁴ auf, Bulgarien, Georgien und die Türkei die niedrigsten²⁵. Eine Sonderrolle nimmt die FIR Portugal Santa-Maria ein, die für den portugiesisch-atlantischen Luftraum verantwortlich ist. Eine Dienstleistungseinheit wurde hier mit $10,89\text{ €}$ berechnet (EUROCONTROL 2016d).²⁶

$$r_i = d_i \cdot \sqrt{\frac{MTOW}{50}} \cdot t_i = \frac{600}{100} \cdot \sqrt{\frac{70}{50}} \cdot 82,68\text{ €} = 586,97\text{ €}$$

Die unterschiedlichen Gebühren können auf unterschiedlich hohe Kosten der jeweiligen Flugsicherungen zurückgeführt werden. Insbesondere die stark frequentierten Lufträume erfordern einen höheren Personaleinsatz, besseres Equipment hinsichtlich Hard- und Software sowie ein aufwändigeres Training der Fluglotsen. Ein weiterer Grund sind die in Europa unterschiedlichen Lohnsätze sowie der Einfluss von Gewerkschaften. Dadurch unterscheiden sich die Kosten pro Lotsenstunde signifikant. Die unterschiedlichen Gebührensätze weisen den entscheidenden Nachteil auf, dass u. U. den Fluggesellschaften Anreize induziert werden, Umwege zu fliegen und damit ihre Kosten zu minimieren (vgl. Abschnitt 2.5). Beispielsweise ist der kroatische Luftraum ein oft genutztes Substitut des italienischen Luftraumes. Der Gebührensatz der Croatia Control ist nur etwa halb so hoch wie jener der italienischen Flugsicherung ENAV. Die negativen externen Effekte durch diese Umwege sind vor allem höhere Emissionen und ein höherer Treibstoffverbrauch.

Die Gebühren sind lediglich vom IFR-Verkehr zu entrichten. Die Leistungen der Streckenkontrolle sind für den VFR-Verkehr gebührenfrei. Zudem werden für medizinische Transporte sowie die Flüge von Staats- und Regierungsoberhäuptern keine Gebühren erhoben. Für die militärische Luftfahrt gilt ebenfalls das Kostendeckungsprinzip. Der Gesamtaufwand der Dienstleistungen wird durch die DFS ermittelt und von der Bundeswehr erstattet. Das Central Route Charges Office (CRCO) von EUROCONTROL berechnet und fakturiert die Streckengebühr. Das heißt, der Luftraumnutzer zahlt einmalig pro durchgeführten Flug. Diese Einnahmen werden dann anteilig an die jeweiligen Staaten bzw. Flugsicherungen überwiesen (EUROCONTROL 2019j). Die Gebühren werden jedoch nicht entsprechend der im Flugplan²⁷ aufgegebenen Route berechnet, sondern nach der tatsächlich geflogenen.

Neben den Streckengebühren fallen zudem die An- und Abfluggebühr L an. Diese richten sich lediglich nach einer Gewichtskomponente sowie dem Gebührensatz t_i . Im Gegensatz zu den Streckengebühren sind die An- und Abfluggebühren auch vom VFR-Verkehr zu entrichten. Für das obige Beispiel (A320-200) ergibt sich für einen Gebührensatz von $159,23\text{ €}$ ²⁸ eine An- und Abfluggebühr von je $201,52\text{ €}$. Die dargestellten Gebühren stellen Bruttowerte dar.

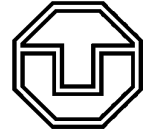
²⁴ Die Wechselkurse wurden berücksichtigt. Die Werte sind in € angegeben.

²⁵ Die Gebührensätze entsprachen etwa ein Viertel von jenen der Schweiz.

²⁶ Stand: Juni 2016.

²⁷ Der Flugplan umfasst neben Start- und Zielflughafen auch Wegpunkte entlang der Strecke (Route).

²⁸ Stand: Juni 2016.



$$L = \left(\frac{MTOW}{50}\right)^{0,7} \cdot t_i \quad (2-3)$$

$$L = \left(\frac{70}{50}\right)^{0,7} \cdot 159,23 \text{ €} = 201,52 \text{ €} \quad (2-4)$$

Der Vorteil der Gebühren-Finanzierung ist, dass die Nutzer des Luftraumes auch für die Kosten der Bewirtschaftung aufkommen. Allerdings führen die teilweise stark voneinander abweichenden Gebührensätze zu Ausweichenreizen für Airlines. Eine Herausforderung für die Flugsicherungen stellt die Anreizregulierung hinsichtlich der „Cost Efficiency“ dar, die zu einer Senkung der Gebührensätze führt (siehe Abschnitt 3.2.1). Alternative Finanzierungskonzepte für die Flugsicherungsdienstleistungen adressieren sowohl die Gebührenberechnung als auch deren Verteilung. Ziel ist die Vermeidung von Umwege-Anreizen und eine effektivere Allokation der finanziellen Mittel an die ANSPs. Hier sei auf eine Studie von Verbeek und Visser (2016) verwiesen.

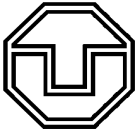
2.5 Herausforderungen für das Luftverkehrsmanagement

Während sich die vorherigen Kapitel mit der Bereitstellung der Dienstleistung beschäftigten, soll nun die Nachfrage der Luftraumnutzer diskutiert werden. Abbildung 2-8 zeigt die geplanten Trajektorien für einen Tag (beispielhaft der 15.10.2016). Auf den dunkelrot markierten Routen sind pro Tag mehr als 100 Flugbewegungen geplant. Die Nachfrage kann dabei höher sein als die abgebildeten Flüge. Dies ist abhängig von Kapazität der durchflogenen Sektoren (vgl. Abschnitt 2.2).

Es ist zu sehen, dass die Trajektorien (und damit die Flüge) nicht gleichmäßig im Raum verteilt sind. Ähnlich dem Straßenverkehr gibt es hochfrequentierte Routen, aber auch Regionen, wo sehr wenig oder gar kein Verkehr stattfindet. Die hochfrequentierten Routen liegen vor allem im Zentrum Europas, speziell Deutschland, Großbritannien, Frankreich, den Beneluxstaaten und der Schweiz. Hier befinden sich auch die größten Flughäfen, welche zu einer hohen Verkehrsdichte und Komplexität beitragen. Dies führt dazu, dass 55 % aller kontrollierten europäischen Flüge in den Lufträumen dieser Länder stattfinden (FABEC 2018a).

Auch wenn die Kapazitätsplanung maßgeblich auf den geplanten Trajektorien aufbaut, können kurz-, mittel- und langfristige Effekte dazu führen, dass tatsächliche Verkehrsmengen und -verteilungen abweichen. Mögliche Ursachen sind Wetterereignisse, Naturereignisse (z. B. Vulkanausbruch), politische Krisen oder Streiks (FABEC 2018c).

Laut dem Functional Airspace Block Europe Central (FABEC) sehen sich Flugsicherungen einer zunehmenden Verkehrsvolatilität gegenüber (FABEC 2018b; Standfuss et al. 2019b). Dabei kommt es meist zu einer Abweichung zwischen geplanter und tatsächlich geflogener Route. Luftfahrzeuge, die ungeplant den Luftraum durchfliegen (sogenannte „Intruder“), erhöhen die Nachfrage innerhalb des Sektors. Wird dadurch die Kapazitätsgrenze überschritten, kommt es zu erhöhten Kosten durch Ver-



spätungen, Umwege oder Mehrbedarf an Ressourcen. Im entgegengesetzten Fall, also der Nichtnutzung des Luftraumes, spricht man von „Runaways“. In diesem Fall wird zu viel Kapazität bereitgestellt, was durch die dadurch höheren Stückkosten wiederum ineffizient ist (Standfuss et al. 2018c).

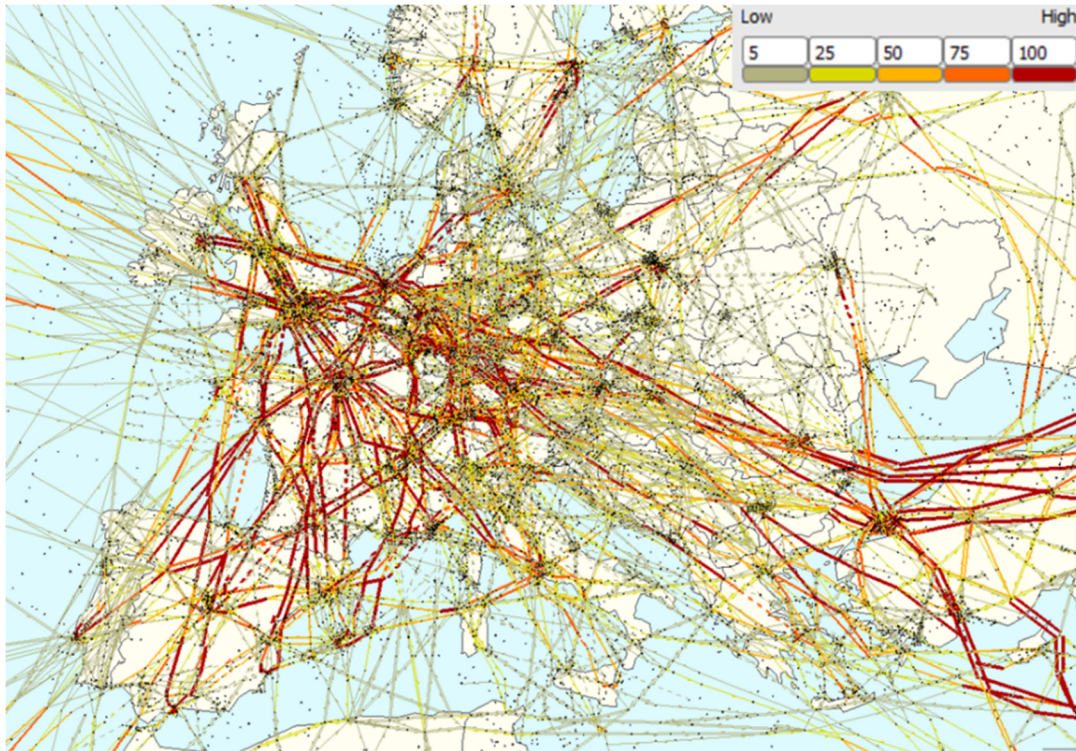


Abbildung 2-8: Density Plot der geplanten Trajektorien
NEST, 05.10.2016, (EUROCONTROL 2018c)

Die Abweichungen zwischen geplanter und geflogener Flugroute können sowohl vertikale als auch horizontale Komponenten beinhalten, wie in Abbildung 2-9 veranschaulicht ist. Das Beispiel zeigt einen Flug von Figari nach Berlin Schönefeld. Dabei repräsentiert die rote Route den aufgegebenen IFR-Flugplan, die grüne Route den tatsächlichen Flugverlauf. Es ist zu sehen, dass sowohl lateral als auch vertikal die beiden Routen abweichen. In diesem Fall führt es auch dazu, dass andere Sektoren durchflogen wurden als ursprünglich geplant.

Ausweichrouten können jedoch auch durch die unterschiedlichen Gebührensätze hervorgerufen werden. Die Airline verfolgt in ihrer Zielfunktion eine kostenminimale Route. Diese richtet sich neben anderen Umweltfaktoren auch nach Treibstoffkosten und Flugsicherungsgebühren. Das zeigt das Beispiel eines Fluges von Mailand nach Korfu (Abbildung 2-10). Blau eingezeichnet ist die direkte Verbindung beider Flughäfen. Die geplante bzw. geflogene Route sind rot bzw. grün dargestellt. Ähnliche Ausweichverkehre sind auch an der polnisch-schwedischen Luftraumgrenze beobachtbar.

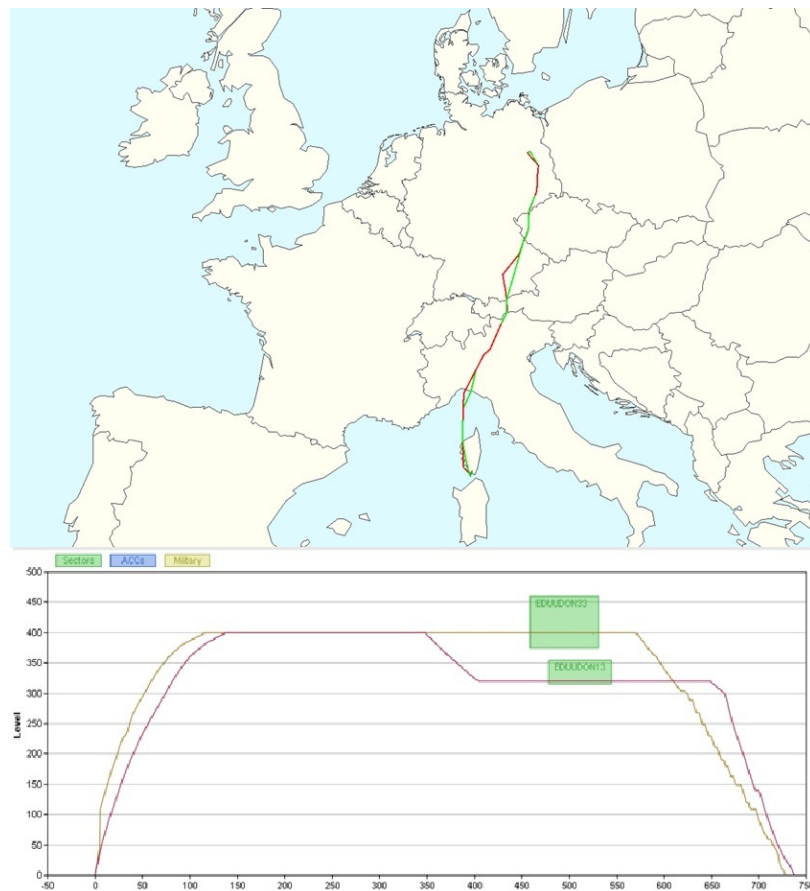
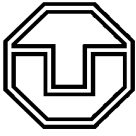


Abbildung 2-9: Abweichung zwischen geplanter und tatsächlich geflogener Route (Beispiel) NEST (EUROCONTROL 2018c), basierend auf EUROCONTROL-Daten



Abbildung 2-10: Umwege durch unterschiedlich hohe Gebührensätze der ANSPs (Beispiel) NEST (EUROCONTROL 2018c), basierend auf EUROCONTROL-Daten

Der dynamischen Verkehrsnachfrage steht ein (relativ) starrer Luftraum gegenüber. Zwar bestehen einige Möglichkeiten, die Kapazität an die Nachfrage anzupassen, jedoch sind diese betrieblich begrenzt. Der Luftraum gehört rechtlich zu einem spezifischen Staatsgebiet. Diese Situation führt dazu,



dass das Air Traffic Management in der Europäischen Union stark fragmentiert ist. Dies bezieht sich speziell auf die Luftraumstruktur, aber auch auf die verwendeten Systeme und Verfahren (FABEC 2019; Standfuss et al. 2018c).

Diese Fragmentierung wird als ineffizient und kostenintensiv betrachtet. EUROCONTROL schätzte bspw. die Mehrkosten des nationalstaatlichen Air Traffic Managements auf bis zu 5 Mrd. € pro Jahr (European Commission 2012c). Die Voraussetzung für eine höhere Kapazität und Effizienz, mehr Sicherheit zu geringeren Flugsicherungskosten, wird allgemein in einer verstärkten Zusammenarbeit zwischen den Staaten gesehen. Die Europäische Kommission verfolgt diese Ziele innerhalb des Single European Sky (European Commission 2004a, 2004b). In dessen Folge soll der Luftraum in Europa neu strukturiert und organisiert werden. Ein zentrales Element dieser Restrukturierung stellt die Organisation der Lufträume in Funktionalen Luftraumblöcken dar (EUROCONTROL 2018b).

Der Ansatz, ANSPs in FABs zu organisieren, geht davon aus, dass europäische Flugsicherungen unter steigenden Skalenerträgen agieren und durch Konsolidierungen Größenvorteile (Economies of Scale) realisieren können. Durch die Schaffung dieser Luftraumblöcke soll u. a. die Vermeidung von Koordinationsaufwänden eine Steigerung der Kapazität hervorrufen. Zudem sollten sich die Lufträume an den Verkehrsströmen orientieren, um so einen funktionalen Nutzen zu erfüllen. Abbildung 2-11 zeigt die konsolidierten Lufträume. Die Staaten behalten dabei ihre nationalen Lufthoheitsrechte, jedoch sollen Verfahren und technische Ausstattungen vereinheitlicht werden (EUROCONTROL 2017a).

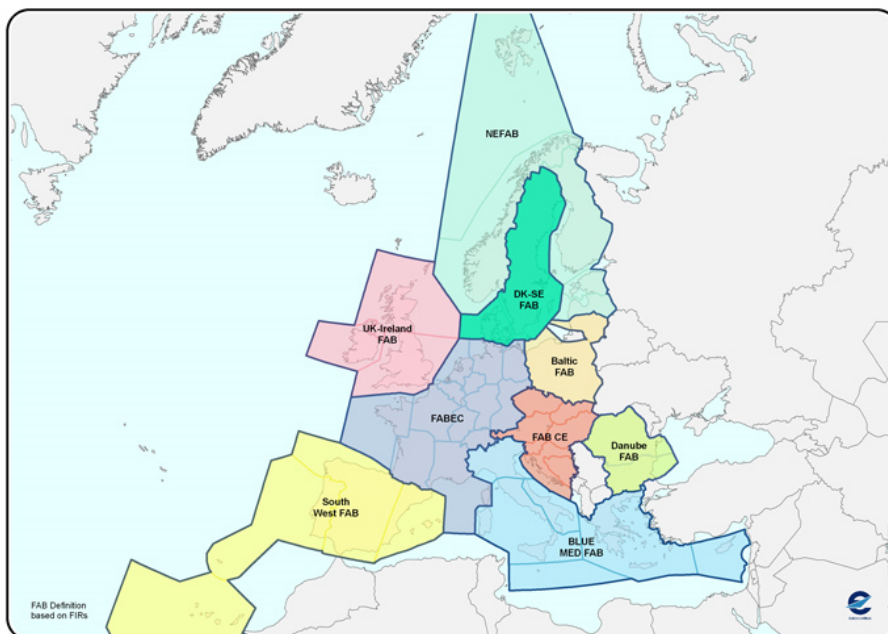
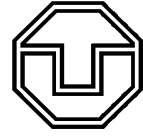


Abbildung 2-11: Functional Airspace Blocks (FABs)
(EUROCONTROL 2017a)

Inwieweit die vermuteten Größenvorteile tatsächlich realisiert werden können, ist ein Bestandteil der Analysen dieser Arbeit und wird in Abschnitt 4.5 diskutiert. Dabei wurde untersucht, welche ANSPs potentiell Größenvorteile realisieren können (Standfuss et al. 2017b, 2019a). Dieser Analyseansatz berücksichtigt, dass eine bessere Allokation der ANSPs zu funktionalen Luftraumblöcken prinzipiell möglich sein könnte, welche von der momentan existenten Zuordnung abweicht.



3 Performance-Vergleich europäischer Flugsicherungen

3.1 Ökonomische Grundlagen

3.1.1 Produktivität und Effizienz

3.1.1.1 Leistungsbewertung von Unternehmen

Die ökonomische Leistungsbewertung von Unternehmen ist ein zentrales Element in der Volks- und Betriebswirtschaftslehre. Zahlreiche Studien zu Methoden (Fried et al. 2008; Lovell 1992; Stepan und Fischer 2014) und Anwendungen (Ahmada et al. 2017; Albrecht et al. 2012; Hoffmann 2006) in den verschiedenen Wirtschaftszweigen wurden publiziert. Das Performance Benchmarking im Luftverkehr, speziell im ANSP-Kontext, ist dagegen eine noch junge Disziplin. Einen umfangreichen Überblick zum Thema Performance Benchmarking bietet u. a. Bogetoft (2013).

Unter Produktivität versteht man im Allgemeinen einen Quotienten aus der Menge an produzierten Gütern und der dafür verwendeten Menge an Produktionsfaktoren (Cantner et al. 2007). Werden lediglich ein Output und ein Input berücksichtigt, wird dies als Index-Ziffer-Methode bezeichnet. Im Falle einer Nutzung von mehreren Inputs und/oder mehreren Outputs werden zunächst Indizes gebildet, welche dann in die Berechnung der Produktivität eingehen. Dies ermöglicht sowohl die Berechnung der Totalen Faktorproduktivität (TFP), d. h. die Bewertung der Gesamtleistung eines Unternehmens (oder einer Einheit), als auch die Berechnung der Partiellen Faktorproduktivität, welche den Einfluss eines einzelnen Faktors auf die Gesamtproduktivität bestimmt (Fried et al. 2008).

Die Effizienz beschreibt die Relation aus der gemessenen Produktivität eines Unternehmens und dem beobachteten Maximum an Produktivität innerhalb einer gegebenen Technologie. Die Technologie beinhaltet die Summe aller sowohl theoretisch möglichen als auch praktisch beobachteten Input-Output-Kombinationen, welche eine ökonomische Einheit (respektive Decision Making Unit, DMU) nach den aktuellen technischen Möglichkeiten realisieren kann.

Bei einer relativen Effizienzmessung wird angegeben, zu wieviel Prozent ein Unternehmen effizient arbeitet. Je nach Betrachtungsweise wird entweder bei fixiertem Input der Output maximiert oder der fixe Output mit einem minimalen Einsatz an Produktionsfaktoren realisiert. Diese Grundlagen werden u. a. in Bielecki (2011) ausführlich beschrieben.

In den vergangenen Jahren haben sich verschiedene Methoden zum Vergleich der Leistungsfähigkeit von Unternehmen bzw. Betrieben etabliert. Aufgrund der charakteristischen Spezifika (vgl. Abschnitt 2.2) werden in dieser Arbeit sowohl qualitative als auch quantitative Methoden eingesetzt und entsprechend dem Analysezweck angepasst. Der aktuelle Abschnitt fokussiert die quantitative Analyse. Eine Übersicht über mögliche Verfahren bietet Abbildung 3-1 (Bielecki 2011; Last und Wetzel 2009).

Das Ergebnis eines Leistungsvergleiches ist einer Rangliste der Unternehmen, wobei das leistungsfähigste Unternehmen den ersten Rang repräsentiert. Hier können ökonomische (z. B. Stückkos-



ten) oder technische Kennzahlen zum Einsatz kommen. Letztere setzen keine Preisinformationen voraus und basieren auf Produktionsfunktionen (siehe Abschnitt 3.1.1.2). Diese kann vorab bekannt oder unbekannt sein. Ist sie unbekannt, kann sie über parametrische oder nichtparametrische Methoden bestimmt bzw. geschätzt werden (Cantner et al. 2007).

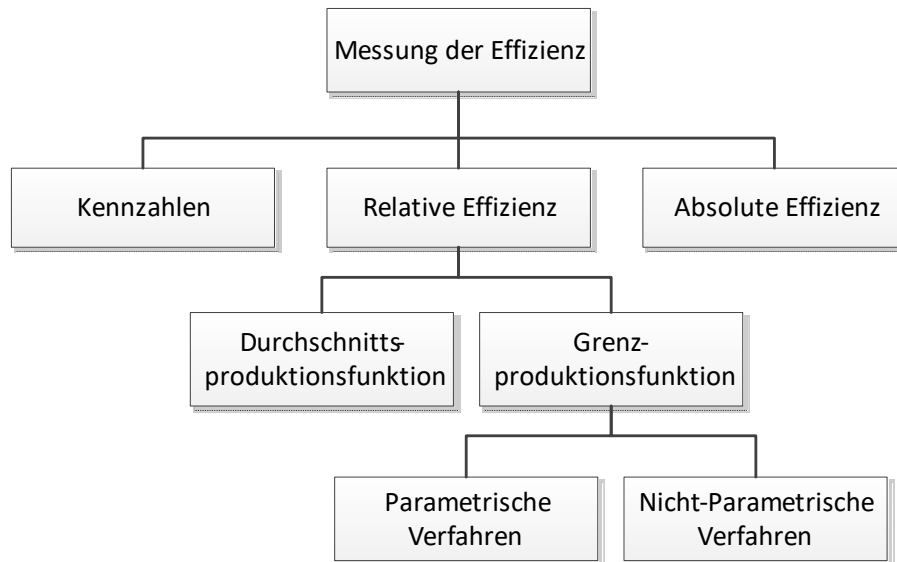
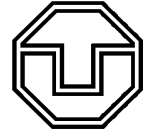


Abbildung 3-1: Effizienzmessverfahren
Angelehnt an Hirschhausen et al. (2007) und Bielecki (2011)

Kennzahlen bzw. Index-Ziffern repräsentieren im Allgemeinen einen Quotienten aus Output (z. B. Flugstunden) und Input (z. B. ATCO-Stunden). Das heißt, die erbrachte Leistung wird in das Verhältnis zu den verwendeten Ressourcen gestellt. Je größer dieser Quotient ist, desto höher ist auch die Produktivität bzw. Effizienz des Unternehmens. Ein Ranking unterschiedlicher Unternehmen oder deren Einheiten wird somit ermöglicht. Der Vorteil einer solchen Darstellung ist die intuitiv einfache Interpretation durch die absoluten Produktivitätswerte und dem darauf basierenden Ranking. Diese Methodik wird u. a. in den offiziellen EUROCONTROL-Berichten (EUROCONTROL 2019a, 2019h) genutzt (vgl. Abschnitt 3.2.1).

Die Berücksichtigung multipler Inputs und Outputs ist in der absoluten und relativen Effizienzmessung durch die Generierung von Produktionsfunktionen möglich. Da ein funktionaler Zusammenhang zwischen Ressourcen und produzierten Gütern in den meisten Fällen nicht herleitbar ist (Bielecki 2011), werden Randproduktionsfunktionen entweder parametrisch oder nicht-parametrisch über eine empirische Approximation bestimmt (Fried et al. 2008). Bei parametrischen Methoden müssen ex ante Annahmen über die funktionale Form der Produktionsfunktion getroffen werden. Bei nicht-parametrischen Methoden sind keine ex ante Annahmen über funktionale Zusammenhänge oder Störgrößen notwendig, diese werden stattdessen mittels mathematischer Programmierung bestimmt (siehe Abschnitt 3.1.2). Effizienzbeeinflussende Faktoren werden in dem Fall über ökonometrische Methoden, z. B. Regressionsanalysen, als zweiter Schritt in die Untersuchung integriert (Banker und Natarajan 2008; Hoff 2007; Simar und Wilson 2007).



Auch die Ergebnisse nicht-parametrischer Ansätze sind abhängig von der Modellierung bzw. den inhärenten Annahmen. Zunächst kann das Modell input- oder outputorientiert sein. Die Skalenerträge können als konstant, variabel, nicht steigend oder nicht sinkend angenommen werden. Des Weiteren wird in deterministische und stochastische Produktionsfunktionen unterschieden (Cantner et al. 2007). In der ökonomischen Forschung finden insbesondere die Data Envelopment Analysis und die Stochastic Frontier Analysis Anwendung, welche in den Abschnitten 3.1.2 und 3.1.3 näher beschrieben werden.

3.1.1.2 Produktionsfunktionen

Zur ökonomischen Erfassung des Prozesses der Leistungserstellung werden bevorzugt Kosten- oder Produktionsfunktionen genutzt. Aufgrund der flugsicherungsspezifischen Charakteristika (vgl. Abschnitt 2.3) ist die Nutzung einer Produktionsfunktion besser geeignet als die einer Kostenfunktion. Sie stellt einen funktionalen Zusammenhang zwischen Input(s) und Output(s) her. Empirisch beobachtete Produktionsverfahren der Unternehmen können als Actual-Practice-Produktionsfunktionen dargestellt werden.

In der funktionalen Darstellung unterscheidet man zwischen einer Durchschnittsproduktionsfunktion (DPF), die alle Input-Output-Kombinationen mittels einer Ordinary Least Squares (OLS) Regression verbindet, und einer Randproduktionsfunktion (RPF), welche die Einheiten mit den maximalen Verhältnissen aus Output zu Input verbindet. Die Fläche unter dieser Randproduktionsfunktion bildet die Technologiemenge. Die RPF wird daher auch als „Effizienter Rand“ oder „Efficient Frontier“ bezeichnet. Die zugehörigen methodischen Ansätze werden als „Frontier Analysis“ bezeichnet (Bielecki 2011; Cantner et al. 2007).

Abbildung 3-2 stellt die oben genannten Produktionsfunktionen gegenüber. Die schwarzen Punkte repräsentieren dabei die Output-Input-Verhältnisse der betrachteten Unternehmen. Die Durchschnittsproduktionsfunktion ist orange dargestellt. Die Randproduktionsfunktionen unterscheiden sich je nach den implizierten Skalenerträgen. Im Falle konstanter Skalenerträge (Constant Returns to Scale, CRS) ist die Funktion linear (blaue Linie), bei variablen Skalenerträgen (Variable Returns to Scale, VRS) konvex (rote Linie). Die linear-konvexe Form wird durch das Pareto-Prinzip bzw. durch das Pareto-Koopmans-Kriterium hergeleitet (Cantner et al. 2007), siehe auch Abschnitt 3.1.2.2. Das heißt, der Bereich der Produktionsmöglichkeiten wird von der abschnittsweisen linearen Verbindung der effizienten DMUs begrenzt, wobei der Anstieg der RPF in einem Abschnitt maximal so hoch sein kann wie der Anstieg im Abschnitt zuvor, mindestens jedoch gleich 0. Diese Konvexitätsbedingung schließt dadurch auch aus, dass sich die implizierten Skalenerträge ändern (die RPF kann bspw. nicht durch Punkt B verlaufen). Zudem wird der Technologie-Bereich durch einen möglichen Schlupf (Slack) bei Input (Parallele zur Abszisse) und Output (Parallele zur Ordinate) begrenzt.

Für die Effizienzbetrachtung ist ein Vergleich zwischen der Performance einer Einheit und jener der bestmöglichen Leistungsfähigkeit notwendig. Die OLS-Methodik, respektive die Nutzung einer DPF, hat diesbezüglich den entscheidenden Nachteil, dass Einheiten oberhalb der Regressionsgerade als „zu effizient“ gelten würden (Punkte A, B, C und D). Das heißt, für die Zwecke des Performance Benchmarkings ist dieser Ansatz nicht zielführend. Durch die Nutzung der RPF werden alle DMUs mit dem



oder den beobachteten Maxima²⁹ verglichen (best practice). Einheiten, welche nicht auf dem effizienten Rand liegen, werden als ineffizient identifiziert und bewertet.

Die Erzeugung dieser RPF ist abhängig von der gewählten Methodik und kann deterministisch (jede Abweichung wird als Ineffizienz gewertet) oder stochastisch (Abweichung ist eine Kombination aus Ineffizienz und Zufallseinflüssen) geprägt sein (Bielecki 2011; Cantner et al. 2007; Kumbhakar und Lovell 2003). Auf Basis dieser Überlegungen können verschiedene Effizienzmaße abgeleitet werden, die im folgenden Abschnitt diskutiert werden.

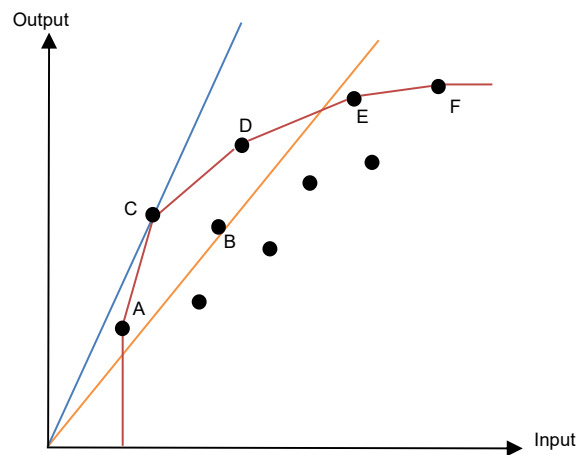


Abbildung 3-2: Produktionsfunktionen im Vergleich

Es ist jedoch zu berücksichtigen, dass nur ein Teil der konvexen RPF empirisch beschrieben ist. Im Falle von Abbildung 3-2 entspricht dies den Bereich zwischen den Unternehmen A und F. In der Literatur werden jene Abschnitte außerhalb dieser Beobachtungsmenge durch einen parallelen Verlauf der RPF zur Abszisse bzw. Ordinate dargestellt. Dies führt zu der (teils kontraintuitiven) Implikation, dass trotz eines Inputs > 0 kein Output erzeugt wird (untere Grenze, Schnittpunkt mit der Abszisse). Zudem wird impliziert, dass oberhalb von Punkt F auch zusätzliche Input-Einheiten zu keiner weiteren Steigung des Outputs führen.

Tatsächlich repräsentieren diese Parallelverläufe jedoch Extremwerte bzw. Grenzfälle. Der Anstieg der RPF unterhalb von A kann maximal unendlich sein (wie abgebildet) oder marginal höher als der Anstieg zwischen A und C. Analog kann der Anstieg der RPF oberhalb von F minimal 0 sein (wie abgebildet), oder marginal kleiner als jener zwischen E und F.

Die Messung der (relativen) Effizienz anhand der Randproduktionsfunktion eignet sich insbesondere zur Identifikation und Quantifizierung von Verbesserungspotentialen. Durch die Analyse von Produktionsfunktionen ist die Dekomposition einer ökonomischen (In-) Effizienz in eine technische, skalare und eine allokativen Komponente möglich. Diese sollen im Folgenden näher erklärt werden. Dabei wird, sofern nicht explizit auf das Maximalprinzip verwiesen wird, das Minimal-Prinzip betrachtet.

²⁹ Maxima bezieht sich auf Performance der DMUs innerhalb der betrachteten Technologie.

3.1.1.3 Effizienzmaße

Ein Unternehmen gilt als technisch effizient, wenn es einen gegebenen Output (z. B. Flugbewegungen) mit dem minimal möglichen Einsatz an Produktionsfaktoren (z. B. Fluglotsen) produziert. Die technisch effizienten Output-Input-Quotienten der jeweiligen DMUs bilden eine Teilmenge der Technologie, welche durch die Randproduktionsfunktion abgebildet wird. Die effizienten DMUs werden in der relativen Effizienzmessung mit 1 bzw. 100 % bewertet.

DMUs, die sich nicht auf der Randproduktionsfunktion befinden, werden mit einem Wert zwischen 0 und 1, respektive 0 % und 100 %, bewertet. Dieser Grad an Effizienz wird mittels Distanzfunktionen³⁰ errechnet und besagt, um wieviel Prozent die betrachtete Einheit ihren Input bzw. ihre Inputs senken müsste, um als effizient zu gelten. Gemäß Abbildung 3-3 ist dies die Differenz zwischen der orange dargestellten actual-practice Funktion F der DMU B und der blauen best practice Funktion F^* der Technologie. Um Effizienz zu erreichen, müsste die DMU bei fixiertem Output y' den Input auf x' reduzieren (Punkt B' , Minimalprinzip) oder mit konstantem Input x^* ihren Output auf y^* erhöhen (Punkt B^* , Maximalprinzip) (Cantner et al. 2007).

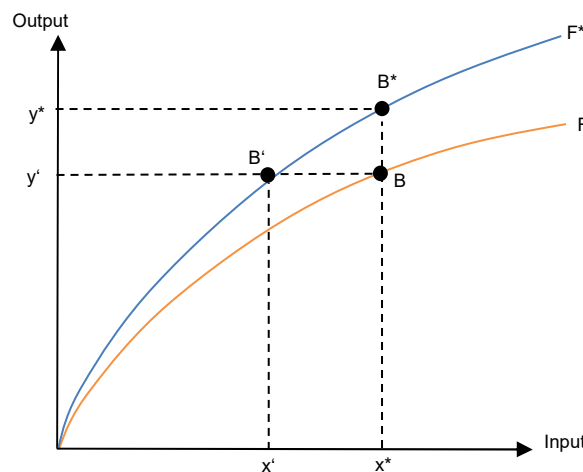


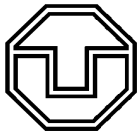
Abbildung 3-3: Berechnung der Effizienz über Distanzfunktionen
Angelehnt an Cantner et al. (2007)

Die relative Ineffizienz wird durch die radiale Distanz zwischen der beobachteten Einheit und dem effizienten Rand der Technologie³¹ definiert. Dadurch entsteht ein dimensionsloses Maß, welches als Grad an Effizienz interpretiert werden kann. Eine DMU gilt als effizient, wenn entweder keine Erhöhung eines Outputs ohne die Reduktion eines anderen Outputs bzw. der Erhöhung eines Inputs möglich ist oder keine Senkung eines Inputs ohne die Erhöhung eines anderen Inputs bzw. die Reduktion eines Outputs möglich ist (Cooper et al. 2011).

Abbildung 3-4 skizziert die Messung der Effizienz mittels radialer Distanz an einem Beispiel für zwei Inputs (x_1 und x_2) sowie einem Output (y). Die rote Kurve stellt die Isoquante des Outputs dar, welche durch die effizienten Einheiten A und B aufgespannt wird. Sie bilden die Peer Group für DMU C, welche als ineffizient ausgewiesen wird. Um Effizienz zu erreichen, müsste sich C auf dem effizienten

³⁰ Die Nutzung von Distanzfunktionen bietet den Vorteil, dass multiple Inputs und Outputs berücksichtigt werden können.

³¹ Summe aller möglichen Input-Output-Kombinationen.



Rand befinden (C^*). Die radiale Distanz ist gleich dem Quotienten der Distanzen $\overline{OC^*}$ und \overline{OC} . Die blauen Pfeile repräsentieren die Gewichte des Optimierungsmodells. Sie geben den relativen Abstand von C^* zu den Peers A und B an.

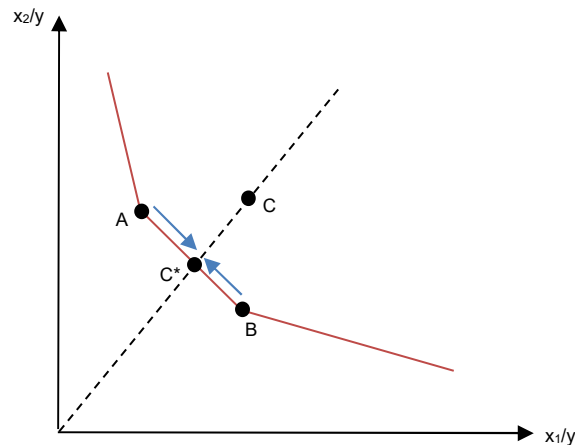


Abbildung 3-4: Messung der Effizienz mittels radialer Distanz zu einer deterministischen Grenze

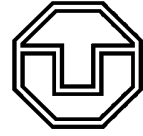
Ein Berechnungsbeispiel ist in Tabelle 3-1 für den Fall von 8 DMUs dargestellt, welche zwei Inputs verwenden, um einen Output zu produzieren. Unter der Annahme konstanter Skalenerträge wird der Output auf 1 normiert und die Inputs dementsprechend angepasst. Die Einheiten mit dem geringsten Faktoreinsatz gelten dann als (technisch) effizient. Durch die Projektion eines Ursprungsstrahles kann nun die Effizienz aller Unternehmen (DMUs) berechnet werden.

DMU 1 weist den geringsten Faktoreinsatz von Input 1 auf, DMU 2 von Input 2. Daher bilden diese Einheiten den effizienten Rand. Auch DMU 6 nutzt den geringsten Faktoreinsatz für Input 1. Dennoch könnte das Unternehmen 0,5 Einheiten weniger an Input 1 verwenden um den gleichen Output zu erzeugen. Diese Eigenschaft wird als Schlupf oder Slack (S) bezeichnet. DMU 8 weist eine Effizienz von 72,8% auf. Um auf den effizienten Rand zu gelangen, müsste die DMU demnach ihre Inputs um je 27,2% senken, ohne den Output zu verändern.

Tabelle 3-1: Berechnungsbeispiel für die Effizienz im Fall multipler Inputs

DMU	1	2	3	4	5	6	7	8
Input 1	1	2	1,5	1,4	1,25	2,5	2	1,6
Input 2	0,9	0,3	0,8	0,7	1,1	0,3	1	1,1
Output	1	1	1	1	1	1	1	1
Effizienz	100%	100%	88,2%	97,4%	81,1%	100% (S)	68,2%	72,8%

Es ist unter Umständen möglich, dass zwei Unternehmen technische Effizienz (TE) erreichen, jedoch die Produktivität unterschiedlich hoch ist. In diesem Fall befinden sich zwar beide Einheiten auf dem effizienten Rand der Produktion, ein Ursprungsstrahl zu beiden Unternehmen hätte jedoch eine unterschiedliche Steigung. Je höher diese Steigung ist, desto höher ist die Produktivität. Diese geringere Produktivität kann, neben dem bereits diskutierten Slack, auf eine Skalenineffizienz (SE) zurückzuführen sein.



Wie bereits in Abschnitt 3.1.1.2 diskutiert, werden durch eine lineare Form der Produktionsfunktion konstante Skalenerträge impliziert. Eine Technologie mit variablen Skalenerträgen führt zu einer konvexen RPF. Der Vergleich beider Funktionen ermöglicht eine Analyse der DMU-individuellen Skaleneffizienz sowie deren Skalenerträge. Die Skaleneffizienz wird bestimmt, indem für jede DMU i der technische Effizienzwert je einmal unter Annahme konstanter (TE_{ik}) und variabler Skalenerträge (TE_{iv}) bestimmt wird. Die Skaleneffizienz wird über den Quotienten in Formel (3-1) berechnet.

$$SE_i = \frac{TE_{ik}}{TE_{iv}} \quad (3-1)$$

In Abbildung 3-5 (hier dargestellt ist eine Output-Maximierung) sind die DMUs A, C und D nach variablen Skalenerträgen technisch effizient. Der Ursprungsstrahl bei Punkt C ist gleichzusetzen mit der Randproduktionsfunktion unter Annahme konstanter Skalenerträge. Im (technisch effizienten) Punkt A hat dieser Strahl einen geringeren Anstieg. Obwohl beide Einheiten auf dem effizienten Rand der Technologie liegen, ist die Produktivität von Unternehmen C somit höher. Die Differenz beschreibt die Skaleneffizienz.

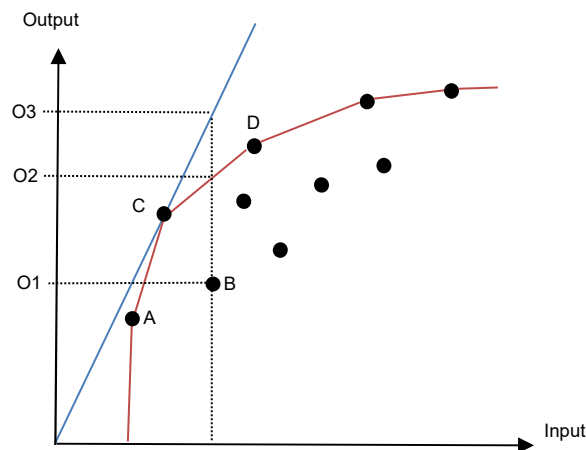


Abbildung 3-5: Skaleneffizienz
Angelehnt an Pascoe et al. (2003)

In der abgebildeten Situation ist lediglich DMU C sowohl technisch- als auch skaleneffizient, die DMUs A und D technisch effizient, jedoch nicht skaleneffizient. DMU B ist weder skaleneffizient noch technisch effizient. Der Grad an Ineffizienz setzt sich für dieses Unternehmen zusammen aus der technischen Komponente (Distanz $O_1 O_2$) sowie der skalaren Komponente (Distanz $O_2 O_3$).

Unternehmen D ist, wie Unternehmen A, skaleneffizient und technisch effizient. Die Einheiten unterscheiden sich jedoch in ihren Skalenerträgen. Wird eine DMU als skaleneffizient identifiziert, ist die Unternehmensgröße entweder zu groß oder zu klein. Hier kann zusätzlich bestimmt werden, ob die DMU unter steigenden³² oder fallenden³³ Skalenerträgen arbeitet (vgl. auch Abschnitte 3.1.2.1 und 3.1.2.2).

³² Increasing Returns to Scale (IRS).

³³ Decreasing Returns to Scale (DRS).



Bei steigenden Skalenerträgen wächst der Output überdimensional stark bei einer Erhöhung des Inputs. Die Durchschnittskosten sinken demnach bei der Erhöhung der produzierten Menge. Arbeitet das Unternehmen unter fallenden Skalenerträgen, so ist eine überdimensionale Steigerung des Inputs notwendig, um den Output zu steigern. Diese Größeneffekte (Economies of Scale, EOS) sind jedoch abzugrenzen von Economies of scope³⁴ und Economies of density³⁵ (Cento 2009; Hutzschenreuter 2015; Vasigh et al. 2008).

Während bei der technischen und skalaren Effizienz Bezug zu Mengen hergestellt wird, sind zur Berechnung der allokativen Effizienz zusätzlich Preisinformationen der Produktionsfaktoren notwendig. Alternativ zur entsprechenden Kostenfunktion können auch Erlös- oder Gewinnfunktionen verknüpft werden. Dadurch kann die kostenminimale Input-Kombination zur Produktion des Outputs bestimmt werden (Coelli et al. 2005).

Abbildung 3-6 zeigt ein Beispiel für zwei Inputs (x_1 und x_2) und einen Output (y). Die rote Kurve repräsentiert die Isoquante des Outputs, aufgespannt von den beiden technisch effizienten DMUs A und B. Die blaue Gerade stellt die Isokostenlinie der Inputs dar. Verdeutlicht wird dadurch, dass DMU A zwar technisch effizient arbeitet, jedoch nicht die kostenminimale Inputkombination gewählt hat. Dagegen kann DMU B sowohl technische als auch allokativen Effizienz nachweisen. Das Produkt aus technischer und allokativer Effizienz ist die Kosteneffizienz (Cantner et al. 2007).

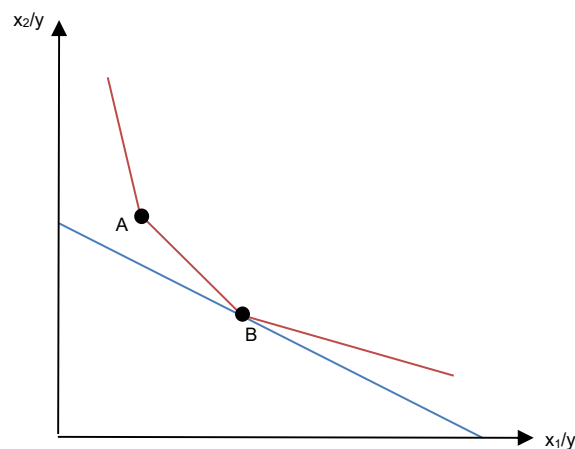
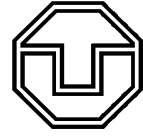


Abbildung 3-6: Bestimmung der allokativen Effizienz

Voraussetzung für die Berechnung der allokativen Effizienz sind vollständige Informationen über die Preise bzw. Kosten der Produktionsfaktoren. Zudem kann sie lediglich dann berechnet werden, wenn mehrere Inputs verwendet werden. Im Flugsicherungssektor unterscheidet sich jedoch die Allokation der Kosten zwischen den Unternehmen. Diese fehlenden oder unvollständigen Informationen führen dazu, dass die allokativen Effizienz in dieser Arbeit nicht berücksichtigt werden kann.

³⁴ Verbundvorteile führen zu niedrigeren Stückkosten, wenn Ressourcen in mehreren Projekten oder Produktionsketten verwendet werden können.

³⁵ Dichtevorteile entstehen bei einer räumlichen Konzentration von Anbietern und/oder Nachfragern. Ein Beispiel sind die Hub-and-Spoke Netze von Airlines.



Die bisher dargestellten Leistungskennzahlen beschreiben einen statischen Zustand der Technologie zu einem spezifischen Zeitpunkt. Produktionsprozesse bleiben zeitraumübergreifend jedoch nicht konstant. Technische Innovationen oder betriebliche Optimierungen können zu einer Erhöhung dieser Kennzahlen führen. Grafisch entspricht dies einer Verschiebung oder Drehung der Produktionsfunktion nach oben oder links.³⁶ Das heißt, dass in der Folgeperiode der gleiche Output mit einem geringeren Input realisiert werden kann, oder durch den konstanten Einsatz von Produktionsfaktoren mehr Güter und/oder Dienstleistungen produziert werden.³⁷ Die Verschiebung oder Drehung der Produktionsfunktion wird als technischer Fortschritt (TF) bezeichnet (Cantner et al. 2007).

Der technische Fortschritt ist im mehrperiodischen Fall von der Änderung in den Effizienzmaßen abzugrenzen. Die Inputreduzierung bzw. Output-Erweiterung eines Unternehmens kann bspw. auch auf eine Verbesserung der technischen Effizienz zurückzuführen sein. Er ist zudem von preisbedingten Substitutionseffekten der Inputfaktoren (allokative Effizienz) sowie der Skaleneffizienz abzugrenzen. Eine Separation der Skaleneffekte vom technischen Fortschritt ist jedoch in der empirischen Analyse nicht immer eindeutig.

Effizienzverbesserungen und/oder technischer Fortschritt führen zu einer Veränderung der Totalen Faktorproduktivität. Diesbezügliche Untersuchungen sind innerhalb einer Panel-Analyse möglich. Umfasst der zugrunde liegende Datensatz mehrere Jahre, so kann unter anderem die parametrische Malmquist-DEA (vgl. Abschnitt 4.3.2) angewendet werden, eine spezielle Form der im folgenden Abschnitt beschriebenen Data Envelopment Analysis.

3.1.2 Data Envelopment Analysis

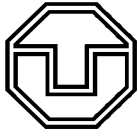
3.1.2.1 Konstante Skalenerträge

Die DEA zählt zu den nichtparametrischen, deterministischen Verfahren (Cantner et al. 2007). Daher sind keine Annahmen über funktionale Zusammenhänge zwischen den Produktionsfaktoren und den Outputs notwendig. Stattdessen wird auf Basis der beobachteten Input-Output Kombinationen aller DMUs mittels mathematischer Programmierung eine umschließende Hülle erzeugt. Über ein Optimierungsmodell wird für jede DMU die minimale Distanz zur RPF berechnet und die jeweiligen Gewichte, Peers und Schlupfvariablen ausgegeben. Dafür werden zunächst alle Inputs zu einem virtuellen Input aggregiert und alle Outputs zu einem virtuellen Output. Die Gewichtung ist Teil des Optimierungsproblems.

Das mathematische Grundmodell der Data Envelopment Analysis wurde zunächst von Charnes, Cooper und Rhodes für konstante Skalenerträge entwickelt (Charnes et al. 1978). Im Folgenden soll das nach den Autoren benannte ‚CCR-Modell‘ sowie die daraus erhaltenen ökonomischen Informationen erörtert werden. Dabei wird aufgrund der Modellierung innerhalb der Dissertation die Inputorientierung verfolgt.

³⁶ Die Steigung des Ursprungstrahles ist in $t+1$ größer.

³⁷ Im Falle einer Produktivitätsverschlechterung verschiebt sich die RPF nach unten bzw. rechts.



Um den mathematischen Hintergrund zu erläutern werden folgende Annahmen getroffen:

- Die Technologie beinhaltet $j = \{1 \dots n\}$ Unternehmen
- Sie produzieren $\{r = 1 \dots s\}$ Outputs y_{rj}
- Sie nutzen dafür $\{i = 1 \dots n\}$ Inputs x_{ij}

Das Grundmodell der DEA optimiert zwei Aggregationsgewichte u und v derart, dass die Effizienz E der Einheit $k \in j$ maximiert wird (3-2). Danach werden die Inputs und Outputs aller übrigen DMUs mit den für k optimalen Aggregationsgewichten bewertet. Erreicht k die maximale Produktivität, so wird sie mit 100% bewertet (3-3). Dieser Vorgang wird für alle n DMUs ausgeführt. Beide Gewichte müssen größer als oder gleich 0 sein (3-4).

$$\max E_k = \frac{\sum_{r=1}^s u_r y_{rk}}{\sum_{i=1}^n v_i x_{ik}} \quad (3-2)$$

u. d. N.

$$\frac{\sum_{r=1}^s u_r y_{rk}}{\sum_{i=1}^m v_i x_{ik}} \leq 1 \quad (3-3)$$

$$u \geq 0 ; v \geq 0 \quad (3-4)$$

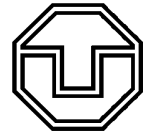
Neben dem Effizienzmaß werden auch die Aggregationsgewichte bestimmt. Diese können als Schattenpreise bzw. Opportunitätskosten interpretiert werden. Über die Charnes-Cooper-Transformationen (siehe Anhang B1) kann dieses Quotientenprogramm in ein lineares Optimierungsproblem überführt werden³⁸ (Cantner et al. 2007; Stepan und Fischer 2014).

Basierend auf dem „Dualitätssatz der Linearen Programmierung“ wird das mathematische Modell schließlich in die Envelopment-Form (Minimierungsproblem) überführt (Coelli et al. 2005). Diese stellt eine aus den Daten ermittelte Randproduktionsfunktion dar. Sie ermöglicht die Identifikation von Unternehmen mit ähnlichen Produktionscharakteristiken (Cantner et al. 2007).

In der Envelopment-Form wird der radiale Abstand der k -ten DMU zum effizienten Rand minimiert (3-5). Die Dualgewichte λ_j geben Auskunft über das Ausmaß der Gewichtung. Durch den Index an den Dualvariablen können zudem die zu Einheit k zugeordneten Peer units ausgegeben werden.³⁹ Dieses Programm wird n mal gelöst. Nebenbedingung (3-6) stellt sicher, dass die Inputs innerhalb des möglichen (d. h. beobachteten) Datensets bleiben. Für die Outputs steht analog Formel (3-7). Die Dualgewichte dürfen nicht negativ sein (3-8). Alternativ kann man die Dualform auch in Vektorschreibweise überführen. Die dadurch aufgespannte Hülle wird als abschnittsweise linear angenommen. Der Inputvektor schneidet diese Hülle gemäß Abbildung 3-4. Dieser Schnittpunkt ist eine Linearkombination der zu k zugeordneten Peers (Coelli et al. 2005).

³⁸ Diese Darstellungsweise wird auch „Multiplifier-Form“ genannt.

³⁹ D. h. mit welchen DMUs die k -te Einheit verglichen wird.



Durch diese Umformung können nun sowohl die Peers als auch die optimalen Faktoreinsätze bestimmt werden. Dafür stehen zwei Möglichkeiten zur Verfügung: Zum einen die Multiplikation des jeweiligen Inputs mit dem Effizienzwert θ , zum anderen die Multiplikation der Dualgewichte mit den Faktoreinsätzen der zugehörigen Peers.

$$\min \theta \tag{3-5}$$

u. d. N.

$$\theta \cdot x_{ik} \geq \sum_{j=1}^n \lambda_j x_{ij} \tag{3-6}$$

$$y_{rk} \leq \sum_{j=1}^n \lambda_j y_{rj} \tag{3-7}$$

$$\lambda_j \geq 0 \tag{3-8}$$

Das Prinzip wird anhand eines fiktiven Beispiels in Abbildung 3-7 dargestellt. Die DMUs A und B bilden die Peers für die beiden ineffizienten Einheiten C und D. Der optimale Faktoreinsatz von D kann nun über die Multiplikation von 0,8 mit den jeweiligen Inputs erfolgen. Alternativ können die Dualgewichte mit den Inputs von DMU A und B multipliziert und deren Summe gebildet werden. Für Input 1 gilt dann:

$$x_1^*(D) = 0,4 \cdot x_1(A) + 0,6 \cdot x_1(B) \tag{3-9}$$

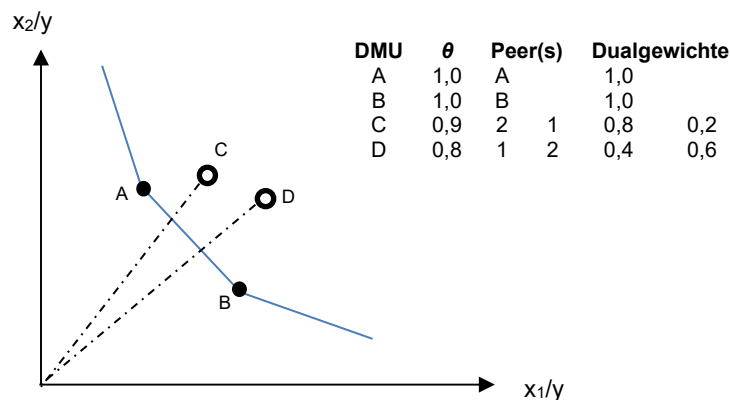


Abbildung 3-7: Peers und Gewichte des dualen Modells

3.1.2.2 Variable Skalenerträge

Die bisherigen Betrachtungen setzen eine Technologie mit konstanten Skalenerträgen voraus. Im Ein-Input-Ein-Output-Fall entspricht dies dem Ursprungsstrahl durch die Einheit mit der höchsten beobachteten Produktivität. Abszisse und Ursprungsstrahl begrenzen die Technologie und damit alle möglichen Output-Inputkombinationen. Im Umkehrschluss bedeutet dies, dass jede Einheit als ineffizient bewertet wird, die nicht auf diesem Strahl liegt. Eine Berücksichtigung der Unternehmensgröße findet nicht statt.



In der Ökonomie sind konstante Skalenerträge jedoch selten, zumeist ist die Technologie durch variable Skalenerträge gekennzeichnet. Insbesondere kleine Unternehmen können Größenvorteile, u. a. durch Synergieeffekte, realisieren. Bei zu großen Unternehmen ist unter Umständen das Gegenteil zu beobachten. Für eine detailliertere Beschreibung sowie zu den mathematischen Hintergründen siehe u. a. Knieps (2008) und Varian (2016).

Um variable Skalenerträge in der nicht parametrischen Effizienzanalyse zu berücksichtigen, entwickelten Banker, Charnes und Cooper das nach ihnen benannte BBC Modell (Banker et al. 1984). Die Unternehmensgröße wird dabei durch das Output-Volumen repräsentiert. Der Zusammenhang zwischen Unternehmensgröße und Effizienz wird über Skaleneffizienz und Skaleneffekte abgebildet. Auf Basis des Pareto-Koopmans-Kriterium wird dies im BCC Model grafisch durch die lineare Verbindung aller effizienten Einheiten erreicht (Cantner et al. 2007). Je nach Anstieg der Geraden kann auf steigende, konstante oder fallende Skalenerträge geschlossen werden (vgl. Abschnitt 3.1.1.2).

Das mathematische Grundmodell ist in den nachfolgenden Formeln abgebildet. Es unterscheidet sich nur in geringem Maße von der Matrixschreibweise der Envelopment-Form auf Basis der konstanten Skalenerträge (Multiplier-Form nach Charnes-Cooper-Transformation). Der radiale Abstand θ zwischen DMU und Randproduktionsfunktion wird dabei minimiert (3-10). Zur Herleitung der Envelopment-Form sowie zu alternativen Darstellungen siehe Cantner et al. (2007). Neben den Effizienzwerten θ werden auch die Gewichtungsfaktoren λ ausgewiesen. Die Technologiemenge wird nun durch Konvexkombinationen gebildet. Dadurch wird die Summe aller Gewichte 1 (3-13).

$$\min \theta \tag{3-10}$$

u. d. N.

$$Y\lambda \geq y_i \tag{3-11}$$

$$\theta x_i - X\lambda \geq 0 \tag{3-12}$$

$$1^T \lambda = 1 \tag{3-13}$$

$$\lambda \geq 0 \tag{3-14}$$

Die Formeln dienen zur Bestimmung der technischen Effizienz. Durch die Implikation variabler Skalenerträge besteht die Möglichkeit, die Skaleneffizienz zu bestimmen. Die gesamte Ineffizienz besteht demnach aus technischer und skalarer Ineffizienz. Die Skaleneffizienz wird aus dem Quotienten aus technischer Effizienz nach konstanten Skalenerträgen und technischer Effizienz nach variablen Skalenerträgen berechnet (siehe Abbildung 3-5). Diese ist ebenfalls im Intervall (0,1] skaliert. Eine Ineffizienz ist hier durch eine nicht-optimale Betriebsgröße bedingt. Diese ist entweder zu groß (fallende Skalenerträge) oder zu klein (steigende Skalenerträge). Dies erfolgt durch den Vergleich der Unternehmensgröße mit der „most productive scale size“ (MPSS)⁴⁰ (Cantner et al. 2007).

⁴⁰ Gesamteffizienz, technische Effizienz und Skaleneffizienz sind für entsprechende Unternehmen gleich 1.

3.1.2.3 Sonderformen und Erweiterungen der DEA

Die Annahme über die Form der Produktionsfunktionen waren in den vorherigen Abschnitten entweder linear (konstante Skalenerträge) oder konvex (variable Skalenerträge). Darüber hinaus sind Mischformen möglich, z. B. nicht-steigende oder nicht-sinkende Skalenerträge. Dies kann im DEA Modell implementiert werden, indem Formel (3-13) angepasst wird (Cantner et al. 2007). Grafisch verläuft die Produktionsfunktion im ersten Fall zunächst linear und danach konvex (rote Linie), im zweiten Fall vice versa (blaue Linie), wie in Abbildung 3-8 dargestellt.

Über die Einführung von Schlupfvariablen in die Dualform ist es möglich Slacks zu bestimmen (vgl. Abschnitt 3.1.1.3). Das Optimierungsproblem wird dann in zwei Phasen gelöst, indem zuerst der Optimalwert für θ bestimmt und im zweiten Schritt dieser als Basis für die Ermittlung der Schlupfvariablen verwendet wird. Ist eine der Schlupfvariablen positiv, so liegt ein Input- oder Outputslack vor (Stepan und Fischer 2014).⁴¹ Im DEA-Modell können diese zwar berechnet werden, jedoch beeinflussen die Slacks nicht den spezifischen Effizienzwert.

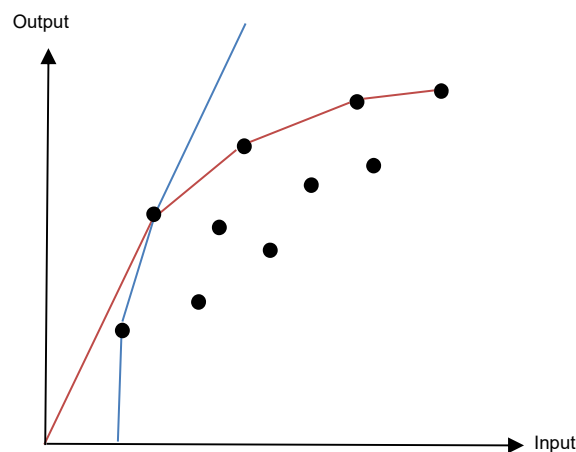
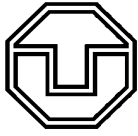


Abbildung 3-8: Nicht-sinkende und nicht-steigende Produktionsfunktion

DEA-Modelle sind primär Input- oder Output-orientiert. Darüber hinaus existieren jedoch additive bzw. multiplikative Modelle, die beide Ansätze kombinieren, also simultan Inputsenkungen und Outputsteigerungen berücksichtigen (Zhu 2014). Ein Beispiel für ein additives Modell ist die Slack Based DEA. Hier werden die Slacks in die Effizienzbewertung einbezogen. Daher kann der Ansatz als ganzheitliche Effizienzbewertung betrachtet werden. Dieser Ansatz kann jedoch aufgrund der mathematischen Modellierung lediglich zwischen effizienten und ineffizienten Einheiten unterscheiden, jedoch nicht per se den Grad an Ineffizienz bestimmen (Tone 2001).

Eine weitere Spezifikation ist der Vergleich effizienter Einheiten. Anhand des DEA-Modells werden alle effizienten DMUs mit 1 bewertet. Eine weitere Differenzierung effizienter Einheiten ist dadurch nicht möglich. Nutzt man eine hohe Anzahl an Inputs und Outputs bei einer relativ geringen Beobachtungszahl, so werden aufgrund der DEA-Methodik relativ viele Einheiten als effizient ausgewiesen

⁴¹ Siehe auch „Satz vom Dualen Schlupf“.



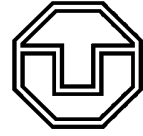
(Standfuss et al. 2018b). Über Modifikationen können diese Einheiten gesondert betrachtet und verglichen werden. Zum Beispiel schlugen Andersen und Petersen (1993) vor, die jeweilig betrachtete Einheit i aus der Bildung der Technologiereferenzmenge auszuschließen. Dadurch kann die Output-Input-Kombination der Einheit theoretisch auch über der Randproduktionsfunktion liegen. Der zugehörige Effizienzwert liegt in diesem Fall im Intervall $(0, \infty)$. Alle ineffizienten Einheiten erhalten den selben Score wie im ursprünglichen Modell (Chen und Du 2015). Jedoch werden nun verschiedene Frontierfunktionen für die Bewertungen genutzt, da für die effizienten DMUs der Abstand zu einer gesonderten Randproduktion gemessen wird (Cantner et al. 2007). Diese Art der DEA repräsentiert ein sogenanntes Supereffizienz-Modell und kann insbesondere dazu genutzt werden, Extremwerte zu identifizieren.

Eine Einschränkung bei der Nutzung der DEA ist die Sensitivität gegenüber Extremwerten in den Daten. Eine Möglichkeit zur Erhöhung der Robustheit wird durch eine stochastische DEA eröffnet. Hier wird eine Abweichung von der Produktionsfunktion nicht automatisch als Ineffizienz gesehen, sondern auch als Resultat von Störgrößen (z. B. Messfehler). Die Anwendung ermöglicht zudem das Berechnen von Konfidenzintervallen (Banker 1986; Cantner et al. 2007; Sengupta 1982).

Die DEA hat, als nicht-parametrischer Ansatz, den Nachteil, dass keine Aussage über die Güte des Gesamtmodells angegeben werden kann. Aus diesem Grund wurde die Bootstrap-DEA entwickelt. Unter Bootstrapping versteht man im Allgemeinen eine Methode zur Berechnung der Genauigkeit von statistischen Schätzern. Dabei wird zufällig aus einer Stichprobe eine beliebige Anzahl Datenpunkte gezogen (mit Zurücklegen). Die neue Stichprobe (Bootstrap-Stichprobe) hat genau so viele Elemente wie der ursprüngliche Datensatz. Das heißt, Datenpunkte sind evtl. mehrfach oder gar nicht vorhanden. Dieses Sampling wird mehrfach wiederholt und die statistischen Maße (Median, Standardabweichung, etc.) werden verglichen. Des Weiteren können Konfidenzintervalle abgeleitet werden (Bogetoft und Otto 2011). Dieses Verfahren kann auf das DEA-Modell angewendet werden. Dabei kann das Bootstrapping entweder auf Basis der Effizienzwerte erfolgen, oder die DEA wird auf Basis von Bootstrapping-Samples angewendet. Die erste Methode setzt voraus, dass die Effizienzwerte unabhängig voneinander sind. Bei Anwendung der DEA auf die Stichproben können für jede Einheit Mittelwert und Varianz bestimmt werden. Der Bootstrap-Algorithmus kann allerdings dazu führen, dass eine DMU nicht in einer Stichprobe und damit nicht in der Technologiemenge enthalten ist. Dies führt insbesondere bei effizienten Einheiten (im Standardmodell) zu Effizienzwerten (bootstrapped) größer 1. Diese Problematik kann man über eine Glättung der Stichprobe beheben und diese in den DEA-Algorithmus überführen (Bogetoft und Otto 2011). Durch die Anwendung wird eine „korrigierte“ Randproduktionsfunktion generiert.

3.1.3 Stochastic Frontier Analysis

Das Ziel der Stochastic Frontier Analysis ist, wie bei der DEA, ein Effizienzranking einzelner Unternehmen bzw. DMUs durch die Messung eines radialen Abstandes zur Randproduktionsfunktion. Das Verfahren wurde parallel von Meeusen und Broeck (1977) sowie Aigner et al. (1977) entwickelt. Des Weiteren basiert die Methode auf der Arbeit von Battese und Corra (1977). Grundlagen, Weiterentwicklungen und Anwendungen des Modells sind u. a. in Kumbhakar und Lovell (2003) und Kumbhakar et al. (2015) übersichtlich dargestellt.

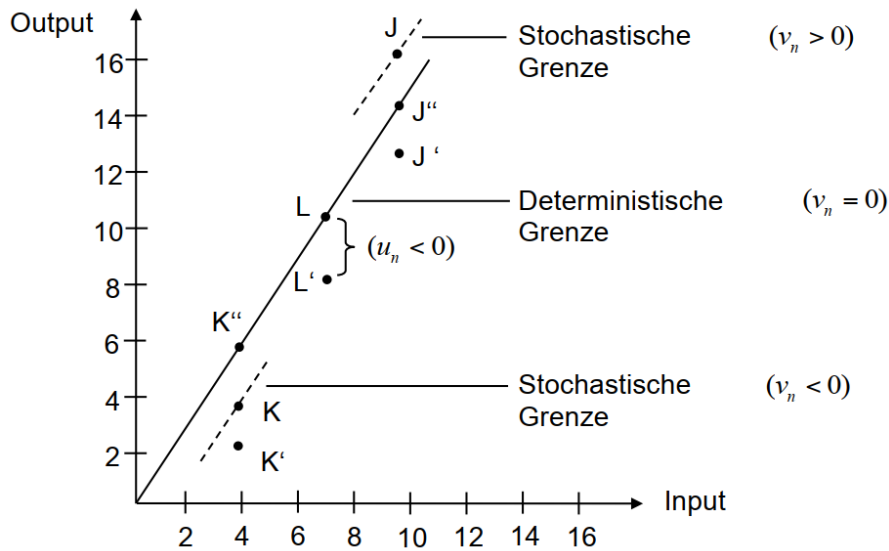


Die SFA ist ein regressionsanalytisches Verfahren, bei dem eine stochastische Randproduktionsfunktion erzeugt wird. Im Gegensatz zu deterministischen Verfahren wird eine Abweichung vom effizienten Rand nicht mehr ausschließlich als systematische Ineffizienz interpretiert, sondern beinhaltet auch einen Zufallsanteil (Aigner et al. 1977). Einen weiteren wesentlichen Unterschied stellt die Optimierung dar, welche nicht wie bei der DEA individuell erfolgt, sondern ganzheitlich für die Technologie. Das bedeutet, dass die genutzten Faktorgewichte für alle DMUs einheitlich sind (Bielecki 2011).

Während bei der DEA jeder ineffizienten Einheit mehrere Benchmark-DMUs zugewiesen werden, ist dies bei der SFA in der Regel lediglich jene Einheit mit dem maximalen Output-Input- oder minimalen Input-Output-Verhältnis. Durch diese DMU (maximal positives Residuum) wird die effiziente Grenze (vorab definierte Produktionsfunktion) geschätzt. Diese kann als deterministische Randproduktionsfunktion verstanden werden (Bielecki 2011).

Die stochastische Grenze (Zieloutput) wird für jede Einheit individuell bestimmt (siehe Abbildung 3-9) und repräsentiert den funktionalen Zusammenhang zwischen Output und Input, basierend auf einem Fehlerterm ε . Dieser wird in eine Ineffizienz-Komponente μ und die Störungskomponente ν untergliedert. Dabei müssen a priori Annahmen über die Verteilungsfunktion getroffen werden, z. B. Halb-Normalverteilung oder Gamma-Verteilung (Zhang und Garvey 2008). Während die Ineffizienz-Komponente immer negativ ist, kann der Störterm positive oder negative Werte annehmen. Der Fehlerterm wird für eine spezifische Einheit n aus der Differenz beider Komponenten gemäß Formel (3-15) gebildet.

$$\varepsilon_n = v_n - \mu_n \tag{3-15}$$



DMU	μ_n	ν_n	Zieloutput	Ineffizienz	Stoch. Effekt
J'	< 0	> 0	J	J'J''	J''J
K'	< 0	< 0	K	K'K	KK''
L'	< 0	0	L	L'L	-

Abbildung 3-9: Stochastische Randproduktionsfunktionen (Bielecki 2011)



Eine SFA kann sowohl auf Querschnittsdaten als auch auf Paneldaten angewendet werden (Kumbhakar und Lovell 2003). Die hier aufgezeigte Beschreibung des mathematischen Grundmodells erfolgt auf Basis von Querschnittsdaten und ist in Formel (3-16) dargestellt. Aus illustratorischen Gründen wird eine lineare Funktion angenommen, andere funktionale Zusammenhänge können jedoch impliziert werden (z. B. eine Cobb-Douglas oder Translog-Produktionsfunktion). Dabei entspricht y_n dem Output der DMU n und i allen Inputfaktoren. Des Weiteren ist der kombinierte Fehlerterm (siehe Formel (3-17)) Bestandteil der Formel. β_0 repräsentiert die Konstante der linearen Gleichung und β_i die zu schätzenden Inputfaktor-Koeffizienten.

$$y_n = \beta_0 + \sum_{i=1}^I \beta_i x_{i,n} + \varepsilon_n \quad (3-16)$$

Ein wesentlicher Vorteil der SFA ist die Qualitätsüberprüfung des verwendeten Modells. Dem gegenüber steht die Einschränkung, dass im Grundmodell lediglich ein Output berücksichtigt werden kann (Coelli et al. 2005). Alternative, regressionstheoretische Ansätze diskutieren u. a. Collier et al. (2011). Eine Erweiterung der SFA durch die Nutzung von Copulae⁴² (Embrechts et al. 2001; Hofert et al. 2019) wird von Carta und Steel (2012) vorgeschlagen. Des Weiteren stellt die Schätzung der funktionalen Zusammenhänge sowie jene der Fehlerterme eine Herausforderung dar. Bei insuffizienten Modellannahmen sind die Effizienzergebnisse möglicherweise fehlerbehaftet (Carree 2002). Diese umfangreichen Annahmen müssen jedoch a priori erfolgen. Eine detaillierte Übersicht zu beiden Verfahren inklusive Vor- und Nachteilen bieten Bogetoft und Otto (2011).

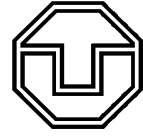
3.1.4 Methodenauswahl für die Analyse im ATM

In Abschnitt 3.1.1 wurden verschiedene Ansätze zur Bestimmung der Leistungsfähigkeit diskutiert. Die Indexziffer-Methode stellt dabei die einfachste Form eines Performance-Indikators dar, welcher einen Output ins Verhältnis zu einem Input stellt. Die Nutzung dieser Methodik hat jedoch bedeutende Nachteile. Durch Substitution eines oder beider Parameter (Input und/oder Output) verändert sich auch das Ranking (Buxbaum und Standfuss 2014; Hammerschmidt 2006). Somit wird eine Betrachtung der Gesamteffizienz, speziell bei großen Datenmengen, erschwert (Nyhan und Martin 1999). Des Weiteren werden bei Indexziffern lediglich ein Input und ein Output betrachtet, sodass Produktionsprozesse mit multiplen Inputs und/oder Outputs unzureichend abgebildet sind (Gladen 2014). Da die Modellierung der Produktionsprozesse im ATM mehrere, parallel verwendete Input- und Output Parameter berücksichtigen muss⁴³, ist die Nutzung dieses Ansatzes nicht zielführend. Zwar besteht die Möglichkeit, mehrere Parameter zu aggregieren, jedoch setzt dies eine geeignete Gewichtung voraus (Bielecki 2011).

Eine mögliche Alternative repräsentieren ökonometrische Methoden, deren bekanntesten Vertreter im Benchmarking die Stochastic Frontier Analysis darstellt. Dieser Ansatz ist robust gegenüber Extremwerten und trennt, im Unterschied zu deterministischen Ansätzen, eine Abweichung vom Optimum

⁴² Mittels Copulae (oder Copulas) können stochastische Abhängigkeiten modelliert werden. Sie repräsentieren multivariate Verteilungsfunktionen. Eine Copula stellt einen funktionalen Zusammenhang zwischen der Marginalverteilung von Zufallsvariablen und der Wahrscheinlichkeitsverteilung her.

⁴³ Wenngleich dies abhängig von der betrachteten betrieblichen Ebene ist. Referenziert ist die ANSP Ebene.



in eine Ineffizienz- und eine Störkomponente. Allerdings setzt die Anwendung umfangreiche a priori Annahmen hinsichtlich der Modellierung einer Produktions- bzw. Kostenfunktion im ANSP-Sektor voraus. Durch die Gefahr von Fehlannahmen wächst auch die Möglichkeit fehlerhafter Ergebnisse sowie deren Interpretation (siehe auch Abschnitt 3.2.2). Da Leistungsindikatoren und die dadurch abgeleiteten Verbesserungspotentiale unter Umständen auch als Grundlage für ökonomische Regulierungen dienen, sind Unsicherheiten aufgrund der politischen Tragweite zu vermeiden (Standfuss et al. 2020).

Die DEA setzt keine solchen a priori Annahmen voraus. Allerdings bedingt sie eine gute Qualität der Daten, da die berechnete Effizienz sensibel auf Extremwerte reagiert. Dies ist insbesondere dann der Fall, wenn die Randproduktionsfunktion aufgrund eines Fehlers verschoben wird und dadurch alle im Abschnitt befindlichen DMUs auf Basis der fehlerhaften Technologiegrenze bewertet werden. Dieses Risiko kann allerdings durch eine intensive Datenanalyse minimiert werden. Zudem werden viele Nachteile der Grund-DEA durch die in Abschnitt 3.1.2.3 diskutierten Modifikationen vermieden oder zumindest abgeschwächt.

Prinzipiell kann es als sinnvoll erachtet werden, sowohl die DEA als auch die SFA anzuwenden und die Ergebnisse komparativ zu diskutieren. Durch die methodischen Unterschiede sind die Ergebnisse jedoch kaum vergleichbar. Dies kann dazu führen, dass beide Ergebnisse signifikant abweichen, wie bspw. European Commission (2018) sowie Neiva (2014) zeigen. Stattdessen werden innerhalb der vorgelegten Arbeit sowohl verschiedene DEA-Modelle als auch verschiedene Modifikationen des Grundmodells angewendet und deren Ergebnisse verglichen. Des Weiteren wird die Anwendbarkeit der jeweiligen Methode auf den vorhandenen Datensatz diskutiert.

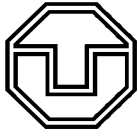
3.2 Status Quo der Leistungsbewertung im ATM

3.2.1 EUROCONTROL

3.2.1.1 Single European Sky

Die Bewertung der Performance der einzelnen Flugsicherungsdienstleister ist ein wesentlicher Bestandteil des Single European Sky Konzeptes mit dem Ziel, kostengünstige und effizientere Flüge sowie eine Reduktion der Verspätungen zu realisieren. Diese Bestrebungen werden seitens der Europäischen Kommission im Single European Sky Konzept verfolgt, welches die zwei Verordnungs-Pakete SES I und SES II umfasst. Das Konzept beinhaltet neben strukturellen Maßnahmen auch technologische Neuerungen im Rahmen von SESAR. Darin werden u. a. neue Verfahren oder Tools konzipiert, um das europäische Air Traffic Management zu optimieren (European Commission 2020b). Ziel ist es einen freien und funktionsfähigen Wettbewerb ermöglichen

Das SES I Paket umfasst zunächst die Trennung von Aufsichts- und Durchführungsaufgaben. Diese ist in der Verordnung Nr. 549/2004, Art. 4 Abs. 1 und 2 der Europäischen Union festgeschrieben. Dafür ist für alle Mitgliedsstaaten die Einrichtung einer vom Flugsicherungsdienstleister unabhängigen Aufsichtsbehörde bindend. In Deutschland übernimmt diese Aufgabe das Bundesaufsichtsamt für Flug-



sicherung (BAF) in Langen. Zur Umsetzung wurde Artikel 87 des Grundgesetzes geändert. Dies beinhaltet auch, dass ausländische Organisationen nach Zulassung in der Europäischen Union die Dienstleistungen einer Flugsicherung übernehmen dürfen (Deutscher Bundestag 2009).

Eine Wettbewerbsgrundlage wurde durch die Einrichtung eines Zertifizierungssystems für Flugsicherungsdienste geschaffen. Das erteilte Zertifikat beinhaltet Rechte und Pflichten eines Flugsicherungsunternehmens und wird von der nationalen Aufsichtsbehörde bei Erfüllung der Kriterien erteilt. Dadurch wird es ermöglicht, ein Unternehmen unabhängig vom Ort der Zertifizierung mit den Dienstleistungen einer Flugsicherung zu beauftragen. In Deutschland wird diese Aufgabe vom BAF wahrgenommen. Den rechtlichen Rahmen bildet die Verordnung Nr. 550/2004 der Europäischen Union. Das SES I Paket umfasst weiterhin eine Neuordnung des Luftraumes durch Implementierung funktionaler Luftraumblöcke gemäß Verordnung Nr. 551/2004 der EG sowie die Verordnung Nr. 552/2004, welche eine Interoperabilität, respektive Harmonisierung aller Systeme adressiert. Hierbei wird vor allem die betriebliche Optimierung angestrebt.

Auf Basis der durch SES I gewonnenen Erkenntnisse wurden für SES II Regularien zur Verbesserung von Leistungsfähigkeit und Nachhaltigkeit angepasst und ergänzt. Die Verordnungen des SES II Paketes umfassen vier Kernpunkte:

- Leistungsregulierung
- Technologie
- Sicherheit
- Kapazität am Boden

Die Leistungsregulierung umfasst zunächst die Bewertung der Flugsicherungen durch eine unabhängige Behörde. Das implementierte Performance Scheme umfasst je nach Referenzperiode verschiedene Kennzahlen hinsichtlich Sicherheit, Kosteneffizienz, Kapazität und Umweltaspekten. Die Zielerreichung der Key Performance Areas (KPAs) wird durch die Key Performance Indicators (KPIs) analysiert. Des Weiteren wird die Implementierung der FABs adressiert und unterstützt. Die dritte Komponente umfasst die Netzwerkmanagementfunktion. Hier ist vor allem das Slot Management sowie das Design des Streckennetzes verankert.

Für das SES II Paket wurden nicht nur die Leistungsmetriken, sondern auch das Bepreisungsschema angepasst. Ziel war der Abbau von Ineffizienzen einer reinen Vollkostendeckung, die bis ins Jahr 2011 als Grundlage der Flugsicherungsgebühren diente (European Commission 2016). Die Einführung einer gemeinsamen Gebührenregelung ist in der Verordnung 1191/2010 der Europäischen Union verankert und tangiert primär die KPA Cost Efficiency. Sie umfasst zum einen die ökonomische Regulierung, welche eine Risikoteilung (zwischen ANSP und Luftraumnutzern) hinsichtlich Kosten und Verkehr beinhaltet (EUROCONTROL 2020j). Zum anderen ist in der Verordnung eine Anreizregulierung hinsichtlich des Streckenfluges implementiert. Diese werden innerhalb der FABs umgesetzt und umfassen u. a. Pönale bzw. Boni für Abweichungen von den Performancezielen (FABEC 2012).

Die Verknüpfung von Gebühren- und Leistungssystem ist in Abbildung 3-10 illustriert. Das Leistungssystem definiert Kosteneffizienzziele auf europäischer Ebene, welche mit jenen auf der nationalen

Ebene gekoppelt sind. Diese Ziele werden unter Berücksichtigung des prognostizierten Verkehrs und der geplanten Kosten aufgestellt. Die Verkehrsprognose wird mit dem tatsächlich stattfindenden Verkehr abgeglichen, welches im Verfahren zur Verkehrsrisiko- und Kostenrisikoteilung gemäß Verordnung 1191/2010 der Europäischen Union implementiert ist. Sie umfasst Szenarien, welche die Abweichung zwischen prognostiziertem und tatsächlichem Verkehr beschreiben und zu treffende Maßnahmen hinsichtlich Kostenübertragung bei zusätzlichem Gewinn oder Verlust festlegen.

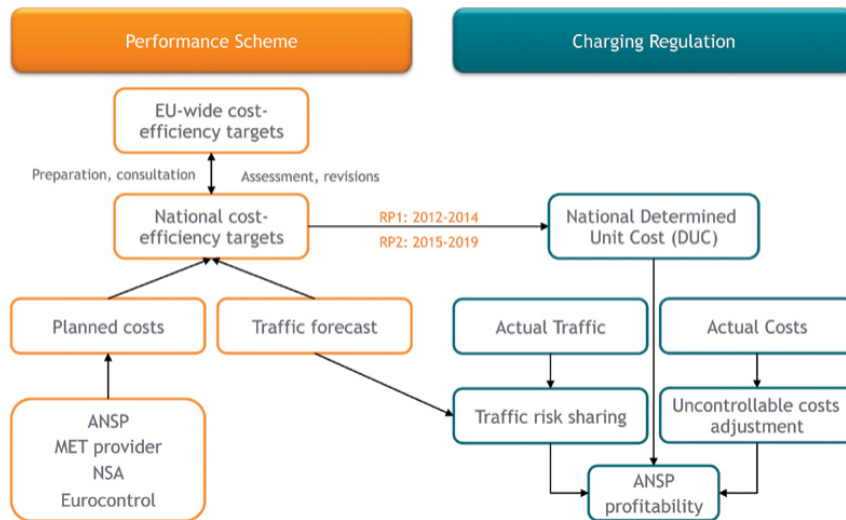


Abbildung 3-10: Gebühren- und Leistungssystem der europäischen Flugsicherungsunternehmen (Helios 2013)

Die Basis zur Leistungsbewertung bilden die Verordnungen Nr. 691/2010, Nr. 1794/2006 und 1194/2010 der Europäischen Union. Die Umsetzung erfolgt in Referenzperioden (Reference Period, RP)⁴⁴. Darin werden für die Mitglieder bindende Ziele für einzelne Leistungsindikatoren vorgeschlagen, geprüft, revidiert und beschlossen. Die Flugsicherungen stellten ihrerseits nationale oder übernationale Performancepläne auf, die zur Realisierung der Vorgaben beitragen sollen. Diese Verbindlichkeiten gelten auch für Drittstaaten, die sich um eine Mitgliedschaft bewerben (European Commission 2016). Die Bewertung wird von Flugsicherungsunternehmen unabhängigen Gremien innerhalb der EUROCONTROL durchgeführt und publiziert.

3.2.1.2 Performance Scheme

Im August 2008 initiierte die Performance Review Commission (PRC) ein Diskussionspapier, welches EU-weit an die entsprechenden Stakeholder gesendet wurde. Daraus entwickelte sich ein erstes Performance Scheme für Flugsicherungsdienstleister. Aus diesen Bestrebungen gingen mehrere Verordnungen für die Referenzperiode 1 (RP1, 2012 bis 2014) hervor^{45,46}.

⁴⁴ Seit 2020 ist die Referenzperiode 3 (RP3) aktuell.

⁴⁵ Regulation 549/2004.

⁴⁶ Regulation 691/2010.



RP1 umfasste vier KPAs, in denen acht KPIs überwacht wurden (EUROCONTROL 2010):

- Cost Efficiency
 - Terminal air navigation service costs and unit rates
- Capacity
 - Total ATFM delay attributes to terminal and airport ANS
 - Additional time in the taxi out phase
 - Additional time for Arrival Sequencing and Metering Area (ASMA)
- Safety
 - Safety management effectiveness
 - Percentage of application of severity classification of the risk analysis tool
 - Minimum level of measure of just culture
- Environment
 - Effective use of civil/military airspace structures

Zudem wurden für einige Performance Indikatoren auch Zielwerte festgesetzt. Sie umfassten eine Reduktion der geflogenen Umwege um 0,75 %, eine Reduktion des ATFM-bedingten Enroute-Delay auf maximal 0,5 Minuten pro Flug sowie maximale Kosten pro Dienstleistungseinheit von 53,92€⁴⁷ (European Commission 2011). Die KPA Safety wurde nicht explizit berücksichtigt, spielte jedoch eine übergeordnete Rolle, da keine Maßnahme das Sicherheitsniveau reduzieren durfte.

Referenzperiode 2 (RP2) umfasste die Jahre 2015 bis 2019 und adressierte innerhalb des Performance Plans 24 Indikatoren in vier KPAs (EUROCONTROL 2013a). Diese sind in Abbildung 3-11 dargestellt. Dabei bilden die Vierecke die KPAs. KPIs werden abgerundet dargestellt.

Die Unterschiede zur Leistungsbewertung im RP1 betreffen zwei wesentliche Punkte. Zum einen wurden weitere KPIs ergänzt oder anderen KPAs zugeordnet⁴⁸. Zum anderen setzte die Europäische Kommission Zielwerte für bis dahin nur überwachte Indikatoren fest. Zielwerte („Targets“, grün markiert) und Beobachtung („Monitoring“, gelb markiert) obliegt entweder dem ANSP selbst, ihrem zugehörigen Functional Airspace Block oder der Europäischen Union⁴⁹. Die Regelungen und Performance-Targets der zweiten Referenzperiode sind unter anderem in European Commission (2012b) und European Commission (2014b) festgehalten.

Seit 2020 ist die dritte Referenzperiode (RP3) aktiv. Die Durchführungsverordnung (EU) 2019/317 der Kommission vom 11. Februar 2019 zur Festlegung einer Leistungs- und Gebührenregelung im einheitlichen europäischen Luftraum und zur Aufhebung der Durchführungsverordnungen (EU) Nr. 390/2013 und (EU) Nr. 391/2013 beinhaltet verschiedene Änderungen der wirtschaftlichen Regulierung des ATM / der ANS. Die Durchführungsverordnung (EU) Nr. 2019/317 der Kommission enthält Änderungen im Vergleich zu den Durchführungsverordnungen (EU) Nr. 390/2013 und (EU) Nr. 391/2013,

⁴⁷ Die Werte beziehen sich auf das Jahr 2014 im Verhältnis zum Ursprungsjahr 2009.

⁴⁸ Bspw. ist der KPI „Additional Time in Taxi-Out Phase“ vom KPA Capacity zum KPA Environment transferiert worden.

⁴⁹ Existent sind des Weiteren Mischformen, bei denen mehrere Ebenen das Monitoring übernehmen, sowie das Monitoring der Flughäfen.

z. B. das Recht der nationalen Aufsichtsbehörden, entweder auf nationaler Ebene oder auf der Ebene der funktionalen Luftraumblöcke, Leistungspläne zu erstellen. Zudem beinhaltet sie die Anwendung einer vereinfachten Gebührenregelung (European Commission 2018, 2020a).

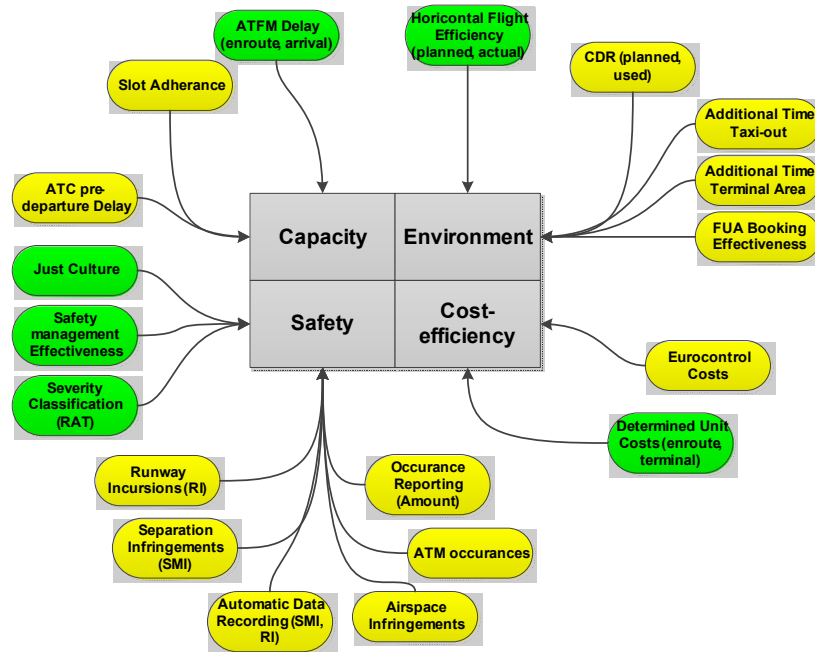


Abbildung 3-11: Key Performance Areas und Indicators in RP2
Angelehnt an EUROCONTROL (2013a)

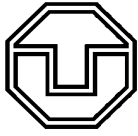
3.2.1.3 ATM Cost-Effectiveness Report

Mittels des seit 2002 jährlich erscheinenden ATM Cost-Effectiveness (ACE) Reports, im Folgenden kurz ACE-Report genannt, publiziert die Performance Review Commission, unterstützt durch die Performance Review Unit (PRU), flugsicherungs-spezifische Daten. Dabei fokussiert sich der Bericht sowohl auf Leistungsparameter des gesamten europäischen Luftraumes als auch der einzelnen ANSPs und ACCs. Des Weiteren sind Änderungen der Indikatoren über mehrere Zeitperioden dargestellt. Datengrundlage für den ACE-Report bilden Parameter und deren Ausprägung, die durch die PRU erfasst wurden. Die Akquise der Daten erfolgt über Excel-Tabellen, die von jedem Mitgliedsstaat ausgefüllt und an die PRU übersandt werden. Je nach Struktur der Betriebsebenen umfasst die Datenerhebung ca. 230 Parameter. Für jeden dieser Parameter existieren Vorgaben zur Erfassung und Kalkulation, sodass eine international konsistente Erhebung gewährleistet sein soll. Die Daten werden in „Enroute“, „Terminal“ und „Gate-to-Gate“⁵⁰ Operations untergliedert.

Der letzte für diese Arbeit herangezogene ACE-Report erschien im Mai 2019 (EUROCONTROL 2019a) und beinhaltet die Daten von 38 Flugsicherungsunternehmen für das Jahr 2017. Er umfasst zunächst die Entwicklung der Kosten hinsichtlich verschiedener Komponenten (u. a. Personalkosten) sowie den pan-europäischen Vergleich (z. B. hinsichtlich der ATM/CNS⁵¹ Kosten). Zudem werden mögliche exogene und endogene Einflüsse auf die Performance benannt (u. a. Lebenshaltungskosten und

⁵⁰ Gate-to-Gate ist Summe aus Enroute und Terminal Operations.

⁵¹ Communication, Navigation, Surveillance.



Verkehrsvariabilität). Als vergleichender Leistungsindikator wird u. a. Kosten je Composite Flight Hour (CFH⁵²) dargestellt (siehe Abbildung 3-12). Diese umfassen neben den ATM/CNS Kosten auch die Kosten der Verspätungen. Analysen zu Verspätungen repräsentieren einen umfangreichen Teil des ACE-Reports. Es wird jedoch auch das Ranking ohne Verspätungskosten⁵³ publiziert.

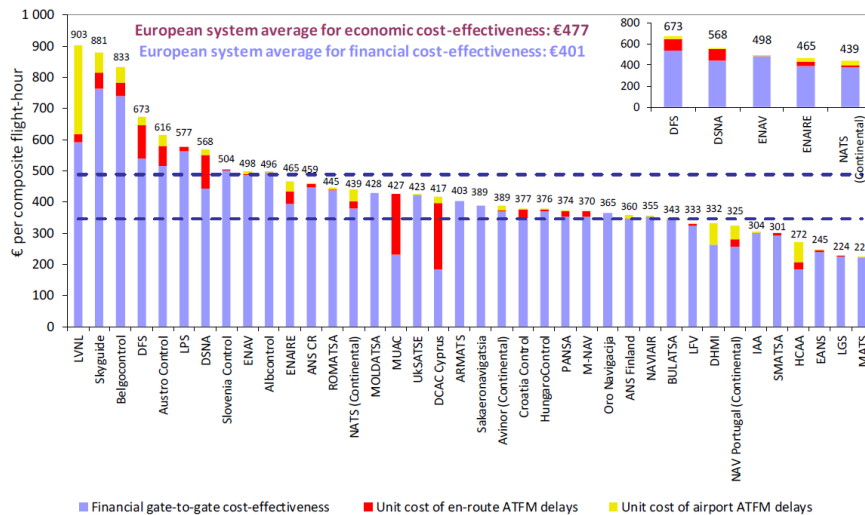


Abbildung 3-12: Gate-to-Gate Kosteneffizienz, gemessen in Kosten pro CFH (EUROCONTROL 2019a)

Neben diesen kostenbasierten Analysen sind auch betriebliche Indikatoren im ACE-Report enthalten. So wird bspw. für jede ANSP die Lotsenproduktivität bestimmt und in ein Ranking überführt. Dieses Produktivitätsranking ist für das Jahr 2017 in Abbildung 3-13 dargestellt. Signifikant ist der hohe Produktivitätswert von MUAC, der etwa 13-mal höher ist als die Produktivität der moldawischen Flugsicherung MOLDATSA.

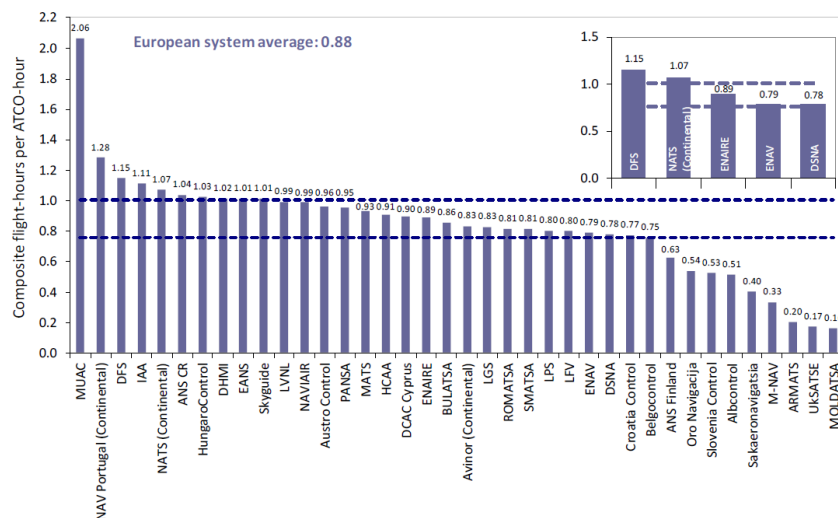
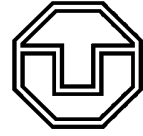


Abbildung 3-13: Lotsenproduktivität 2017, gemessen in CFH pro Lotsenstunde (EUROCONTROL 2019a)

⁵² Die Composite Flight hours dienen als Näherungswert für den Gesamtoutput. Dafür werden die Flugstunden und Flugbewegungen am Flughafen aggregiert. Diese werden entsprechend ihrem pan-Europäischen Stückkostenanteil gewichtet.

⁵³ „Financial Cost-Effectiveness“.



Der Bericht umfasst neben den Leistungsindikatoren auch Vergleiche hinsichtlich der Kosten pro Lotsenstunde. In 2017 kostete eine Lotsenstunde die DFS (Rang 1) mit 232€ etwa das 19-fache im Vergleich zur ukrainischen UKSATSE (Rang 38). Dies kann bspw. in den unterschiedlichen absoluten Lohnniveaus begründet liegen. Weitere Analysen betreffen u. a. den Vergleich der Arbeitskosten und Supportkosten pro CFH. Ergänzt sind außerdem sogenannte „Fact Sheets“ für jede ANSP. Diese fassen alle wesentlichen Determinanten (Komplexität, Variabilität, Produktivität, etc.) zusammen. Die zugrundeliegenden Daten sind, bis auf wenige Ausnahmen, online verfügbar (EUROCONTROL 2020h, 2020i).

3.2.1.4 Performance Review Report

Der Performance Review Report (PRR) wird, wie der ACE-Report, von der Performance Review Commission publiziert. Die Veröffentlichungen sind seit 1998 verfügbar. Die Berichte adressieren neben allgemeinen Verkehrsdaten die Themen Safety, Capacity, Environment und Cost-Efficiency. Die nachfolgenden Ausführungen beziehen sich auf den Report vom Mai 2019 (EUROCONTROL 2019h), welcher die Daten von 41 Staaten⁵⁴ für das Jahr 2018 (2017 bei Kosteneffizienz) gegenüberstellt.

Der PRR analysiert zunächst die Entwicklung der Verkehrsnachfrage. Im Vergleich zu 2017 stieg die Anzahl der Flüge für alle betrachteten ANSPs an, durchschnittlich um 2,8%. Auch die Dienstleistungseinheiten (im Bericht Total Service Units, TSU) stiegen durch das höhere durchschnittliche Startgewicht der Luftfahrzeuge sowie die längere Distanz pro Flug. Dem gegenüber stand die Verringerung der Service-Qualität, die in dem Bereich durch die Verspätungsminuten repräsentiert wird.

Der erste Schwerpunkt des PRR fokussiert verkehrsspezifische Informationen, u. a. den Verkehr differenziert nach Flughöhe sowie den Vergleich zwischen prognostizierter und tatsächlicher Nachfrage. Die Entwicklung der Nachfrage wird zudem nach Marksegment⁵⁵, Art des Fluges⁵⁶ und der genutzten Metrik⁵⁷ differenziert. Auch räumliche (Verteilung des Wachstums über die ANSPs) und zeitliche Aspekte (Verkehr je Wochentag) werden adressiert. Einen weiteren Schwerpunkt bilden Betrachtungen zur Komplexität, differenziert nach zeitlicher Entwicklung, Flugfläche und ACC.

Der zweite Teil beinhaltet Informationen hinsichtlich Umweltauswirkungen, z. B. CO₂-Emissionen, Treibstoffeffizienz und Lärm. Es wird davon ausgegangen, dass 6% der Kohlenstoffdioxidemissionen durch die ANS beeinflussbar sind. Dies umfasst alle Flugphasen⁵⁸, die innerhalb des Reports auch näher diskutiert werden.

Als dritter Schwerpunkt beschäftigt sich der PRR mit der KPA Safety und vergleicht die Entwicklung hinsichtlich sicherheitsrelevanter Vorfälle, wie Un- und Zwischenfälle, Staffelungsunterschreitungen (Separation Minima Infringements, SMIs) und Runway Incursions (RIs). Bis auf zwei Ausnahmen stiegen diese Vorfälle in Relation zum Vorjahr an.

⁵⁴ Inklusive Israel und Marokko, siehe Anhang, Abbildung A 2.

⁵⁵ Z. B. Low Cost Carrier.

⁵⁶ Z. B. Überflüge.

⁵⁷ Flüge, Flugstunden und Distanz.

⁵⁸ Die Phasen beinhalten Taxi, Steigflug, Streckenflug, Sinkflug.



Die Teile vier und fünf stellen die ANS Performance für den Streckenverkehr und den Betrieb an Flughäfen gegenüber. Die Verspätungsminuten waren in 2018 mehr als doppelt so hoch wie im Vorjahr. Das ATFM-Delay wird zudem in Ursachen untergliedert, z. B. ATC Kapazität, Personal oder Wetter. Seit 2013 verzeichnet die PRC einen Anstieg des ATFM-Delay, welcher jedoch monatlich stark schwankt. So sind von Juni bis August die Werte signifikant höher als bspw. in November und Januar. Zudem wird die „Horizontale Flugeffizienz“ (Horizontal Flight Efficiency, HFE)⁵⁹ bewertet. Diese blieb über die vergangenen Jahre nahezu konstant bei 97,3 %. In diesem Zusammenhang werden die Konzepte Free Route Airspace und Flexible Use of Airspace diskutiert (EUROCONTROL 2015, 2019e). Die Verspätungsminuten spielen ebenfalls eine zentrale Rolle bei der Bewertung des Flughafenbetriebes. Zunächst wird die Nachfrage der 30 größten europäischen Flughäfen gegenübergestellt. Die Performance wird dann mithilfe des ATFM delays sowie der ASMA Additional Time bewertet. Des Weiteren betrachtet der Bericht die Einhaltung der ursprünglich geplanten Slots (slot adherence). Einen weiteren Schwerpunkt bildet die Implementierung des Airport Collaborative Decision Making (A-CDM). In früheren Berichten wurden zudem der Continuous Descent Approach (CDA) diskutiert.

Den letzten Teil des PRR repräsentiert die ANS Kosteneffizienz. Diese wird in Kosten pro Dienstleistungseinheit dargestellt. Da die DLEn über die vergangenen Jahre anstiegen und die Gesamtkosten relativ konstant blieben, sind die Stückkosten insgesamt gesunken. Diese Kosten werden weiter differenziert, zum einen nach Art (z. B. Personalkosten) und zum anderen nach ANSP. Die Untersuchungen und Ergebnisse sind sowohl für Enroute- als auch Terminal Services verfügbar. Für beide Indikatoren wird zudem eine Prognose für die Entwicklung bereitgestellt.

3.2.1.5 Weitere Berichte und Studien

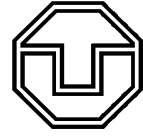
Für die in Referenzperiode 1 festgesetzten KPAs und KPIs erfolgten jährliche Berichte über die Leistungen der einzelnen Flugsicherungen. Außerdem sind geplante Maßnahmen der Mitgliedsstaaten zur Erfüllung der Vorgaben integriert. Weiterhin werden Empfehlungen an die Europäische Kommission formuliert, welche Maßnahmen zur Erreichung der in RP1 gesetzten Ziele beinhalten (EUROCONTROL 2013b, 2014).

Monatlich veröffentlicht wird der Network Operations Report. Dieser umfasst im Wesentlichen die Verkehrsnachfrage im Vergleich zu den Vormonaten sowie eine besondere Betrachtung des ATFM-Delays nach Kategorien. Außerdem werden die Gründe für Verspätungen adressiert, bspw. Strukturierungsmaßnahmen oder Wetterlagen (EUROCONTROL 2016g, 2018e, 2020g).

Neben diesen Berichten publiziert EUROCONTROL zahlreiche Daten und Analysen im Flughafenbereich. Diese betrachten meist die Kapazitäten an Flughäfen sowie den Einfluss der Flugsicherungsdienstleistungen. Sie repräsentieren eine Spezialisierung zu den Benchmarking Reports und können daher als tiefgreifende Ergänzung gesehen werden (EUROCONTROL 2011).

Zwei weitere, regelmäßig erscheinende Berichte vergleichen das europäische und das amerikanische Luftverkehrsmanagement. Sie adressieren die operative (EUROCONTROL und FAA 2019a)

⁵⁹ Abweichung der Trajektorie von der Großkreisdistanz.



und finanzielle (EUROCONTROL und FAA 2019b) Heterogenität zwischen beiden ATM Systemen. Neben der Identifikation der Unterschiede werden auch Performancekennzahlen der Flugsicherungen verglichen. Die Vergleichbarkeit beider Systeme ist jedoch teilweise nicht gegeben. So weisen Standfuss und Whittome (2019a) nach, dass die unterschiedliche Performance maßgeblich auf die operative Heterogenität zurückzuführen ist. Auch die Finanzierung der ANS Dienstleistungen unterscheidet sich signifikant (Whittome und Standfuss 2018).

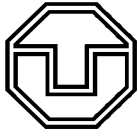
3.2.2 Überblick zu akademischen Studien

In den vergangenen Jahren ist das Interesse an einer Leistungsbewertung im Flugsicherungsbe-
reich gestiegen. Neben den offiziellen Benchmarking Reports beschäftigten sich daher auch akademi-
sche Studien mit der Analyse der Leistungsfähigkeit europäischer Flugsicherungsdienstleister. In die-
sem Abschnitt werden jene Papiere zusammengefasst, deren Analysen auf der Nutzung multipler Inputs
und/oder Outputs basieren. In der Literatur hat sich die Nutzung der Data Envelopment Analysis
und/oder Stochastic Frontier Analysis etabliert. Im Folgenden wird auf eine Auswahl dieser Studien
eingegangen.

Eine erste Studie zum Thema ANSP-Benchmarking, welche multiple Faktoren berücksichtigt, re-
präsentiert (NERA 2006). Auf Basis der SFA wurde eine Cobb-Douglas-Kostenfunktion geschätzt und
über Fixed und Random Effects Modelle ausgewertet (siehe auch Abschnitt 4.3). Das zeitinvariante
Random Effects Modell lieferte die besten Ergebnisse hinsichtlich Robustheit, Signifikanz und Interpre-
tierbarkeit. Es wird jedoch darauf hingewiesen, dass durch fehlende Varianz in den exogenen Einfluss-
größen das Modell Ineffizienzen möglicherweise überschätzt.

Eine Abhandlung über die potentiellen ökonomischen Vorteile von FABs veröffentlichten Button
und Neiva (2013). Mittels einer Bootstrap VRS-DEA wurde ein Modell getestet, welches aus zwei kos-
tenbasierten Inputs und drei Outputs bestand (Flugstunden, Bewegungen am Flughafen und Verspätun-
gen). Die DEA Werte wurden je FAB aggregiert. Eine OLS-Regressionsanalyse testet den Einfluss ver-
schiedener Faktoren auf die Effizienzwerte. Das Verfahren ist nur oberflächlich erklärt und begründet.
So fehlt z. B. die Angabe, ob die DEA input- oder outputorientiert angewendet wurde. Es findet kein
Robustheitstest hinsichtlich verschiedener DEA-Modelle statt. Auch wird nicht erklärt, warum eine OLS-
Regression verwendet wird (siehe hierzu auch Abschnitt 4.2). Die Nutzung von zwei Input- und drei
Outputgrößen ist bei der Größe des verwendeten Datensatzes kritisch zu sehen. Die gleichen Kritik-
punkte gelten für das Papier von Button und Neiva (2014), welches in Methodik und Modellierung über-
einstimmt. Lediglich die Koeffizienten der Regressionsanalyse wurden der Zielvorgabe, die ökonomi-
sche Effizienz einzelner ANSPs zu untersuchen, angepasst.

Basierend auf der DEA Methode untersuchen Čujić et al. (2015) die Effizienz europäischer Flug-
sicherungen in den Jahren 2009-2011. Sie nutzen ein Modell mit zwei kostenbasierten Inputs und drei
Outputs (Verspätungen, CFH und Gesamteinnahmen). Die Ergebnisse sind jedoch weder hinsichtlich
der Effizienzwerte noch der Rankings robust. So wird die Anzahl der effizienten Einheiten im betrachte-
ten Zeitraum verdoppelt. Es wird nicht erklärt, ob die DEA unter Annahme von konstanten oder variablen
Skalenerträgen durchgeführt wurde. Die Ergebnisse sind zudem kontraintuitiv (u. a. hohe Effizienz von



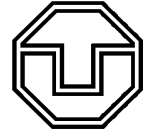
UkSATSE, MoldATSA und NATA Albania). Außerdem weist die Modellierung erhebliche Verbesserungspotentiale auf. Zum einen sind die Inputs rein kostenbasiert und für die Outputs werden teils finanzielle, teils operative Werte verwendet. Zum anderen ist die parallele Verwendung von Einnahmen und Composite Flight Hours kritisch zu sehen, da Flugstunden und Distanzen hoch korrelieren, die Distanzen einen wesentlichen Einfluss auf die Service Units haben und diese wiederum die Einnahmen der Flugsicherung bestimmen. Wie bei Button und Neiva (2013) sowie Button und Neiva (2014) führt die parallele Nutzung von fünf Faktoren in einer DEA aufgrund der wenigen Beobachtungen zu vielen effizienten Einheiten.

Im Jahr 2014 erschien die Dissertation von Rui Neiva (Neiva 2014). Darin wird umfangreich die Regulierungstheorie und -geschichte beschrieben. Er wendet für das Performance Benchmarking sowohl die DEA als auch die SFA an. Dabei werden sowohl FAB- als auch ANSP-level abgedeckt. Die Modellierung unterscheidet sich dabei nicht von Button und Neiva (2013) bzw. Button und Neiva (2014). Zudem umfasst die Dissertation eine qualitative SWOT⁶⁰-Analyse, den ‚Differences-in-Differences‘ Ansatz sowie eine räumliche Regression. Wesentliche Schwächen sind die fehlende Datenanalyse, die nur rudimentäre Beschreibung der Methodik und deren Auswahlkriterien, sowie die fehlende Ableitung von Maßnahmen aus der Ursachenanalyse. Weiterhin fehlen Robustheitstests. Zwar nutzt der Autor zwei Methoden, diese sind allerdings aufgrund der unterschiedlichen mathematischen Modelle kaum vergleichbar. Die Ergebnisse weichen daher signifikant voneinander ab.

Arnaldo et al. (2014) nutzen ebenfalls die DEA zur Effizienzmessung 35 europäischer ANSPs in den Jahren 2001-2011. Die Autoren liefern eine umfangreiche Beschreibung der verwendeten Daten und der daraus erfolgten Modellierung. Sie testen verschiedene Ansätze hinsichtlich Orientierung (Input vs. Output) sowie Skalenerträgen (fix versus variabel). Die Ergebnisse zeigen, dass zu viele Faktoren innerhalb eines DEA-Modells gleichzeitig betrachtet wurden: Fast die Hälfte der DMUs wird bei der VRS-DEA als effizient ausgewiesen. Dadurch sind die Ergebnisse scheinbar robust. Sie unterscheiden sich primär zwischen den unterstellten Skaleneffekten: Die Ergebnisse der beiden Modelle unter Annahme konstanter Skalenerträge sind identisch.

Bilotkach et al. (2015) berechnen Kosteneffizienz sowie Produktivität der europäischen ANSPs mittels eines Malmquist-DEA-Modells. Für die Untersuchung standen Daten der Jahre 2002 bis 2011 zur Verfügung. Die Autoren nutzen ‚Kontrollierte Flugstunden‘ und ‚Flugbewegungen am Flughafen‘ als Outputs sowie die Gate-to-Gate ATM/CNS-Kosten als Input. Des Weiteren werden zu Bestimmung der allokativen Effizienz die Input-Preise für Lotsen, sonstiges Personal, Kapital und sonstige Ressourcen implementiert. Die berechneten Effizienz-Scores werden mittels Bootstrapping Algorithmus „korrigiert“ und in den Ergebnissen das zugehörige 95 %-Konfidenzintervall angegeben. Die Autoren konstatieren, dass die Kosteneffizienz signifikant zwischen den einzelnen ANSPs variiert und im Durchschnitt ein (korrigierter) Score von lediglich 57 % erreicht wird. Es wird angenommen, dass dies auf Lohneffekte

⁶⁰ Strength Weaknesses Opportunities Threats



zurückzuführen ist. Tatsächlich unterscheiden sich die Kosten pro Lotsenstunde innerhalb Europas signifikant. Diese Kosten sind jedoch größtenteils exogen und damit nicht von der Flugsicherung beeinflussbar. Es kann daher kritisch hinterfragt werden, ob eine Performance-Analyse auf Basis dieser Kosten zielführend ist. Die im Paper vorhandenen Grafiken lassen zudem darauf schließen, dass das Konfidenz-Intervall teilweise zu einer sehr großen Spannweite⁶¹ des Effizienz-Scores führt. Insbesondere bei der technischen Effizienz wurde u. a. für DFS, DSNA, MUAC und Avinor eine Spannweite von über 30 % berechnet. Robustheitstests oder Ursachenanalysen sind im Paper nicht enthalten.

Mehrere SFA-Modelle werden von Blondiau et al. (2016) im Projekt COMPAIR angewendet. Dafür standen die Daten von 37 ANSPs für elf Jahre (2004-2014) zur Verfügung. Die geschätzten Kostenfunktionen werden nach Enroute- und Terminal Services differenziert. Im Enroute-Bereich sind lediglich drei der zehn getesteten Koeffizienten statistisch signifikant. Dies ist möglicherweise in der inhomogenen Allokation der Kosten zu Enroute oder Terminal begründet. Dies betrifft insbesondere Approach-Einheiten. Als weitere potentielle Schwächen sehen die Autoren einen zu geringen Datenumfang und das Fehlen eines qualitativen Outputs.

Adler et al. (2017) adressieren in ihrem Paper den Zusammenhang zwischen Leistungsfähigkeit und Eigentumsverhältnis der ANSP. Dabei differenzieren sie, ob es sich u. a. um staatliche Organisationen oder teilprivatisierte Unternehmen handelt. Die Autoren verwenden die SFA zur Schätzung der Produktions- und Kostenfunktion von 37 Flugsicherungen auf Basis der Daten von neun Jahren. Im Ergebnis erreichten die ANSPs mit öffentlich-privater Eigentumsform mit Beteiligung der Interessengruppen statistisch signifikant höhere Produktivität und Kosteneffizienz als ein staatliches Unternehmen oder eine staatliche Behörde. Des Weiteren erkennen die Autoren u. a. „erhebliche Ineffizienz auf dem europäischen Flugsicherungsmarkt“⁶². Methodisch ist jedoch zu berücksichtigen, dass die SFA sehr fehleranfällig auf unpräzise a priori Annahmen reagiert. Zudem ist eine Differenzierung der Eigentumsverhältnisse schwierig, da lediglich zwei ANSPs als teilprivatisiert gelten können. Die konkreten Zuordnungen sind u. a. im ACE-Report zu finden.

Die folgenden Absätze beziehen sich auf eine Studie der „Academic Group“ zur Festlegung von Performance Zielen für die europäischen ANSPs (European Commission 2018). Da diese Analyse einen hohen Einfluss auf politische Entscheidungen hatte, wird diese Studie intensiver betrachtet. Dabei werden sowohl methodische als auch inhaltliche Schwachpunkte hinsichtlich Modellierung und Ergebnisinterpretation fokussiert.

Die ‚Single European Sky‘-Initiative wurde seit dem letzten Jahrzehnt in mehreren Schritten (Referenzperioden) umgesetzt, die sich u. a. hinsichtlich der wirtschaftlichen Regulierung unterscheiden (vgl. Abschnitt 3.2.1). Für die Referenzperiode 3 (RP3) beauftragte der Performance Review Body (PRB) eine „Academic Group“ mit der Durchführung einer Leistungsbewertung, um Kostenziele für jeden einzelnen europäischen Flugsicherungsdienst für den Zeitraum von RP3 festzulegen. Die Aufgabe

⁶¹ Mit „Spannweite“ wird im Folgenden die Differenz zwischen maximaler und minimaler Ausprägung beschrieben.

⁶² Übersetzt durch den Verfasser.



der Academic Group bestand darin, die PRB bei der Zielsetzung der Kosteneffizienz zu beraten. Der Bericht wurde am 3. Juni 2018 veröffentlicht (European Commission 2018).

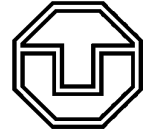
Die Academic Group nutzt für ihre Untersuchung ein DEA- und ein SFA-Modell. Die Effizienzergebnisse zeigen ein teilweise homogenes und teilweise heterogenes Bild, wie eine Auswahl an DMUs in Tabelle 3-2 zeigt. Das geschätzte jährliche Einsparpotenzial für die Bereitstellung von Enroute Services variiert zwischen einer Milliarde Euro (dies entspricht einer durchschnittlichen Kostenreduzierung von rund 17 %) und 1,7 Mrd. €, (30 % Einsparung). Für die Terminal Services wird ein Einsparpotenzial von 290 Mio. €. (18 %) bis 515 Mio. € (33 %) ausgewiesen. Diese Studie legt daher nahe, dass die Flugsicherungsdienste durchschnittlich 25 bis 30 % der Gesamtkosten einsparen können. Es gibt jedoch erhebliche Unterschiede hinsichtlich der möglichen Kosteneinsparungen in den einzelnen Flugsicherungsunternehmen.

Tabelle 3-2: Ergebnisse RP3 Target Setting für ausgewählte ANSPs

ANSP	Land	Enroute Services		Terminal Services	
		DEA	SFA	DEA	SFA
BULATSA	Bulgarien	0,36	0,63	0,26	0,68
DCAC Cyprus	Zypern	0,87	0,50	0,62	0,73
DSNA	Frankreich	0,86	0,84	0,63	0,72
DFS	Deutschland	0,95	0,92	0,97	0,87
ENAV	Italien	0,67	0,66	0,66	0,75
LVNL	Niederlande	0,57	0,85	0,78	0,82
Avinor	Norwegen	1,00	0,70	0,52	0,69
PANSA	Polen	0,51	0,71	0,49	0,67
ENAIRE	Spanien	0,53	0,58	0,38	0,72
NATS	UK	0,93	0,88	0,93	0,81

Die Anwendung von Data Envelopment Analysis und Stochastic Frontier Analysis ist für das Benchmarking im ATM üblich. Dennoch weist die Studie einige signifikante Schwächen auf. Ein erster Kritikpunkt umfasst die ökonomische Modellierung sowie die dafür genutzten Daten (Inputs und Outputs). Die Hauptressourcen eines ANSP werden durch Personal und Kapital repräsentiert (Standfuss et al. 2018b), siehe auch Abschnitt 3.3.2. Aufgrund der methodischen Einschränkungen einer DEA ist die Anzahl der parallel betrachteten Inputs und Outputs begrenzt. Da der Datensatz lediglich 30 ANSPs umfasst, kann die parallele Verwendung von fünf Parametern innerhalb eines Modells zu verzerrten Ergebnissen führen (vgl. Abschnitt 3.4.2).

Die Autoren nutzen mit Komplexität und Variabilität des Verkehrs exogene Faktoren innerhalb des DEA-Modells. Der Output wird durch diese Gewichtung künstlich vergrößert oder verkleinert. Es kann hier kritisch hinterfragt werden, ob eine Einsparung basierend auf exogenen Determinanten zielführend ist, da der ANSP weder auf Komplexität, noch auf Variabilität einen Einfluss hat. Darüber hinaus weisen beide Parameter in sich mehrere Mängel auf, die sie für eine Integration im Modell als ungeeig-



net erscheinen lassen. So ist die Variabilität lediglich die Differenz zwischen maximaler und durchschnittlicher Nachfrage⁶³. Sinnvoller scheint ein Maß, welches die gesamte Verkehrsschwankung abdeckt (Standfuss et al. 2019b).

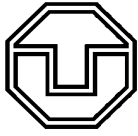
Für die Komplexität (im Sinne von PRU) werden potentielle Wechselwirkungen und nicht die tatsächlichen Wechselwirkungen berücksichtigt. Standfuss und Rosenow (2020) konnten zudem zeigen, dass die Berechnung der Komplexitätswerte signifikante Schwächen aufweist, u. a. da wesentliche Einflussfaktoren nicht berücksichtigt werden. Daher ist die Eignung dieses Maßes zur Gewichtung des Outputs fragwürdig. Die Nutzung der Anzahl an Sektoren als Output bedeutet, dass ein stark fragmentierter Luftraum zu einer Verbesserung der Leistungsfähigkeit führt. Tatsächlich erhöht es den Koordinierungsaufwand und führt zu Produktivitätsverlusten (Standfuss et al. 2018c).

Wie in Abschnitt 3.1.2 diskutiert und in Abschnitt 3.4.2 gezeigt wurde, ist die DEA Methodik sensibel gegenüber Ausreißern. Daher wäre eine intensive Datenanalyse hinsichtlich des Erfassungsprozesses und der Qualität notwendig. Eine für diese Dissertation durchgeführte Analyse im Rahmen der FABEC PMG Performance Benchmarking Taskforce hat gezeigt, dass die Nutzung von Kostendaten problematisch ist. Dies hat zwei wesentliche Gründe: Zum einen werden die Kosten teilweise willkürlich den Services (Enroute bzw. Terminal) zugeordnet und zum anderen werden Lohneffekte als Ineffizienz ausgewiesen (bspw. ist eine ATCO-Stunde in der Schweiz teurer als in Albanien). Dies wird zwar partiell berücksichtigt, indem die zugehörigen Variablen durch die Kaufkraft-Paritäten (Purchasing Power Parities, PPP) der einzelnen Länder geteilt werden, die Unterschiede werden dadurch jedoch nur z. T. korrigiert. Zudem erfolgt keine Berücksichtigung der Kapitalkosten innerhalb des Modells.

Für den deterministischen Ansatz fehlt eine quantifizierte Ursachenanalyse. Exogene Effizienztreiber oder -inhibitoren werden dadurch nicht umfassend betrachtet. In der Folge kommt es zu einem erheblichen Informationsverlust und die (In-) Effizienzwerte können zu fehlerhaften Schlussfolgerungen führen. Darüber hinaus ist keine Analyse der Skaleneffizienz und der Skalenerträge integriert. Unberücksichtigt bleibt auch die Vorhersagbarkeit des Verkehrs.

ANSPs planen auf der Grundlage der Verkehrsprognose ihre Kapazität und ihre Kosten. Die Bereitstellung von Kapazität ist in der Regel eine mittel- bis langfristige Aufgabe (Einstellung zusätzlicher Lotsen, Erweiterung des ATM-Systems, usw.). Eine Abweichung der Vorhersage vom tatsächlichen Verkehr wirkt sich auf die Effizienz aus (Standfuss und Whittome 2019b). Die Studie enthält zudem keine Informationen über die Servicequalität der ANSPs. Somit werden Effizienzziele entwickelt, ohne die Servicequalität der verschiedenen Flugsicherungsdienste zu berücksichtigen. Die Autoren führen zudem keine Robustheitstests durch. Sensitivitätsanalysen, z. B. die Substitution von Inputs und Outputs, wären notwendig, um die Eignung des Modells zu überprüfen. Eine alternative Modellierung kann zu anderen Ergebnissen führen, die abweichende Maßnahmen und politische Implikationen erfordern würde. Ergänzende Kritikpunkte wurden in Standfuss et al. (2020) publiziert.

⁶³ Nachfrage i.S.v. der tatsächlichen Anzahl an Flügen.



3.2.3 Zusammenfassende kritische Würdigung

Das prognostizierte Wachstum der Nachfrage sorgte insbesondere im vergangenen Jahrzehnt für ein gesteigertes Interesse an einer Leistungsbewertung im Luftverkehrsmanagement. Offizielle Berichte und akademische Studien nutzen verschiedene Ansätze, Methoden und Modelle um die Produktivität und Effizienz der ANSPs zu bestimmen, Wirkungsketten zu verstehen und effizienzbeeinflussende Effekte nachzuweisen. Insbesondere die EUROCONTROL-Berichte stützen ihre Analysen auf Produktivitätskennwerte und das dadurch entstehende Ranking. Ein absoluter Wert ist einfach interpretierbar und anhand der genutzten In- und Outputs können Verbesserungspotentiale abgeleitet werden. Die Nachteile einer solchen Indexziffer-Methode wurden bereits in Abschnitt 3.1.4 diskutiert. Zudem findet keine quantitative Ursachenanalyse statt.

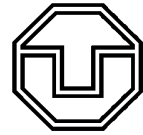
Im Sinne einer Analyse von Verbesserungspotentialen eignet sich eine Effizienzanalyse, bei der man den Wert einer DMU mit dem besten der Gruppe vergleicht und somit ein relatives Maß für die spezifische Leistungsfähigkeit erhält (Hirschhausen et al. 2007). Akademische Studien nutzen zur Effizienzbestimmung daher überwiegend die Data Envelopment Analysis und die Stochastic Frontier Analysis. Die zitierten Studien umfassten die Daten mehrerer Jahre, was die Anwendung von Panelanalysen ermöglichte. Eine umfangreiche Datenanalyse fand jedoch dort bislang nicht statt. Insbesondere bei der DEA ist die Datenerhebung und -qualität jedoch aufgrund der Sensitivität gegenüber Extremwerten und der deterministischen Randproduktionsfunktion entscheidend. Es findet zudem keine Überprüfung statt, ob die von der PRU bereitgestellten Indikatoren (z. B. Komplexität oder Variabilität) überhaupt zielführend sind. Dies ist eine mögliche Ursache dafür, dass insbesondere die SFA oft zu insignifikanten und/oder kontraintuitiven Ergebnissen führt.

Ein weiterer Kritikpunkt adressiert die Modellierung. Die meisten Studien verwenden kostenbasierte Ressourcen (z. B. Personalkosten). Dies bedeutet, dass Ineffizienzen bspw. auf Lohneffekte zurückzuführen sind. Auf diese Effekte hat ein ANSP jedoch keinen Einfluss. Sinnvoller ist daher eine Betrachtung betrieblicher Faktoren wie bspw. Personal (Anzahl der Lotsen) oder gebundenes Kapital (Anzahl der Tower und/oder ACCs).

Alle akademischen Studien führen eine Leistungsbewertung auf ANSP Ebene durch. Das heißt, dass insbesondere die betrieblichen Daten aus mehreren kleineren Einheiten (Sektoren) aggregiert werden und damit Durchschnittswerte darstellen. Wird bspw. die Produktivität der deutschen Flugsicherung höher bewertet als die der französischen DSNA, so kann dennoch eine der französischen Kontrollzentralen (z. B. Reims) produktiver arbeiten als eine der deutschen (z. B. Langen).⁶⁴ Im Sinne eines holistischen Ansatzes ist es demnach notwendig, die Performance und deren beeinflussende Effekte auch auf disaggregierter Ebene zu untersuchen.

Das wesentliche Ziel eines Benchmarkings ist es, Stellschrauben zur Verbesserung der Leistungsfähigkeit zu identifizieren und zu quantifizieren. In der SFA werden diese Effekte implementiert, bei der DEA wird in einer zweiten Stufe eine Regression durchgeführt. Einige Autoren nutzen hierfür

⁶⁴ Der Vergleich repräsentiert nur ein fiktives Beispiel und keine tatsächlich beobachteten Werte.



eine OLS Regression, obwohl dies durch andere Studien als unzureichend bewertet wird (Simar und Wilson 2007). In anderen Analysen wird stattdessen eine Tobit Regression verwendet (Ahmada et al. 2017; Spaho 2015), welche jedoch ebenfalls als Fehlspezifikation kritisiert wird (Hoff 2007). In der Leistungsbewertung von Flugsicherungen hat eine Analyse zur Anwendung der verschiedenen Regressionsformen bis dato nicht stattgefunden. Gleiches gilt für die Anwendung auf Paneldaten (Fixed Effect Model⁶⁵ versus Random Effects Model⁶⁶ versus Pooled OLS Regression). Eine vergleichende Analyse fehlt zudem hinsichtlich der Anwendung verschiedener DEA-Typen. In den Papieren von Button und Neiva wird eine Bootstrapped DEA verwendet, andere Autoren nutzen das Grundmodell. Auch das Supereffizienz-Modell findet Anwendung. Eine Slack Based DEA wurde im ANSP Bereich bis dato nicht angewendet. Die Papiere wenden jeweils lediglich eine Methode an, um die Effizienz zu bestimmen, vergleichen die Ergebnisse jedoch nicht mit jenen anderer DEA-Modelle.

Schließlich werden weder in akademischen Studien noch in den offiziellen Berichten Robustheitstests durchgeführt. Es wird lediglich ein ökonomisches Modell (Ressourcen und Outputs) verwendet. Die Anwendung von zwei Methoden (DEA und SFA) kann ggf. als Robustheitstest interpretiert werden, jedoch sind die Ergebnisse aufgrund der verschiedenen Grundmodelle meist nicht vergleichbar. Es ist daher empfehlenswert, mehrere ökonomische Modelle mittels DEA zu testen, indem Inputs und/oder Outputs ceteris paribus substituiert oder (dis)aggregiert und die Ergebnisse der Effizienzanalyse verglichen werden.

3.3 Datengrundlage

3.3.1 Vorgehen

Die Literaturrecherche hat gezeigt, dass ein wesentlicher Schwachpunkt bisheriger Analysen in der Nutzung eines ungeeigneten ökonomischen Modells (Anzahl und Auswahl der In- und Outputs) begründet liegt. Unabhängig von der Methode ist die Auswahl der genutzten Inputs und Outputs entscheidend für die Validität der daraus resultierenden Leistungskennzahlen. In der Ursachenanalyse spielt zudem eine umfassende Betrachtung exogener Faktoren eine zentrale Rolle, diese wird in Abschnitt 4.1 adressiert. Abbildung 3-14 stellt das Vorgehen grafisch dar.



Abbildung 3-14: Prozesskette der Leistungsbewertung

Um die wesentlichen ökonomischen Determinanten der Dienstleistungen abzubilden, müssen zunächst die wesentlichen Prozesse bestimmt werden (Abschnitt 3.3.2.1). Dies beinhaltet v. a. die Diskussion relevanter Inputs und Outputs (Abschnitte 3.3.2.2 und 3.3.2.3), die je nach Analyseziel abhängig von der betrachteten operativen Ebene sein können. Dieses Kapitel fokussiert die ANSPs, in späteren Abschnitten werden jedoch auch ACCs und Sektoren einbezogen.

⁶⁵ FEM.

⁶⁶ REM.



In einem zweiten Schritt wird die Datenverfügbarkeit und -qualität geprüft (Abschnitt 3.3.3). Dies wurde durch eine umfangreiche Zusammenarbeit innerhalb der 2015 gegründeten Performance Benchmarking Taskforce (PBTF) der Performance Management Group (PMG) ermöglicht. Nachdem zusammen mit den operativen Experten relevante Faktoren bestimmt wurden, können diese in das ökonomische Modell überführt werden (Abschnitt 3.4.1). Zur Überprüfung der Robustheit werden dabei verschiedene Kombinationen aus Inputs und Outputs verwendet. Finaler Schritt ist die Anwendung der ausgewählten Methodik auf die ökonomischen Modelle (Abschnitt 3.4). Wie vorab diskutiert, stehen dafür verschiedene Verfahren zur Verfügung. In der vorgelegten Dissertation basiert das Benchmarking maßgeblich auf der Data Envelopment Analysis.

3.3.2 Wertschöpfungsprozess

3.3.2.1 Grundmodell

Zur Leistungsbewertung von Flugsicherungsdienstleistern werden die maßgeblichen Prozesse und Einflussgrößen auf die Effizienz modelliert. Eine Herausforderung dabei stellen die Unterschiede zwischen den betrachteten Unternehmen dar. Die Flugsicherungen unterscheiden sich hinsichtlich der angebotenen Dienstleistungen, der Rechtsform (z. B. Aktiengesellschaft) oder den Besitzverhältnissen (staatlich oder teilprivatisiert). Des Weiteren sind einige Flugsicherungen für große Lufträume über Gewässern zuständig (Oceanic Airspace), andere kontrollieren lediglich den oberen Luftraum (Maastricht).

Zunächst wird der ANSP aus Performance-Sicht modelliert. Dieses Grundmodell ist in Abbildung 3-15 dargestellt und illustriert die wesentlichen Prozesse und Interdependenzen zwischen den einzelnen Faktoren. Der eigentliche Produktivitäts- bzw. Effizienzwert wird gemäß Definition durch das Verhältnis aus Output und den dafür eingesetzten Produktionsfaktoren gebildet. Der dadurch entstehende Score gibt entweder absolut oder relativ die Leistungsfähigkeit der Flugsicherung wieder.

Die Outputseite umfasst mit den ANS-Leistungen das Kerngeschäftsfeld der Flugsicherung. Hier können verschiedene betriebliche (z. B. Flugstunden) und finanzielle Parameter (z. B. Umsatz) verwendet werden. Auf Inputseite sind insbesondere Personal, technisches Equipment, eingesetzte Software sowie das gebundene Kapital relevant. Dies kann differenziert werden nach Center (Enroute) und Tower (Terminal) Operations.

Um die Leistungserstellung zu optimieren, werden organisatorische Einheiten gebildet, die wiederum spezifische Zielfunktionen aufweisen und durch unterschiedliche Umwelteinflüsse charakterisiert werden (vgl. Abbildung 2-7). Um Größenvorteile zu realisieren, wurde das bereits besprochene FAB Konzept und die damit verbundene engere Kooperation initiiert. Ein weiterer Unterschied zwischen den Flugsicherungen sind die Prozeduren, bspw. innerhalb der Anflugkontrolle. Diese Unterschiede können die Performancewerte beeinflussen, ohne dass sie unmittelbar in die Leistungskennzahlen eingerechnet werden.

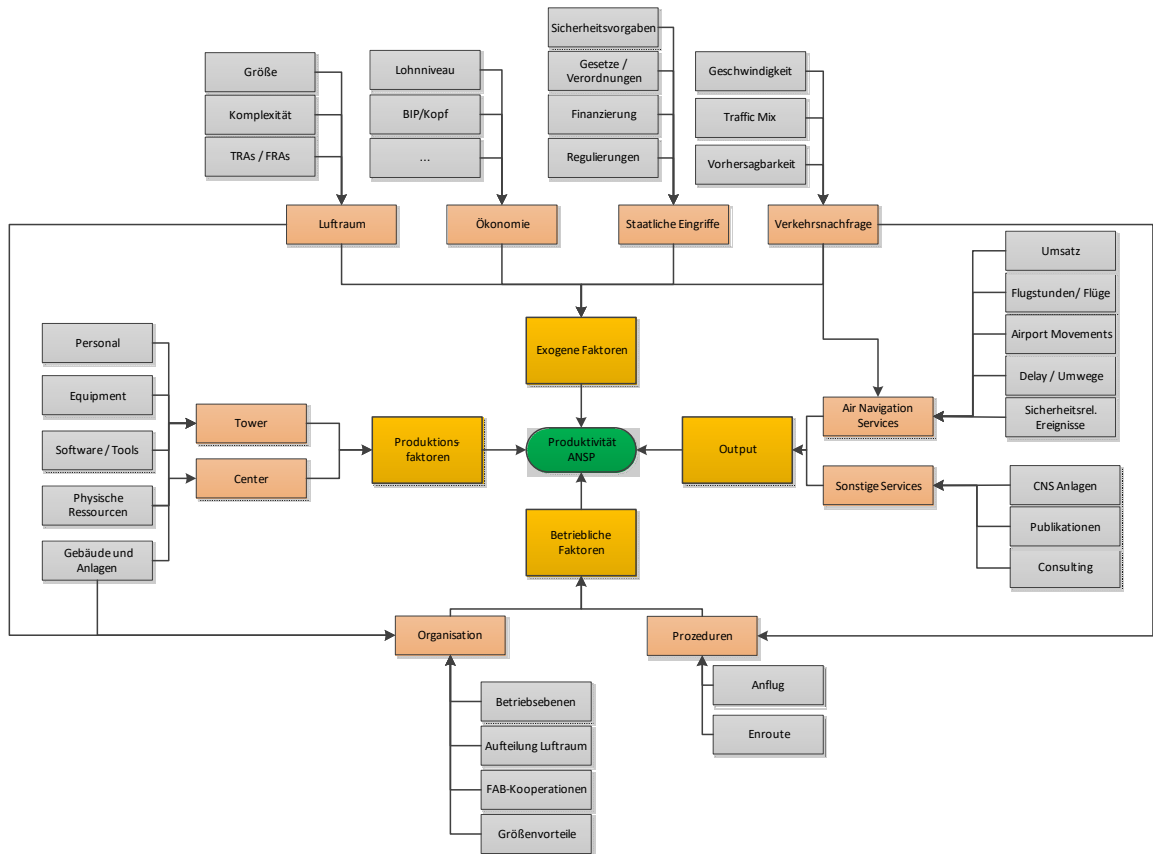
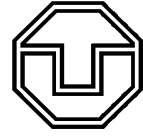
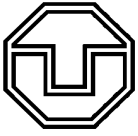


Abbildung 3-15: Vereinfachtes Performance-Modell einer ANSP

Eine strikte Trennung zwischen bzw. eine klare Aufteilung der einzelnen Faktoren ist nicht möglich. So weist bspw. das gebundene Kapital Interdependenzen zu gesetzlichen Grundlagen eines Staates bzw. zu Europarecht auf. Des Weiteren spielen Abhängigkeiten zwischen den Verkehrscharakteristiken und der Aufteilung des Luftraumes eine bedeutende Rolle. Verfahren (z. B. Holdings) sind primär abhängig von der Verkehrsnachfrage. So beeinflusst die Komplexität der Verkehrsströme⁶⁷ die Kapazität eines Luftraumes und damit die Anzahl der durchführbaren Flüge hinsichtlich Einflüge pro Stunde und dem „Occupancy“-Wert⁶⁸. Auf diese Nachfrage hat ein ANSP keinen Einfluss, daher wird sie als exogen betrachtet. Der ANSP hat lediglich die Möglichkeit, Kapazitäten anzupassen und dadurch Flüge bzw. Flugstunden zu beeinflussen. Bei einer Reduktion der Kapazität kann dies jedoch zu Verspätungen führen. Weitere exogene Effekte sind die juristischen bzw. regulativen Vorgaben der Europäischen Union oder der jeweiligen Staaten. Auch Luftraumcharakteristiken und sozioökonomische Determinanten können einen Einfluss auf die Produktivität haben.

⁶⁷ Z. B. Traffic Mix, Verkehrsdichte, usw.

⁶⁸ Occupancy bedeutet, wie viele Flugzeuge sich gleichzeitig in einem definierten Bereich (z. B. Sektor) befinden.



3.3.2.2 Inputs

Produktionsfaktoren umfassen im Wesentlichen Personal, Material, Kapital, Energie und eingekaufte Dienstleistungen⁶⁹. Der Ressourcenaufwand kann entweder in reinen Mengen oder in den dafür anfallenden Kosten angegeben werden. Aufgrund der bereits diskutierten Nachteile einer finanziellen Bewertung der Ressourcen wird in der Dissertation die Nutzung physischer Mengen bevorzugt.

In der ATM-Leistungsbewertung wird das Personal primär durch die Fluglotsen repräsentiert. Aufgrund unterschiedlicher Arbeitszeitmodelle wird die Anzahl der Lotsen in Full Time Equivalents (FTEs) angegeben. Das bedeutet, dass die Anzahl der Lotsen über ein Korrekturverfahren zu Vollzeitstellen zusammengefasst wird. Neben dieser reinen Anzahl an (Vollzeit-)Stellen sind auch Arbeitsstunden erfasst (ATCO-hours). Die Fluglotsen können neben der eigentlichen Lotsentätigkeit u. a. auch an Projekten mitarbeiten oder als Trainer fungieren. Daher werden beide Inputs in weitere Unterkategorien differenziert. „ATCOs in OPS“⁷⁰ umfassen alle FTE, die neben der Lotsentätigkeit auch andere Aufgaben wie Schulungen (Refresher Training) oder die Betreuung auszubildender Lotsen umfassen. Nicht zugehörig sind dagegen ATCOs in Projektarbeit, Lehrertätigkeiten an einer Akademie oder am Simulator. Diese Tätigkeiten werden zu „ATCOs not in OPS“ zusammengefasst.

Basierend auf den ATCOs in OPS wird auch die zugehörige Arbeitszeit in „on board“, „on duty“, und „not on duty“ differenziert. Die „on board“-Zeit gibt an, wie viele Stunden ein Lotse tatsächlich am Radarschirm verbracht hat (eigentliche Lotsentätigkeit). Diese Zeit ist eine Teilmenge der „on duty“-Stunden, welche Überstunden und Pausen berücksichtigt. Die „ATCO-hours not on duty“ erfassen zudem die Fehlzeiten eines Lotsen, bspw. aufgrund von Krankheit oder Weiterbildungen. Alle Tätigkeiten außerhalb des eigentlichen Lotsenbetriebes (z. B. Projektarbeit) werden in den „ATCO-hours on other duties“ erfasst (EUROCONTROL 2008b, 2012; ICAO 2018).

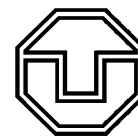
Neben den Fluglotsen beschäftigt eine Flugsicherung auch Personal für die Verwaltung, zur Wartung technischer Anlagen und in diversen anderen Bereichen. Auch hier gilt die Differenzierung zwischen der Anzahl an FTE und den aufgewendeten Arbeitsstunden.

Das gebundene Kapital umfasst Immobilien, technische Geräte und Anlagen sowie andere Investitionsgüter. Sie hängen maßgeblich von der Organisationsstruktur des ANSP ab. Betreibt eine Flugsicherung mehrere ACCs, so hat dies Auswirkungen auf die Instandhaltungskosten der Immobilien, deren Energiebedarf und Personalaufwand. Auch der Betrieb von Tovern hat einen äquivalenten Einfluss auf das gebundene Kapital.

Die Kosten der Produktionsfaktoren unterscheiden sich innerhalb Europas erheblich. Im Jahr 2017 beliefen sich die jährlichen Beschäftigungskosten pro Fluglotsen zwischen 17.894 € in der Ukraine und 277.629 € bei der niederländischen LVNL (Abbildung 3-16). Das heißt, dass die jährlichen Kosten

⁶⁹ Je nach Wirtschaftszweig können auch weitere bzw. andere Faktoren berücksichtigt werden, z. B. Boden.

⁷⁰ Operations.



pro Fluglotse in den Niederlanden etwa 16-mal höher sind als die der ukrainischen UksATSE. Die Kosten pro ATCO-Stunde⁷¹ lagen zwischen 12€ (UksATSE) und 232€ (DFS / Deutschland), siehe Anhang B2. Die unterschiedlichen Kosten liegen insbesondere in den Kaufkraftunterschieden zwischen den europäischen Ländern begründet.

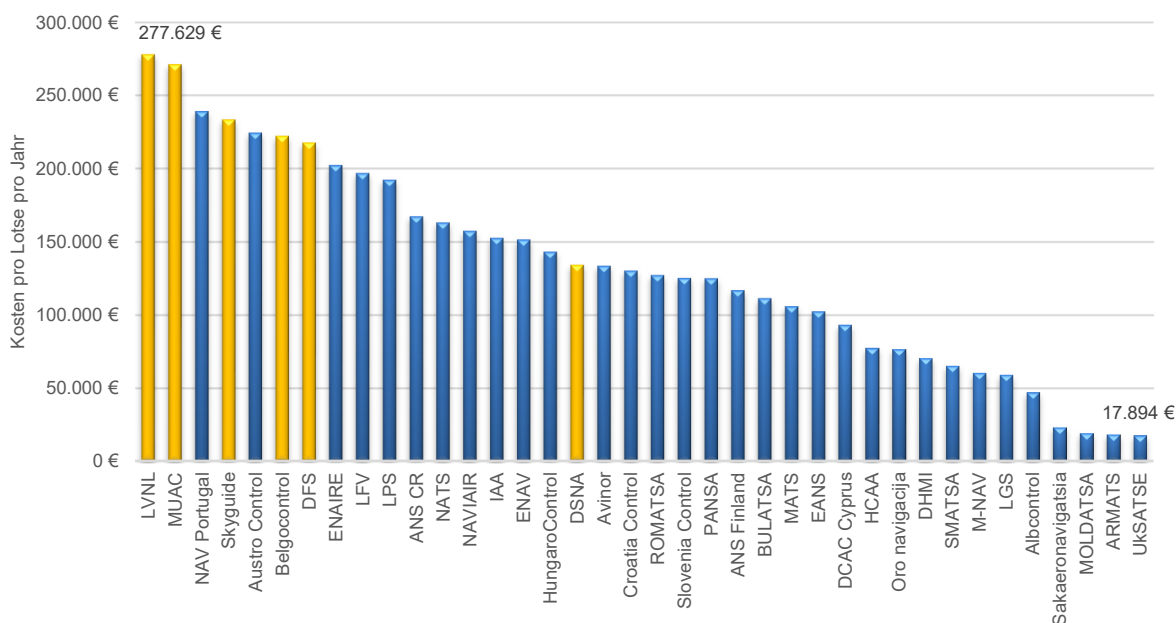


Abbildung 3-16: Kosten pro Fluglotse pro Jahr (2017)

Gemäß Abbildung 3-16 sind insbesondere die Staaten des europäischen Kerngebiets durch hohe jährliche Lotsenkosten gekennzeichnet (orange Säulen repräsentieren die FABEC-Staaten). Die zugehörigen Lufträume gehören durch die geographische Lage zu den am stärksten frequentierten. Dadurch zeichnen sie sich durch hohe Komplexitätsbewertungen aus. Folglich ist der Bedarf an Investitionen sowohl in ATM- und CNS-Equipment, aber auch in Personal, höher als in Staaten mit geringerem Verkehr. Eine Erhöhung des Inputs „Arbeit“ hat jedoch unterschiedliche Auswirkungen auf die Gesamtkosten. Laut „The ATM Policy Institute“ kostet die Ausbildung eines ATCOs mindestens 100.000€, die International Federation of Air Traffic Controllers’ Associations (IFATCA) rechnet in Europa mit Kosten von bis zu 600.000€. Die Zahlen sind ein Hinweis darauf, dass auch die Ausbildungskosten zwischen den jeweiligen ANSPs erheblich variieren können⁷². In den o. g. Kosten ist weiterhin nicht berücksichtigt, dass während der Ausbildungszeit eine gewisse Ausfall- bzw. Abbruchquote besteht (laut CANSO im Median 13 %, Maximum 50 %). Daher sind die Kosten für ein erfolgreiches ATCO-Training höher.

Ein wesentlicher Unterschied ist zudem die Jahresarbeitszeit eines ATCOs. Laut PRU-Daten von 2017 variieren die durchschnittlichen ATCO-Stunden pro Jahr zwischen 938 Stunden (DFS) und 1.949 Stunden (DCAC Cyprus), siehe auch Abbildung 4-22. Unter der Annahme einer durchschnittlichen jähr-

⁷¹ On duty.

⁷² Auch auf Nachfrage standen keine Information bezüglich der ANSP-individuellen Ausbildungskosten pro Lotsen zur Verfügung.



lichen Arbeitszeit von 1.812h pro ATCO (Wert für die USA) belaufen sich die europaweiten Zusatzkosten aufgrund der Arbeitszeitunterschiede auf 730 Mio. € (Standfuss und Whittome 2018). Es ist jedoch nicht bekannt, inwiefern Teilzeiteffekte einen Einfluss auf die unterschiedlichen Jahresarbeitszeiten haben.

3.3.2.3 Outputs

Die angebotenen Dienstleistungen der ANSPs werden primär von den Fluggesellschaften nachgefragt. Der Output kann dabei in monetären oder betrieblichen Größen bewertet werden. Betriebliche Outputs sind wiederum entweder mengen- oder zeitbasiert. Potentielle Größen sind:

- Anzahl der betreuten IFR Flüge oder Flugstunden
- geflogene IFR-Kilometer
- IFR Flugbewegungen an den Flughäfen
- Anzahl der Composite Flight hours
- Umsatz

Die Dienstleistungen der Flugsicherung umfassen sowohl jene für den Streckenflug (Enroute), als auch im Bereich der Terminals (Starts und Landungen). Die korrespondierenden Outputs werden beispielsweise durch die kontrollierten IFR Flugstunden sowie die IFR Flugbewegungen an den Flughäfen repräsentiert. Um einen Leistungsindikator für eine Gate-to-Gate Betrachtung zu berechnen, wurde ein kombiniertes Maß für beide Outputparameter eingeführt: Die Composite Flight Hours (EUROCONTROL 2020c). Dieser Wert wird als gewichtete Summe aus beiden Output-Parametern berechnet. Der Gewichtungsfaktor w für die Flugbewegungen am Flughafen (Airport Movements, AM) berechnet sich aus dem gesamteuropäischen Verhältnis zwischen Terminal- und Enroute-Stückkosten, welches in Formel (3-17) exemplarisch für das Jahr 2014 veranschaulicht wird. Flugstunden (Total Controlled Flight Hours, TCFH) werden mit dem Wert 1 gewichtet, Flugbewegungen mit 0,27. Die PRU nutzt diese Werte, um die Flugsicherungen nach Kosteneffizienz (€ pro CFH) oder Produktivität (CFH pro ATCO-Stunde) zu bewerten (EUROCONTROL 2019a).

$$w(2014) = \frac{\frac{\text{Terminal ATM/CNS costs}}{AM}}{\frac{\text{Enroute ATM/CNS costs}}{TCFH}} = \frac{\frac{1.735 \text{ M€}}{14.964.958}}{\frac{6.064 \text{ M€}}{14.060.091}} = \frac{115,98 \text{ €}}{431,33 \text{ €}} = 0,27 \quad (3-17)$$

Wenngleich die Nutzung einer kombinierten, einheitlichen Output-Größe einige Vorteile bietet (siehe Abschnitt 3.4.1), so repräsentiert sowohl der Gewichtungswert als auch der dadurch entstehende Output eine künstliche Größe. Dieser Wert kann lediglich als Approximation des Gesamtoutputs interpretiert werden. Aufgrund der Heterogenität der ANSPs könnte ein pan-europäischer Wert von 0,27 nicht zielführend sein. Erstens aggregiert der Parameter zwei sehr unterschiedliche Outputgrößen. Die Differenzierung nach Enroute und Terminal Services zeigt deutliche Unterschiede in den Rankings von Produktivität und Kosteneffizienz. Zweitens weisen die disaggregierten, d. h. individuellen Gewichte eine starke Streuung auf: gemäß Abbildung 3-17 zwischen 0,10 und 0,68 (Standfuss et al. 2018b). Die rote Linie repräsentiert den verwendeten Gewichtungswert von 0,27.

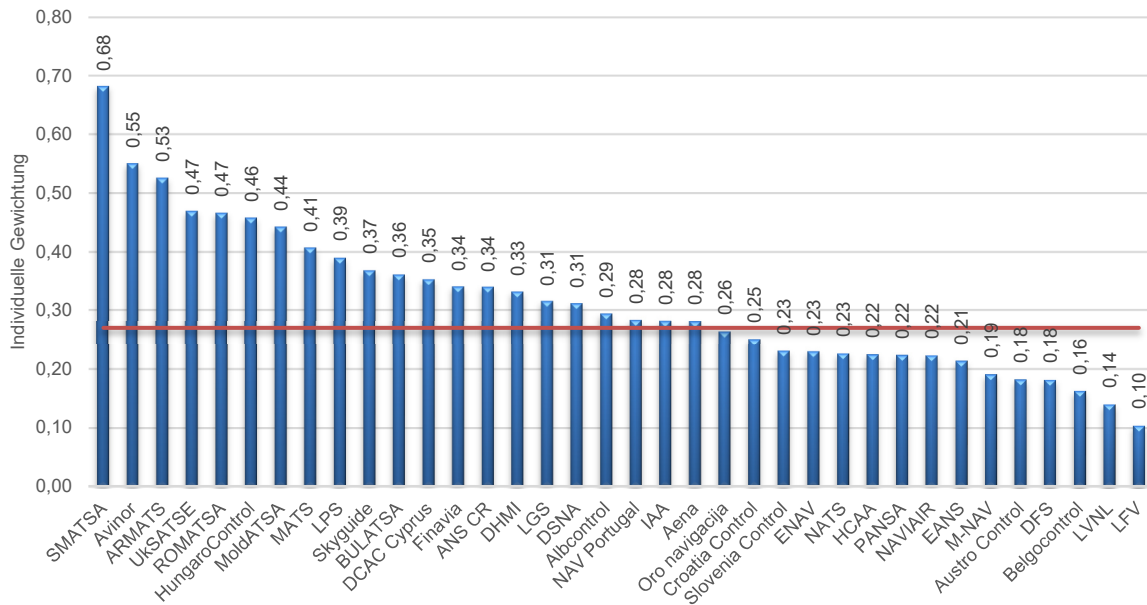
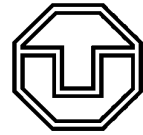


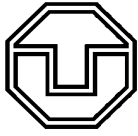
Abbildung 3-17: Individuelle Gewichte der ANSPs 2014

Mittels einer Sensitivitätsanalyse kann der Effekt der Gewichtung auf die ATCO-Produktivität und Kosteneffizienz getestet werden. Dafür wurde einmal der niedrigste und einmal der höchste beobachtete Gewichtungswert verwendet und im Anschluss das Ranking verglichen. Für die ATCO-Produktivität ist eine durchschnittliche Verschiebung um fünf Ränge zu beobachten, während sich sechs ANSPs um zehn oder mehr Ränge veränderten. Lediglich sechs Flugsicherungen wiesen eine Konstanz im Ranking auf. Eine ähnliche Auswirkung des Gewichtungsfaktors ist auf die Kosteneffizienz zu beobachten. Obwohl lediglich zwei Flugsicherungen um mehr als zehn Ränge variieren, wiesen nur vier Flugsicherungsorganisationen keine Änderungen im Ranking auf (Standfuss et al. 2018b).

Um diese Sensitivität zu validieren, wurde im Zuge der Dissertation eine Monte-Carlo-Simulation durchgeführt und in R implementiert (siehe Anhang B3). Dabei wurden die Gewichtungsfaktoren zufällig generiert und deren Auswirkung auf Output-Maß, Produktivität, Kosteneffizienz und DEA-Scores (siehe auch Abschnitt 3.4) überprüft. Diese bestätigte, dass sowohl die Performance-Kennwerte als auch das darauf basierende Ranking sensitiv auf den Gewichtungsfaktor reagieren. Die geringste Streuung (Maximaler Rang – Minimaler Rang) wiesen dabei die DEA-Scores auf.

Trotz der Sensitivität der Gewichtung kann es aus methodischen Gründen notwendig sein, aggregierte Parameter zu verwenden. Die Ergebnisse zeigen, dass eine fixe Gewichtung nach Kosten oder deren Komponenten kritisch zu betrachten ist. Die ANSP-individuellen Gewichte sollten daher in der Modellierung berücksichtigt und deren Ergebnisse mit jenen einer fixen Gewichtung verglichen werden. Eine potentielle Verbesserung wäre die Gewichtung auf Basis betrieblicher Determinanten.

Neben den Indikatoren oder Approximationen zur Verkehrsnachfrage können auch Kriterien der Servicequalität, wie bspw. Verspätungen oder von der Flugsicherung verursachte Umwege einen, in dem Falle negativen, Output darstellen. Ein weiteres Kriterium ist die Anzahl der sicherheitsrelevanten Zwischenfälle. In der luftseitigen Sicherheitsbetrachtung unterscheidet man in Luftfahrzeugnäherung



und Staffelungsunterschreitungen (STUs) bzw. Separation Minima Infringements. Die erste Kategorie ist dabei qualitativ und individuell bewertet. Bei einer STU werden vorher festgelegte Grenzwerte für die Separation von Luftfahrzeugen unterschritten. Daher stellen diese ein quantitatives Sicherheitskriterium dar. Sie betrifft nur Luftfahrzeuge im IFR-Verkehr. Sicherheitsrelevante Kennzahlen sind jedoch nicht als Output-Maß für eine Effizienzbetrachtung geeignet. Dies liegt zum einen begründet in der Datenverfügbarkeit und zum anderen daran, dass die Sicherheitskriterien gesetzlich vorgeschrieben und damit als exogen betrachtet werden können (vgl. Abbildung 3-15).

3.3.3 Datenanalyse

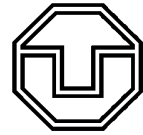
Im vorigen Abschnitt wurden die relevanten Wertschöpfungsprozesse skizziert und potentiell relevante Ressourcen und Outputs diskutiert. Bevor konkrete Modelle für die Leistungsbewertung abgeleitet werden, wird in diesem Abschnitt die Datenverfügbarkeit und -qualität untersucht. Die Analyse fand in enger Kooperation mit der FABEC PMG PBTF statt und fokussiert dementsprechend die ANSPs DFS, LVNL, MUAC, DSNA, skyguide und Belgocontrol⁷³. Die Analyse adressiert primär Daten auf ANSP Ebene und umfasste neben Inputs und Outputs auch exogene Faktoren.

Die Basis der Untersuchung bildeten die PRU Daten. Diese erfassen etwa 60 betriebliche und monetäre Indikatoren und teilen diese in Enroute, Terminal und Gate-to-Gate Services. Über das Onesky-Portal sind diese Daten abrufbar, welche auch für die Benchmarking-Reports der EUROCONTROL (EUROCONTROL 2019a, 2019h) verwendet werden. Die PRU-Datenbanken umfassen (zum Zeitpunkt der Analyse) bis zu 37 ANSPs für die Jahre 2003-2016. Tabelle 3-3 zeigt eine deskriptive Statistik des Jahres 2016 für einige ausgewählte Daten. Sie verdeutlicht, dass der europäische Luftraum stark heterogen ist, u. a. da sich die jeweiligen Lufträume prinzipiell an den Landesgrenzen orientieren. Hinzu kommt, dass die Nachfrage unterschiedliche Charakteristiken aufweist, was an Parametern wie dem Anteil an Überflügen deutlich wird. MUAC repräsentiert einen Sonderfall mit 100 % Überflügen und 0 Bewegungen am Flughafen. Spaniens Luftraum ist etwa doppelt so groß wie der zweitgrößte (Frankreich) und etwa 107-mal größer als der kleinste Luftraum (Slowenien). Diese Unterschiede wirken sich teilweise auf die Gesamtnachfrage aus. Hier ist Frankreich der ANSP mit den meisten kontrollierten Flugstunden und Flügen: Die Nachfrage ist hinsichtlich der Flüge 77-mal höher als in Armenien, bezogen auf Flugstunden sogar 242-mal größer als jene von Moldawien.

Um die Datenqualität zu prüfen, wurden zwei Ansätze gewählt. Zum einen wurden die Daten mehrerer Jahre miteinander verglichen und statistisch sowie graphisch ausgewertet. Über Spannweite⁷⁴, Standardabweichung und Variationskoeffizient konnten so fehlerhafte Daten identifiziert und ggf. korrigiert werden. Beispielweise stieg 2009 die Anzahl der Tower in Italien von 13 auf 28 an und sank im nächsten Jahr wieder auf 12. Allerdings ist nicht jede signifikante Abweichung der Daten auf eine mangelhafte Datenqualität zurückzuführen. Wie Abbildung 3-18 zeigt, sank die Nachfrage der ukrainischen UksATSE 2014 und 2015 sprunghaft. Dies ist durch die politische Situation begründet und der daraus resultierenden Schließung einiger Lufträume. Aufgrund der geographischen Lage ist auch

⁷³ Belgocontrol wurde kürzlich in „skeyes“ umbenannt. Innerhalb der Dissertation wird jedoch der alte Name geführt.

⁷⁴ Maximalwert – Minimalwert.



Moldawien von einer äquivalenten Nachfragentwicklung betroffen. Die Ausweichverkehre nutzten insbesondere die Lufträume der Balkanstaaten. Abbildung 3-18 zeigt dies anhand der Nachfrage in Bulgarien.

Tabelle 3-3: Deskriptive Statistik PRU Daten 2016

	Min	Median	3. Quartil	Max	Std.Abw
Luftraumgröße (km ²)	20.400	151.500	560.250	2.190.000	428.848
ACCs	1	1	2	5	1
Tower	0	6	16	77	14
IFR Flüge	38.968	611.342	860.928	3.015.153	746.008
Distanz der IFR Flüge	6.036.705	186.399.176	293.686.451	1.647.941.888	358.521.614
IFR Flugstunden	9.442	237.314	426.481	2.287.512	511.051
Flugbewegungen an Flughäfen	0	161.381	479.862	2.017.084	545.210
Composite Flight Hours	15.302	286.178	545.122	2.777.883	648.698
Anteil Überflüge	10 %	59 %	82 %	100 %	26 %
Anteil Inlandsflüge ⁷⁵	0 %	2 %	12 %	50 %	10 %
Anteil Int. Flüge ⁷⁶	0 %	37 %	49 %	87 %	20 %
Personal (FTE)	146	772	1.560	7.669	1.814
Lotsenstunden	88.065	318.425	670.369	3.671.472	749.165
Lotsen (FTE)	52	223	451	2.859	623
Lotsenkosten (€'000)	1.220	31.250	64.066	397.351	102.828
Personalkosten (€'000)	2.912	68.454	136.688	842.016	199.101
Gesamtkosten (€'000)	8.564	121.691	197.303	1.529.680	336.492
Gesamtumsatz (€'000)	7.658	125.129	218.231	1.637.439	370.339

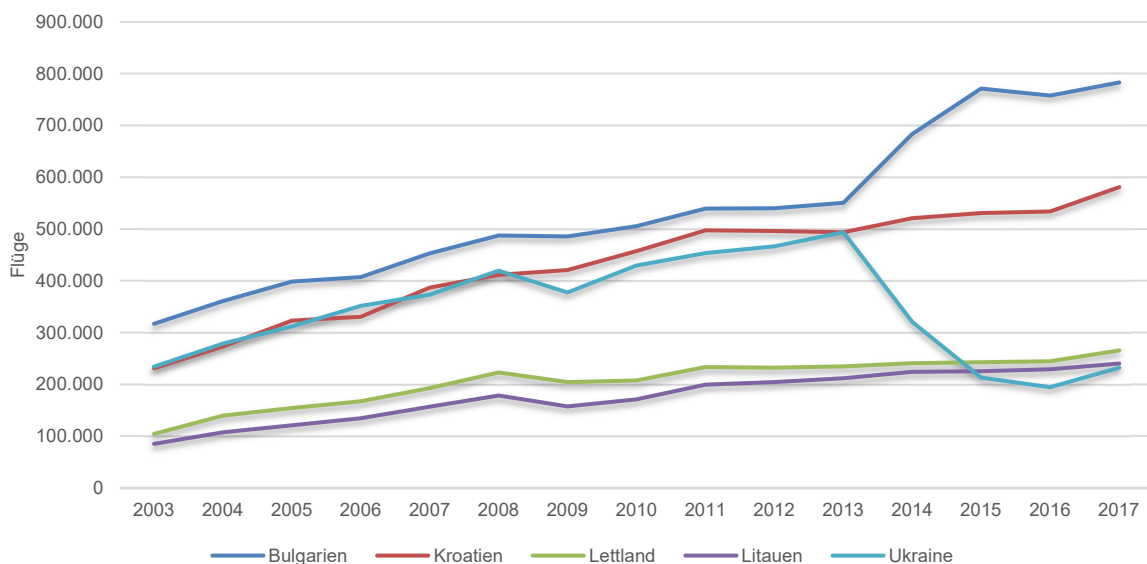


Abbildung 3-18: Entwicklung der Nachfrage 2003-2017

⁷⁵ Gemeint sind Flüge, die in einem spezifischen Land starten und landen.

⁷⁶ Gemeint sind Flüge, die in einem spezifischen Land starten oder landen.



Einige Schwankungen sind jedoch nicht plausibel erklärbar. Daher wurde in einem zweiten Schritt die Datenerfassung und -verrechnung analysiert. Zudem umfassen die von der PRU bereitgestellten Daten nicht alle von den ANSPs übersendeten Indikatoren. Zwischen Mai 2015 und März 2016 wurde daher monatlich eine WebEx-Konferenz mit den Mitgliedern der FABEC ANSPs durchgeführt. Ziel dabei war es, die Daten zu konsolidieren und validieren⁷⁷.

Abweichungen und Erfassungs-Unterschiede waren insbesondere bei den Personaldaten zu beobachten und betrifft u. a. Fehlzeiten wegen Krankheit oder Mutterschaft. Inkonsistenzen gab es auch in der Erfassung. So werden Daten teilweise automatisch erfasst, andere ANSPs nutzen Schätzungen oder fixe Werte für einen Parameter und berechnen andere darauf basierend. Das heißt, diese Parameter spiegeln unter Umständen keine tatsächlichen Werte wider. Unterschiede gibt es weiterhin bei der Erfassung von Trainingszeiten⁷⁸. Die Analyse wurde sowohl auf ANSP Ebene als auch auf ACC Ebene durchgeführt. Zudem fand ein Produktivitätsvergleich statt. Ausgehend davon wurde die Notwendigkeit einer disaggregierten Effizienzanalyse bestätigt, insbesondere um die Ursachen für Kapazitätsunterschiede zu eruieren (siehe Abschnitt 4.4). Vorgehen und Ergebnisse sind in FABEC (2015) zusammengefasst.⁷⁹

Durch die Analysen konnten u. a. zwei Effekte festgestellt werden. Zum einen unterscheidet sich die Erfassung und Berechnung der Daten zwischen den ANSPs. Durch den Abgleich der PRU-Daten mit anderen Datenquellen, u. a. NEST (EUROCONTROL 2018c) oder dem Leistungsinformationszentrum (LIZ) der Deutschen Flugsicherung konnten fehlerhafte Daten jedoch rückwirkend korrigiert und in die Analysen aufgenommen werden. Zum anderen wurde beobachtet, dass sich die Datenqualität seit 2008 verbessert hat. Ob dies einen Einfluss auf die verwendeten Metriken hat, ist Gegenstand der Untersuchung (siehe u. a. Abschnitt 4.3).

3.4 Anwendung DEA

3.4.1 Modellierung

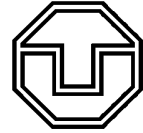
Nachdem in den vorherigen Abschnitten die wesentlichen Prozesse der Flugsicherungen sowie die dazu verfügbaren Daten und deren Qualität diskutiert wurden, werden diese nun zu einem ökonomischen Modell zusammengefasst. Die Betrachtung aller Inputs und Outputs in einem Modell ist nicht möglich, zum einen aus methodischen Gründen, zum anderen, weil die Faktoren miteinander korrelieren. Zudem werden jene Faktoren, welche in der Datenanalyse bereits als unpräzise identifiziert wurden, nicht im ökonomischen Modell verwendet. Basierend auf der Datenanalyse und den diskutierten Umwelteinflüssen ist die Auswahl der Inputs und Outputs durch fünf wesentliche Aspekte begründet:

1. Das ökonomische Modell sollte die für die Leistungserstellung maßgeblichen Inputs und Outputs beinhalten.
2. Die Inputs sollen alle wesentlichen Kostenfaktoren enthalten.

⁷⁷ Dies ist insbesondere notwendig, da einige Indikatoren aggregiert sind und diese Verrechnung u. U. nicht homogen stattfindet.

⁷⁸ Wird als „Refresher Training“ angegeben.

⁷⁹ Die Ergebnisse wurden aufgrund der Sensibilität der Daten nicht veröffentlicht, jedoch FABEC-intern verwendet.



3. Parallel verwendete Inputs oder Outputs sollten eine geringe Korrelation aufweisen.⁸⁰
4. Fehlerhafte bzw. unpräzise PRU-Daten sind innerhalb der DEA-Modelle zu vermeiden, da dies zu einer Verschiebung der Produktionsfunktion und/oder zu fehlerhaften Effizienzwerten führen kann.
5. Operative Größen werden bevorzugt. Monetäre Werte sind zu vermeiden.

Für die ökonomische Leistungsbewertung wurden mehrere Input- und Output-Kombinationen getestet.⁸¹ Diese Vorstudien umfassten u. a. auch die Trennung von Enroute und Terminal Services oder die Eliminierung bestimmter ANSPs. Die dadurch erlangten Erkenntnisse resultierten in der Formulierung von vier Modellen, um die Flugsicherungsdienstleister zu bewerten.

Das Basismodell umfasst die zwei wesentlichen Outputs eines ANSP für Terminal (Airport Movements) und Enroute Services (Total Controlled IFR-Flight Hours). Als Input wurden aufgrund der hohen Heterogenität der Bedingungen in den jeweiligen Staaten, die sich u. a. auch auf Lohnsätze der ATCOs auswirkt, monetäre Werte so weit wie möglich vermieden.⁸² Des Weiteren implizieren betriebliche Optimierungen im Rahmen einer Inputminimierung ohnehin Kosteneinsparungen. Neben dem operativen Personaleinsatz sollte auch der Produktionsfaktor Kapital berücksichtigt werden. Maßgeblich sind hier die Infrastruktureinrichtungen Tower und Area Control Center zu integrieren.

Zur Erhöhung der Genauigkeit der Ergebnisse der DEA ist im Falle von lediglich 38 betrachteten DMUs eine Reduzierung der Zahl der betrachteten Produktionsfaktoren empfehlenswert. Um die Zahl der Inputs zu reduzieren, wurde ein „Composite Infrastructure Unit“ Indikator (CIU) eingeführt. Dieser kombiniert die Anzahl der Tower und ACCs zu einer Kennzahl, die als Input im Basismodell verwendet wird. Dabei erfolgt die Gewichtung zunächst in Anlehnung an die Vorgehensweise der PRU mit dem Faktor 0,27. Das Basismodell (Modell 1) besteht damit aus drei Inputs und zwei Outputs. Als weiteren Input wurden die Betriebskosten abzüglich der Kosten für Lotsen, da diese bereits in die ATCO-Stunden berücksichtigt wurden, verwendet. Die Approach Einheiten wurden nicht integriert, da diese sich in den meisten Fällen entweder im ACC oder dem Kontrollturm befinden.

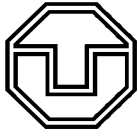
Aufgrund der spezifischen Charakteristika der Flugsicherungsdienstleistungen kann angenommen werden, dass die Ressourcen angepasst werden können, während der Output durch die Nachfrage der Luftraumnutzer vorgegeben ist. Aufgrund der begrenzten Luftraumkapazität kann diese nicht unendlich steigen. Maßnahmen zur Kapazitätserweiterung, z. B. durch computergestützte Systeme, sind wiederum den Ressourcen zuzurechnen. Daher wird die Effizienzbewertung inputorientiert durchgeführt.

Ausgehend von diesem Basismodell wurden drei Untermodelle entwickelt. Modell 2 kombiniert die beiden Outputs gemäß der PRU zu den Composite Flight Hours. In Modell 2A werden Inputs und Outputs nicht verändert, jedoch wird aufgrund der stark abweichenden Dienstleistungscharakteristik Maastricht UAC nicht in die Analyse integriert. Modell 2B berücksichtigt die identifizierten Nachteile der

⁸⁰ Eine vorab durchgeführte Korrelationsanalyse identifiziert die parallel verwendbaren Indikatoren.

⁸¹ U. a. der Einsatz von Flügen oder Flugstunden als Output-Maß.

⁸² Der Anteil der Lotsenkosten an den Gesamtkosten beträgt bspw. je nach ANSP zwischen 37,4 und 85,0 %.



PRU-Berechnung der CFH. Daher wird hier für die Gewichtung kein pan-europäisches Stückkostenverhältnis genutzt, sondern das ANSP-individuelle. Die gleiche Gewichtung wird auch für die CIU verwendet. Die Indikatoren haben den Index „var“ zur Abgrenzung von jenen mit fixer Gewichtung. Tabelle 3-4 fasst die vier Modelle zusammen.

Die Verwendung mehrerer Modelle hat den Vorteil, dass Robustheitstests nicht nur hinsichtlich verschiedener Bezugsjahre durchgeführt werden können, sondern auch modellübergreifend. Dadurch kann u. a. festgestellt werden, ob die Eliminierung von MUAC die Randproduktionsfunktion wesentlich verändert. Des Weiteren wird durch den Vergleich von Modell 2 und 2B die Anwendbarkeit des fixen Gewichtungswertes für die CFH analysiert. Zum Zeitpunkt der Effizienzanalyse standen die Daten für 38 ANSPs zur Verfügung.

Tabelle 3-4: DEA-Modelle

	1	2	2A	2B
Inputs	ATCO-hrs OPS Costs CIU			ATCO hrs OPS Costs CIU _{var}
Outputs	Flight Hours Airport Movements		CFH	CFH _{var}
ANSPs	38		37 ⁸³	38

Da nur die Flüge nach Instrumentenflugregeln eine Gebühr entrichten müssen, umfassen die potentiell für das Modell genutzten Outputs lediglich diesen Verkehr⁸⁴. Der VFR-Verkehr bindet zwar Lotsenkapazität, wird jedoch nicht in den Daten berücksichtigt. Zudem ist der VFR-Anteil sehr gering und findet nur in bestimmten Lufträumen, kontrolliert durch FIS-Lotsen, statt.

Auf eine Integration des Delay wurde in den Modellen verzichtet. Zwar wären Verspätungen als negativer oder reziproker Output im ökonomischen Modell abbildbar, innerhalb der Dissertation wurde jedoch aus mehreren Gründen davon abgesehen. Erstens ist die Erfassungsmethode kritisch zu betrachten. Verspätungen werden immer jenem Sektor zugeordnet, der am meisten zum Gesamt-Delay beiträgt⁸⁵, unabhängig davon, ob auch in anderen Einheiten daran beteiligt waren (EUROCONTROL 2020f; European Commission 2010b). Zweitens wurde bei den Feldstudien (siehe Abschnitt 4.4) deutlich, dass Verspätungen nicht aktiv in Kauf genommen werden um Ressourcen zu sparen. Alle betrachteten ACCs gaben an, dass die Regulierung⁸⁶ eines Sektors lediglich die letztmögliche Maßnahme ist. Drittens kann es zu Verspätungen kommen, da der tatsächliche Verkehr von dem geplanten (oder prognostizierten) Aufkommen abweicht.

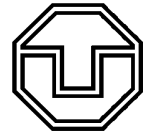
In den folgenden Abschnitten sind die Ergebnisse der Effizienzbewertung zusammengefasst. Dabei werden zunächst konstante, später variable Skalenerträge unterstellt. Die Ergebnisse beziehen sich auf das Jahr 2016. Die technische Effizienz der einzelnen ANSPs wird modellübergreifend verglichen und mögliche Abweichungen in Score und Ranking diskutiert. In einem weiteren Analyseschritt wird der

⁸³ Datensatz exclusive MUAC.

⁸⁴ Dies ist auch bedingt durch die Datenverfügbarkeit.

⁸⁵ „Most penalizing Sector“.

⁸⁶ Damit werden Maßnahmen zur Verkehrssteuerung bezeichnet, siehe auch Abschnitt 4.4.3.



DEA-Score im Zeitverlauf beobachtet. Für die Jahre 2003-2016 wurden für alle Modelle Effizienz und Rang der ANSPs berechnet. Da die Anzahl der ANSPs über die Jahre nicht konstant ist, wurden die Analysen je einmal auf Basis eines homogenen⁸⁷ und eines inhomogenen⁸⁸ Datensatzes durchgeführt. Die Effizienzwerte und Rankings wurden dann auf Fluktuationen und Trends untersucht, um so eine zeitliche Robustheit der Ergebnisse zu bestimmen.

3.4.2 Ergebnisse

3.4.2.1 Konstante Skalenerträge

Durch Annahme konstanter Skalenerträge werden keine Größeneffekte berücksichtigt. Das heißt, dass jede Abweichung von der Randproduktionsfunktion als technische Ineffizienz interpretiert wird. Tabelle 3-5 fasst die Ergebnisse modellübergreifend zusammen. Die meisten effizienten Einheiten wurden für das Modell M1 ausgewiesen. In diesem Modell werden zwei Outputs betrachtet und somit insgesamt fünf Faktoren (statt vier in den anderen Modellen) im ökonomischen Modell integriert. Dies führt durch die DEA Methodik zu einer höheren Anzahl an effizienten Einheiten. In M2 werden nur zwei ANSPs als effizient bewertet (MUAC und die estnische EANS). Auch der durchschnittliche Effizienz-Score sinkt von 97 % auf 59 %. Die Vermutung, dass MUAC die Produktionsfunktion maßgeblich verschiebt, bestätigt M2A. Sowohl der durchschnittliche Score als auch die Anzahl der technisch effizienten DMUs steigen im Vergleich zu M1. Modell 2 und 2B unterscheiden sich nur geringfügig. Dies ist nicht verwunderlich, da die beiden Modelle sich lediglich in der Gewichtung der Output- und der Infrastrukturkomponenten unterscheiden.

Tabelle 3-5: Modellvergleich konstante Skalenerträge

	M1	M2	M2A	M2B
Durchsch. Score	0,799	0,590	0,745	0,607
Effiziente Einheiten	9	2	6	2

Durch die Unterschiede in den betrachteten Inputs sowie die Eliminierung von MUAC in Modell 2A sind die Ergebnisse modellübergreifend nicht robust. Dies zeigt auch eine Analyse der individuellen Effizienzscores und den darauf basierenden Rankings. Der maximale und minimale Score eines ANSP variiert im Durchschnitt um 25 Prozentpunkte, das Ranking um 12 Plätze. Insbesondere die niederländische LVNL und die belgische Belgocontrol bilden hier Extremwerte. Sie erreichen in Modell 1 einen hohen Effizienzscore (100 % bzw. 92 %), in Modell 2 und 2B jedoch lediglich Werte zwischen 28 % und 38 %. Beide ANSPs sind für den unteren Luftraum zuständig, sodass die Reduktion des Outputs auf CFH den Einfluss der Airport Movements reduziert⁸⁹. Für 14 der 28 betrachteten ANSPs sind keine oder nur geringfügige Änderungen zu beobachten. Dies ist insbesondere bei MUAC, EANS, Albcontrol (Albanien), MATS (Malta) und M-NAV (Mazedonien) zu beobachten. Eine Verbindung zur Betriebsgröße oder der durchschnittlichen (modellübergreifenden) Effizienz kann nicht identifiziert werden.

⁸⁷ Die Analyse nutzt lediglich jene ANSPs, deren Daten in jedem betrachteten Jahr verfügbar waren.

⁸⁸ Der Datensatz umfasst jeden ANSP, für die in einem spezifischen Jahr Daten zur Verfügung standen.

⁸⁹ Beide ANSPs haben etwa dreimal mehr Airport Movements als kontrollierte IFR-Flugstunden.

Parallel zu den Scores weisen auch die Ränge der ANSPs Unterschiede auf. Hier ist wiederum die LVNL besonders betroffen: Zwischen Modell 1 und M2 bzw M2B beträgt die Änderung 36 Ränge. Auch ANS Finnland, DFS, Belgocontrol und Avinor weisen eine Abweichung von mehr als 20 Plätzen auf. Dagegen sind ANSPs wie EANS, ENAIRE, MUAC und Oro navigacija (Litauen) modellübergreifend robust. Bei einer Analyse der Ränge muss jedoch berücksichtigt werden, dass durch die DEA mehrere ANSPs Platz 1 erreichen können. Der erste nicht-effiziente ANSP erhält somit nicht Rang 2, sondern einen Rang entsprechend der Anzahl an effizienten Einheiten. Eine direkte Vergleichbarkeit ist insbesondere bei einer stark unterschiedlichen Anzahl effizienter Einheiten nicht gegeben.

Dieser Effekt wird dadurch bestätigt, dass die wesentlichen Unterschiede in den Scores und Rankings durch Modell 1 verursacht werden. Eliminiert man dieses Modell aus den Robustheitstests, so schwankt der Score lediglich um 18,7% und die Rankings um acht Plätze. Dennoch gibt es auch in diesem Fall starke Unterschiede, insbesondere bei ANSPs die durch die Eliminierung von MUAC wesentlich höhere Effizienzwerte erreichen konnten (z. B. DFS und NATS). Zwischen Modell 2 und 2B sind die Abweichungen nur minimal (durchschnittlich 4% bzw. drei Plätze). Ausnahmen bzw. Extremwerte bilden hier Avinor, MoldATSA und ARMATS. Alle drei ANSPs hatten bei der Gewichtung der CFH einen signifikant höheres individuelles Stückkostenverhältnis als 0,27 (siehe Abbildung 3-17).

Im Folgenden werden die individuellen Effizienzwerte näher beleuchtet. Eine ausführliche Darstellung aller Modelle befindet sich im Anhang, Tabelle A 2. Abbildung 3-19 vergleicht die DEA-Scores der einzelnen ANSPs für das Jahr 2016 nach Modell 2A (links) und 2 (rechts). Aus illustratorischen Gründen sind die ANSPs entsprechend den Grenzen ihrer zugehörigen Länder dargestellt. Zudem ist Maastricht nicht abgebildet.

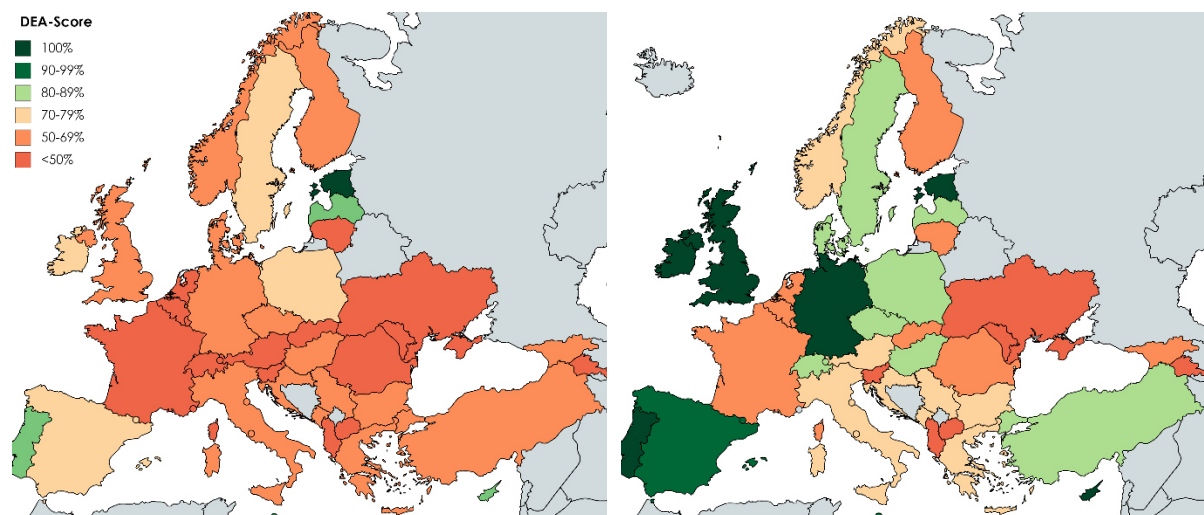
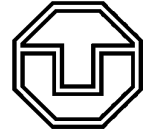


Abbildung 3-19: DEA-Scores 2016, Modell 2 und 2A, konstante Skalenerträge

Abbildung 3-19 verdeutlicht die Unterschiede in den Effizienz-Scores zwischen beiden Modellen. In Modell 2 sind die meisten europäischen ANSPs durch einen Effizienzwert von unter 70% gekennzeichnet. Ausnahmen bilden lediglich Portugal, Malta, Zypern, Estland und Lettland sowie Maastricht (nicht dargestellt). Diese Ergebnisse widersprechen, mit Ausnahme Maastricht, der Erwartung, dass große ANSPs effizienter sind als kleine (vgl. Abschnitt 2.5). Sie bestätigen jedoch die Annahme, dass



die besondere Charakteristik der Dienstleistung von MUAC einen wesentlichen Einfluss auf die Randproduktionsfunktion hat.

Dies wird durch die Ergebnisse in 2A bestätigt. Insbesondere Länder mit einer hohen Verkehrsnachfrage weisen, mit Ausnahme von Frankreich, hohe Effizienzwerte ($>80\%$) auf. Die Randproduktionsfunktion ist somit durch einen wesentlich geringeren Anstieg gekennzeichnet. Dadurch reduzieren sich die radialen Abstände, insbesondere bei DMUs mit hoher Nachfrage. Zudem gelten mehrere ANSPs als effizient, welche in Modell 2A als signifikant ineffizient ausgewiesen werden, u. a. DFS und NATS. Belgien und die Niederlande erreichen in beiden Modellen einen Score von unter 70% , was ebenfalls in der Luftraumcharakteristik und dem verwendeten Outputmaß begründet liegen kann (siehe Fußnote 89).

Abbildung 3-19 zeigt weiterhin, dass vor allem ANSPs mit geringer Nachfrage einen niedrigen Effizienzwert erreichen. Im Falle von Mazedonien, Moldawien, Slowakei und Albanien gilt dies für beide Modelle. Die niedrigen Effizienzwerte von MoldATSA und UkSATSE können zudem in der politischen Lage und den daraus resultierenden Luftraumsperrungen in der Ostukraine sowie dem daraus folgenden Ausweichverkehr begründet sein. Überraschenderweise können Bulgarien und Rumänien, die primär von diesem Verkehr profitieren, keine hohen Scores nachweisen. Nichtsdestotrotz ist es möglich, dass die ANSPs ihre Effizienz im Vergleich zu den Vorjahren verbessert haben. Diese Effekte werden in Abschnitt 4.3 mittels einer Panel-Analyse untersucht.

Die Untersuchung zeigt, dass die Effizienzanalyse mit Annahme konstanter Skalenerträge zu teils kontraintuitiven Ergebnissen führt. Insbesondere ist dies von der Wahl der Inputs und Outputs sowie der berücksichtigten DMUs abhängig. Werden Einheiten mit stark abweichenden Charakteristika einbezogen, so kann das zu einer signifikanten Drehung der Randproduktionsfunktion führen und die Effizienzwerte der anderen DMUs dementsprechend negativ beeinflussen. Zudem kann die Aggregation des Outputmaßes zu nachteiligen Scores von DMUs führen.

Der wesentliche Nachteil ist jedoch, dass bei konstanten Skalenerträgen die Betriebsgröße nicht berücksichtigt wird und somit jede Abweichung von der RPF als technische Ineffizienz bewertet wird. Dadurch wird impliziert, dass jede Einheit bei gleicher relativer Inputsteigerung auch die gleiche relative Outputsteigerung realisieren kann. Dies ist jedoch aufgrund von Sättigungseffekten des Luftraumes sowie der erhöhten Komplexität durch höhere Nachfrage unrealistisch. Es soll daher untersucht werden, ob die Analyse auf Basis variabler Skalenerträge robuster ist und zu höheren Effizienzwerten führt.

3.4.2.2 Variable Skalenerträge

Durch die Implikation variabler Skalenerträge wird die Gesamteffizienz aus einer technischen und einer skalaren Komponente gebildet (vgl. Abschnitt 3.1.2.2). Die Skaleneffizienz wird in Abschnitt 4.5 untersucht. Tabelle 3-6 fasst die Ergebnisse modellübergreifend zusammen. Sowohl die durchschnittlichen DEA-Scores als auch die Anzahl der effizienten Einheiten ist erwartungsgemäß höher als bei jenen unter konstanten Skalenerträgen. Auch die generellen Effekte auf Anzahl und Scores aufgrund der Modellspezifikationen sind äquivalent.

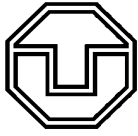


Tabelle 3-6: Modellvergleich variable Skalenerträge

	M1	M2	M2A	M2B
Durchschn. Score	0,906	0,780	0,881	0,800
Effiziente Einheiten	15	9	18	9

Ein wesentlicher Unterschied ist, dass Modell 2A die höchste Anzahl an effizienten Einheiten aufweist: Mit 18 ANSPs werden fast die Hälfte der DMUs als effizient bewertet. Trotz dieser Unterschiede streuen die durchschnittlichen Scores nur gering: Maximal- und Minimalwert weichen um lediglich 13 Prozentpunkte ab. Analog der Betrachtungen zu konstanten Skalenerträgen ist Modell M2 durch den niedrigsten Durchschnitts-Score gekennzeichnet, M1 durch den höchsten. M2 und M2A unterscheiden sich um lediglich zwei Prozentpunkte.

Die individuellen DEA-Scores sind in Anhang Tabelle A 3 dargestellt. Sie unterscheiden sich untereinander erwartungsgemäß in einem geringeren Umfang als bei Annahme konstanter Skalenerträge. Die mittlere Abweichung beträgt modellübergreifend 17 Prozentpunkte. Berücksichtigt man nur die Modelle mit einem Output (M2x), sinkt diese Abweichung auf 13, bei homogenem Datensatz um vier Prozentpunkte. Die Rankings variieren in ähnlichem Ausmaß wie bei konstanten Skalenerträgen, jedoch ist hier insbesondere die bereits diskutierte Problematik der Nummerierung gegeben. ANSPs, die in Modell 2A als effizient gelten (Rang 1), weisen in Modell 2 oder 2A geringe Ineffizienzen auf (z. B. Albcontrol 98,5 %). Durch die hohe Anzahl an effizienten Einheiten in M2A sinkt der Rang von 1 auf 11. Durch die Implikation variabler Skalenerträge gibt es jedoch auch eine höhere Anzahl an DMUs, die in jedem Modell als effizient ausgewiesen werden, z. B. DHMI (Türkei), NATS (UK) und ENAIRE (Spanien).

Abbildung 3-20 stellt die Ergebnisse von Modell 2 und 2A gegenüber. Alle ANSPs, die in Modell 2 als effizient bewertet wurden, erreichen dies auch in Modell 2A. Die Eliminierung von MUAC sorgt zudem dafür, dass einige DMUs mit einem hohen Grad an Effizienz (M2) nun ebenfalls einen Score von 100 % (M2A) erreichen (z. B. NAV Portugal und die irische IAA). Da diese nun den Benchmark bilden und damit Peers für die sonstigen ANSPs, ändert sich auch die Form der Produktionsfunktion. Dies wirkt sich positiv auf die Effizienz-Scores aus: Keine DMU hat in Modell 2A einen niedrigeren Effizienzwert als in Modell 2. Einige ANSPs können jedoch den radialen Abstand zur RPF nur in geringem Umfang reduzieren, wie bspw. ENAV (Italien), Belgocontrol (Belgien) oder UksATSE (Ukraine).

Abbildung 3-20 zeigt, dass insbesondere ANSPs mit hoher Nachfrage im Zentrum und Westen Europas als technisch effizient oder schwach ineffizient bewertet werden. Dies entspricht den Erwartungen, da auf den Score keine Größeneffekte einwirken. Belgien und die Niederlande weisen aufgrund ihrer Luftraumcharakteristik Ineffizienzen auf, die bereits im vorigen Abschnitt diskutiert wurden. Ähnliche Effekte sind für die Schweiz sichtbar. Den höchsten Effizienzwert erreicht die ANSP, wie LVNL und Belgocontrol, in Modell 1. Dem gegenüber stehen ANSPs, welche trotz ihrer geringen Größe als effizient bewertet werden, wie u. a. die moldawische und armenische Flugsicherung. Durch die Implikation variabler SKE wird die RPF für kleinere Betriebsgrößen feingliedriger, sodass mehrere ANSPs mit relativ kleiner Betriebsgröße den Benchmark bilden können. Dies wird durch den Vergleich mit Abbildung 3-19 bestätigt. Hier kann also eine mögliche Ineffizienz durch nicht realisierte Größenvorteile verursacht werden (siehe Abschnitt 4.5).

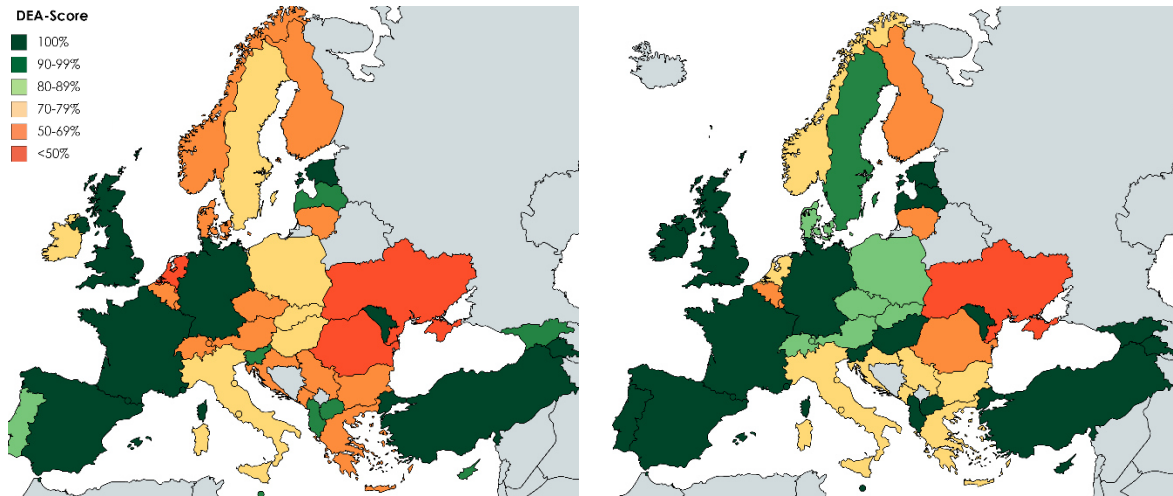


Abbildung 3-20: DEA-Scores 2016, Modell 2 und 2A, variable Skalenerträge

Des Weiteren ist eine unterschiedliche Bewertung der nordskandinavischen ANSPs Avinor, LfV und ANS Finnland sichtbar, insbesondere bei Modell 2A. Die Luftraumcharakteristiken sind relativ ähnlich: Alle drei ANSPs befinden sich in der Peripherie Europas. Der Anteil an transkontinentalen Überflügen ist gering. Im Hoheitsgebiet befindet sich ein Großflughafen im Süden des Landes sowie einige kleinere Flugplätze im Norden. Dadurch sind Ineffizienzen plausibel erklärbar. Im Vergleich zu Schweden und Norwegen ist in Finnland das Verkehrsaufkommen deutlich niedriger (ca. 30 %, gemessen in CFH). Zudem wird der norwegische Luftraum in drei Kontrollzentralen überwacht, Schweden nutzt lediglich zwei ACCs.

Die Ergebnisse zeigen, dass je nach Modell 9 bis 18 ANSPs ihren Output mit dem geringstmöglichen Einsatz von Ressourcen realisieren können. Die Implikation variabler SKE führt zu einer höheren Robustheit, weist allerdings verhältnismäßig viele effiziente Einheiten aus. Dieser Effekt kann durch die geringe Menge an DMUs in Verbindung mit der (relativ) hohen Zahl an Inputs und Outputs begründet werden. Es muss jedoch berücksichtigt werden, dass die technische Effizienz lediglich eine Komponente darstellt und nicht die gesamte, ökonomische und betriebliche Leistungsfähigkeit widerspiegelt.

Trotz der größeren Robustheit unterscheiden sich die Modelle hinsichtlich der Plausibilität der Ergebnisse. Die Modellauswahl ist zudem abhängig von den unterstellten Skalenerträgen. Daraus folgt, dass das Benchmarking der ANSPs nicht nur vom ökonomischen Modell (Auswahl und Anzahl der Inputs und Outputs) und den berücksichtigten DMUs abhängt, sondern auch davon, ob konstante oder variable Skalenerträge impliziert werden.

3.4.2.3 Zeitliche Entwicklung, kompletter Datensatz

Nach dem modellübergreifenden Vergleich für das Jahr 2016 soll in einem zweiten Schritt die zeitliche Entwicklung untersucht werden. Zum Zeitpunkt der Analyse standen die PRU Daten für die Jahre 2003 bis 2016 zur Verfügung. Dementsprechend ist es möglich, die Entwicklung der DEA-Scores über 14 Jahre zu beobachten. Dabei ist zu berücksichtigen, dass diese Entwicklungen lediglich relative Veränderungen des radialen Abstandes in Bezug auf die Randproduktionsfunktion widerspiegeln und keine Änderungen in der Leistungsfähigkeit selbst.



Je nach Betrachtungsjahr variiert die Anzahl der DMUs. So stieg die Anzahl der im PRU Datensatz integrierten ANSPs innerhalb des betrachteten Zeitraumes von 34 auf 38 an. Die Datensätze für Serbien, Georgien, Polen und Armenien sind somit nicht für alle Jahre verfügbar. Daher wird zunächst die technische Effizienz für jedes Jahr, jedes Modell und beide Annahmen bezüglich Skalenerträgen auf Basis aller in dem Jahr verfügbaren ANSPs berechnet. Getrennt nach Modell und implizierten SKE wird dann der Score über den Zeitraum beobachtet und auf Robustheit geprüft.

Die Robustheit wird zunächst durch Streumaße (Max-Min Spanne, Standardabweichung) überprüft. Dabei wird differenziert in die Streuung, welche über den Gesamtzeitraum beobachtet werden kann sowie jener, die zwischen 2008 und 2016 auftritt. Eine hohe Streuung kann auf eine Fluktuation innerhalb der Scores hindeuten, jedoch auch auf einen Trend⁹⁰. Daher wird auch das Bestimmtheitsmaß (R^2) angewendet. Dabei muss berücksichtigt werden, dass R^2 ein Gütemaß für lineare Zusammenhänge ist, jedoch die Scores auch einem Trend mit abweichend funktionaler Form folgen können. Daher wird bei hohem Streumaß und niedrigem R^2 die spezifische DMU gesondert betrachtet.

Abbildung 3-21 stellt die durchschnittlichen DEA-Scores bei konstanten (links) und variablen SKE (rechts) zwischen 2003 und 2016 dar. Es ist zu beobachten, dass die Scores nur geringen, zeitlichen Änderungen unterliegen. Dies gilt insbesondere für Modell 1 und 2A. Bei Annahme variabler Skalenerträge sind die Scores höher, die Streuung ist geringer. Gemäß den Beobachtungen im vorigen Abschnitt kann davon ausgegangen werden, dass die Ergebnisse auf Basis variabler SKE robuster sind.

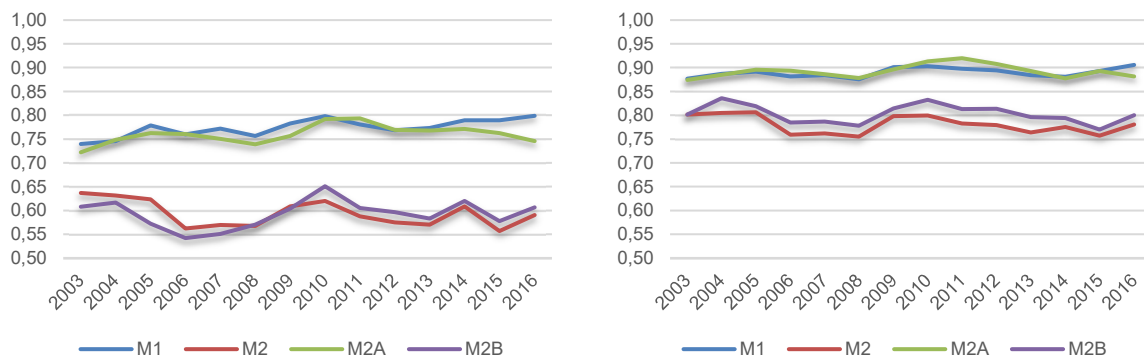
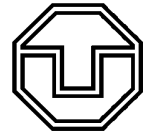


Abbildung 3-21: Durchschnittlicher DEA-Score 2003-2016, kompletter Datensatz
Konstante SKE (links), variable SKE (rechts)

Die ANSP-individuelle Varianz ist bei variablen SKE modellübergreifend gering. Die Spannweite über den Gesamtzeitraum beträgt im Maximum 20 Prozentpunkte (Modell 2B). In Ausnahmefällen übersteigt der individuelle Wert 40 Prozentpunkte. Ein interessanter Aspekt dabei ist, dass die Extremwerte in Streuung bzw. Spannweite durch unterschiedliche ANSPs verursacht werden:

- In M1: HCAA (Griechenland) und DHMI (Türkei)
- In M2: Avinor (Norwegen) und HCAA
- In M2A: BULATSA (Bulgarien)
- In M2B: Croatia Control (Kroatien), ANS Finland (Finnland) und HCAA

⁹⁰ Z. B. dass der radiale Abstand zur RPF sukzessive steigt oder sinkt.



Es ist jedoch zu berücksichtigen, dass die Scores dieser ANSPs zwar Extremwerte darstellen, jedoch die Ausprägung von Modell zu Modell unterschiedlich ist. Bspw. beträgt die maximale Spannweite in Modell 2A 32 Prozentpunkte, in Modell 2 52 Punkte. Wie bereits aus Abbildung 3-21 hervorgeht, nehmen die Schwankungen nach 2008 ab, insbesondere in Modell 2 und 2B. Dies ist auch für die Streuung der individuellen ANSPs zu beobachten.

In den meisten Fällen spiegelt der zeitliche Verlauf eine Tendenz wider. Abbildung 3-22 zeigt anhand von Modell 2A, dass es dennoch Ausnahmen gibt, deren Spannweite durch jährliche Schwankungen verursacht werden. Obwohl alle vier betrachteten ANSPs ähnliche Standardabweichungen bzw. Spannweiten (20-22 Prozentpunkte) aufweisen, unterscheidet sich der Verlauf signifikant. Dies kann auch anhand des Bestimmtheitsmaßes abgelesen werden. Das höchste R^2 wird für ANS Finland ausgewiesen (0,71), das kleinste für die schwedische LFV (0,03).

Änderungen in den Scores können jedoch darin begründet liegen, dass die Anzahl der ANSPs nicht konstant ist. Dadurch wird die Technologiemenge u. U. von anderen DMUs aufgespannt. So wird bspw. Avinor in Modell 2B so lange als effizient bewertet bis die polnische PANSa in den Datensatz integriert wurde. Es sollte demnach kritisch hinterfragt werden, ob die Ergebnisse im Zeitverlauf auf Basis des heterogenen Datensatzes vergleichbar sind und, daraus folgend, diese Analyse einen adäquaten Robustheitstest darstellt. Daher wird im nächsten Schritt die zeitliche Entwicklung auf Basis eines homogenen Datensatzes durchgeführt. ANSPs, deren Daten nicht über den gesamten Zeitraum verfügbar sind, werden dementsprechend aus dem Datensatz entfernt.

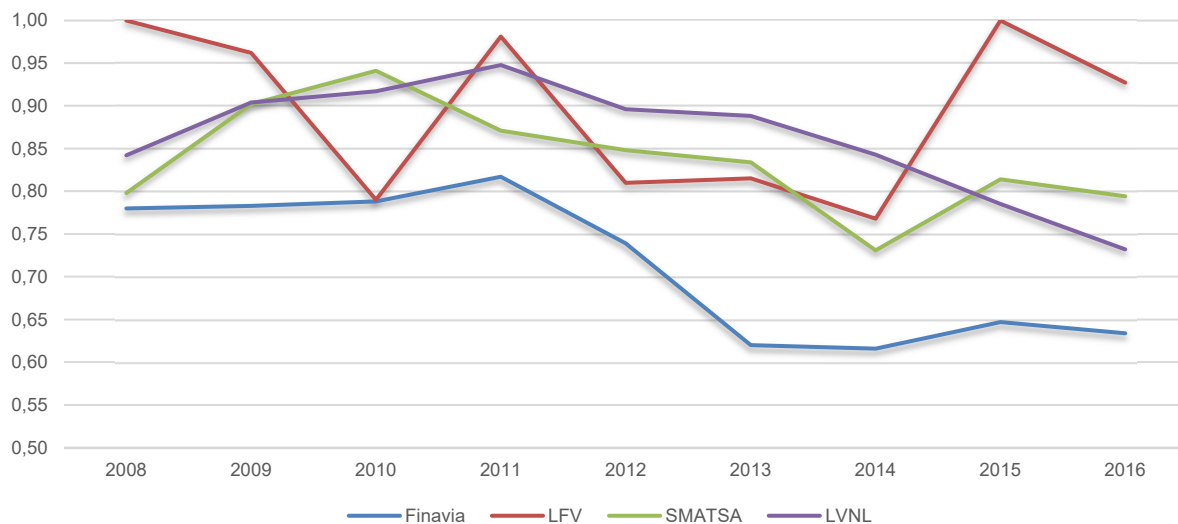
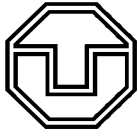


Abbildung 3-22: Durchschnittlicher DEA-Score, Modell 2A, kompletter Datensatz

3.4.2.4 Zeitliche Entwicklung, homogener Datensatz

Die in Abschnitt 3.4.2.3 durchgeführte Analyse zeigt, dass ein wesentlicher Teil der Streuung auf die Effizienzwerte der Jahre 2003-2005 zurückzuführen ist. Für den Robustheitstest auf Basis des homogenen Datensatzes wurde daher der Zeitraum 2006-2016 gewählt. Er umfasst 36 ANSPs. Die Flugsicherungsunternehmen von Georgien und Armenien bleiben jahresübergreifend unberücksichtigt. Es ist zu erwarten, dass die Streuung geringer ist als im Falle des kompletten Datensatzes.



Die zeitliche Entwicklung der durchschnittlichen DEA-Scores ist in Abbildung 3-23 dargestellt. Analog zu den Erkenntnissen aus dem vorigen Abschnitt sind die Ergebnisse unter variablen Skalenerträgen (rechts) robuster als jene bei Annahme konstanter SKE (links). Modell 1 und Modell 2A weisen wiederum eine höhere Robustheit auf als die Modelle 2 und 2B. Die Streuung ist jedoch nicht modellübergreifend geringer als bei den Effizienzbewertungen auf Basis des kompletten Datensatzes. Eine Verringerung ist für die Modelle 1 und 2 (konstante SKE) sowie 1 und 2B (variable SKE) zu beobachten. Diese Unterschiede sind jedoch nur marginal ausgeprägt. Im Folgenden werden lediglich die Modelle unter variablen SKE näher diskutiert.

Die ANSP-individuellen Streuungen liegen in den meisten Fällen unter jenen des kompletten Datensatzes. Die mittlere Abweichung beträgt 9 bis 15 Prozentpunkte. Einzelne Abweichungen über 30 Prozentpunkte bilden die Ausnahme: HCAA (Griechenland) und NAV Portugal in Modell 2, Avinor (Norwegen), Croatia Control (Kroatien), ANS Finland (Finnland), HCAA, LPS (Slowakei), NAV Portugal und Oro navigacija (Litauen) in Modell 2B. Robustheit hinsichtlich des zeitlichen Verlaufs der DEA-Scores können daher insbesondere für die Modelle 1 und 2A nachgewiesen werden.

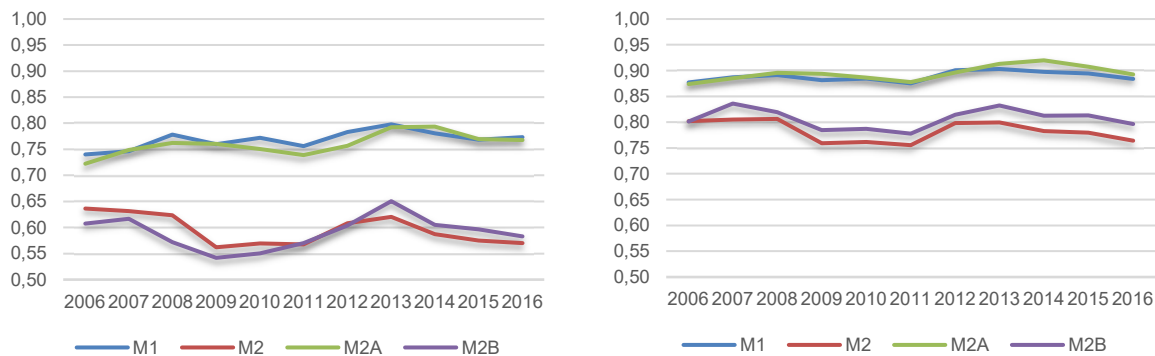


Abbildung 3-23: Durchschnittlicher DEA-Score 2003-2016, homogener Datensatz
Konstante SKE (links), variable SKE (rechts)

Die DMUs, welche in Modell 2B eine hohe Streuung aufweisen, sind in Abbildung 3-24 im Vergleich dargestellt. Auch hier kann man nach trend- und fluktuationsbasierter Streuung unterscheiden. ANS Finland, LPS und NAV Portugal sind durch hohe Bestimmtheitsmaße gekennzeichnet ($R^2 > 0,7$). Dies lässt auf einen Trend schließen, was die Grafik bestätigt. Zwischen 2006 und 2010 konnte der finnische ANSP hohe Effizienzwerte nachweisen, seit 2011 nahmen diese jedoch kontinuierlich ab. Im Gegensatz dazu steigen DEA-Scores für den slowakischen und portugiesischen ANSP fast stetig an. In diesem Fall ist die Streuung demnach kein Indiz für eine fehlende Robustheit des DEA-Scores.

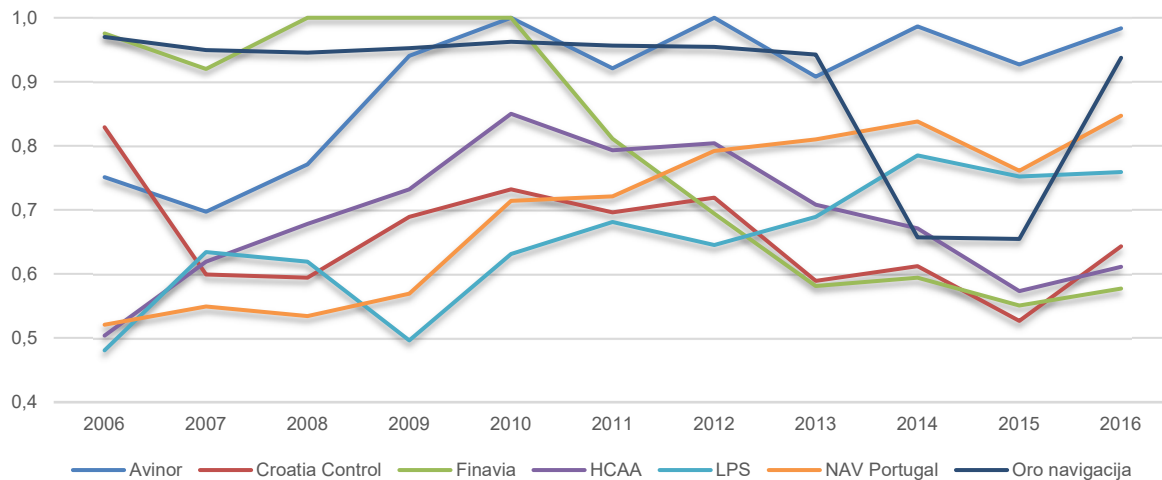
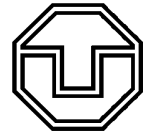


Abbildung 3-24: Durchschnittlicher DEA-Score, Modell 2B, homogener Datensatz

Die griechische HCAA kann zunächst relative Effizienzgewinne verzeichnen, der Score sinkt jedoch ab 2012 wieder. Dadurch ist das R^2 klein, der Wertebereich wird jedoch nicht durch kontinuierlich fluktuierende Scores verursacht. Eine Fluktuation ist insbesondere bei den ANSPs Oro Navigacija und Croatia Control zu sehen. Beide zeichnen sich durch eine geringe Ausprägung des Bestimmtheitsmaßes aus ($R^2 < 0,31$). Bei dem litauischen ANSP wird dies durch die niedrigen Scores in den Jahren 2014 und 2015 verursacht. Derartige Abweichungen können u. a. durch einen Datenfehler in den Inputs und Outputs oder durch eine Zuordnung zu einer anderen Peer Group verursacht werden. Im Falle von Oro Navigacija liegt die Abweichung jedoch in der Gewichtung für die Infrastrukturkomponente CIU begründet, welche in diesen beiden Jahren aufgrund des individuellen Stückkostenverhältnisses höher ist als in den Vorjahren. Dies wird dadurch bestätigt, dass in Modell 2 die DEA-Scores des ANSP robust sind.⁹¹ Die DEA-Scores für dieses Modell sind im Zeitraum von 2007 bis 2016 in Tabelle A 5 (Anhang) dargestellt.

Im Falle variabler Skalenerträge können für einige ANSPs in allen Modellen und allen Jahren Effizienz nachgewiesen werden. Gemäß Tabelle 3-7 sind dies NATS, EANS, MoldATSA und DSNA. Gleiches gilt für MUAC, wenngleich diese DMU nicht in Modell 2A berücksichtigt wurde. Im Vergleich zu den Ergebnissen bei konstanten Skalenerträgen ist jedoch zu beobachten, dass jeweils nur 1-4 DMUs im gesamten Zeitraum als effizient bewertet werden. Modellübergreifend trifft dies lediglich für MUAC (mit Ausnahme bei M2A) zu.

Die französische DSNA und die moldawische MoldATSA sind bei Annahme konstanter Skalenerträge jahres- und modellübergreifend ineffizient, obwohl sie bei variablen Skalenerträgen stets als effizient bewertet wurden. Diese Unterschiede in den DEA-Scores liegen darin begründet, dass Ineffizienzen nicht technischer Natur sind, sondern ausschließlich durch Größenvor- oder -nachteile verursacht werden. Diese (Dis-) Economies of scale werden in Abschnitt 4.5 ausführlich analysiert.

⁹¹ Die Schwankungsbreite beträgt lediglich 3 Prozentpunkte.



Tabelle 3-7: Zeitübergreifend effiziente DMUs

ANSP	Land	Variable SKE				Konstante SKE			
		M1	M2	M2A	M2B	M1	M2	M2A	M2B
Albcontrol	Albanien			X					
DCAC Cyprus	Zypern			X					
DFS	Deutschland	X				X			
DSNA	Frankreich	X	X	X	X				
EANS	Estland	X	X	X	X	X	X	X	
ENAIRES	Spanien				X				
Hungaro Control	Ungarn			X					
LGS	Lettland			X					
LVNL	Niederlande	X				X			
MATS	Malta			X					
M-NAV	Mazedonien			X					
MoldATSA	Moldawien	X	X	X	X				
MUAC	Maastricht	X	X		X	X	X		X
NATS	UK	X	X	X	X			X	
Slovenia Control	Slowenien			X					

Es ist zu konstatieren, dass die DEA-Scores jahresübergreifend robust sind. Fluktuationen der Scores bilden eine Ausnahme. Es empfiehlt sich, diese ANSPs tiefergehend zu analysieren und die Ursache der Fluktuationen zu finden. Dies gilt insbesondere dann, wenn der zeitübergreifende Extremwert 100 % annimmt und dadurch die RPF verschiebt. Dies war jedoch in keinem der betrachteten Modelle der Fall. Des Weiteren ist die Entscheidung wichtig, ob die Analyse auf Basis des kompletten Datensatzes durchgeführt oder eine homogene Datenbank genutzt wird. Gemäß den Ergebnissen ist die mittlere und individuelle Streuung bei einem homogenen Datensatz geringer. Allerdings werden möglicherweise wichtige DMUs nicht in die Kalkulation integriert. Auch hier empfiehlt sich eine Vorabanalyse, welcher Trade-off zwischen Jahren und betrachteten DMUs als sinnvoll erscheint (siehe auch Abschnitt 4.3).

3.4.2.5 Räumliche Autokorrelation

Zu den Aufgaben der Fluglotsen gehört die operative Koordination mit benachbarten Sektoren und/oder ANSPs. Diese setzt nicht nur eine ausführliche Planung voraus, sondern stellt auch hohe Anforderungen an die operativen Prozesse. So kann zwar durch die Abkürzung von Flugwegen mittels sogenannter „directs“ der Durchsatz und damit die Produktivität erhöht werden, jedoch sinkt dadurch die Vorhersagbarkeit und damit die Planbarkeit aus der Sicht der anschließend durchflogenen Kontrollzonen („downstream sectors“). Vice versa können Verspätungen unter Umständen nicht bis zum Zielort abgebaut werden. Daraus folgt, dass die Effizienz oder Ineffizienz einer Einheit auch Auswirkungen auf benachbarte bzw. nachgelagerte Sektoren haben kann. Diese räumlichen Interdependenzen können mittels Autokorrelationsmodellen bestimmt werden.

Im Gegensatz zu Zeitreihenkorrelationen, die bspw. über Fixed- oder Random Effect Modelle gerechnet werden können, wirken räumliche Interdependenzen bi- oder multilateral. Dabei wird davon ausgegangen, dass die Ähnlichkeit zwischen zwei geographischen Einheiten mit zunehmender Distanz

sinkt. Diese Verhältnismäßigkeiten müssen modelliert und in eine Gewichtungsmatrix überführt werden. Mittels räumlicher Autokorrelation kann dann bestimmt werden, ob eine Variable räumlich positiv oder negativ abhängig ist oder nicht (Schulze 1995).

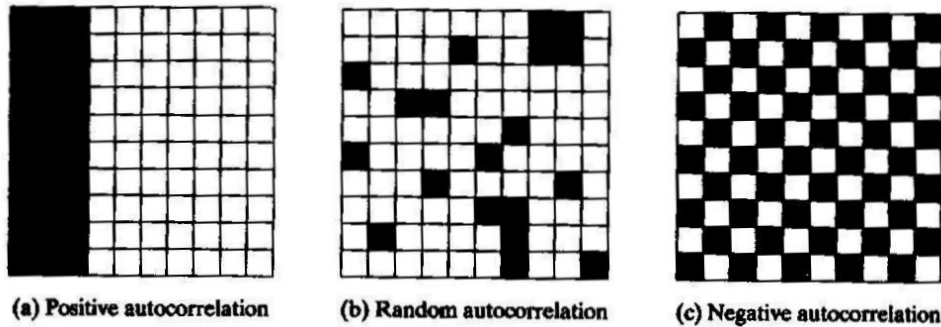


Abbildung 3-25: Beispiele für räumliche Autokorrelation (Modell)
(Lo und Yeung 2006)

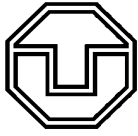
Zur Bestimmung der Räumlichen Autokorrelation stehen im Wesentlichen zwei Indizes zur Verfügung: Der Gearys-Index und der Moran-Index. Vor- und Nachteile beider Methoden werden ausführlich u. a. in Yamagata und Seya (2019) beschrieben. Da sich beide Methoden stark ähneln und der Moran Index im ATM Kontext plausibler interpretiert werden kann, wird diese Kennzahl angewandt. Für zwei Flugsicherungen r und s erfolgt die Berechnung der Autokorrelation C gemäß Formel (3-18). Dabei repräsentiert R die Anzahl aller betrachteten Flugsicherungen, respektive räumlichen Regionen. Außerdem wird die Ordnung, d. h. die räumliche Komponente, durch k definiert. Der Faktor w repräsentiert binäre Gewichtungskoeffizienten, auf die im Folgenden eingegangen wird.

$$C^k = \frac{R \sum_{r=1}^R \sum_{s=1}^R (y_r - \bar{y})(y_s - \bar{y}) \cdot w_{rs}^k}{R_K \sum_{r=1}^R (y_r - \bar{y})^2} \quad (3-18)$$

In den Ansätzen von Geary und Moran werden binäre Gewichtungsfaktoren angenommen. Verläuft eine Grenze zwischen beiden betrachteten Einheiten, so wird der entsprechende Wert 1, ansonsten 0. Dieses Maß ist jedoch sehr ungenau. Weitere Ansätze integrieren eine Komponente für die Distanz d zwischen 2 spezifischen Punkten der Region sowie den Anteil der Grenze zwischen beiden Regionen b im Verhältnis zur Gesamtgrenzlänge B der betrachteten Region. Als ein weiterer Gewichtungsfaktor könnte dabei, insbesondere bei geökonomischen Betrachtungen, die Fläche F der Region dienen, wie in Formel (3-19) dargestellt ist.

$$w_{rs} = d_{rs} \cdot F \cdot \frac{b_{rs}}{B} \quad (3-19)$$

Dieses Grundmodell muss jedoch auf die Anwendung im Flugsicherungsbereich angepasst werden. Insofern ist es fraglich, inwieweit Fläche, Grenzlänge und Distanz bewertet oder substituiert werden müssen. Insbesondere da Flüge über mehrere Landesgrenzen hinweg durchgeführt werden, und somit



auch mehrere ANSPs verantwortlich sind, können Ineffizienzen weiträumig weitergegeben werden. Als Beispiel sei ein Flug genannt, der bereits in der ersten Region deutlich Verspätung hat. Diese kann, auch auf langen Strecken, unter Umständen nicht vollständig abgebaut werden, was sich wiederum auf Kapazitätsplanungen und, daraus resultierend, Performancewerte auswirkt.

Zur Überprüfung der technischen Effizienz auf räumliche Autokorrelation wurde der Moran Index verwendet. Die Gewichtung erfolgt zunächst über eine Nachbarschaftsmatrix. Diese wurde für alle vier Modelle implementiert. Eine weitere Analyse fokussiert die Gewichtung mittels Distanzmatrix. Diese misst die Entfernung zwischen zwei Punkten der räumlichen Einheit. Je höher die Distanz, desto geringer ist die Abhängigkeit der Effizienzwerte. Prinzipiell könnte diskutiert werden, ob ein Abbruchkriterium notwendig ist. Jedoch sind auch bei Langstreckenflügen die Abweichungen vom Flugplan in Downstream Sektoren sichtbar. Daher wurde auf dieses Kriterium verzichtet.

Die Nullhypothese des Moran Index besagt, dass keine räumliche Autokorrelation vorliegt. In diesem Fall sind die Beobachtungen statistisch unabhängig. Um dies zu testen, wird zunächst ein Erwartungswert (expected, Exp)⁹² unter Annahme der Nullhypothese berechnet. Der beobachtete Wert (observed, Obs) wird mittels Formel (3-18) berechnet. Liegt der beobachtete Wert signifikant über dem Erwartungswert, so liegt eine positive räumliche Autokorrelation vor. Ist ‚Obs‘ signifikant kleiner als ‚Exp‘, so sind die Einheiten negativ korreliert. Bei einer Differenz um 0 ist die räumliche Autokorrelation zufällig, d. h., die Nullhypothese kann in diesem Fall nicht abgelehnt werden.

Für alle vier Modelle wurde die Überprüfung auf Autokorrelation sowohl mit einer Gewichtung per Nachbarschafts- als auch per Distanzmatrix durchgeführt. Die Ergebnisse sind in Tabelle 3-8 und Tabelle 3-9 dargestellt. Für die Untersuchung wurde das R-Paket „ape“ verwendet.

Tabelle 3-8: Ergebnisse der räumlichen Autokorrelation mit Nachbarschaftsmatrix

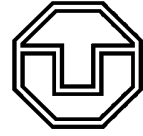
	Modell 1	Modell 2	Modell 2A	Modell 2B
Obs	0,20	-0,06	-0,03	-0,16
Exp	-0,03	-0,03	-0,03	-0,03
SD	0,11	0,11	0,10	0,11
p-Wert	0,04	0,75	0,95	0,25

Tabelle 3-9: Ergebnisse der räumlichen Autokorrelation mit Distanzmatrix

	Modell 1	Modell 2	Modell 2A	Modell 2B
Obs	0,04	0,05	0,04	0,08
Exp	-0,03	-0,03	-0,03	-0,03
SD	0,03	0,03	0,03	0,03
p-Wert	0,03	0,44	0,69	0,15

Bei der Gewichtung nach Nachbarschaftsmatrix kann die Nullhypothese für die Modelle 2 und 2A nicht abgelehnt werden. Die Werte von ‚Obs‘ und ‚Exp‘ unterscheiden sich nicht (Modell 2A) oder nicht

⁹² $Exp = -1/(N-1)$. Bei $N=38$ Beobachtungen folgt daraus: $Exp = -1/(38-1) = 0,03$.



signifikant (Modell 2). Dagegen wird für Modell 1 eine positive sowie für Modell 2B eine schwach negative Autokorrelation indiziert. Es ist jedoch zu konstatieren, dass lediglich Modell 1 statistisch signifikant ist.

Unter Verwendung der Distanzmatrix kann für alle vier Modelle eine sehr schwach ausgeprägte, positive Autokorrelation abgeleitet werden. Aufgrund des geringen Abstandes zwischen geschätzten und beobachteten Werten kann diese jedoch, insbesondere für die Modelle 1, 2 und 2A, auch als zufällig interpretiert werden. Zudem ist auch hier lediglich Modell 1 statistisch signifikant.

3.4.3 Grenzen der Anwendung

Die letzten Abschnitte haben gezeigt, dass die DEA im ANSP-Kontext anwendbar und relativ robust ist. Die Ergebnisse sind jedoch abhängig vom ökonomischen Modell, der unterstellten Skalenerträge und der integrierten DMUs. Die Effizienz-Scores entsprechen größtenteils den erwarteten Werten, unplausible Ergebnisse können methodisch erklärt werden. Sie sind zudem abhängig von endogenen und exogenen Faktoren, die in Kapitel 4 untersucht werden.

Der deterministische Ansatz hat jedoch einige, bereits beschriebene Einschränkungen. So berücksichtigt bspw. das DEA Grundmodell keine Datenfehler, sondern interpretiert jede Abweichung von der RPF als Ineffizienz. In der Literatur werden daher oft alternative DEA Ansätze gewählt, welche bereits in Abschnitt 3.1.2.3 diskutiert wurden. Im aktuellen Abschnitt wird überprüft, ob die Anwendung eines alternativen Optimierungsmodells zu weiteren Erkenntnissen oder plausibleren Ergebnissen führt und dieses Verfahren daher bevorzugt im ATM-Benchmarking eingesetzt werden sollte. Zudem wird der Free-Disposal-Hull (FDH) Ansatz auf Anwendbarkeit geprüft.

Die Implementierung erfolgt über ein R-Skript unter Verwendung der R-Pakete „Benchmarking“, „additive DEA“ und „DJL“. Im ersten Schritt werden die Daten (Inputs und Outputs) eines spezifischen Jahres eingelesen. In Schritt zwei wird der Effizienzscore für jeden ANSP auf Basis von fünf verschiedenen DEA-Varianten berechnet (vgl. auch Abschnitt 3.1.2.3):

- Grundmodell
- Supereffizienz
- Slack Based
- Slack Based Supereffizienz
- Bootstrapped DEA

Die Berechnung findet sowohl unter der Annahme konstanter als auch variabler Skalenerträge statt. Auf Basis der Scores wird im dritten Schritt das jeweilige Ranking abgeleitet. Schritt 4 umfasst die statistische Auswertung, wie die Streuung der Scores und Rankings sowie die Abweichung vom DEA-Score des Grundmodells. Die Analyse umfasst alle Modelle für das Jahr 2016. Zur Überprüfung der Ergebnisse wurden auch andere Jahre getestet, die diskutierten Erkenntnisse beziehen sich jedoch primär auf die Daten von 2016.

Die Supereffizienz-DEA identifiziert Extremwerte bei DMUs. Diese sind durch Scores > 100 % gekennzeichnet. Für alle im Grundmodell identifizierten ineffizienten Einheiten bleibt der Score dagegen



konstant. Besonders signifikant sind die Abweichungen für MUAC. Für diese DMU wird, je nach Modell und SKE, ein Effizienzwert von nahezu 400 % ausgewiesen. Andere effiziente DMUs weisen einen deutlich geringeren Abstand zur 100 %-Marke auf, wie Tabelle 3-10 zeigt. Das Ergebnis bestätigt die Annahme, dass MUAC die Randproduktionsfunktion wesentlich beeinflusst und damit in einem Benchmarking auf ANSP Ebene nicht integriert werden sollte. Möglicherweise ist dies auf einer Subebene jedoch sinnvoll, da vergleichbare Einheiten bei Kontrollzentralen⁹³ oder Sektoren existieren (vergleiche Abschnitt 4.4).

Tabelle 3-10: Supereffizienz ausgewählter ANSPs

ANSP	Konstante SKE				Variable SKE			
	M1	M2	M2A	M2B	M1	M2	M2A	M2B
DHMI	0,879	0,648	0,872	0,654	1,026	1,022	1,048	1,044
ARMATS	0,626	0,373	0,373	0,476	1,240	1,240	1,240	1,240
EANS	1,071	1,064	1,069	0,976	1,165	1,165	1,165	1,158
LVNL	1,124	0,371	0,647	0,287	1,192	0,495	0,732	0,623
MUAC	3,520	2,706	-	2,435	3,964	3,370	-	3,133
NATS	1,047	0,560	1,135	0,545	1,223	1,128	1,137	1,239

Die Anwendung des Slack Based DEA Modells führt nicht zu einer Veränderung der generierten Randproduktionsfunktion. In Folge dessen sind die effizienten Einheiten gleich jenen des Grundmodells. Ausnahmen bilden DMUs, die einen Slack aufweisen (vgl. Abschnitt 3.1.1.3). Bei diesen wird der Effizienzwert nach unten korrigiert. Der additive Ansatz führt zu einer generellen Veränderung der Bewertung ineffizienter Einheiten (vgl. Abschnitt 3.1.2.3). Dadurch weisen alle DMUs modellübergreifend einen niedrigeren Score als im Grundmodell auf.

Die mittlere Abweichung beträgt bei konstanten SKE je nach Modell 10-20 Prozentpunkte. In den Modellen 1, 2 und 2B sind die individuellen Effizienzscores signifikant niedriger als in Modell 2A, bei dem zudem die mittlere Abweichung am geringsten ist. Dies gilt auch für die Ergebnisse bei variablen Skalenerträgen, jedoch sind die Abweichungen höher. Auch die Rankings unterscheiden sich zwischen Slack-Based und Standard-DEA, im Mittel um 2 bis 6 Ränge. Die individuellen Abweichungen liegen insbesondere bei Slack-basierten Ineffizienzen über dem Durchschnitt.⁹⁴

Im Allgemeinen wäre der Slack-basierte Ansatz der DEA jenem des Grundmodells vorzuziehen, da er einen ganzheitlichen Ansatz widerspiegelt. Jedoch ist er im ANSP-Kontext nicht empfehlenswert. In allen Modellen reduzieren sich die Scores, z. T. sind diese Abweichungen signifikant. Das führt zu zahlreichen unplausiblen Ergebnissen. Um effizient zu werden, müsste bspw. die Niederlande etwa 72 % ihrer Inputs reduzieren (Modell 2). Diese Effekte sind zwar in Modell 2A, vermutlich durch die Eliminierung MUACs, weniger stark ausgeprägt, allerdings dennoch vorhanden (u. a. im Falle von Belgicocontrol, Belgien). Auch die Slacked-Based Supereffizienz-DEA brachte keine neuen Erkenntnisse, sondern bewertet lediglich die bereits identifizierten Extremwert-DMUs mit einem Score > 100 %.

⁹³ Z. B. Karlsruhe UAC.

⁹⁴ Bsp.: Albanien wird in 2A mit 100 % bewertet (Grundmodell), auf Basis der slack-basierten Methode lediglich mit 49 % (Rang 31).

Als vierte Methode wird die Bootstrap-DEA angewendet. Gemäß dem im Abschnitt 3.1.2.3 beschriebenen Verfahren wird dabei die Randproduktionsfunktion „korrigiert“. Dadurch sind alle Scores im Vergleich zum Grundmodell niedriger. Im R-Skript wurde die Methode für ein Konfidenzintervall von 95% und 1000 Replika implementiert. Die Analyse zeigt, dass insbesondere effiziente Einheiten oder DMUs mit einer überdurchschnittlichen Betriebsgröße durch die Verschiebung der RPF an gemessener Effizienz verlieren. Dies ist insbesondere bei der Annahme konstanter SKE der Fall. Geringer ist diese Ausprägung bei ANSPs, die bereits im Grundmodell einen sehr niedrigen Effizienzscore aufweisen (z. B. UkSATSE). Dies liegt darin begründet, dass der Anstieg der RPF nach Bootstrapping signifikant höher ist als im Grundmodell (siehe auch Abbildung 3-26).

Es wurde eine unterschiedliche Anzahl an Replika getestet. Abbildung 3-26 fasst die Entwicklung zusammen. Y repräsentiert den Output (CFH), die Inputs werden aus darstellerischen Gründen zu einem aggregierten Wert zusammengefasst. Es ist zu sehen, dass sich unabhängig von der Anzahl an Replika die RPF nur marginal ändert (rote, gelbe und grüne Linie). Diese Randproduktionsfunktionen weichen jedoch, insbesondere bei höheren Outputs, signifikant von der ursprünglichen RPF (schwarze Linie) ab. Diese Beobachtung war modellübergreifend konsistent und bestätigt die Ergebnisse der statistischen Analyse.

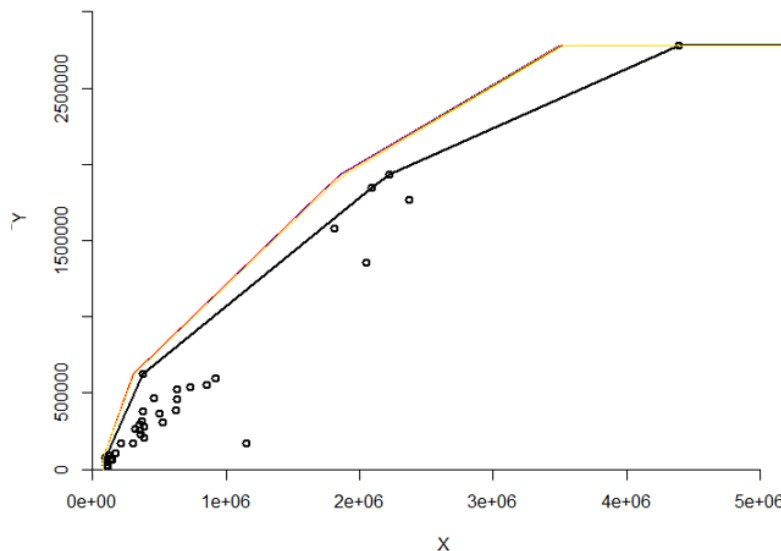


Abbildung 3-26: Bootstrapping DEA Frontier mit verschiedener Anzahl an Replika, Modell 2

Der Vorteil der Bootstrapping-Methodik ist, dass auch für die in der Standard-DEA als effizient bewerteten Einheiten potentielle Verbesserungen ausgewiesen werden. Allerdings repräsentiert der Algorithmus eine künstliche Verschiebung der RPF, was für die individuellen Einheiten zu einer Überschätzung der Ineffizienz führen kann. Beispielsweise wird für MUAC in Modell 2 ein Effizienzgrad von lediglich 79,6% berechnet. Albanien, Georgien und Mazedonien werden in selbigem Modell mit signifikant höheren Scores bewertet. Eine mögliche Ursache ist, dass der Datensatz sehr heterogene Einheiten enthält. Dadurch kommt es zu einem überdimensionalen „Penalizing“ großer DMUs. Eine weitere Begründung kann in der geringen Anzahl an Einheiten gesehen werden.

Schließlich wurde die Anwendbarkeit der FDH-Produktionsfunktion getestet. Sie hat die in Abbildung 3-27 illustrierte Stufenform (gestrichelte Linie). Es ist zudem möglich, lokal konstante Skalenerträge zu implizieren. Das bedeutet, dass in einem begrenzten Bereich um eine effiziente DMU ein linearer Anstieg konstruiert wird (punktierter Linie). Die nachfolgende Abbildung illustriert neben beiden FDH-Produktionsfunktionen auch die RPF unter variablen Skalenerträgen (durchgehende Linie) für Modell 2.

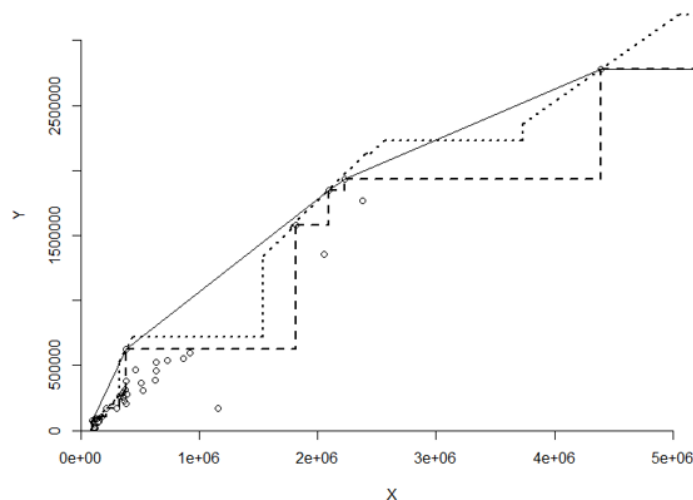


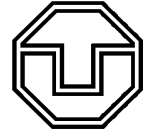
Abbildung 3-27: Free Disposal Hull (Modell 2)

Abbildung 3-27 zeigt, dass auf Basis einer FDH-Analyse viele Einheiten als effizient bewertet werden, insbesondere weil der Datensatz viele kleine DMUs beinhaltet. Der heterogene Datensatz kann zudem dazu führen, dass die Effizienz dieser DMUs überschätzt wird, da die RPF (teilweise) parallel zur x-Achse verläuft. Daher ist zu konstatieren, dass im ATM-Kontext ein Benchmarking auf Basis dieser Methodik nicht empfehlenswert ist. Da auch für die anderen Sonderformen und Weiterentwicklungen des DEA-Grundmodells signifikante Nachteile identifiziert wurden, werden zur weiteren Analyse die ursprünglichen Scores (Abschnitte 3.4.2.1 bis 3.4.2.4) verwendet.

3.5 Zwischenfazit

In diesem Abschnitt sollen die wesentlichen Ergebnisse und Erkenntnisse aus dem Performance Benchmarking zusammengefasst werden. Grundlage der Leistungsbewertung stellt zunächst ein ökonomisches Modell dar, welches alle wesentlichen Wertschöpfungsprozesse und Einflussfaktoren innerhalb eines ANSP abbildet (Abbildung 3-15). Dieses Modell muss jedoch an die jeweils betrachtete operative Ebene angepasst werden (Abbildung 2-7). Die disaggregierte Analyse wird in Abschnitt 4.4 adressiert.

Zur Überführung des ökonomischen (Grund-) Modells in die Leistungsbewertung sind zwei wesentliche Analyseschritte notwendig. Erstens sind viele Elemente durch eine hohe Korrelation gekennzeichnet (z. B. Anzahl Flugstunden und Flugstrecke). Aus methodischen Gründen können diese nicht parallel innerhalb eines Benchmarking-Modells betrachtet werden. Zweitens wurde überprüft, welche



Datenquellen verfügbar sind. Eine umfangreiche Analyse der Daten sowie deren Validierung ist obligatorisch für die Überführung in die genutzte Benchmarking-Methodik.

Durch die intensive Zusammenarbeit innerhalb der FABEC PMG PBTF konnte festgestellt werden, dass die PRU-Daten seit 2008 eine vergleichsweise höhere Qualität für FABEC-ANSPs aufweisen. Die Qualität der Daten ist elementar für den Benchmarking-Prozess und sollte daher auch bei der Modellierung berücksichtigt werden. Verzerrungen der Daten können insbesondere bei deterministischen Ansätzen zu unpräzisen Effizienzwerten führen (siehe Abschnitt 3.1.4). Bei multiperiodischen Betrachtungen sollten daher die Datensätze ab 2008 genutzt werden.

Zur Berechnung von Produktivitäts- und Effizienzkennziffern steht eine Vielzahl von Methoden zur Verfügung. Die Beschreibung des ökonomischen Modells sowie die in Abschnitt 2.2 diskutierten Charakteristiken des europäischen ATM führen zu einer Anwendung von Methoden, welche multiple Inputs und/oder Outputs berücksichtigen. Es konnte nachgewiesen werden, dass die DEA für eine Leistungsbewertung im Flugsicherungsbereich auf ANSP Ebene anwendbar ist.

Aufgrund der (teil)exogenen Natur der Nachfrage wurde eine Inputorientierung, d. h. Ressourcenminimierung impliziert. Auf Basis von Vorstudien wurden vier Modelle ausgewählt und ausgewertet. Zwischen den Modellen weichen die Effizienzwerte teilweise ab. Dies liegt in der Sensitivität der Methodik gegenüber Extremwerten und der Anzahl der parallel verwendeten Inputs und Outputs begründet. Eine zusätzliche Analyse mittels SFA ist möglich, jedoch sind aufgrund der methodischen Unterschiede die Ergebnisse nicht vergleichbar.

Im Falle variabler Skalenerträge ist die Berücksichtigung von MUAC empfehlenswert, da andernfalls möglicherweise zu viele Einheiten als „effizient“ bewertet werden. Dadurch können für eine Vielzahl an DMUs keine Verbesserungspotentiale mehr abgeleitet werden. Auch eine Ursachenanalyse würde bei einer hohen Anzahl effizienter Einheiten unpräzise Ergebnisse liefern.

Bei der Implikation konstanter Skalenerträge ist entweder die Betrachtung beider Outputgrößen notwendig⁹⁵ oder die Eliminierung von MUAC empfehlenswert. Eine weitere Option wäre die getrennte Betrachtung von Terminal- und Enroute Services. Da die Allokation der Ressourcen zu den einzelnen Komponenten jedoch nicht homogen⁹⁶ ist, wird dieser Ansatz nicht verfolgt.

Sonderformen der DEA sind anwendbar, führen jedoch in der Regel zu keiner Verbesserung der Aussagekraft bzw. zu keinem Zusatzgewinn an Informationen. Zudem ist die Vergleichbarkeit u. U. nicht gegeben, da bspw. die Slack Based DEA ein additives Modell darstellt. Die Anwendung dieser Methode führt größtenteils zu unplausiblen Ergebnissen. Lediglich die Supereffizienz-DEA identifiziert signifikante Ausreißer, im ANSP Bereich insbesondere durch MUAC repräsentiert. Dies spiegelte sich auch in der Modellbildung (2A) wider.

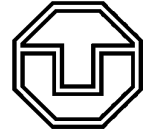
⁹⁵ Bei der geringen Anzahl der Beobachtungen ist die parallele Verwendung jedoch kritisch zu betrachten.

⁹⁶ Dies betrifft im Besonderen die Zuordnung der Approach-Units.



Das Bootstrapping der DEA-Scores der ANSPs wurde insbesondere durch Button und Neiva (2014) sowie Neiva (2014) als obligatorisch angenommen. Es konnte jedoch nachgewiesen werden, dass bei diesem Ansatz die Steigung der RPF steiler ist und damit insbesondere DMUs mit einer großen Betriebsgröße benachteiligt werden. Dies könnte auch die großen Spannbreiten der Scores bei Bilotkach et al. (2015) erklären. Die Beobachtungen bestätigen die kritische Haltung gegenüber dem Bootstrap-Algorithmus in Verbindung mit der DEA durch Coelli et al. (2005). Die Autoren weisen darauf hin, dass der Algorithmus prinzipiell ein Verfahren ist, um die Stichprobenvariabilität zu testen. Bei dem Ansatz wird die RPF generell nach außen verschoben, jedoch wäre bei der Annahme zufälliger Fehler auch eine Verschiebung nach innen theoretisch möglich. Des Weiteren wurde das Bootstrapping-Verfahren für Schätzdaten entwickelt. Die DEA basiert jedoch auf empirisch beobachteten Werten und stellt somit Zähldaten dar. Daher schlagen Coelli et al. (2005) vor, das Bootstrapping lediglich als Sensitivitätsanalyse zu verwenden.

Es ist empfehlenswert, dass die DEA-Methodik auch in offiziellen Benchmarking Reports von EUROCONTROL genutzt wird. Sie bietet zahlreiche Vorteile gegenüber der derzeit verwendeten Indexziffermethode, insbesondere die Berücksichtigung mehrerer Ressourcen und Outputs innerhalb eines ökonomischen Modells. Des Weiteren werden die Einheiten (ANSPs) lediglich mit der korrespondierenden Peer-Group verglichen. Dies ist insbesondere aufgrund der pan-europäisch heterogenen Struktur ein wesentlicher Vorteil.



4 Ursachenanalyse

4.1 Ansatz

Die Produktivität bzw. Effizienz eines ANSP kann durch eine Vielzahl an Faktoren beeinflusst werden. Diese unterscheidet man zunächst in endogene (von dem ANSP beeinflussbar) und exogene (nicht beeinflussbar) Effekte. Die Komplexität der Prozesse setzt eine Berücksichtigung verschiedener Aspekte voraus, u. a.:

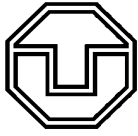
- Skaleneffekte
- Effekte auf der Mikroebene (Sektoren, ACCs)
- Zeiteffekte

In einem ersten Schritt ist die Erstellung einer Liste potentieller Einflussfaktoren notwendig sowie eine Diskussion über deren Wirkung auf die Effizienz der jeweiligen Untersuchungseinheit. Die Menge und Eigenschaften der Faktoren setzen insbesondere auf der disaggregierten Ebene eine Aufteilung in qualitative und quantitative Elemente voraus. Dies wiederum bedingt unterschiedliche Analysemethoden. Für die quantitative Analyse werden vor allem Regressionsmodelle angewendet. Die genauen Spezifikationen sind in den entsprechenden Abschnitten detaillierter beschrieben.

Die qualitative Analyse (Feldstudien, Stakeholder-Konsultationen) deckt vor allem unterschiedliche Arbeitsweisen, Systeme und Berechnungsmethoden ab, aber auch Personalplanung und -flexibilität. Hier wurde eine starke Heterogenität festgestellt, obwohl lediglich die Kontrollzentralen im FABEC Raum untersucht worden sind. Diese Heterogenität ist jedoch auch auf unterschiedliche Nachfragestrukturen zurückzuführen. Des Weiteren spielen zahlreiche (teil-) exogene Effekte, wie bspw. der Einfluss der Gewerkschaften eine wichtige Rolle. Überprüft werden Faktoren, welche sich nicht messen bzw. quantifizieren lassen.

Dem gegenüber stehen die quantifizierbaren Faktoren. Die folgende Liste enthält einige Beispiele potentieller Faktoren:

- Luftraum
 - Größe (Fläche, Volumen)
 - Typ (Oberer, Unterer, etc.)
 - Anzahl und Größe der Temporary Reserved Airspaces (TRAs)
 - Existenz großer Flughäfen
- Nachfrage
 - Verteilung (räumlich, zeitlich)
 - Komplexität
- Juristische, politische und organisatorische Kriterien
 - Eigentümer des Unternehmens, Betriebsmodell
 - Gewerkschaften
 - Regulierungen (z. B. Nachtflugverbote)



- Umweltfaktoren
 - Geographische Besonderheiten (Berge, Ozeane)
 - Wetter (Nebel, Wind, konvektive Wetterlagen)
 - Vulkanaschewolken
- Sozio-Ökonomisch
 - Wohlstand
 - Technologie

Bei der quantitativen Analyse werden zwei Ansätze verfolgt. Zum einen wird über die sukzessive Variablenreduktion ein Modell zur bestmöglichen Prädiktion der abhängigen Variable (Effizienz-/ Produktivitätsmaß) aufgestellt. Dadurch können sowohl Intensität als auch Richtung und statistische Signifikanz der Faktoren untersucht werden. Zum anderen kann über eine schrittweise Betrachtung endogener, teil-exogener und exogener Faktoren die Güte des Effizienzmaßes beurteilt werden. Haben die exogenen Variablen einen signifikanten Einfluss auf die Koeffizienten, so sollten diese in einem Vergleich der Leistungsfähigkeit einbezogen werden. Innerhalb der Dissertation wird dies mittels vier bzw. fünf Iterationen überprüft. In der ersten Iteration werden nur endogene Faktoren berücksichtigt. Iteration 2 und 3 ergänzen Luftraumdaten sowie Nachfrage-Charakteristiken (und somit teil-exogene Faktoren). In der letzten Iteration werden alle Faktoren berücksichtigt. Nach jedem Zwischenergebnis wird ein VIF⁹⁷-Test durchgeführt, sodass Multikollinearität vermieden wird.

Auf makroskopischer Ebene besteht zudem die Herausforderung, dass es relativ wenige Beobachtungen, jedoch viele potentielle Faktoren gibt. Dieses Verhältnis ist auf mikroskopischer Ebene umgekehrt. Daher werden Methoden zur Dimensionsreduktion, insbesondere Cluster-Analyse und Hauptkomponentenanalyse genutzt und auf Durchführbarkeit überprüft (Backhaus et al. 2018). Da die Daten für mehrere Jahre zur Verfügung stehen, werden zudem Panel-Modelle angewendet.

Durch die Ursachen-Analyse können Trends der Effizienzbeeinflussung abgeleitet werden. Dies ermöglicht es, konkrete Maßnahmen vorzuschlagen. Diese Maßnahmen können sowohl operativer als auch institutioneller Natur sein. Die zugrundeliegenden Analyseverfahren werden u. a. in Wooldridge (2002), Hsiao (2014) sowie Welc und Esquerdo (2017) adressiert.

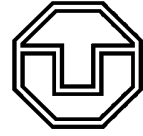
4.2 Regressionsanalyse auf ANSP Ebene

4.2.1 Methoden

4.2.1.1 Regressionsanalyse

Bei der Nutzung der Data Envelopment Analysis (Abschnitt 3.4) wurden auf Basis endogener Faktoren Effizienz- bzw. Produktivitätswerte bestimmt. In einem zweiten Schritt soll untersucht werden, welche (insbesondere exogene) Faktoren die Effizienz eines Flugsicherungsdienstleisters beeinflussen. Dafür können Variablen aus unterschiedlichen Bereichen getestet werden. Das Ziel ist die Erklärung

⁹⁷ Variance Inflation Factor.



der Effizienzwerte durch ein Regressionsmodell. Das Modell untersucht den funktionalen Zusammenhang zwischen einer abhängigen Variablen y (z. B. DEA Effizienzscore) und einer oder mehreren unabhängigen Variablen (z. B. Umwelteinflüsse). Dabei sollte das Modell weder überflüssige noch vernachlässigbare Variablen beinhalten, da sonst die verbleibenden Variablen unpräzise (verzerrt) geschätzt werden (Omitted Variable Bias, OVB) (LeSage und Pace 2009).

In der Regel wird eine lineare Form des funktionalen Zusammenhanges zwischen abhängiger und unabhängigen Variablen angenommen. Sollte dies durch die theoretischen Annahmen des Modells ausgeschlossen werden, sind auch andere Formen möglich (z. B. polynomisch). Des Weiteren ist der Einsatz von Dummy-Variablen gebräuchlich, z. B. wenn eine Beobachtung eine bestimmte Bedingung erfüllt. Im ANSP-Kontext wäre das z. B. die Zugehörigkeit zu einem spezifischen FAB. Zudem sollten Multikollinearität sowie Heteroskedastizität ausgeschlossen werden (Auer und Rottmann 2015).

Wie bereits in Abschnitt 4.1 beschrieben, kann das Verhältnis⁹⁸ von (potentiell) erklärenden Variablen zu den beobachteten Effizienz-Scores dazu führen, dass lediglich eine Auswahl an Faktoren je Regressionsanalyse berücksichtigt werden kann. Dies erhöht jedoch die Wahrscheinlichkeit von OVB. Ein Ansatz ist der Test verschiedener Modelle und die Überprüfung der Modellgütekriterien. Durch Substitution, Eliminierung oder Hinzufügen von Variablen kann sich so einem Modell genähert werden, welches die Einflüsse auf die Effizienz-Scores möglichst präzise abbildet. Dieses Verfahren hat jedoch den Nachteil, dass eine sehr große Anzahl potentieller Möglichkeiten getestet werden müsste. Daher empfiehlt es sich, stattdessen eine Faktoranalyse durchzuführen, um die Anzahl der Variablen zu verringern.

4.2.1.2 Principle Component Analysis

Die Principle Component Analysis (PCA), oder Hauptkomponentenanalyse, ist ein multivariates Verfahren zur Dimensionsreduktion mittels Analyse der Beziehungen zwischen quantitativen Variablen (Schendera 2010). Dabei werden mittels orthogonaler Transformation eine Menge von (korrelierenden) Variablen in eine Menge (nicht-korrelierender) Variablen, die Hauptkomponenten, umgewandelt. Die Variablen werden dabei so reduziert, dass ein möglichst großer Anteil der Varianz aller Variablen erklärt wird. Dies erfolgt über Linearkombinationen.

Bei einer Linearkombination wird explorativ ein lineares Muster in den Daten gesucht, welches den vorhandenen Datensatz bestmöglich beschreibt, d. h. die meiste Varianz erklärt. Dadurch wird der Datensatz um eine Dimension reduziert. Das Verfahren wird solange wiederholt, bis eine gegebene Anzahl von Hauptkomponenten extrahiert ist. Zu beachten ist dabei, dass die Hauptkomponenten nicht korrelieren und die Varianzen je Iteration abnehmen (Gehrke 2019).

Die Zuordnung der Variablen zu den Hauptkomponenten kann (im Falle von zwei Hauptkomponenten) bspw. mittels Biplot dargestellt werden. Um die Interpretierbarkeit der Ergebnisse zu erhöhen, wird im Allgemeinen eine Rotation des Koordinatensystems vorgenommen. Ein gängiges Verfahren ist

⁹⁸ Insbesondere wenn mehr Faktoren untersucht werden sollen als Beobachtungen zur Verfügung stehen, vgl. Abschnitt 4.2.2.1.



hier die Varimax-Rotation. Dadurch können die Variablen präziser den Hauptkomponenten zugeordnet und damit die größtmögliche Varianz aufgeklärt werden (Kaiser 1958).

Die Eignung des Datensatzes kann durch mehrere Tests überprüft werden. Es empfiehlt sich zunächst eine Korrelationsmatrix zu erstellen. Über das Kaiser-Meyer-Olkin (KMO) Kriterium werden Partialkorrelationen überprüft. Dabei sollte mindestens ein Wert von 0,5 erreicht werden (Kaiser 1974). Des Weiteren überprüft der Bartlett-Test, ob sich die Korrelationsmatrix von einer Einheitsmatrix unterscheidet (Bartlett 1951; Snedecor und Cochran 1989). Das Measure of Sampling Adequacy (MSA) Kriterium überprüft wiederum die Eignung einzelner Variablen. Das Kriterium ähnelt dem KMO-Test und sollte ebenfalls Werte größer 0,5 annehmen (Cerny und Kaiser 1977).

Die Anzahl der Hauptkomponenten hängt von den Eigenschaften des Datensatzes ab. Prinzipiell sollte der Eigenwert einer Hauptkomponente größer als 1 sein (Kaiser-Guttman-Kriterium) (Bortz 2013). Zur optischen Darstellung kann zudem ein Scree-Plot genutzt werden, welcher die Eigenwerte nach Größe sortiert⁹⁹ darstellt. Hier greift das sogenannte „Elbow-Kriterium“, welches auf den „Knick“ in der Eigenwertdarstellung hinweist. Es empfiehlt sich, alle Hauptkomponenten links des Knickes zu berücksichtigen. Gibt es mehrere Knicke, so wird die Hauptkomponente vor dem letzten Knick verwendet (Gehrke 2019).

Die durch die PCA identifizierten Komponenten können in die Regressionsanalyse überführt und deren Einfluss auf Produktivität bzw. Effizienz getestet werden. Wichtig ist dabei, dass die Komponenten auch inhaltlich sinnvoll kombiniert werden. Dies kann bspw. durch Erhöhung der vorgegebenen Anzahl an Hauptkomponenten realisiert werden

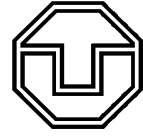
4.2.1.3 Beschränkte und zensierte Modelle

Der funktionale Zusammenhang zwischen Produktivitätswerten und Einflussfaktoren kann mittels einer OLS Regression und den in Abschnitt 4.1 diskutierten Anpassungen analysiert werden. Dieses einfache Regressionsmodell ist anwendbar, falls die abhängige Variable nicht nach oben beschränkt ist. Dagegen können die (relativen) Effizienzscores der DEA lediglich einen Wert zwischen 0 und 1 annehmen.¹⁰⁰ Die Anwendung einer OLS-Regression kann zu verzerrten Ergebnissen führen, da *„die Wirkung der erklärenden Variablen nicht linear ist und die Varianz abnimmt, wenn der Mittelwert sich einer der Grenzen nähert“* (Buis 2010).

Gängige Praxis ist die Anwendung einer Tobit-Regression (Ahmada et al. 2017; Spaho 2015). Dieser Regressionstyp ermöglicht die Begrenzung von Werten auf einen oberen und einen unteren Rand, erfordert jedoch, dass sowohl der obere als auch der untere Begrenzungswert prinzipiell möglich ist. Da ein DEA-Wert von 0 nicht möglich ist, kann ein Tobit-Modell mit oberer und unterer Schranke als Fehlspezifikation angesehen werden (Standfuss und Schultz 2018). Zudem wies Hoff (2007) nach, dass eine OLS-Regression ähnlich gute Ergebnisse liefert, wie die Anwendung von Tobit- oder Papke-Wooldridge-Modellen.

⁹⁹ Beginnend mit dem höchsten Eigenwert.

¹⁰⁰ Wertebereich entspricht]0;1].



Einen Beitrag zur methodischen Auswahl lieferten auch Banker und Natarajan (2008). Die Autoren nutzten sowohl OLS als auch MLE¹⁰¹ als Second-Stage-Methode und überprüften, ob diese Optionen jenen von parametrischen Ansätzen überlegen sind. Auf Basis von Monte-Carlo-Simulationen konnten sie nachweisen, dass die zweistufigen Ansätze (DEA + MLE/OLS/Tobit) eine höhere Performanz erzielen als die getesteten parametrischen Modelle.

Eine Analyse verschiedener Regressionstechniken führten Simar und Wilson (2007) durch. Die Autoren diskutieren in dem Paper verschiedene Möglichkeiten, u. a. die Anpassung der DEA Werte (Log-Transformation) in Verbindung mit einer OLS-Regression, die Anwendung eines Tobit-Modells oder die Durchführung einer Truncated Regression. Sie empfehlen für die Regression basierend auf einer DEA ein „gestütztes“ Modell. Im Allgemeinen kann es hilfreich sein, verschiedene Arten von Regressionsanalysen durchzuführen und die Ergebnisse zu vergleichen. Insbesondere da die Anzahl der Beobachtungen sehr gering ist wird angenommen, dass die Unterschiede zwischen den verschiedenen Regressionsmodellen gering sind.

4.2.2 Daten und Einflussgrößen

4.2.2.1 Auswahl und Gruppierung potentieller Faktoren

Zu Beginn einer Ursachenanalyse steht die Überlegung, welche Faktoren die Leistungsfähigkeit einer Flugsicherung beeinflussen können. Eine erste Orientierung bieten bereits publizierte Analysen, z. B. Button und Neiva (2014). Für die vorliegende Arbeit wurden jedoch bevorzugt Experten-Konsultationen¹⁰² für die Identifikation genutzt und überprüft. Dieser Abschnitt gibt zunächst einen Überblick über potentielle, quantifizierbare Einflussgrößen auf die Performance der Flugsicherungen. Die einzelnen Faktoren werden in den Abschnitten 4.2.2.2 bis 4.2.2.4 ausführlich beschrieben und deren Einfluss auf die Performance diskutiert.

Neben der Identifikation potentieller Faktoren stellt auch die Akquise¹⁰³ der dafür notwendigen Daten einen wesentlichen Bestandteil der Untersuchungen dar. Dies ist insbesondere im Flugsicherungskontext relevant, da Daten teilweise sensibel¹⁰⁴ sind und/oder nicht veröffentlicht werden. Dies betrifft zum Beispiel militärische Aktivitäten. Andere Daten sind nicht vollständig verfügbar, sei es hinsichtlich des Untersuchungszeitraums oder der betrachteten ANSPs.

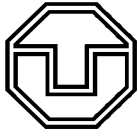
Für die Regressionsanalyse wurde eine Datenbank erstellt, welche die Variablen der einzelnen ANSPs konsolidiert. Diese umfasst sowohl verkehrsspezifische als auch sozioökonomische Daten, welche aus verschiedenen Quellen stammen. Auf ANSP Ebene sind dies insbesondere die PRU (EUROCONTROL 2019g), EUROSTAT (EUROSTAT 2019) und Worldbank (Worldbank 2019). Die Liste aller potentiellen quantifizierbaren Faktoren umfasste 51 Elemente (long list). Aufgrund der begrenzten Zahl an Beobachtungen wurde daher zunächst eine Variablenreduktion mittels Hauptkomponentenanalyse durchgeführt und anschließend in eine Regression überführt. Da die einzelnen Faktoren

¹⁰¹ Maximum Likelihood Estimation.

¹⁰² Diese fanden im Rahmen der FABEC PMG Benchmarking Taskforce statt.

¹⁰³ Dies beinhaltet auch die Prüfung der Verfügbarkeit.

¹⁰⁴ Gemeint sind Daten, welche bspw. nur FABEC-intern verwendet werden dürfen.



nicht sinnvoll zu Komponenten zusammengefügt werden konnten, waren die Modellgüte der Regression¹⁰⁵ sowie die statistische Signifikanz der Komponenten sehr niedrig (vgl. Anhang Abbildung A 5). Trotz diverser Anpassungen und Prüfung von Alternativen, u. a. der Erhöhung der Zahl der Hauptkomponenten und der Substitution von Parametern, konnte keine Verbesserung des Resultates erzielt werden. Daher wurde eine Regressionsanalyse unter Nutzung der originären Faktoren durchgeführt.

Um Singularität und Multikollinearität zu vermeiden, wurden auf ANSP Ebene maximal 21 Faktoren parallel untersucht. Tabelle 4-1 fasst die Ausprägungen der verwendeten Einflussgrößen (exklusive Dummy-Variablen) für das Jahr 2016 zusammen. Die große Streuung der Werte ist bedingt durch die Heterogenität der betrieblichen, geografischen und sozioökonomischen Umweltfaktoren der einzelnen ANSPs. Für den Luftraum kann der Maximalwert von 2.190.000 km² (Spanien) als Extremwert interpretiert werden. Extremwerte sind zudem für die Faktoren „Anteil Inlandsflüge“ sowie „Wohlstand“ sichtbar.

Tabelle 4-1: Deskriptive Statistik der Faktoren (2016)

Einflussgröße	Min	MW	95%Q	Max	STABW
Luftraumgröße (km ²)	20.400	346.350	986.200	2.190.000	428.848
Anteil Überflüge	10 %	59 %	90 %	100 %	26 %
Anteil Inlandsflüge	0 %	7 %	22 %	50 %	10 %
GINI-Koeffizient	2,5 %	10,4 %	19,9 %	25,5 %	5,1 %
Anteil Nicht-ATCO FTEs	38 %	64 %	79 %	87 %	12 %
Arbeitszeit pro Lotse (h/a)	934	1.434	1.901	1.990	237
Kosten pro Lotsenstunde (€)	11	91	189	225	54
Verkehrsdichte	0,69	6,26	11,00	11,47	3,34
Vertikaler Score	0,04	0,15	0,29	0,38	0,09
Horizontaler Score	0,27	0,46	0,61	0,63	0,11
Geschwindigkeitsscore	0,04	0,19	0,36	0,45	0,09
Koordination	2	6	10	11	2
Großflughäfen	0	1	2	3	1
Wohlstand	2.074	29.864	72.306	181.647	32.457
Technologie-Proxy	0,26	13,33	45,20	95,76	18,34

Neben diesen Faktoren werden Dummy-Variablen verwendet. Sofern ein ANSP ein gewisses Merkmal aufweist, nehmen sie den Wert 1 an, ansonsten 0. Dies betrifft vor allem die Organisationsform und die Zugehörigkeit zu FABs. Bei den FABs wurde sowohl die individuelle Zugehörigkeit überprüft (FABEC, FABCE, etc.) als auch die generelle Zugehörigkeit (FAB ja/nein).

Die unterschiedlichen Charakteristiken der Faktoren, insbesondere inwieweit sie durch den ANSP beeinflusst werden können, führt zu einer Untergliederung in endogene, teil-exogene und exogene Elemente. Die Zuordnung der Faktoren zu den drei Gruppen ist jedoch nicht in jedem Fall eindeutig. Zudem werden die teilexogenen Effekte hinsichtlich Luftraumstruktur und Nachfragecharakteristiken untergliedert. Dies ist insbesondere bei der iterativen Regressionsanalyse relevant.

¹⁰⁵ Es wurde dafür eine Varimax-Rotation angewendet.

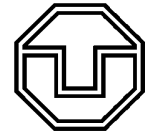


Tabelle 4-2 fasst die in der Regressionsanalyse verwendeten Parameter (Akronyme) zusammen, welche die beschriebenen Einflussgrößen aus Tabelle 4-1 inklusive Dummy-Variablen reflektieren. Um Größeneffekte zu vermeiden, wurden Faktoren mit hohen absoluten Ausprägungen der Werte (z. B. Luftraumgröße) vorab logarithmiert. Dies wird durch ein „(l)“ hinter dem Faktor gekennzeichnet. Des Weiteren werden die Maßeinheiten der Parameter sowie der erwartete Einfluss auf die Produktivität bzw. Effizienz angegeben. Dabei steht ein „+“ für einen positiven Einfluss auf die Effizienz, ein „-“ für einen negativen. Ein „?“ bedeutet, dass die Auswirkung auf die Leistungsfähigkeit auf der Basis der theoretischen A-priori-Analyse positiv oder negativ ausgeprägt sein kann.

Tabelle 4-2: Effizienzbeeinflussende Faktoren in der Regressionsanalyse (ANSP)

	Parameter	Maß	Bedeutung	Einfluss
Endogen	NONA	%	Anteil der Nicht-Lotsen an den gesamten FTE	-
	TIME (l)	h	Arbeitszeit pro Lotsen pro Jahr	+
	DELATM	0/1	Delegated ATM ¹⁰⁶	?
	MET	0/1	MET Services	+
	AIRP	0/1	Beteiligung an bzw. Management von Flughäfen	+
	JSC	0/1	Joint Stock Company	?
	STATE	0/1	State-Owned	
Teil-exogen	SIZE (l)	Km ²	Größe des Luftraumes	+
	OCEAN	0/1	Luftraum enthält einen Oceanic Airspace	?
	COORD	Stück	Koordinations-Proxy – zählt die angrenzenden Lufträume	-
	L_AIRP	Stück	Anzahl der Großflughäfen (Flugbewegungen > 200.000 p.a.)	-
	NOFAB ¹⁰⁷	0/1	Der ANSP gehört zu keinem FAB	?
	OVER	%	Anteil der Überflüge	+
	DOM	%	Anteil der Inlandsflüge	-
	GINI	%	Saisonale Verteilung des Verkehrs	-
	DENSITY	Score	Verkehrsdichte (Komplexitäts-Faktor)	-
	VI	Score	Vertikale Interaktionen (Komplexitäts-Faktor)	
	HI	Score	Horizontale Interaktionen (Komplexitäts-Faktor)	
SI	Score	Geschwindigkeitsinteraktionen (Komplexitäts-Faktor)		
Exog.	COSTS	€	Arbeitskosten pro Lotsenstunde	+
	RES	Stück	Technologie-Proxy (Patente pro 100.000 Einwohner)	+
	WEALTH (l)	€	Wohlstand des Landes in BIP/Kopf	+

Die erwarteten Einflüsse sowie deren Allokation zu endogenen, teil-exogenen und exogenen Faktoren werden detailliert in den folgenden Abschnitten diskutiert. Eine Übersicht zur Korrelation zwischen den verwendeten Parametern bietet Abbildung A 7 in Anhang C3.

4.2.2.2 Endogene Faktoren

Wie im vorigen Abschnitt diskutiert, kann die Leistungsfähigkeit eines Unternehmens kurz-, mittel- oder langfristig durch verschiedene Faktoren beeinflusst werden. Innerhalb des Unternehmens sind dies bspw. die Flexibilität der Personaleinsatzplanung oder der Anteil des Personals, welcher unmittelbar am

¹⁰⁶ Der ANSP lässt einen Teil des Luftraumes durch einen anderen ANSP (gegen einen finanziellen Ausgleich) kontrollieren.

¹⁰⁷ Es werden auch einzelne FABs getestet. Diese sind dann ebenfalls mit 0 (nicht zugehörig) oder 1 (zugehörig) bewertet.



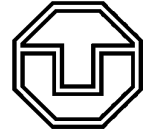
Wertschöpfungsprozess beteiligt ist. Auch verwendete Systeme und Tools können einen Einfluss auf die Leistungsfähigkeit der Unternehmen haben. Des Weiteren sind Betriebsform bzw. Besitzverhältnisse (staatlich vs. privatisiert) sowie das Portfolio an Leistungen zu überprüfen. In diesem Abschnitt werden die endogenen Faktoren (kursiv geschrieben) adressiert.

In der Regressionsanalyse werden zwei Variablen integriert, welche den strukturellen Personaleinsatz beschreiben. *NONA* gibt jenen Anteil des Personals an, welcher keine Lotsentätigkeit ausführt und damit nicht (unmittelbar) am Wertschöpfungsprozess des ANSP im Kernbereich beteiligt ist. Die Variable kann somit als Indikator für die Größe des Overheads bzw. der Verwaltung gesehen werden. Da die dorthin zugeordneten Ressourcen die Kosten tendenziell erhöhen, ist zu erwarten, dass die Variable mit zunehmender Ausprägung einen negativen Einfluss auf die Produktivität (und Kosteneffizienz) eines Unternehmens hat. Allerdings ließe sich auch umgekehrt argumentieren, dass die Non-ATCOs u. a. Supportaufgaben für die ATCOs wahrnehmen, was sich positiv auf deren Produktivität auswirken könnte.

Die zweite Personal-Variable beschreibt die durchschnittliche Arbeitszeit pro Lotsen pro Jahr (*TIME*). Wie in Tabelle 4-1 beschrieben, sind hier pan-europäisch starke Unterschiede sichtbar. Eine geringe Jahresarbeitszeit pro Lotsen bedeutet, dass für den gleichen Bedarf an Lotsenstunden (Input im Produktivitäts- und Effizienzmaß) mehr Lotsen (i.S.v. FTEs) benötigt werden. Daher ist anzunehmen, dass sich eine hohe Jahresarbeitszeit auch positiv auf die Performance auswirkt. Natürlich bestünde die Möglichkeit, dass aufgrund von „Überlastung“ auch eine negative Auswirkung auf die Produktivität zu beobachten ist. Geht man von einer Jahresarbeitsleistung¹⁰⁸ von 1.680 Stunden (40-Stunden Woche) bzw. 1.596 Stunden (38-Stunden Woche) aus, so weisen lediglich drei bzw. fünf europäische ANSPs höhere Jahresarbeitsstunden pro Lotsen auf.

DELATM und *MET* spiegeln klassische „Make or buy“ Entscheidungen wider. Es kann für einen ANSP vorteilhaft sein, Teile des ATMs (insbesondere ATC) an einen anderen ANSP auszulagern, u. a. aufgrund spezifischer Verkehrsflüsse. So wird bspw. ein Teil des Luftraums des ACC Rom (ENAV) an die maltesische MATs delegiert. Dieser Luftraum wird insbesondere von Flügen von und nach Malta frequentiert. Findet für diese Übertragung ein finanzieller Ausgleich statt, so wird der Dummy des zugehörigen ANSP „1“. Als Beispiel ist skyguide für einen Teil des französischen Luftraums in der Nähe von Genf verantwortlich. Beide ANSPs können so von operativen Vorteilen profitieren, allerdings wirken sich diese Allokationen auch auf die Kosten der ANSPs aus. Eine eindeutige Erwartung bezüglich des Einflusses auf die Produktivität ist daher nicht möglich. Zudem reflektiert der Dummy nicht, wie groß der delegierte Luftraum ist, bzw. wie viele Flüge davon betroffen sind. Gemäß (EUROCONTROL 2012) soll die Größe des Luftraumes um jene der delegierten Gebiete korrigiert werden. Ob dies in der Praxis in allen ANSPs stattfindet, ist jedoch nicht bekannt.

¹⁰⁸ Die angegebenen Werte sind jedoch u. a. abhängig von den Urlaubstagen und stellen damit Modellwerte dar.



Bei der Variable *MET* wird ein positiver Einfluss angenommen, da insbesondere große ANSPs von einem internen Wetterbüro profitieren. Somit sind relevante Daten des Wetters, welches einen direkten Einfluss auf Menge und Verteilung des Verkehrs haben kann, schneller verfügbar. Dadurch erhöht sich die Planbarkeit der Verkehrsströme. Zudem bestehen bei externen Anbietern ggf. Einschränkungen hinsichtlich Übertragung von Daten und der Einspeisung in die notwendigen Tools (Schnittstellen). Eine bessere Approximation wäre ggf. über die MET-Kosten möglich, jedoch sind diese Daten unvollständig.

Mit *AIRP* wird angegeben, ob die Flugsicherung auch das Flughafenmanagement übernimmt. Dies ist z. B. bei Avinor in Norwegen der Fall. Die Beteiligung an bzw. der Besitz von Flughäfen durch die Flugsicherung kann die Koordinierung zwischen beiden Instanzen erleichtern. Allerdings erfasst der Dummy nicht die unterschiedlichen Ausprägungen. Bspw. erlaubt er keine Differenzierung, ob er primär Groß- oder Kleinflughäfen abbildet.

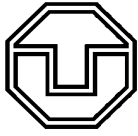
Die Effizienz eines Unternehmens kann auch durch das Besitzverhältnis beeinflusst werden. In ökonomischen Analysen wird dabei in staatliche, privatisierte und teilprivatisierte Unternehmen differenziert. Es wird angenommen, dass die Privatisierung staatlicher Unternehmen, auch im Verkehrswesen, zu einer höheren Leistungsfähigkeit führt (Button 2010; Schwedes 2018). Allerdings konnte Liebert (2011) nachweisen, dass im Flughafenbereich Mischformen zu einer geringeren Effizienz führen. Im ANSP-Kontext ist eine Unterscheidung zwischen staatlichen und privatisierten Unternehmen schwierig, da lediglich die britische NATS und die schweizerische skyguide als teilprivatisiert gelten (EUROCONTROL 2019a). Hierbei handelt es sich jedoch um keine Privatisierungen im Sinne des Eigentums, sondern in beiden Fällen eher um eine formelle Privatisierung. Der ACE Report unterscheidet sechs verschiedene Unternehmensformen, u. a. „state owned“ (*STATE*) oder „Joint Stock Company“¹⁰⁹ (*JSC*). Beide Formen repräsentieren in Summe 75% der ANSPs. Aufgrund der geringen Anzahl an Beobachtungen für die weiteren vier Unternehmensformen wurde auf ergänzende Dummy-Variablen verzichtet. Gemäß der induzierten Effizienzsteigerung durch Privatisierung, wie in der Literatur angegeben, wird für *STATE* ein negativer Einfluss erwartet, während für *JSC* eine positive Konnotation vermutet wird. Die Unternehmensform wird als endogener Faktor betrachtet, wengleich durch staatliche Einflüsse (z. B. Regulierungen) theoretisch auch exogene Effekte einwirken.

4.2.2.3 Teilexogene Faktoren

Viele der identifizierten Faktoren (Tabelle 4-1) sind, zumindest teilweise, exogen. Als erster Unterbereich wird der Luftraum betrachtet. Zwar haben ANSPs gewisse Gestaltungsmöglichkeiten in der Strukturierung des Luftraumes, jedoch ist die eigentliche Größe (*SIZE*) durch die staatlichen Hoheitsgebiete vorgegeben. Dabei kann ein positiver Zusammenhang zwischen der Größe und der Leistungsfähigkeit vermutet werden, da große Lufträume flexibler unterteilt werden können¹¹⁰. Dadurch wird die Reaktionsfähigkeit der Flugsicherung auf Verkehrsschwankungen (z. B. durch Störungen) erhöht. Auch im Falle einer geringen Auslastung (z. B. nachts) ist durch die Zusammenlegung der Sektoren eine

¹⁰⁹ Aktiengesellschaft.

¹¹⁰ Teilung und Zusammenlegung von Sektoren, je nach Verkehrsnachfrage.



große Fläche durch die gegebene Anzahl an Ressourcen (im Normalfall zwei ATCOs) möglich. Diese Möglichkeit ist in kleinen Lufträumen begrenzt (vgl. Abschnitt 4.5).

Das Hoheitsgebiet des Staates kann Lufträume über Ozeanen bzw. Meeren umfassen (z. B. Spanien, Portugal und Irland). Die Dummy-Variable *OCEAN* gibt an, ob der Luftraum des korrespondierenden ANSP „Oceanic Airspaces“ beinhaltet. Bei dieser Variablen wird kein eindeutiger Zusammenhang erwartet. Zum einen müssen große Flächen (Lufträume) abgedeckt werden, in denen teilweise wenig oder kein Verkehr stattfindet. Dies würde sich negativ auf die Leistungsfähigkeit auswirken. Zum anderen ist die Komplexität in diesen Lufträumen gering, was sich positiv auf die Leistungsfähigkeit auswirkt. Die Größe der Oceanic-Lufträume kann sich stark unterscheiden, jedoch wurden Daten zur Größe des Luftraumes bzw. dessen Anteil am Gesamtluftraum bis zum Zeitpunkt der Untersuchung nicht gesondert ausgewiesen.

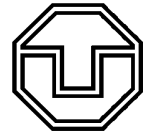
Die geographische Lage bedingt zudem den Koordinierungsaufwand (*COORD*). Je mehr Länder, respektive ANSPs an das eigene Hoheitsgebiet angrenzen, desto höher ist der Koordinationsaufwand. Ggf. ist auch eine Interoperabilität zwischen Systemen nicht gewährleistet, insbesondere wenn die benachbarten ANSPs in unterschiedlichen FABs organisiert sind und/oder sich signifikant in ihren Determinanten¹¹¹ (und den damit verbundenen Anforderungen an Systemen und Tools, vgl. Abschnitt 4.4) unterscheiden. Daher wird ein negativer Zusammenhang erwartet.

Ein weiterer potentieller Faktor, welcher die Leistungsfähigkeit beeinflusst, stellt die Anzahl der Großflughäfen dar. Dieser ist durch die Variable *L_AIRP* definiert. Der Verkehr konzentriert sich (räumlich) in den Regionen um die Großflughäfen. Das erhöht die Komplexität, respektive die Verkehrsdichte und damit den Taskload der Lotsen. Daher ist hier ein negativer Zusammenhang zu vermuten. Gezählt werden alle Flughäfen, die mehr als 200.000 Flugbewegungen pro Jahr aufweisen. Diese Grenze entspricht den Top 20¹¹² Flughäfen Europas. Eine hohe Anzahl an Großflughäfen kann sich auch positiv auf die Performance auswirken, bspw. wenn die vorhandene Luftraumkapazität effizienter ausgenutzt wird. Dies wird jedoch u. a. durch die Entfernung zwischen den einzelnen Großflughäfen beeinflusst. Ein weiterer Faktor ist die Komplexität des spezifischen Bahnsystems (Gelhausen et al. 2019; Kern 2019). Ein unabhängiges Parallelbahnsystem wäre in diesem Falle die optimale Situation. Allerdings ist dies, insbesondere bei Großflughäfen wie Amsterdam, Zürich oder Frankfurt, nicht der Fall.

Die Zusammenarbeit mehrerer ANSPs in einem FAB soll zu einem positiven Einfluss auf die Leistungsfähigkeit des europäischen ATMs führen (EUROCONTROL und FAA 2016; EUROCONTROL 2018b). Aufgrund der geographischen Lage ist die Allokation eines Flugsicherungsdienstleisters zu einem FAB nicht signifikant beeinflussbar. Es wäre jedoch bspw. möglich, die Anzahl der ANSPs pro FAB anzupassen. Des Weiteren hat der ANSP Gestaltungsspielräume hinsichtlich der Kooperation, z. B. ob eine Anpassung der Systeme stattfindet. Das Ziel eines FAB ist die Realisierung von Größenvorteilen. Daher kann von einem positiven Einfluss ausgegangen werden. Im Umkehrschluss bedeutet dies einen

¹¹¹ Z. B. Art des Luftraums (Upper / Lower), Geographie oder Anzahl und Lage militärischer Gebiete.

¹¹² Gilt für 2016. In anderen Jahren kann diese Anzahl geringfügig abweichen.



Nachteil für ANSPs, welche nicht in einem FAB koordiniert sind (*NOFAB*). In Abschnitt 4.5 wird dieser Effekt näher untersucht. Dabei kann gezeigt werden, dass die proklamierten Größenvorteile nicht von allen ANSPs realisiert werden können (Standfuss et al. 2017b, 2019a). Innerhalb der Regressionsanalyse wird daher sowohl die allgemeine Zugehörigkeit zu einem FAB untersucht als auch die Allokation des ANSP in den spezifischen FABs (*FABEC*, *FABCE*, *DANUBE*, *UKIRL*, *NEFAB*, *SWFAB*, *BALTIC*, *BLUEMED*, *DKSEFAB*).

Für die Nachfragecharakteristiken gelten ähnliche Annahmen wie für den Luftraum. Zwar haben ANSPs Steuerungsmöglichkeiten, jedoch sind Menge und Verteilung des Verkehrs größtenteils exogen. In der Regression werden im Wesentlichen drei Elemente untersucht: Typ des Fluges, Variabilität und Komplexität. Der Nachfragetyp beschreibt, ob ein Flug lediglich im Inland stattfindet, Start oder Landung außerhalb des Hoheitsgebiets erfolgt oder ein Überflug ist und damit im betreffenden Staat weder landet noch startet. Dabei repräsentiert die Variable *OVER* den Anteil an Überflügen, die Variable *DOM* den Anteil an Inlandsflügen. Eine Berücksichtigung der internationalen Flüge mit Start oder Landung im Hoheitsgebiet findet aufgrund der Kollinearität nicht statt. Es ist anzunehmen, dass die Variable *OVER* positiv auf die Effizienz einwirkt, da der Streckenflug aufgrund des geringen Vertikalverkehrs¹¹³ eine geringere Komplexität aufweist. Vice versa gilt für die Variable „DOM“ die Annahme einer negativen Auswirkung auf die Leistungsfähigkeit.

Eine weitere Einflussgröße sind Verkehrsschwankungen über die Zeit (z. B. Saisonalität). In Gebieten mit großen Fluktuationen muss mehr Personal vorgehalten werden, was den Ressourcenbedarf erhöht. Im Falle eines Mangels an Flexibilität in der Personaleinsatzplanung kann dies zu Effizienzverlusten führen. Verstärkt wird dieser Effekt dadurch, dass Lotsen nicht jedem beliebigen Sektor zugewiesen werden können. Im ANSP Kontext hat sich für diese Verkehrsschwankungen der Begriff Volatilität etabliert. Definition, Berechnung und Auswirkung dieser Schwankungen werden u. a. in Deltuvaité (2018), FABEC (2018b), FABEC (2018c), Laursen (2018) sowie Standfuss et al. (2019b) adressiert. Für die Dissertation wurden mehrere Ansätze getestet, u. a. die Standardabweichung der Zahl an monatlichen Flügen, der zugehörige Variationskoeffizient sowie der GINI-Koeffizient.

Der Begriff Volatilität stammt ursprünglich aus dem Finanzwesen und bezeichnet die Fluktuation einer Zeitreihe um ihren Mittelwert oder Trend. Die Volatilität repräsentiert somit eine Approximation für das Risiko (Daly 2008; Hafner 2013). Der Volatilitätswert σ einer Einheit errechnet sich gemäß Formel (4-1) durch Standardabweichung der Änderungsrate x in der Zeiteinheit i sowie dem arithmetischen Mittel μ über alle n Zeiteinheiten. Im ANSP-Kontext werden für x tatsächliche Werte (Flüge) verwendet. Diese stützen sich auf Daten der PRU, welche die täglichen Flüge je ANSP veröffentlicht (EUROCONTROL 2020i).

Der Volatilitätswert σ ist jedoch skalenabhängig. Folglich sind die Volatilitätswerte für große Betriebseinheiten im Vergleich zu kleineren ANSPs tendenziell höher. Daher wurde auch der Variationskoeffizient (CV) gemäß Formel (4-2) berechnet.

¹¹³ Steigender oder sinkender Verkehr.



$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (4-1)$$

$$CV = \frac{\sigma}{\mu} = \frac{100}{\mu} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (4-2)$$

Eine Quantifizierung von saisonalen Schwankungen ist zudem über Maße für Ungleichverteilungen möglich. Oft genutzte ökonomischen Messgrößen sind der GINI-Koeffizient und der Herfindahl-Hirschman-Index (HHI), siehe z. B. Rhoades (1993). Der GINI-Koeffizient misst die relative Konzentration, z. B. der Verkehrsnachfrage über das Jahr, basierend auf der Anzahl der Beobachtungen n , dem Beobachtungsindex (z. B. Monat) i und der entsprechenden Nachfrage x_i , siehe Formel (4-3). Je höher der Koeffizient, desto höher ist die Konvexität der Lorenzkurve, desto größer ist die Ungleichverteilung (Pflaumer et al. 2017). Dieses Prinzip ist auch auf die Nachfrage nach Flugsicherungsdienstleistungen anwendbar.

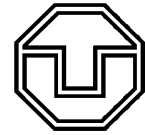
$$GINI = \frac{2 \cdot \sum_{i=1}^n i \cdot x_i}{n \cdot \sum_{i=1}^n x_i} - \frac{n+1}{n} \quad (4-3)$$

Der HHI wird häufig zur Berechnung der Marktanteile von Unternehmen verwendet. Der Index spiegelt die Summe der quadrierten Marktanteile der einzelnen Beobachtungen wider (Formel (4-4)). Bei der Bestimmung der Saisonalität wird auf dieser Basis der Verkehrsanteil der einzelnen Monate bestimmt. Der Index wird meist in einem zweiten Schritt normalisiert (Intervall $[0, 1]$) (Bem et al. 2020). Je größer der Index, desto größer ist die saisonale Konzentration des Verkehrs.

$$HHI = \sum_{i=1}^n \left(\frac{x_i}{\sum_{i=1}^n x_i} \right)^2 \quad (4-4)$$

Auf der Grundlage der PRU-Dashboard-Daten wurden die Volatilitätswerte für jede DMU und jedes Jahr unter Verwendung der Formeln (4-1) bis (4-4) berechnet. Tabelle 4-3 zeigt die Volatilitäts-Scores für eine Auswahl an ANSPs. Daraus lassen sich mehrere Erkenntnisse ableiten. Der absolute Volatilitätswert σ ist skalenabhängig, daher sind insbesondere ANSPs mit einer relativ großen Betriebsgröße durch hohe Volatilitätswerte gekennzeichnet. In der Folge eignet sich das Maß insbesondere für den Vergleich von ANSPs mit ähnlichem Verkehrsaufkommen. Im Gegensatz dazu ist ein Vergleich von heterogenen Einheiten, wie in der Regressionsanalyse notwendig, nicht sinnvoll. Hierfür eignet sich Variationskoeffizient CV, der die Volatilität σ in ein Verhältnis zur Betriebsgröße (i.S.v. Anzahl an Flügen) setzt.

Der HHI ist für die Berechnung der Volatilität wenig aussagekräftig. Die Scores unterscheiden sich nicht signifikant zwischen den ANSPs. Dies liegt darin begründet, dass der HHI weniger empfindlich auf die Unterschiede bei den Flügen reagiert, d. h. die „Marktanteile“ der Monate sich nicht signifikant



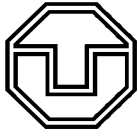
unterscheiden. Daher nimmt der HHI immer einen Wert von knapp über 1/12, respektive 8,3%, an¹¹⁴. Eine bessere Annäherung liefert der GINI-Koeffizient. Da dieser eine hohe Korrelation zum Variationskoeffizienten aufweist, ist eine Substitution beider Indizes möglich. In weiteren Vorstudien konnte der GINI-Koeffizient, u. a. hinsichtlich der Aggregierbarkeit auf FAB Ebene, wesentliche Vorteile gegenüber den alternativen Metriken nachweisen. Daher wurde er auch als Saisonalitäts-Komponente *GINI* in die Regression integriert. Da Verkehrsschwankungen bzw. Ungleichverteilungen zu einer Erhöhung der Vorhaltekosten führen, wird ein negativer Einfluss auf die Leistungsfähigkeit angenommen.

Tabelle 4-3: Volatilitäts-Scores ausgewählter ANSPs für das Jahr 2019

ANSP	Land	σ	CV	HHI	GINI
Albcontrol	Albanien	6.736	36,6%	9,4%	19,7%
ANS Finland	Finnland	1.170	5,2%	8,4%	2,8%
DFS	Deutschland	33.153	12,8%	8,5%	6,8%
DSNA	Frankreich	48.413	17,6%	8,6%	9,5%
EANS	Estland	2.504	13,2%	8,5%	7,2%
ENAIRES	Spanien	31.069	17,3%	8,6%	9,4%
ENAV	Italien	36.114	23,7%	8,8%	12,9%
HCAA	Griechenland	28.969	39,3%	9,5%	21,1%
IAA	Irland	7.869	14,6%	8,5%	7,9%
LVNL	Niederlande	5.130	9,7%	8,4%	5,1%
M-NAV	Mazedonien	7.941	42,8%	9,7%	23,0%
MOLDATSA	Moldawien	1.239	26,9%	8,9%	14,4%
MUAC	International	16.045	10,3%	8,4%	5,6%
NATS (Continental)	UK	28.649	13,6%	8,5%	7,3%
Oro Navigacija	Litauen	3.662	16,4%	8,5%	8,8%
Belgocontrol	Belgien	6.670	12,5%	8,5%	6,7%
Skyguide	Schweiz	18.067	16,5%	8,5%	8,9%

Relevant ist für die Effizienzbetrachtung auch die Komplexität. Hier existieren verschiedene Ansätze zur Berechnung, u. a. Djokic (2014), EUROCONTROL und NATS (2000), EUROCONTROL (2008a) sowie Majumdar und Ochieng (2007). In der vorliegenden Arbeit wurde aufgrund der Datenverfügbarkeit die von der PRU bereitgestellte Metrik genutzt (EUROCONTROL 2006, 2020m). Da es sich bei diesem „Score“ um eine aggregierte Größe handelt, wurden die Einzelkomponenten für die Regressionsanalyse verwendet. Dabei beschreibt *DENSITY* die relative Verkehrsdichte, *HI* die horizontalen und *VI* die vertikalen Interaktionen. Unterschiedliche Flugzeugtypen („Type Mix“) berücksichtigt die PRU über den sogenannten Geschwindigkeitsscore *SI* (aufgrund der unterschiedlichen Luftfahrzeuggrößen und -antriebe). Annahmen, Ansatz und Methodik sind ausführlich in EUROCONTROL (2006) beschrieben. Im Allgemeinen wird durch eine hohe Komplexität eine höhere Arbeitsbelastung für den Lotsen impliziert. Dies hat einen negativen Einfluss auf die Kapazität des Sektors und damit auf die Produktivität bzw. Effizienz. Daher wird für alle vier Komponenten ein negativer Zusammenhang erwartet.

¹¹⁴ Ein Wert von 8,3% würde eine Gleichverteilung des Verkehrs über die Zeiteinheiten (Monate) bedeuten.



4.2.2.4 Exogene Faktoren

Exogene Effekte sind durch die Unternehmen nicht beeinflussbar und umfassen u. a. sozioökonomische Faktoren, z. B. die Größe der Volkswirtschaft oder den Wohlstand eines Landes. Weitere potentielle Faktoren sind Wetter, geografische Lage oder andere Umweltfaktoren. Die zugrundeliegende Annahme ist, dass sich der Wohlstand (*WEALTH*) eines Landes auch in der Effizienz des zugehörigen ANSP widerspiegelt. (z. B. hinsichtlich Investitionspotentialen). Allerdings weisen Größen wie BIP oder BIP/Kopf hohe Korrelationen zu anderen verwendeten Variablen auf. Daher konnte die Variable *WEALTH* nicht in der Regressionsanalyse berücksichtigt werden.

Die Variable *COSTS* beschreibt die Arbeitskosten pro Lotsenstunde. Aufgrund der Annahme, dass hohe Lotsenkosten zu einem effizienteren Ressourceneinsatz führen, wird bei dieser Variablen ein positiver Einfluss auf die Effizienz erwartet. In einer früheren Studie wurden die Kosten zusätzlich durch die individuellen Kaufkraft-Paritäten geteilt¹¹⁵ (European Commission 2018). Auf diese Anpassung wurde verzichtet, da insbesondere die Berücksichtigung dieser Lohneffekte abgebildet werden sollte. Zudem ist die Nutzung des PPP-Indexes nicht unumstritten (Eibner 2010; Ong 2003).

Effizienz kann zudem von dem Technologie-Stand eines Landes abhängig sein. Diese Größe wird in der Variable *RES* approximiert, welche die Patente pro 100.000 Einwohner widerspiegelt. Auch hier kann ein positiver Einfluss auf die Leistungsfähigkeit angenommen werden. Alternative Metriken wurden getestet, jedoch aufgrund mangelnder Datenverfügbarkeit verworfen. Zudem mussten die Werte einiger Länder inter- oder extrapoliert werden, da sie nicht für alle Jahre verfügbar sind.

Die fehlende Datenverfügbarkeit stellt in der Regressionsanalyse allgemein ein wesentliches Problem dar. So können potentiell einflussreiche Faktoren, z. B. militärischer Luftverkehr oder Anteil des General Aviation Verkehrs, nicht berücksichtigt werden. Beide tragen wesentlich zur Komplexität des Luftraumes bei, werden allerdings im Maß der PRU ebenfalls nicht berücksichtigt (EUROCONTROL 2006). Auch Wettereinflüsse können nicht zielführend quantifiziert werden und bleiben daher unberücksichtigt.

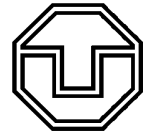
4.2.3 Einfluss auf die Produktivität

In den offiziellen EUROCONTROL Berichten wird insbesondere die ATCO-Produktivität untersucht. Diese wird durch den Quotienten aus kontrollierten IFR-Flugstunden und den dafür eingesetzten ATCO-Stunden berechnet (vgl. Abschnitt 3.2.1) (EUROCONTROL 2018a). Ziel ist zunächst, die Einflussgrößen auf diese Metrik zu bestimmen. Dafür wird das Regressionsmodell mittels Variablenreduktion angepasst, sodass die abhängige Variable möglichst gut auf Basis der verwendeten Variablen vorhergesagt werden kann. Als Abbruchkriterium wurde ein p-Wert von 0,33¹¹⁶ genutzt. Da die abhängige Variable nicht nach oben beschränkt ist, wird eine OLS-Regression gewählt.

Die Ergebnisse sind in Tabelle 4-4 dargestellt. Neben den Modellgütekriterien (Akaike und Korrigiertes R^2) ist auch die Anzahl der Beobachtungen N angegeben. Aus Gründen der Übersichtlichkeit

¹¹⁵ Dies fand jedoch im Effizienzmodell statt.

¹¹⁶ Willkürliche Auswahl, für alle optimierten Regressionen verwendet.



werden lediglich Faktoren dargestellt, die einen p-Wert von unter 0,33 aufweisen. Der VIF Test auf Kollinearität war negativ. Das Ergebnis der Regressionsanalyse entspricht größtenteils den Erwartungen und erreicht mit einer aufgeklärten Varianz von 0,86 eine hohe Modellgüte. Der Anteil der Nicht-Lotsen, delegiertes ATM, staatlicher Besitz, der Anteil an Inlandsflügen, der GINI-Koeffizient und der Technologie-Proxy haben einen negativen Einfluss auf die ATCO-Produktivität. Letzteres entspricht nicht der Erwartung, ist jedoch statistisch insignifikant.

Tabelle 4-4: Regressionsmodell zur Prädiktion der abhängigen Variable PRU-Produktivität

INT	-2,69 (1,684)
NONA	-0,454 (0,328)
TIME (I)	0,213 (0,178)
DELATM	-0,138 (0,071)*
AIRP	0,576 (0,142)***
JSC	0,129 (0,073)*
STATE	-0,065 (0,065)
SIZE (I)	0,125 (0,038)***
OCEAN	0,211 (0,095)**
COORD	0,019 (0,018)
OVER	0,587 (0,15)***
DOM	-1,8 (0,568)***
GINI	-3,069 (0,597)***
DENSITY	0,026 (0,013)*
HI	0,619 (0,361)
COSTS	0,002 (0,001)**
RES	-0,003 (0,002)
Korr. R ²	0,86
Akaike	-36,38
N	38

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%** , 99%***

Nicht erwartungskonform sind die Komplexitätsindikatoren als positiv, nicht signifikant oder nicht relevant ausgewiesen. Dies kann jedoch ein Hinweis darauf sein, dass die Berechnungsmethodik der Komplexitätswerte durch die PRU wesentliche Aspekte nicht berücksichtigt. Des Weiteren ist der positive Zusammenhang zwischen Koordination und Produktivität kontraintuitiv, jedoch ist diese Interdependenz nicht signifikant. Die Luftraumgröße wirkt sich wie erwartet positiv auf die Leistungsfähigkeit aus. Hier wurden auch parabolische und hyperbolische Zusammenhänge getestet, was jedoch zu keinen signifikanten Verbesserungen führte. Zusätzlich wurde die Inkludierung einzelner FABs überprüft. Dies wirkte sich jedoch negativ auf die Ergebnisse hinsichtlich der Modellgütekriterien sowie der statistischen Signifikanzen der Parameter aus.



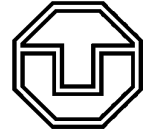
In einem zweiten Schritt wird das Regressionsmodell sukzessiv aufgebaut, wie in Abschnitt 4.1 beschrieben. Zunächst testet das Modell lediglich endogene Variablen, bevor Luftraum- und Nachfragedaten sowie exogene Faktoren integriert werden. Hier ist die Zuordnung nicht immer eindeutig. Bspw. sind die Kosten pro Lotsen (FTE) nicht vollständig exogen. Da hier aber Lohneffekte eine große Rolle spielen, wird dieser Faktor erst im letzten Schritt berücksichtigt. Tabelle 4-5 fasst die Ergebnisse zusammen.

Tabelle 4-5: OLS Regression ATCO-Produktivität

	2015				2016			
	I1	I2	I3	I4	I1	I2	I3	I4
INT	4,788 (2,352)*	3,934 (2,847)	-4 (2,928)	-2,067 (3,38)	4,656 (2,597)*	3,474 (3,029)	-3,293 (2,336)	-1,256 (2,387)
NONA	-1,037 (0,464)**	-0,233 (0,536)	-0,499 (0,438)	-0,536 (0,47)	-1,048 (0,492)**	-0,221 (0,551)	-0,53 (0,364)	-0,584 (0,366)
TIME (I)	-0,443 (0,322)	-0,482 (0,353)	0,243 (0,277)	0,16 (0,3)	-0,422 (0,354)	-0,45 (0,378)	0,307 (0,228)	0,173 (0,221)
DELATM	-0,082 (0,14)	-0,142 (0,134)	-0,213 (0,088)**	-0,131 (0,097)	-0,07 (0,148)	-0,138 (0,139)	-0,226 (0,074)***	-0,119 (0,076)
MET	-0,126 (0,108)	-0,056 (0,116)	0,01 (0,078)	0,008 (0,075)	-0,137 (0,116)	-0,065 (0,123)	0,017 (0,067)	0,004 (0,06)
AIRP	-0,034 (0,173)	0,031 (0,177)	0,597 (0,21)**	0,626 (0,225)**	-0,082 (0,184)	-0,035 (0,184)	0,57 (0,177)***	0,621 (0,179)***
JSC	-0,008 (0,138)	0,067 (0,155)	0,178 (0,111)	0,139 (0,109)	-0,017 (0,147)	0,048 (0,159)	0,163 (0,092)*	0,131 (0,082)
STATE	-0,146 (0,136)	-0,089 (0,141)	-0,017 (0,097)	-0,035 (0,095)	-0,128 (0,147)	-0,083 (0,146)	-0,069 (0,081)	-0,067 (0,074)
SIZE (I)		0,051 (0,063)	0,188 (0,072)**	0,107 (0,085)		0,073 (0,064)	0,139 (0,059)**	0,071 (0,06)
OCEAN		0,08 (0,201)	0,218 (0,142)	0,173 (0,138)		0,094 (0,21)	0,207 (0,116)*	0,218 (0,103)*
COORD		-0,003 (0,032)	-0,003 (0,028)	0,016 (0,028)		-0,005 (0,034)	0,015 (0,024)	0,032 (0,022)
L_AIRP		0,017 (0,088)	0,051 (0,089)	-0,008 (0,091)		-0,006 (0,093)	0,003 (0,072)	-0,04 (0,067)
NOFAB		-0,453 (0,161)***	0,075 (0,158)	0,062 (0,157)		-0,471 (0,164)***	0,039 (0,12)	0,042 (0,109)
OVER			0,571 (0,33)	0,434 (0,342)			0,437 (0,278)	0,346 (0,262)
DOM			-2,046 (0,81)**	-1,666 (0,834)*			-2,249 (0,69)***	-1,976 (0,667)***
GINI			-2,722 (1,387)*	-3,086 (1,427)**			-4,258 (1,077)***	-4,32 (1,021)***
DENSITY			0,045 (0,016)**	0,03 (0,018)			0,051 (0,013)***	0,031 (0,014)**
VI			0,182 (0,854)	-0,252 (0,855)			0,486 (0,636)	-0,022 (0,591)
HI			1,494 (0,624)**	0,813 (0,698)			0,976 (0,47)*	0,436 (0,468)
SI			0,157 (1,01)	-0,115 (1,001)			-0,856 (0,824)	-0,975 (0,742)
COSTS				0,003 (0,002)*				0,003 (0,001)**
RES				-0,003 (0,003)				-0,003 (0,002)
Korr. R ²	0,1	0,22	0,69	0,71	0,05	0,21	0,8	0,84
Akaike	26,24	23,61	-9,19	-12,91	30,88	26,73	-23,67	-33,66
N	38	38	38	38	38	38	38	38

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%** , 99%***

Auf Basis der Regressionstabelle sind mehrere Effekte zu beobachten. Erstens steigen pro Iteration beide verwendeten Modellgütekriterien an. Das heißt, mit zunehmender Iterationszahl kann die



abhängige Variable besser vorhergesagt werden. Die aufgeklärte Varianz erreicht bis zu 84 % und erreicht dadurch eine ähnlich hohe Modellgüte wie das optimierte Modell.

Zweitens ändern sich die Werte der Koeffizienten je Iterationsschritt teilweise signifikant und beinhalten u. U. auch den Wechsel des Vorzeichens. Diese Änderung ist vor allem bei I3 sichtbar und, im Hinblick auf den Zusammenhang der spezifischen Variable zur Produktivität, in einigen Fällen plausibel (z. B. Speed Interaction), in anderen Fällen jedoch kontraintuitiv (z. B. Koordinationsaufwand).

Drittens ändern sich die Signifikanzniveaus. So ist bspw. in der zweiten Iteration ein statistisch signifikanter, negativer Einfluss auf die Produktivität sichtbar, wenn der korrespondierende ANSP nicht in einem FAB organisiert ist. Dies ist intuitiv plausibel, jedoch ändert sich in den weiteren Iterationsschritten sowohl das Vorzeichen als auch die statistische Signifikanz. Der entgegengesetzte Effekt tritt u. a. bei der Variable AIRP auf, welche erst ab Iteration 3 statistische Signifikanz aufweist.

Die Fluktuationen in den Koeffizienten und Signifikanzen weisen darauf hin, dass die Metrik der PRU zur Bestimmung der Produktivität maßgeblich von (teil)exogenen Faktoren abhängig ist. Eine Prädiktion der Produktivität basierend auf den endogenen Variablen ist nicht möglich. Dies wird auch durch die plausiblen Effekte gestützt. So ist in Iteration 1 und 2 die Arbeitszeit pro Lotse pro Jahr negativ konnotiert. Dies bedeutet, dass die Produktivität mit sinkender jährlicher Arbeitszeit steigen würde. Der gleiche Effekt ist für die Bereitstellung von Wetterdiensten, dem Betrieb von Flughäfen und der Betriebsform „Joint Stock Company“ zu beobachten.

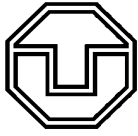
Der Wechsel des Faktoren-Vorzeichens zwischen zwei Iterationsschritten kann auf eine hohe Multikollinearität hinweisen (Auer und Rottmann 2015; Brunberg 2018). Dies konnte jedoch durch den VIF Test ausgeschlossen werden, welcher nach jeder Iteration durchgeführt wurde (vgl. Abschnitt 4.1). Im Falle eines Nachweises von Multikollinearität müsste das Regressionsmodell angepasst werden, z. B. indem der betreffende Faktor aus dem Modell entfernt wird. Alternative Methoden zur Überprüfung auf und Eliminierung von Multikollinearität werden u. a. von Albers et al. (2009) diskutiert.

4.2.4 Einfluss auf die DEA-Scores

4.2.4.1 Vorgehen

In Abschnitt 3.4 wurden mittels DEA Effizienz-Scores für vier Modelle ermittelt, welche die Leistungsfähigkeiten des jeweiligen ANSP beschreiben. Die Einflussgrößen auf die Effizienz werden, wie im vorigen Abschnitt, mittels Regressionsanalyse bestimmt und quantifiziert. Ziel ist es zu bestimmen, welche Faktoren einen Einfluss auf die Leistungsfähigkeit haben, wie groß der Einfluss ist und in welche Richtung (positiv oder negativ) sie wirken.

Als abhängige Variable werden die CRS-Scores eingesetzt, da diese Effizienzwerte keine Größeneffekte beinhalten. Die potentiellen Einflussfaktoren bilden wiederum die unabhängigen Variablen. Im Gegensatz zum vorigen Abschnitt soll die Prädiktion eines relativen Performance-Wertes erfolgen, welcher zwischen 0 und 1, bzw. 0 und 100 % beschränkt ist. Eine OLS Regression könnte hierfür nicht anwendbar sein, ein Regressionsmodell mit beschränkt abhängiger Variable ist vorzuziehen. Ob die Ergebnisse sich signifikant unterscheiden ist Gegenstand dieser Untersuchung.



Die ursprünglich 51 Faktoren wurden mittels PCA zu Hauptkomponenten zusammengefasst und durch Anwendung der Varimax-Rotation angepasst. Die Methodik wurde in die Statistiksoftware R implementiert. Über verschiedene Tests wurde die optimale Zahl an Komponenten bestimmt, jedoch führten die ermittelten Hauptkomponenten aufgrund der heterogenen Zusammensetzung nicht zu einer adäquaten Regressionsgüte (siehe Anhang C1). Auch die Erweiterung der Komponentenzahl brachte nur geringe Verbesserungen. Daher wurde die Regression mit den originären Variablen durchgeführt und mittels statistischer Tests, u. a. VIF-test, auf 21 Variablen reduziert.

Es ist üblich, die DEA-Scores vor der OLS-Regression logarithmisch zu transformieren (siehe u. a. Simar und Wilson (2007)). Damit soll erreicht werden, dass die Scores normalverteilt sind. Dies wird jedoch primär bei großen Datensätzen angewendet und ist für eine geringe Anzahl an Beobachtungen nicht sinnvoll. Auch im Kontext der ANSPs konnte nachgewiesen werden, dass ein Logarithmieren nicht zielführend ist (siehe Anhang, Abbildung A 6).

4.2.4.2 Maximierung der Modellgüte

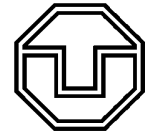
In einem ersten Schritt wird die Modellgüte für drei Regressionsmodelle (OLS-, Tobit- und Trunkated Regression) maximiert. Dafür werden die DEA-Scores von Modell 1, 2 und Modell 2A als abhängige Variable genutzt.

Für die OLS- und Tobit-Regressionsmodelle wurde das Statistikprogramm GRETL genutzt.¹¹⁷ Das Trunkated Regressionsmodell wurde mittels Anwendung der Software STATA analysiert, da das Modell weder in GRETL implementiert ist, noch zum Zeitpunkt der Analyse ein R-Paket zur Verfügung stand, welches eine zweiseitige Begrenzung berücksichtigen konnte. Das Programm berechnet jedoch weder das Akaike-Kriterium, noch das (korrigierte) R^2 . Die Modellgüte kann stattdessen über den Log-Likelihood Wert bestimmt werden (Kjærulff und Madsen 2012).

Tabelle 4-6 illustriert die Ergebnisse nach Variablenreduktion. Für die Reduktion wurde, analog zu den Analysen in Abschnitt 4.2.3, ein Schwellenwert von 0,33 (p-Wert) gesetzt. Bei Auswahl einer identischen Stichprobe an Faktoren ändert sich zwischen OLS- und Tobit-Modell lediglich die statistische Signifikanz und der Standardfehler, nicht jedoch die Koeffizienten. Abweichende Zusammensetzungen sind demnach bedingt durch den Schwellenwert von $p < 0,33$. Das heißt, dass die Koeffizienten zwar zwischen den DEA-Modellen, nicht jedoch zwischen den beiden Regressionstypen variieren. Dies ist dadurch begründet, dass innerhalb der Beobachtungen keine Daten zensiert werden (es gibt weder DEA-Werte größer 1, noch kleiner 0). Daher ändern sich Koeffizienten, Signifikanzen und Standardfehler lediglich bei der optimierten Regressionsformel für DEA-Modell 2A.

Die Ergebnisse der Trunkated Regression unterscheiden sich teilweise deutlich von jenen der beiden anderen Regressionsmodelle. Dies ist insbesondere bei DEA-Modell 2A sichtbar. Zu beobachten ist zudem, dass der FAB-Dummy in keiner der Regressions- und DEA-Modelle relevant ist. Weiterhin ist der Faktor „Oceanic“ lediglich in der Trunkated Regression für Model 2A relevant und signifikant. Umgekehrt sind die „Horizontal Interactions“ als Komplexitätsfaktor nur bei Model 2 im Falle von OLS

¹¹⁷ Die Ergebnisse wurden zusätzlich durch Programmierung und Anwendung eines R-Skripts überprüft.



und Tobit relevant und lediglich bei der zensierten Regression signifikant. Prinzipiell werden Komplexitätsfaktoren, mit Ausnahme der Speed Interactions, mehrheitlich nicht als relevant bzw. signifikant identifiziert. Dies ist in Übereinstimmung mit den Produktivitätskennziffern ein Hinweis für die geringe Eignung der PRU-Metrik.

Tabelle 4-6: Regressionsanalyse für maximierte Modellgüte

	OLS			Tobit			Trunkated		
	M1	M2	M2A	M1	M2	M2A	M1	M2	M2A
INT	2,416 (0,519)***	1,716 (0,182)***	-0,961 (1,504)	2,416 (0,421)***	1,716 (0,15)***	-1,181 (1,186)	2,609 (1,509)*	0,74 (0,366)**	0,235 (0,972)
NONA	-0,602 (0,199)***	-0,765 (0,221)***	-0,333 (0,26)	-0,602 (0,162)***	-0,765 (0,183)***	-0,39 (0,231)*	-1,184 (0,322)***	-0,524 (0,184)***	-0,722 (0,295)**
TIME (I)			0,176 (0,155)			0,192 (0,119)	0,497 (0,438)		0,497 (0,298)*
DELATM						-0,059 (0,048)			-0,06 (0,052)
MET	0,044 (0,041)			0,044 (0,033)			0,099 (0,049)**		0,043 (0,037)
AIRP	0,303 (0,107)***	0,18 (0,133)	0,293 (0,134)**	0,303 (0,087)***	0,18 (0,11)	0,287 (0,104)***	0,222 (0,117)*	0,183 (0,099)*	0,312 (0,065)***
JSC			0,079 (0,062)			0,084 (0,053)			
STATE	-0,078 (0,038)**	-0,116 (0,042)**	-0,073 (0,056)	-0,078 (0,031)**	-0,116 (0,035)***	-0,072 (0,046)	-0,128 (0,045)***	-0,1 (0,035)***	-0,124 (0,037)***
SIZE (I)	-0,036 (0,028)		0,063 (0,036)*	-0,036 (0,023)		0,074 (0,03)**	-0,26 (0,08)***	0,127 (0,05)**	
OCEAN									-0,46 (0,115)***
COORD		0,021 (0,012)*	0,018 (0,014)		0,021 (0,01)**	0,016 (0,012)			
L_AIRP		-0,084 (0,046)*			-0,084 (0,038)**		-0,161 (0,072)**	-0,05 (0,03)*	-0,112 (0,037)***
OVER	-0,547 (0,15)***		0,18 (0,164)	-0,547 (0,122)***		0,203 (0,146)	-1,051 (0,31)***		
DOM	-1,165 (0,398)***	-0,705 (0,446)	-1,16 (0,491)**	-1,165 (0,323)***	-0,705 (0,369)*	-1,273 (0,388)***		-0,817 (0,346)**	
GINI	-3,034 (0,661)***	-2,76 (0,628)***	-2,673 (0,793)***	-3,034 (0,536)***	-2,76 (0,52)***	-2,733 (0,635)***	-4,068 (0,927)***	-2,048 (0,543)***	-3,928 (0,537)***
DENSITY						0,009 (0,008)	0,021 (0,016)		0,011 (0,008)
VI	0,54 (0,392)			0,54 (0,318)*		0,403 (0,375)			
HI		-0,394 (0,285)			-0,394 (0,235)*				
SI	-1,07 (0,48)**	-1,049 (0,334)***	-0,736 (0,483)	-1,07 (0,389)***	-1,049 (0,276)***	-0,866 (0,396)**	-1,446 (0,626)**	-0,727 (0,303)**	-1,582 (0,331)***
COSTS	0,001 (0,001)	0,001 (0,001)	0,002 (0,001)***	0,001 (0)*	0,001 (0,001)	0,001 (0,001)**	0,002 (0,001)**		0,002 (0,001)***
RES	-0,004 (0,001)***	-0,003 (0,002)**	-0,003 (0,002)*	-0,004 (0,001)***	-0,003 (0,001)***	-0,003 (0,001)*	-0,011 (0,003)***	-0,003 (0,001)**	-0,007 (0,002)***
Korr. R ²	0,71	0,6	0,65						
Akaike	-52,05	-46,65	-44,20	-50,05	-44,64	-38,75			
Log-Lik.	39,02	35,32	36,10	39,02	35,32	37,38	36,06	34,75	35,55
N	38	38	37	38	38	37	38	38	37

Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***

Dem gegenüber sind einige Faktoren in jedem Modell relevant und teilweise auch signifikant. Diese Faktoren haben zudem modellübergreifend dasselbe Vorzeichen, sodass die Wirkung als robust eingestuft werden kann. Negativ wirken sich ein geringerer Anteil an Lotsen am Gesamtpersonal, staatliche Kontrolle des ANSP (i.S.v. Nicht-Privatisierung), eine zeitliche Ungleichverteilung des Verkehrs, die Speed Interactions sowie der Technologie-Proxy aus. Letzteres ist kontraintuitiv, jedoch lässt der



niedrige Wert des Koeffizienten den Schluss zu, dass der Einfluss sehr gering ist. Stets positiv wirken sich der Betrieb von Flughäfen und die Kosten pro Lotsenstunde aus, wengleich für den zweitgenannten Faktor nur in acht von neun Fällen Relevanz nachgewiesen werden konnte.

Gemäß des Korrigierten R^2 kann der höchste Anteil an Varianz bei Modell 1 aufgeklärt werden, gefolgt von Modell 2A und Modell 2. Unter Bezugnahme auf das Akaike-Kriterium ist jedoch die Modellgüte von Modell 2 höher als von Modell 2A. Dabei erreichen die OLS-Regressionen generell einen besseren Fit als die korrespondierenden Tobit-Modelle. Die Ergebnisse werden durch den Log-Likelihood-Wert bestätigt, welcher für Modell 1 jeweils am größten ist. Die Modellgüte ist für die Trunkated Modelle geringer als die korrespondierenden OLS- und Tobit-Modelle.

Die Modellgüte steigt unter der Inkludierung von WEALTH an, jedoch konnte mittels VIF-Test Multikollinearität nachgewiesen werden. Des Weiteren sind im Falle einer Substitution der NOFAB Variable durch die einzelnen FABs die jeweiligen Variablen signifikant. Allerdings sank durch den Austausch die Modellgüte und lieferte unplausible Ergebnisse hinsichtlich der Koeffizienten. Dies könnte darin begründet liegen, dass die Zusammensetzung der FABs unterschiedlich ist. So bestehen einige FABs aus lediglich zwei ANSPs, andere setzen sich aus fünf Unternehmen zusammen. Bei Modell 2A ist die Plausibilität am höchsten. Einzige Ausnahme ist die bereits diskutierte Technologie-Variable. In Modell 1 wird bspw. der Anteil der Überflüge negativ konnotiert. Bei Modell 2 wiederum ist die Anzahl an signifikanten Variablen am geringsten.

4.2.4.3 Stufenweiser Aufbau

Analog zu den Produktivitätsbetrachtungen soll in diesem Abschnitt der Einfluss der einzelnen Variablengruppen untersucht werden. In Abschnitt 4.2.4.2 unterscheiden sich die Ergebnisse von OLS- und Tobit-Modellen lediglich bei Signifikanz und Standardfehler, jedoch nicht bei Ausprägung und Vorzeichen der unabhängigen Variablen. Daher werden im weiteren Verlauf lediglich Tobit- und Trunkated Regression verglichen. Anhang Tabelle A 6 stellt die Maximal-Modelle (I4) für die Trunkated Regression dar. Die höchste Modellgüte wird dabei für Modell 2A berechnet, die geringste für Modell 2. Die Basis der Analyse bildet daher das Modell 2A. Tabelle 4-7 fasst die Ergebnisse der einzelnen Iterationen zusammen.

Im Allgemeinen sind die Effekte ähnlich zu den in Abschnitt 4.2.3 diskutierten Einflüssen auf die Produktivität:

- Mit zunehmender Iteration verbessert sich die Modellgüte. Ausnahme: Log-Likelihood-Kriterium bei der Tobit-Regression in I3.
- Die Koeffizienten und teilweise auch deren Vorzeichen variieren in einigen Fällen.
- Die Signifikanzniveaus ändern sich.

Die zwischeniterativen Änderungen (z. B. Vorzeichenwechsel) sind beim Trunkated Modell stärker ausgeprägt als beim Tobit-Modell (siehe beispielhaft den Faktor „Oceanic“). Mit zunehmendem Iterationsschritt steigt zudem die Plausibilität der Ergebnisse. So ist in I1 die Interdependenz zwischen dem Faktor „Arbeitszeit pro Lotsen“ und dem DEA-Score negativ.

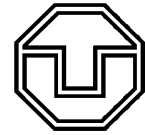
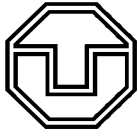


Tabelle 4-7: Tobit und Trunkated Regression für Modell 2A, 2016

	Tobit				Trunkated			
	I1	I2	I3	I4	I1	I2	I3	I4
INT	1,372 (1,303)	-0,752 (1,313)	-2,786 (1,466)*	-1,372 (1,581)	1,712 (2,186)	-0,473 (1,589)	-4,055 (1,821)**	-1,361 (1,553)
NONA	-0,576 (0,25)**	-0,021 (0,238)	-0,281 (0,231)	-0,405 (0,241)*	-0,502 (0,408)	0,22 (0,317)	-0,244 (0,391)	-0,949 (0,388)**
TIME (I)	-0,019 (0,177)	0,097 (0,165)	0,299 (0,141)**	0,197 (0,144)	-0,172 (0,687)	0,025 (0,468)	0,886 (0,447)**	0,725 (0,35)**
DELATM	0,008 (0,074)	-0,043 (0,06)	-0,112 (0,05)**	-0,059 (0,052)	0,025 (0,116)	-0,028 (0,081)	-0,115 (0,067)*	-0,083 (0,053)
MET	-0,129 (0,058)**	-0,071 (0,052)	-0,008 (0,042)	-0,01 (0,04)	-0,095 (0,083)	-0,04 (0,057)	0,017 (0,049)	0,037 (0,038)
AIRP	-0,091 (0,09)	-0,113 (0,078)	0,241 (0,11)**	0,308 (0,116)***	-0,032 (0,133)	-0,001 (0,103)	0,307 (0,152)**	0,435 (0,125)***
JSC	-0,044 (0,072)	0,026 (0,067)	0,099 (0,057)*	0,084 (0,054)	-0,074 (0,112)	0,094 (0,085)	0,129 (0,081)	0,047 (0,065)
STATE	-0,097 (0,072)	-0,054 (0,062)	-0,073 (0,052)	-0,069 (0,051)	-0,099 (0,111)	0,016 (0,079)	-0,051 (0,081)	-0,115 (0,065)*
SIZE (I)		0,073 (0,028)***	0,118 (0,04)***	0,08 (0,041)*		0,192 (0,08)**	0,35 (0,113)***	0,145 (0,102)
OCEAN		0,012 (0,089)	0,041 (0,073)	0,043 (0,069)		-0,228 (0,19)	0,005 (0,283)	-0,196 (0,225)
COORD		0,001 (0,014)	0,008 (0,015)	0,017 (0,015)		-0,004 (0,018)	-0,007 (0,02)	0,001 (0,016)
L_AIRP		0,021 (0,04)	0,03 (0,047)	0,007 (0,047)		0,065 (0,052)	0,053 (0,076)	-0,082 (0,066)
NOFAB		-0,209 (0,069)***	0,015 (0,083)	0,02 (0,082)		-0,245 (0,074)***	0,058 (0,106)	0,1 (0,084)
OVER			0,294 (0,18)	0,228 (0,181)			0,259 (0,231)	0,112 (0,184)
DOM			-1,408 (0,427)***	-1,38 (0,434)***			-1,432 (0,985)	-1,16 (0,785)
GINI			-2,435 (0,679)***	-2,678 (0,672)***			-2,021 (1,024)**	-3,878 (0,927)***
DENSITY			0,016 (0,009)*	0,009 (0,01)			0,018 (0,012)	0,019 (0,011)*
VI			0,47 (0,413)	0,289 (0,397)			-0,304 (0,73)	0,006 (0,583)
HI			0,388 (0,295)	0,146 (0,305)			0,862 (0,405)**	0,286 (0,365)
SI			-0,599 (0,507)	-0,718 (0,484)			-0,086 (0,817)	-1,137 (0,693)
COSTS				0,001 (0,001)*				0,002 (0,001)*
RES				-0,003 (0,001)*				-0,007 (0,002)***
Akaike	-11,74	-15,73	-19,67	-29,42				
Log-Lik.	15,87	29,87	23,83	37,71	14,75	23,63	30,65	37,39
N	37	37	37	37	37	37	37	37

Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***

Im Allgemeinen sind die Ergebnisse der Schritte 3 und 4 plausibel, insbesondere im Vergleich zu den ersten beiden Iterationsstufen. Ausnahmen bilden wiederum der Technologie-Proxy sowie der FAB-Dummy. *NOFAB* ist in den beiden letzten Iterationen positiv bewertet, in I2 jedoch negativ. Betrachtet man überwiegend endogene Faktoren, so ist eine Kooperation innerhalb eines FABs positiv, bei Inkludierung weiterer exogener Variablen schwächt sich dieser Effekt ab bzw. kehrt sich um. Da dieser Effekt jedoch statistisch nicht signifikant ist (im Gegensatz zur negativen Konnotation in I2), ist das Ergebnis insgesamt ebenfalls plausibel. Modell- und iterationsübergreifend robust ist hinsichtlich Koeffizienten und Signifikanz vor allem der GINI-Koeffizient. Dies gilt auch für die Nutzung der verbleibenden DEA-Modelle.



4.3 Panel-Analyse

4.3.1 Daten und Modellierung

Die Ergebnisse der Abschnitte 3.4, 4.2.3 und 4.2.4 basieren auf der Betrachtung eines spezifischen Jahres. Bindet man mehrerer Zeiträume in die Effizienzbewertung ein, ist die Produktionsfunktion nicht konstant, sondern verschiebt sich aufwärts¹¹⁸ oder abwärts¹¹⁹. Daraus folgend können sich die radialen Abstände zwischen DMUs und Produktionsgrenzen mit der Zeit ändern (Coelli et al. 2005).

Zur Durchführung einer Analyse, welche mehrere Zeitperioden umfasst, sind Paneldaten erforderlich. Die PRU veröffentlicht Daten seit dem Jahr 2000. Der Umfang hinsichtlich der Anzahl der einbezogenen ANSPs und der abgedeckten Parameter ist jedoch unterschiedlich. In der Folge gibt es einen Trade-off zwischen Zeitspanne und geografischem Umfang. Für diese Untersuchung wurde daher der Zeitraum zwischen 2003 und 2016¹²⁰ gewählt. Die für diese Untersuchung verwendete Software ist DEAP 2.1 (Coelli 2008).

Die Daten können aufgrund der inhomogenen Erfassungsmethodik bis 2007 einige Verzerrungen aufweisen (Standfuss et al. 2016a). Dies muss bei der Formulierung der DEA-Modelle berücksichtigt werden, indem man lediglich Parameter verwendet, welche keinen signifikanten Ungenauigkeiten unterliegen, wie z. B. Anzahl der ACCs, der kontrollierten Gesamtflugstunden oder die Luftraumgröße.

Auch die europäische Heterogenität hinsichtlich der angebotenen Dienstleistungen muss bei der ökonomischen Modellierung berücksichtigt werden. So ist Maastricht Upper Airspace Control nur im oberen Luftraum aktiv. In Abschnitt 3.4 wurde nachgewiesen, dass diese DMU die Produktionsgrenze erheblich verschiebt. Es bietet sich daher an, eine disaggregierte Malmquist-Analyse nach Art der Dienstleistung (Enroute versus Terminal) vorzunehmen. Um lokale Effekte zu bestimmen, ist zudem eine mikroskopische Analyse einzelner ANSPs sinnvoll.

Für die Gate-to-Gate Betrachtung¹²¹ werden die in Abschnitt 3.4.1 diskutierten Modelle verwendet. Da die PRU ihre Daten nach Enroute und Terminal Services aufschlüsselt, können individuelle Modelle für beide Services aufgestellt werden (siehe Anhang Tabelle A 8 und Tabelle A 9). Da die Ergebnisse jedoch aufgrund der abweichenden Inputs und Outputs (disaggregierte Maße) nicht vergleichbar sind, wird hier lediglich die Gate-to-Gate Betrachtung adressiert.

¹¹⁸ Z. B. durch technischen Fortschritt.

¹¹⁹ Z. B. durch eine wirtschaftliche Rezession.

¹²⁰ Zum Zeitpunkt der Untersuchung waren Daten bis einschließlich 2016 verfügbar.

¹²¹ Umfasst Enroute- und Terminalbetrieb.

4.3.2 Methodik

Die zeitlichen Änderungen der Produktivität bzw. Effizienz werden mit dem sogenannten Malmquist-Index (MI) berechnet. Er beruht auf dem Quotienten der TFP, respektive Distanzfunktionen¹²², zweier Zeitpunkte (t und t+1) und kann, wie die Effizienzanalyse, input- oder outputorientiert sein. Beispielfhaft wird im Folgenden die Berechnung unter Outputorientierung dargestellt. Abbildung 4-1 stellt die Änderung der Technologie T für zwei Zeitperioden dar.

Unter der Annahme konstanter Skalenerträge wird die Veränderung der TFP bzw. Effizienz jeweils auf Basis von Periode 1 und Periode 2 getrennt bestimmt. Diese beiden Indizes sind nur in Ausnahmefällen gleich groß (Färe et al. 1994). Der Malmquist-Index bildet aus beiden Werten das geometrische Mittel. Dadurch soll eine willkürliche Festlegung der Referenztechnologie, also Periode 1 oder 2, vermieden werden (Cantner et al. 2007).

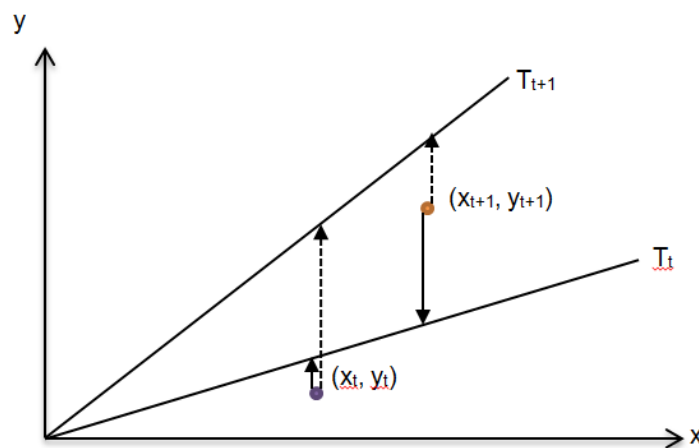


Abbildung 4-1: Totale Faktorproduktivität
(Fried et al. 2008)

Formel (4-5) illustriert die Berechnung der TFP-Änderungsrate zwischen zwei Zeitperioden t und t+1 in Abhängigkeit des Inputvektors x und des Outputvektors y. D steht für die Distanzfunktion zwischen dem beobachteten Unternehmen zum Benchmark der Technologie T der jeweiligen Periode. Aus beiden Quotienten wird das geometrische Mittel gebildet. Ist der berechnete Index M größer als 1, so ist ein Wachstum der Produktivität zu verzeichnen, bei gleich 1 Stagnation und bei < 1 eine Rückgang (Fried et al. 2008).

$$M(x_t, y_t, x_{t+1}, y_{t+1}) = \left(\frac{D_t(x_{t+1}, y_{t+1})}{D_t(x_t, y_t)} \times \frac{D_{t+1}(x_{t+1}, y_{t+1})}{D_{t+1}(x_t, y_t)} \right)^{\frac{1}{2}} \quad (4-5)$$

Entsprechend der diskutierten Effizienzmaße, welche in den vorigen Abschnitten beschrieben wurden, kann auch der Malmquist-Index durch Umformungen in einzelne Komponenten zerlegt werden, um die Ursachen der Produktivitätssteigerung oder –senkung zu analysieren. Durch Umstellung von Formel (4-5) kann der Malmquist-Index in zwei ökonomisch interpretierbare Komponenten aufgesplittet

¹²² Distanz zwischen Unternehmen und dem effizienten Rand (Benchmarking Technologie).



werden (4-6). Der linke Term entspricht der Änderung der technischen Effizienz in Periode 2. Jener Teil, welcher nicht auf diese Änderung der technischen Effizienz zurückzuführen ist, wird im rechten Term zusammengefasst und als technischer Fortschritt bezeichnet.

$$M(x_t, y_t, x_{t+1}, y_{t+1}) = \frac{D_{t+1}(x_{t+1}, y_{t+1})}{D_t(x_t, y_t)} \times \left(\frac{D_t(x_t, y_t)}{D_{t+1}(x_t, y_t)} \times \frac{D_t(x_{t+1}, y_{t+1})}{D_{t+1}(x_{t+1}, y_{t+1})} \right)^{\frac{1}{2}} \quad (4-6)$$

Produktivitätssteigerungen zwischen zwei Perioden können gemäß der Skalentheorie auch durch eine Anpassung der Unternehmensgröße erzielt werden. Um diese Prozesse abzubilden, ist die Änderung der Skaleneffizienz zu berechnen. Hierfür werden die Distanzen zur VRS Produktionsfunktion benötigt. Bei der Implikation konvexer RPF ist die Reziprozität nicht gewährleistet, sodass die Effizienzwerte abweichen können (Cantner et al. 2007). Coelli et al. (2005) zeigten jedoch, dass eine Zerlegung der Änderungsrate der technischen Effizienz möglich ist, solange diese durch Annahme variabler Skalenerträge errechnet wurde.

Die Änderung der Skaleneffizienz (SE) kann durch die Umformung des Effizienz-Terms von Formel (4-6) abgeleitet werden. Formel (4-7) beschreibt diese Zerlegung in die Änderungskomponenten der reinen technischen Effizienz (PE, links), welche relativ zu der VRS Grenzproduktionsfunktion gemessen wird, sowie der Skaleneffizienz (rechts). Technologien mit variablen Skalenerträgen werden mit dem Index ‚vrs‘ versehen, jene mit konstanten Skalenerträgen mit dem Index ‚crs‘. Der rechte Term kann als geometrisches Mittel von zwei Änderungen der Skaleneffizienz interpretiert werden, einmal auf Basis der Benchmarks in Periode 1 und einmal auf Basis des Benchmarks in Periode 2.

$$\frac{D_{t+1}(x_{t+1}, y_{t+1})}{D_t(x_t, y_t)} = \frac{D_{t+1}^{vrs}(x_{t+1}, y_{t+1})}{D_t^{vrs}(x_t, y_t)} \times \left(\frac{D_{t+1}^{vrs}(x_{t+1}, y_{t+1})}{D_{t+1}^{crs}(x_{t+1}, y_{t+1})} \times \frac{D_t^{vrs}(x_{t+1}, y_{t+1})}{D_t^{crs}(x_{t+1}, y_{t+1})} \right)^{\frac{1}{2}} \quad (4-7)$$

4.3.3 Ergebnisse

Die Malmquist DEA ermöglicht eine Analyse der Änderung von TFP sowie deren Komponenten für die Jahre 2004 bis 2016. Die durchschnittlichen Wachstumsraten sind in Tabelle 4-8 dargestellt. Dabei wird das ungewichtete, geometrische Mittel über alle ANSPs und alle Jahre genutzt. Wie in Tabelle 4-8 dargestellt liegt das durchschnittliche jährliche Wachstum zwischen 0,6 % und 0,8 %. Lediglich Modell 2A zeigt ein größeres jährliches Wachstum. Dies ist hauptsächlich auf den Ausschluss von MUAC zurückzuführen, da diese Einheit innerhalb der DEA die technologische Grenze verschiebt. Dies unterstützt auch das Argument, MUAC beim Benchmarking mit Gate-to-Gate-Betrachtungen auszuschließen. Diese These wird zudem bei der zeitlichen Differenzierung der Wachstumsraten unterstützt. Die Entwicklung der TFP Wachstumsraten ist modellvergleichend in Abbildung 4-2 dargestellt. Eine Gewichtung der einzelnen ANSPs findet dabei nicht statt. Es ist zu sehen, dass die Verläufe der TFP-Kurven nur geringfügig voneinander abweichen. Lediglich die grüne Kurve, welche Modell 2A repräsentiert, weist eine größere Abweichung auf.

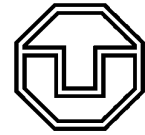


Tabelle 4-8: Durchschnittliche Wachstumsraten, Gate-to-Gate

	TE	TF	PE	SE	TFP
M1	1,009	0,999	1,004	1,005	1,008
M2	0,995	1,011	0,997	0,997	1,006
M2A	1,005	1,009	1,001	1,004	1,014
M2B	1,001	1,005	1,000	1,000	1,006

Auch ein drittes Argument unterstützt die These, MUAC bei Gate-to-Gate Analysen auszuschließen. Vergleicht man die TFP-Änderung einzelner ANSPs, so ist die Streuung zwischen den Modellen gering (durchschnittlich 0,8 %). Nur einige ANSPs, wie die slowakische LPS und das niederländische LVNL, weisen eine Abweichung von mehr als 1 % zwischen den vier Modellen auf. Andere ANSPs zeigen keine oder nur marginale Streuung (<0,2 %), z. B. BULATSA und MUAC. Auf einer jährlichen Basis zeigen die TFP-Änderungen von 2005, 2008 und 2015 einige erhebliche Abweichungen zwischen den Modellen. In einem Fall wurde dies hauptsächlich durch das Modell 4A verursacht, das auf ein kleines Wachstum für das Jahr 2008 hindeutet, während die anderen drei Modelle einen Rückgang der TFP auswiesen.

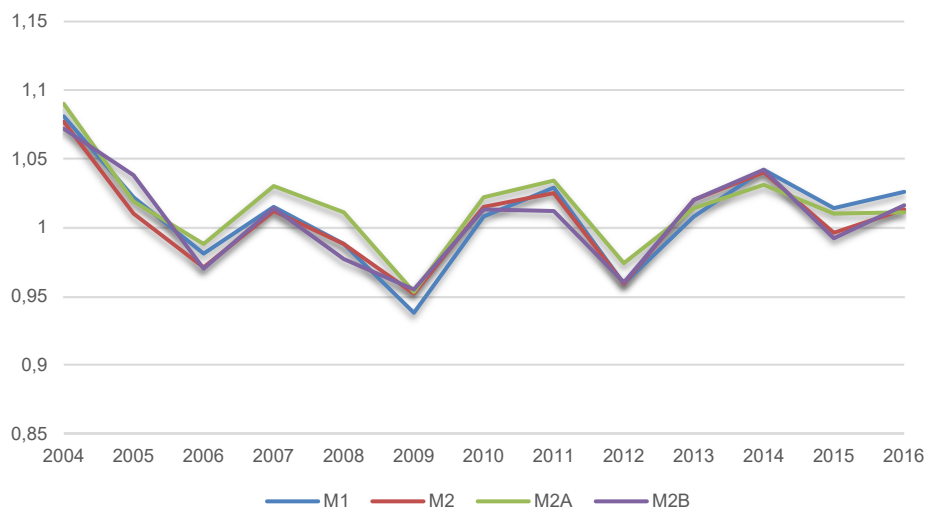
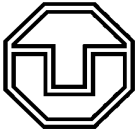


Abbildung 4-2: Modellvergleich TFP Wachstum

Ausgehend von der Sonderstellung von MUAC wird im Folgenden Modell 2A genauer betrachtet. Zunächst wird die jährliche Entwicklung der TFP untersucht. Abbildung 4-2 zeigt, dass lediglich in drei Jahren ein Rückgang der Indizes zu verzeichnen war (grüne Linie). Da die Daten einen hohen Aggregationslevel aufweisen, sind die Gründe dafür in globalen Ereignissen zu suchen. Der höchste Abschwung fand 2009 statt (-5 %), verursacht durch die 2008 beginnende Wirtschaftskrise. Durch die Zerlegung der Kurve in die diskutierten Teilkomponenten konnte festgestellt werden, dass der technische Fortschritt maßgeblich dieses Ergebnis beeinflusst, die Senkung der TFP also nicht durch eine technische Ineffizienz hervorgerufen wird. Ein weiteres negatives Wachstum war im Jahr 2012 zu verzeichnen, wobei hier der Abschwung maßgeblich durch Effizienzverluste verursacht wird. Zu beachten ist zudem, dass im Jahr nach jedem Rückgang, auch 2006/2007, ein signifikantes Wachstum des TFP ausgewiesen wird.



Im Allgemeinen ist für die Mehrheit der Flugsicherungsdienstleister ein Wachstum des TFP zu beobachten. Die größten Produktivitätsgewinne verzeichnen die maltesische MATS, LGS aus Lettland und BULATSA aus Bulgarien. Nur neun Unternehmen sind von einem negativen Wachstum betroffen, vor allem Avinor (Norwegen, -1,3%), HCAA (Griechenland, -1,8%) und LVNL (Niederlande, -2,4%). Dänemark erzielt eine durchschnittliche TFP-Änderung von 0%.

Die räumliche Verteilung der Malmquist-Indizes ist in Abbildung 4-3 dargestellt. Sie zeigt, dass insbesondere die südosteuropäischen Staaten ein hohes Wachstum verzeichnen können. Aufgrund der politischen Krise in der Ostukraine kommt es zu einer Verkehrsverlagerung in den Süden. Insbesondere in Bulgarien steigen die kontrollierten Flugstunden im betrachteten Zeitraum um mehr als das Doppelte. Ein weiterer möglicher Grund ist der zunehmende Verkehr zwischen Europa und den Golfstaaten. In der Folge profitieren einige Balkanstaaten und die Türkei von steigenden Verkehrszahlen. Der Abschwung der französischen DSNAs könnte mit umfangreichen Streikaktivitäten gekoppelt sein. Dagegen steigt die Produktivität in allen Nachbar-ANSPs (ausgenommen der schweizerischen skyguide) zwischen 1% und 2%. Dies könnte auch durch die daraus folgenden Ausweichverkehre begründet sein.

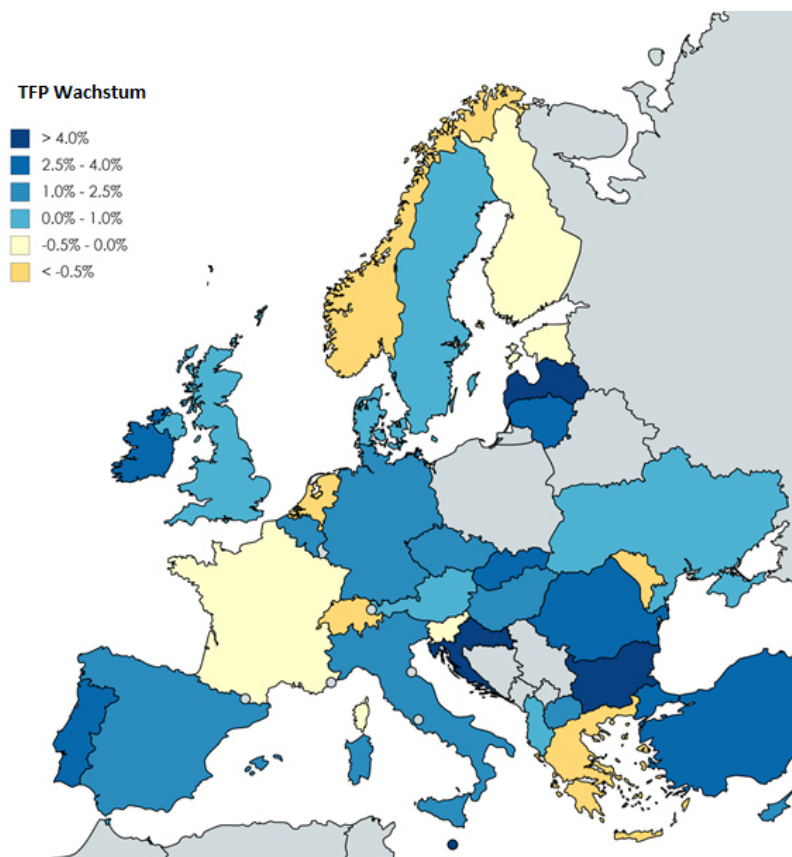
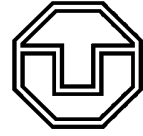


Abbildung 4-3: Durchschnittliche Wachstumsrate einzelner ANSPs (Modell 2A)

Die Ergebnisse sind jedoch teilweise kontraintuitiv, bzw. schwer interpretierbar. So überrascht es bspw., dass auch die Ukraine trotz der politischen Krise ein durchschnittliches jährliches Wachstum von 0,5% erreicht. Dies kann jedoch darin begründet liegen, dass der politische Konflikt erst im Jahr 2013 begann und daher nur drei der betrachteten Jahre umfasst. Es fällt außerdem auf, dass Nachbarländer teilweise stark unterschiedliche Entwicklungen im TFP aufweisen. Dies ist z. B. bei Estland und Lettland,



Belgien und den Niederlanden sowie Norwegen und Schweden der Fall. Hierfür gibt es auf diesem Aggregationslevel keine plausiblen Erklärungen. Eine mögliche Ursache ist, dass die Abbildung lediglich das geometrische Mittel der einzelnen Länder illustriert. In einigen Fällen streuen diese Wachstumsraten jedoch im Betrachtungszeitraum, was sowohl auf globale als auch lokale Effekte zurückgeführt werden könnte (siehe Anhang C5). Hier kann eine Regressionsanalyse helfen, effizienzbeeinflussende Faktoren zu identifizieren und quantifizieren.

4.3.4 Regression

4.3.4.1 Ansatz

Um den Einfluss von Faktoren über eine zeitlich sich verändernde Beobachtungsreihe zu analysieren, stehen im Wesentlichen drei Methoden zur Verfügung: Pooled-OLS-, Fixed-Effects- und Random-Effects-Regression. Für die Auswahl des geeigneten Modelles dienen verschiedene Tests. Es ist gängige Praxis, zunächst eine Pooled-OLS Regression durchzuführen. Über die gemeinsame Signifikanz der verschiedenen Gruppenmittel kann die Eignung eines OLS-Modells bestimmt werden. Wird die Nullhypothese (das gepoolte KQ-Modell ist angemessen) abgelehnt, wird ein Fixed-Effects Modell bevorzugt. Der Breusch-Pagan-Test untersucht ebenfalls, ob die Pooled-OLS-Regression angemessen ist. Die Alternativhypothese besagt, dass ein Random Effects Modell angewendet werden sollte (Albers et al. 2009; Auer und Rottmann 2015). Eine Entscheidung hinsichtlich der Verwendung zwischen Fixed- und Random-Effects ermöglicht der Hausman-Test. Die Nullhypothese besagt, dass ein Random Effects Modell konsistent ist. Muss diese abgelehnt werden, ist ein Fixed-Effects Modell zu bevorzugen (Albers et al. 2009; Auer und Rottmann 2015; Springer 2014). Die Tests¹²³ sind in der genutzten Software GRETL implementiert.

Ein alternatives Vorgehen wurde von Clark und Linzer (2015) vorgeschlagen. Die Autoren weisen darauf hin, dass die Auswahl nach Größe des Datensatzes getroffen werden soll. Ein weiteres Entscheidungskriterium ist, ob die Varianz der Werte primär in den Regressoren oder zwischen den einzelnen Zeiteinheiten beobachtet werden kann. Im ersten Fall sollte ein Random-Effects Modell lediglich dann angewendet werden, wenn Vorhersagen über unbekannte Einheiten getroffen werden sollen. Im zweiten Fall ist ein Fixed Effects Modell nur dann zu bevorzugen, wenn es eine hohe Anzahl an Beobachtungen (≥ 10) und Zeiteinheiten (≥ 5) gibt, oder wenn die Korrelation zwischen Regressoren und den Zeiteffekten hoch ist.

Für die Panel-Analyse wurden die in Abschnitt 4.2.2 diskutierten Faktoren verwendet. Diese wurden wie folgt angepasst:

- Es werden lediglich die Jahre 2011-2016 berücksichtigt.
- Für die sozio-ökonomischen Variablen von MUAC wird der niederländische Wert verwendet.
- In einigen Ländern sind die Patentzahlen nicht für den gesamten Zeitraum verfügbar (u. a. Albanien). Die fehlenden Werte wurden daher inter- oder extrapoliert.

¹²³ Test auf Multikollinearität und Panel-Diagnostik (Hausman-Test u. a.).



- Variablen mit großen Ausprägungen (z. B. Größe des Luftraumes) wurden vor der Regression logarithmiert.

Für die unabhängigen Variablen wurde vorab einer Korrelationsanalyse (siehe Anhang Abbildung A 12) durchgeführt. Das Panel-Modell wurde dabei schrittweise ausgebaut. In der ersten Iteration werden nur endogene Faktoren berücksichtigt. Die zweite Iteration schließt die Jahres-Dummys mit ein. Iteration 3 und 4 umfassen Luftraumdaten sowie Nachfragecharakteristiken. In der letzten Iteration werden alle Faktoren berücksichtigt. Die Iterationen spiegeln das Maß an Einfluss des ANSP wider, da einige Variablen als teil-exogen betrachtet werden können.

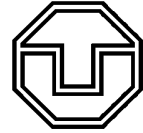
Nach jedem Zwischenergebnis wird ein VIF-Test durchgeführt, sodass Multikollinearität vermieden wird. Zudem werden die Modellgüte-Kriterien angegeben und über die Paneldiagnostik das bevorzugte Regressionsmodell ausgewählt. Hier kann es dazu kommen, dass je nach Iterationsschritt das bevorzugte Modell variiert. In diesem Fall wird das im Iterationsschritt 5 bevorzugte Modell angegeben. Als abhängige Variablen wurden sowohl die im ACE-Report verwendete Produktivität als auch die Efficiency-Scores der vier DEA-Modelle eingesetzt. Bei der Ergebnispräsentation wird die Pooled Regression nicht berücksichtigt.

4.3.4.2 Einfluss auf Produktivität

In einem ersten Schritt wird das Panelmodell für die Produktivitätswerte der PRU-Metrik (Composite Flight Hours per ATCO hour) geschätzt. Variablen für die durchschnittliche Flugstrecke innerhalb der Einheit, BIP sowie Relevanz der Einheit wurden getestet, jedoch aufgrund von Multikollinearität eliminiert. Gleiches gilt für den Jahres-Dummy für 2011. Tabelle 4-9 fasst die Ergebnisse bei Fixed- und Random Effects zusammen. Zudem sind die Modellgütekriterien dargestellt. Je nach Iterationsschritt können bis zu 92% der Varianz aufgeklärt werden (FEM). Insbesondere die Inkludierung der Nachfragecharakteristiken erhöht die Modellgüte beträchtlich. Dies ist modellübergreifend zu beobachten. Einige Effekte sind somit ähnlich zur uniperiodischen Betrachtung.

Die Ergebnisse der Panelanalyse bestätigen in den meisten Fällen die Erkenntnisse aus Abschnitt 4.2. Eine signifikante Ausnahme bilden hier die Faktoren *NONA* und *TIME*. Überraschend ist auch, dass alle Zeit-Dummys negativ bewertet sind. Dagegen sind u. a. die positiven Interdependenzen von *SIZE* sowie die negativen Auswirkungen von *NOFAB* und *STATE* entsprechend dem uniperiodischen Analyseergebnis. Eine wesentliche Änderung ist auch die positive Konnotation des Technologie-Proxys. Sowohl auf Basis des Hausman-Tests als auch der durch Clark und Linzer (2015) vorgeschlagenen Auswahlkriterien wird FEM gegenüber dem REM bevorzugt. Auch ist im Gesamtmodell die Modellgüte bei der Annahme fixer Effekte am höchsten.

Die Ergebnisse sind hinsichtlich der Luftraumcharakteristiken plausibel. Lediglich die positive Interdependenz zwischen Produktivität und Koordinierungsaufwand entspricht nicht der erwarteten Wirkung. Dagegen sind die Ergebnisse der Nachfrageeinflüsse plausibel und in Iteration 5 zudem alle signifikant. Gleiches gilt für die exogenen Variablen. Lediglich der fehlende Einfluss des BIP-Wachstums ist zunächst überraschend. Möglicherweise ist dies jedoch darauf zurückzuführen, dass die abhängige



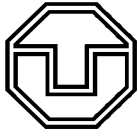
Variable keine Änderung zum Vorjahr, sondern eine für das Jahr spezifische Größe repräsentiert. Diese Hypothese kann anhand der Regression der Malmquist-Indizes überprüft werden.

Tabelle 4-9: FEM und REM für Produktivität gemäß PRU

	Fixed Effects					Random Effects				
	I1	I2	I3	I4	I5	I1	I2	I3	I4	I5
INT	4,732 (1,087)***	5,789 (0,993)***	5,049 (1,386)***	1,813 (1,055)*	-2,508 (0,868)***	4,56 (0,924)***	5,42 (0,845)***	4,61 (1,211)***	1,927 (0,933)**	-1,686 (0,843)**
NONA	-0,814 (0,208)***	-1,02 (0,19)***	-0,674 (0,263)**	-0,374 (0,189)*	0,771 (0,165)***	-0,838 (0,178)***	-1,002 (0,163)***	-0,675 (0,221)***	-0,39 (0,155)**	0,549 (0,157)***
TIME (I)	-0,451 (0,146)***	-0,558 (0,133)***	-0,556 (0,163)***	-0,167 (0,117)	-0,021 (0,09)	-0,425 (0,125)***	-0,51 (0,113)***	-0,496 (0,143)***	-0,197 (0,103)*	-0,02 (0,089)
DELATM	-0,200 (0,06)***	-0,21 (0,054)***	-0,231 (0,055)***	-0,19 (0,032)***	-0,201 (0,028)***	-0,174 (0,051)***	-0,194 (0,046)***	-0,185 (0,048)***	-0,19 (0,028)***	-0,169 (0,026)***
MET	0,043 (0,051)	0,069 (0,046)	0,151 (0,056)***	0,101 (0,033)***	0,018 (0,024)	0,019 (0,042)	0,054 (0,038)	0,101 (0,045)**	0,097 (0,028)***	0,021 (0,023)
AIRP	-0,067 (0,071)	0,011 (0,065)	0,019 (0,075)	0,419 (0,091)***	0,165 (0,073)**	-0,055 (0,064)	0,018 (0,059)	0,03 (0,07)	0,409 (0,078)***	0,202 (0,073)***
JSC	-0,058 (0,068)	-0,141 (0,063)**	-0,2 (0,081)**	0,084 (0,065)	0,158 (0,047)***	-0,065 (0,058)	-0,128 (0,053)**	-0,173 (0,072)**	0,056 (0,056)	0,109 (0,045)**
STATE	-0,205 (0,067)***	-0,193 (0,061)***	-0,231 (0,074)***	-0,012 (0,052)	-0,015 (0,036)	-0,196 (0,057)***	-0,189 (0,051)***	-0,218 (0,063)***	-0,061 (0,043)	-0,057 (0,034)
Y2012		-0,281 (0,068)***	-0,223 (0,074)***	-0,11 (0,044)**	-0,014 (0,033)		-0,276 (0,063)***	-0,213 (0,068)***	-0,106 (0,041)***	-0,02 (0,035)
Y2013		-0,013 (0,066)	-0,012 (0,064)	-0,016 (0,035)	-0,007 (0,024)		-0,013 (0,061)	-0,013 (0,06)	-0,016 (0,034)	-0,006 (0,026)
Y2014		-0,269 (0,068)***	-0,212 (0,074)***	-0,094 (0,044)**	-0,008 (0,033)		-0,264 (0,063)***	-0,202 (0,068)***	-0,094 (0,041)**	-0,019 (0,035)
Y2015		-0,005 (0,066)	-0,004 (0,064)	-0,011 (0,035)	-0,005 (0,024)		-0,006 (0,061)	-0,005 (0,06)	-0,011 (0,034)	-0,003 (0,027)
Y2016		-0,28 (0,068)***	-0,221 (0,074)***	-0,117 (0,043)***	-0,015 (0,033)		-0,274 (0,063)***	-0,212 (0,068)***	-0,114 (0,041)***	-0,026 (0,034)
SIZE (I)			0,049 (0,028)*	0,054 (0,022)**	0,117 (0,018)***			0,051 (0,024)**	0,061 (0,02)***	0,109 (0,017)***
OCEAN			0,222 (0,087)**	0,147 (0,052)***	0,063 (0,039)			0,122 (0,072)*	0,17 (0,043)***	0,066 (0,038)*
COORD			-0,021 (0,016)	0,016 (0,011)	0,025 (0,008)***			-0,026 (0,014)*	0,006 (0,009)	0,01 (0,007)
L_AIRP			-0,003 (0,041)	-0,021 (0,041)	-0,102 (0,031)***			0,008 (0,035)	-0,022 (0,035)	-0,104 (0,03)***
NOFAB			-0,175 (0,086)**	-0,201 (0,059)***	-0,272 (0,052)***			-0,14 (0,076)*	-0,186 (0,051)***	-0,248 (0,051)***
OVER				0,25 (0,105)**	0,531 (0,079)***				0,285 (0,096)***	0,496 (0,079)***
DOM				-1,002 (0,362)***	-0,475 (0,278)*				-0,933 (0,304)***	-0,379 (0,264)
GINI				-5,065 (0,4)***	-2,963 (0,335)***				-4,838 (0,357)***	-3,084 (0,341)***
DENSITY				0,046 (0,007)***	0,015 (0,005)***				0,042 (0,005)***	0,015 (0,005)***
VI				1,056 (0,328)***	0,744 (0,233)***				0,792 (0,276)***	0,58 (0,225)**
HI				-0,082 (0,178)	-0,496 (0,133)***				0,085 (0,152)	-0,283 (0,125)**
SI				-1,581 (0,289)***	-0,834 (0,218)***				-1,251 (0,255)***	-0,629 (0,215)***
COSTS				0,003 (0)***					0,003 (0)***	
GDP_G				0 (0,003)					0,004 (0,003)	
WEALTH (I)				0,137 (0,026)***					0,078 (0,025)***	
RES				0,002 (0,001)*					0,001 (0,001)	
R ² within	0,20	0,37	0,43	0,84	0,92					
Akaike	125,07	85,92	77,87	-164,81	-310,47	65,62	29,22	27,10	-201,16	-299,48
N	204	204	204	204	204	204	204	204	204	204

Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***

In Iteration 3 wurden statt *NOFAB* auch die einzelnen *FAB* Zuordnungen überprüft. Hier konnte eine Verbesserung des Akaike Kriteriums erreicht werden. Auch das korrigierte R² war größer. Mit Ausnahme von *SWFAB* waren alle Koeffizienten positiv, für *FABEC*, *NEFAB* und *SWFAB* wurde jedoch



keine statistische Signifikanz nachgewiesen. Diese Ergebnisse decken sich teilweise mit dem REM, hier sind *FABEC* und *UKIRL* nicht signifikant.

4.3.4.3 Einfluss auf TFP-Change

Als alternative abhängige Variable wird der TFP-Change, berechnet aus der Malmquist-DEA, verwendet. Das heißt, die Einflussfaktoren werden auf eine Produktivitätsänderung verhältnismäßig zum Vorjahr getestet. Nach der Durchführung der Paneldiagnostik wird in jedem der Modelle eine Pooled-OLS-Regression empfohlen.

Gemäß Hausman-Test wird das FEM Modell bevorzugt. Die Ergebnisse decken sich teilweise mit jenen für die Produktivitätswerte. Insbesondere bei den endogenen Variablen kommt es auch hier zu einem negativen Einfluss der Arbeitszeit-Variable und zu einer positiven Abhängigkeit des Anteils an Nicht-Lotsen im Gesamtmodell. Beide Effekte sind nicht plausibel erklärbar. Im Regressionsmodell mit ausschließlich endogenen Variablen sind diese Effekte umgekehrt und damit logisch ableitbar, jedoch ist in diesem Falle das Vorzeichen von JSC und STATE kontraintuitiv.

Die Jahres-Dummys sind iterations- und modellübergreifend positiv bewertet und in den meisten Fällen statistisch signifikant. Des Weiteren liefert die Inkludierung der Luftraumvariablen robuste und erwartete Ergebnisse. Die Inkludierung einzelner FABs verschlechterte die Modellgüte. Kein FAB hatte einen negativen Effekt, es war jedoch auch keine statistische Signifikanz nachweisbar.

Dem gegenüber entsprechen die Ergebnisse für die Nachfragedaten, mit Ausnahme *SI*, nicht der erwarteten Wirkung. So ist es bspw. unplausibel, dass der Anteil an Überflügen sich negativ, der Anteil des Inlandsverkehrs jedoch positiv auf die Änderung der TFP auswirkt. Im Gegensatz zu allen anderen Betrachtungen ist auch der GINI positiv konnotiert, jedoch statistisch nicht signifikant.

Komplett exogene Variablen haben kaum Einfluss auf die Effizienzänderung. Lediglich für das GDP-Wachstum wird ein statistisch signifikanter, positiver Zusammenhang ausgewiesen. Im Gesamtmodell sind zudem die Wirkung einer fehlenden FAB-Zugehörigkeit, vier der fünf Jahres-Dummys sowie die Komplexitätskomponenten VI und SI signifikant. Der positive Einfluss der vertikalen Interaktionen auf die TFP Änderung ist auf die bereits diskutierten Schwächen des PRU-Maßes zurückzuführen.

Hinsichtlich der Modellgüte stellt die Nutzung des Malmquist-Index eine Verschlechterung dar. Im FEM können lediglich 21 % der Varianz durch das Modell erklärt werden. Das Akaike-Kriterium ist bis auf eine Ausnahme relativ konstant bei ca. -470. Auch wenn die anderen Malmquist-Modelle zu ähnlichen Regressionsergebnissen hinsichtlich Koeffizienten und Signifikanzen führten, war hier die Modellgüte unter jener des Modells 2A. Für die Modelle 1 und 2 wurde zudem gemäß Hausman-Test ein Random Effects Modell bevorzugt. Die niedrigen Werte für Signifikanzniveau und Modellgüte lassen den Schluss zu, dass in mehrperiodischen Betrachtungen weitere Faktoren eine Rolle spielen oder nicht alle Effekte auf der aggregierten Ebene erklärt werden können. Daher wird die Analyse auf tiefere operative Ebenen ausgeweitet.

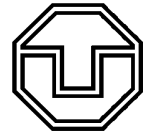


Tabelle 4-10: Fixed und Random Effects Model für Model 2A

	Fixed Effects					Random Effects				
	I1	I2	I3	I4	I5	I1	I2	I3	I4	I5
INT	1,022 (0,257)***	1,032 (0,26)***	0,876 (0,359)**	1,004 (0,459)**	1,011 (0,602)*	0,979 (0,232)***	0,99 (0,234)***	0,917 (0,336)***	0,896 (0,432)**	0,955 (0,529)*
NONA	-0,052 (0,047)	-0,052 (0,047)	0,007 (0,059)	0,117 (0,088)	0,142 (0,126)	-0,058 (0,044)	-0,058 (0,044)	-0,008 (0,055)	0,082 (0,076)	0,094 (0,108)
TIME (I)	0,003 (0,034)	-0,001 (0,035)	-0,021 (0,044)	-0,034 (0,054)	-0,032 (0,058)	0,009 (0,031)	0,006 (0,032)	-0,012 (0,041)	-0,017 (0,05)	-0,018 (0,054)
DELATM	-0,005 (0,014)	-0,009 (0,014)	-0,008 (0,014)	0 (0,015)	0 (0,018)	-0,005 (0,013)	-0,01 (0,013)	-0,009 (0,013)	-0,001 (0,014)	0,004 (0,016)
MET	0,01 (0,012)	0,006 (0,013)	0,012 (0,015)	0,027 (0,018)	0,023 (0,018)	0,003 (0,011)	-0,001 (0,011)	0,005 (0,013)	0,019 (0,016)	0,011 (0,015)
AIRP	-0,005 (0,016)	-0,004 (0,018)	0,012 (0,022)	-0,012 (0,046)	-0,032 (0,055)	-0,003 (0,015)	-0,003 (0,017)	0,003 (0,021)	-0,018 (0,039)	-0,038 (0,049)
JSC	-0,005 (0,015)	-0,004 (0,015)	0,004 (0,021)	-0,025 (0,029)	-0,031 (0,029)	-0,004 (0,014)	-0,003 (0,014)	-0,009 (0,019)	-0,031 (0,026)	-0,046 (0,027)*
STATE	0,007 (0,015)	0,009 (0,015)	0,015 (0,019)	0,002 (0,023)	-0,002 (0,023)	0,012 (0,014)	0,014 (0,014)	0,01 (0,017)	0,001 (0,021)	-0,008 (0,021)
Y2012		0,023 (0,017)	0,03 (0,018)*	0,032 (0,022)	0,039 (0,026)		0,023 (0,017)	0,026 (0,017)	0,029 (0,021)	0,034 (0,025)
Y2013		0,019 (0,018)	0,043 (0,02)**	0,056 (0,027)**	0,062 (0,029)**		0,017 (0,018)	0,029 (0,019)	0,047 (0,025)*	0,045 (0,027)*
Y2014		0,028 (0,017)*	0,028 (0,016)*	0,028 (0,016)*	0,028 (0,016)*		0,028 (0,016)*	0,028 (0,016)*	0,027 (0,016)*	0,029 (0,016)*
Y2015		0,032 (0,017)*	0,039 (0,018)**	0,041 (0,022)*	0,047 (0,026)*		0,032 (0,017)*	0,035 (0,017)**	0,038 (0,021)*	0,041 (0,025)*
Y2016		0,006 (0,018)	0,031 (0,02)	0,045 (0,027)*	0,053 (0,029)*		0,004 (0,018)	0,017 (0,019)	0,035 (0,025)	0,037 (0,027)
SIZE (I)			0,023 (0,007)***	0,02 (0,009)**	0,02 (0,012)			0,016 (0,006)**	0,017 (0,009)**	0,016 (0,01)
OCEAN			-0,002 (0,022)	-0,002 (0,025)	-0,013 (0,028)			0,012 (0,02)	0,012 (0,023)	-0,002 (0,026)
COORD			-0,002 (0,003)	-0,008 (0,005)	-0,008 (0,006)			-0,003 (0,003)	-0,008 (0,004)*	-0,008 (0,005)*
L_AIRP			-0,016 (0,011)	-0,028 (0,022)	-0,039 (0,025)			-0,012 (0,01)	-0,022 (0,019)	-0,034 (0,022)
NOFAB			-0,065 (0,024)***	-0,097 (0,033)***	-0,099 (0,047)**			-0,04 (0,022)*	-0,071 (0,029)**	-0,055 (0,041)
OVER				-0,034 (0,067)	-0,063 (0,075)				-0,024 (0,059)	-0,037 (0,064)
DOM				0,146 (0,202)	0,197 (0,232)				0,164 (0,17)	0,207 (0,199)
GINI				0,054 (0,217)	0,154 (0,264)				0,114 (0,195)	0,234 (0,24)
DENSITY				-0,003 (0,003)	-0,004 (0,004)				-0,001 (0,003)	-0,003 (0,003)
VI				0,253 (0,154)	0,325 (0,168)*				0,162 (0,14)	0,255 (0,153)*
HI				0,012 (0,089)	0,008 (0,093)				0,038 (0,079)	0,016 (0,081)
SI				-0,239 (0,135)*	-0,292 (0,145)**				-0,155 (0,124)	-0,203 (0,134)
COSTS				0 (0)						0 (0)
GDP_G				0,004 (0,002)**						0,006 (0,002)***
WEALTH (I)				-0,005 (0,021)						-0,004 (0,018)
RES				0 (0,001)						0 (0,001)
R ² within	0,03	0,06	0,14	0,18	0,21					
Akaike	-470,93	-467,85	-409,78	-470,39	-469,18	-495,43	-491,69	-490,35	-481,17	-485,93
N	198	198	198	198	198	198	198	198	198	198

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%** , 99%***



4.4 Disaggregierter Ansatz

4.4.1 Motivation und Vorgehen

Die Analyse der Effizienz sowie der beeinflussenden Faktoren auf ANSP Ebene hat zwei wesentliche Nachteile. Zum einen sind die Dienstleister sehr heterogen, sowohl hinsichtlich operativer als auch institutioneller Charakteristiken. Zum anderen beschränkt die Anzahl von maximal 38 Einheiten (ANSPs) die Modellbildung hinsichtlich der DEA. Dies führt dazu, dass nicht alle Effekte auf einer aggregierten Ebene ausreichend bzw. vollständig erklärt werden können. Aus diesem Grund ist im Sinne einer holistischen Betrachtungsweise eine disaggregierte Analyse notwendig, welche auf Basis der operativen Subebenen (vgl. Abbildung 2-6) Kapazitäts- und Effizienzeinflüsse untersucht. Dadurch können hypothetisch mehr vergleichbare Einheiten, z. B. hinsichtlich des Luftraumes, der Nachfrage oder der internen Entscheidungen¹²⁴ analysiert werden. Daher wird zunächst in Abschnitt 4.4.2 eine disaggregierte Ursachenanalyse auf ACC Ebene durchgeführt.

Für die weitere Erhöhung der Granularität wurde in Zusammenarbeit mit der FABEC PMG Benchmarking Taskforce eine Liste potentieller Einflussfaktoren formuliert (Long List). Aufgrund der Charakteristik dieser Faktoren werden zwei wesentliche Ansätze verfolgt. Erstens wird die operative und institutionelle Heterogenität zwischen den einzelnen Einheiten mittels einer qualitativen Analyse erfasst. Diese betrachtet insbesondere jene Unterschiede, die nicht quantifizierbar¹²⁵ sind. Zweitens wird der Einfluss quantifizierbarer Faktoren mittels statistischer Methoden analysiert. Dabei werden sowohl Abhängigkeiten zwischen den einzelnen Einflüssen untersucht als auch deren Wirkung auf die Gesamtproduktivität. Hierfür stehen Daten auf ACC Ebene und Sektor-Ebene zur Verfügung.

Sowohl die qualitative Untersuchung als auch die quantitative Analyse auf Sektorebene basieren auf Daten der FABEC Staaten. Gemäß der PRU-Daten findet in den betreffenden Lufträumen etwa ein Drittel des gesamten europäischen Verkehrs statt (gemessen in Flugstunden), obwohl deren Gesamtfläche lediglich 12% des Luftraumes repräsentiert (FABEC 2019). Dementsprechend hoch sind Verkehrsdichte und -komplexität. Ein wesentlicher Bestandteil der Untersuchung war die Datenakquise und -auswertung (Standfuss et al. 2016a) und Konzeptionierung der Analyse (Standfuss et al. 2017a). Abschnitt 4.4.3 basiert maßgeblich auf diesen Ansätzen, deren Ergebnisse 2019 in einem Bericht zusammengefasst wurden (FABEC 2019).

4.4.2 Quantitative Analyse ACC

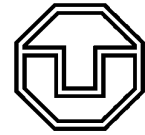
4.4.2.1 Datengrundlage

Die PRU veröffentlicht seit 2011 Daten auf ACC Ebene. Sie werden zusammen mit den ANSP Daten in den „Key-Ops-Data“ Tabellen veröffentlicht und sind über das Onesky-Portal verfügbar. Der Datensatz umfasst, je nach Jahr, 62-63 ACCs¹²⁶. Tabelle 4-11 zeigt eine deskriptive Statistik der veröffentlichten Daten sowie einiger Parameter, die daraus berechnet wurden, für das Jahr 2017. Dabei bildet

¹²⁴ Z. B. Arbeitszeit pro Lotse.

¹²⁵ Z. B. Eingesetzte Tools oder Bestimmung von Kapazitätseckwerten.

¹²⁶ Daten für Simferopol (Ukraine) sind von 2011 bis 2013 verfügbar, für Tiflis (Georgien) zwischen 2015 und 2017.



die Verkehrsdichte einen Quotienten aus Flügen und Luftraumgröße ab. Diese Verkehrsdichte kann nur als grobe Approximation dienen, denn die tatsächliche Verteilung der Verkehre im Luftraum wird dadurch nicht abgebildet. Die Systemrelevanz wurde aus dem relativen Anteil des ACC am Gesamtverkehrsaufkommen berechnet. Arbeitskosten sowie durchschnittliche Arbeitsstunden wurden aus den ANSP-Daten verwendet und sind somit für alle ACCs eines spezifischen ANSP gleich. Die Daten für die Arbeitskosten sind nicht kaufkraftbereinigt (siehe auch Abschnitt 4.2.2.4).

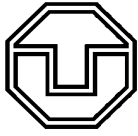
Tabelle 4-11: Deskriptive Statistik ACC Daten, 2017

	Min	Median	95 % Q	Max	Streuung	Rel. Str
Kontrollierte Flugstunden	4.454	186.489	579.492	1.050.311	198.766	87,14 %
Lotsenstunden (on duty)	30.008	166.192	456.862	641.892	136.970	67,40 %
Durchschnittliche Belegungszeit ¹²⁷ (min)	7,69	19,30	34,44	47,54	8,11	39,36 %
IFR Flüge	17.174	595.241	1.408.210	2.058.859	455.813	71,59 %
Luftraumgröße	20.400	147.000	621.700	1.370.000	233.468	104,81 %
Anzahl der Lotsen	22	109	400	477	120	74,96 %
Größe des OPS-Rooms (m ²)	121	800	1.588	1.998	453	56,49 %
Sektoren in der maximalen Konfiguration	1	9	26	36	8	68,65 %
Sektorstunden	8.760	38.500	146.647	153.648	43.074	76,86 %
ATCO-Stunden pro Sektorstunde	2,06	3,80	5,79	7,21	1,10	27,80 %
ATCO-Stunden pro ATCO	785,94	1367,00	1607,00	1972,00	222,01	16,13 %
Proxy für Verkehrsdichte	0,06	3,28	15,70	25,72	5,40	99,48 %
Arbeitskosten pro ATCO (€)	17.894	133.983	233.148	277.629	67.542	48,36 %
Arbeitskosten pro ATCO-Stunde (€)	12	104	232	232	59	55,64 %
Systemrelevanz	0,04 %	1,48 %	3,51 %	5,13 %	1,14 %	71,59 %

Tabelle 4-11 bestätigt die im vorigen Kapitel diskutierte Heterogenität zwischen den einzelnen ACCs. Diese Unterschiede sind sowohl bei Luftraum als auch bei Nachfrage und Betrieb zu sehen. Bei der Anzahl der IFR Flüge unterscheiden sich die Ausprägungen zwischen Minima (Dnipropetrowsk) und Maxima (London AC) um den Faktor 120, bei Flugstunden um Faktor 236 (Maximum Ankara). Ähnliche Unterschiede sind bei Luftraum (Größe) und Betrieb (Arbeitsstunden pro Lotse) sichtbar.

Erwartungsgemäß ist eine hohe Korrelation zwischen einigen Faktoren sichtbar (vgl. Anhang Abbildung A 13). Plausibel ist u. a. der positive Zusammenhang zwischen kontrollierten Flugstunden und eingesetzten Lotsenstunden. Des Weiteren ist eine negative Korrelation zwischen der Anzahl der eingesetzten Lotsen und der durchschnittlichen Arbeitszeit pro Lotse sichtbar. Dagegen ist auf den ersten Blick überraschend, dass die Anzahl der Sektoren nicht mit der Größe des Luftraumes korreliert. Dies ist dadurch bedingt, dass sowohl große Lufträume mit wenig Verkehr als auch kleine mit viel Verkehr Teil des Datensatzes sind. Dies wird auch durch die hohe Streuung der Verkehrsdichte bestätigt.

¹²⁷ Beschreibt die Zeit, in der sich ein Luftfahrzeug im Luftraum befindet.



4.4.2.2 Leistungsbewertung

Basierend auf den in Abschnitt 4.4.2.1 beschriebenen Daten kann die Produktivität der einzelnen Sektoren berechnet werden. Hierfür gibt es mehrere mögliche Metriken (z. B. Flüge pro Lotse). Aufgrund der bereits im ANSP Kontext diskutierten Vorteile sind jedoch zeitbasierte Parameter (z. B. ATCO-Stunden, kontrollierte Flugstunden, etc.) zu bevorzugen. Es wird zudem davon ausgegangen, dass die Leistung der betrieblichen Einheit maßgeblich von Durchsatz und Personaleinsatz bestimmt wird. Da insbesondere die Streckenkontrolle betrachtet wird, werden „Flugstunden“ als Output-Maß genutzt. Eine weitere Möglichkeit wäre, Verspätungsminuten als negativen Output zu ergänzen, um den Trade-off zwischen Kapazität und Delay abzubilden. Allerdings stehen diese Daten nicht für alle betrachteten Einheiten zur Verfügung. Der Personaleinsatz wird über die Lotsenstunden modelliert. Die resultierende Produktivitätskennziffer kann schließlich verglichen werden.

Abbildung 4-4 stellt die Leistungsfähigkeit der zehn ACCs mit der höchsten und niedrigsten ATCO-Produktivität (kontrollierte Flugstunden je ATCO-Stunde) dar. Die höchste Produktivität weist das ACC Warschau auf, welches auch in den Vorjahren ähnlich hohe Werte erreichte. Die FABEC ACCs Maastricht und Karlsruhe belegen die Ränge 3 und 4. Die geringste Produktivität ist für die ukrainischen ACCs sowie das moldawische ACC sichtbar. Für eine Darstellung aller ACCs siehe Anhang Abbildung A 14. Die FABEC ACCs sind jeweils gelb hervorgehoben.

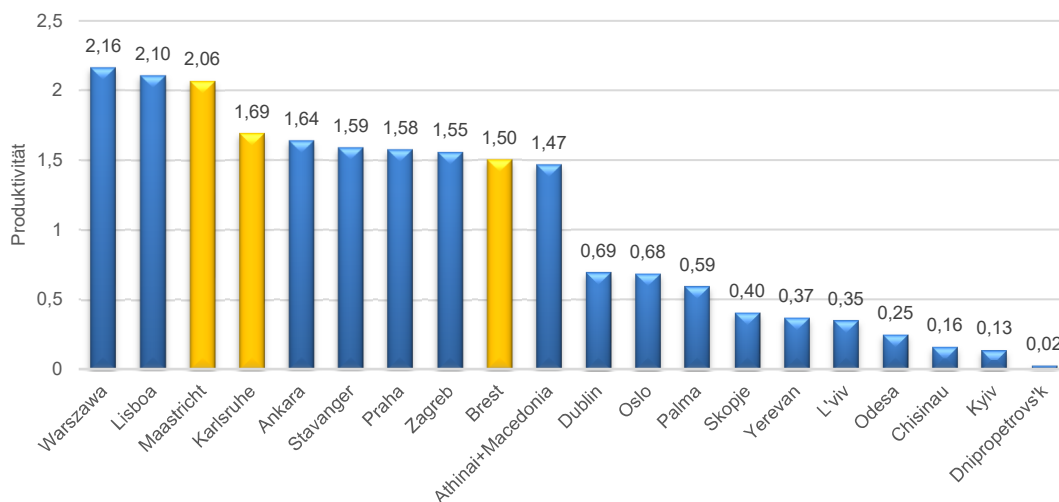
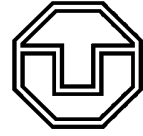


Abbildung 4-4: ACCs mit der höchsten und niedrigsten ATCO-Produktivität 2017

Die Leistungskennzahlen unterliegen, bis auf wenige Ausnahmen, nur geringen jährlichen Schwankungen. Vor allem die ukrainischen und das moldawische ACC weisen eine hohe Streuung zwischen den Werten auf. Dies ist jedoch darauf zurückzuführen, dass seit der Ukrainekrise die Produktivität signifikant rückläufig ist. So sank die Produktivität von Dnipropetrowsk von 0,51 (2010) auf 0,02 (2017). Im Gegensatz dazu stiegen die Kennwerte u. a. für Bukarest, Sofia und Budapest. Dieser Effekt ist auf die Ausweichverkehre zurückzuführen.



4.4.2.3 Cluster-Analyse

Durch die Heterogenität der ACCs ist es nicht ohne weiteres möglich, Zusammenhänge zwischen den Daten herzustellen, z. B. zwischen der Produktivität und potentiell beeinflussenden Faktoren. Eine Möglichkeit ist die Daten zu clustern. Ähnlich der PCA findet dadurch eine Reduktion des Datenumfanges statt. Allerdings werden nicht Faktoren gruppiert, sondern DMUs bzw. ACCs. Grundsätzlich unterscheidet man in partitionierende und hierarchische Verfahren. Eine Einführung in die Clusteranalyse sowie deren mathematischen Background bietet Bacher et al. (2011).

Für die Untersuchungen im ACC Kontext wurden vier¹²⁸ Cluster-Verfahren in R implementiert:

- k-means
- Partitioning Around Medoids (PAM)
- Hierarchisch Top-Down
- Hierarchisch Bottom-up

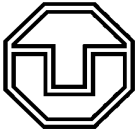
Das k-means Clustering gehört wie PAM zu den partitionierenden Verfahren. Dabei wird vorab eine Anzahl an zu bildenden Clustern (k) bestimmt. Dies kann über sogenannte Scree-Plots¹²⁹ (wie bei der PCA) oder Silhouetten-Plots erfolgen. Beide Ergebnisse können abweichen: Für das k-means Clustering kann mittels Scree-Plot auf neun Cluster geschlossen werden, gemäß des Silhouetten-Plots werden dagegen sieben, zehn oder 14 Cluster empfohlen. Für die Analyse werden alle drei Varianten getestet und auf Plausibilität überprüft. Zudem muss berücksichtigt werden, dass mit steigender Anzahl an Clustern zwar die Homogenität der in einer spezifischen Gruppe enthaltenen Einheiten steigt, sich jedoch die Aussagekraft der Faktoren untereinander reduziert.

Abbildung 4-5 stellt die Gruppierung für 14 Cluster nach k-means dar. Für die Darstellung wurden mittels PCA zwei Hauptkomponenten gebildet, welche x- und y-Achse repräsentieren. Die Aufteilung ist für die gefundenen Cluster plausibel. So sind bspw. die primär im oberen Luftraum agierenden ACCs (Karlsruhe, MUAC und London AC) zu einer Gruppe zusammengefasst. Auch geographische Aspekte sind teilweise interpretierbar (z. B. bei Cluster 11), jedoch sind diese nicht in der Clusteranalyse integriert und daher rein zufällig.

Drei der Cluster bestehen jedoch aus lediglich einem oder zwei ACCs. Diese repräsentieren auf Basis der verwendeten Faktoren einen Sonderfall. Eine Abhängigkeit zwischen Faktoren ist hier, bspw. durch eine Regressionsgerade, kaum ableitbar. Daher wird überprüft, ob eine Reduktion der Cluster einen Vorteil bewirkt. Die Ergebnisse sind in Anhang Abbildung A 17 dargestellt. Auf der linken Seite wurden die ACCs 7 Clustern zugeteilt, rechts 10 Clustern. Die Abbildung zeigt, dass 7 Cluster eine zu geringe Anzahl sind, da u. a. ACCs mit ausschließlich oder überwiegend Streckenverkehr (MUAC, Karlsruhe und London AC) mit anderen ACCs gruppiert werden, die auch einen signifikanten Anteil an Ter-

¹²⁸ Neben den genannten Verfahren gibt es u. a. auch dichte-basierte Verfahren, neuronale Netze oder Fuzzy Clustering.

¹²⁹ Für die Screeplots gilt wieder das Ellbow-Kriterium.



minal-Verkehr aufweisen. Dies wird durch die Gruppierung in zehn Cluster vermieden – die drei benannten ACCs sind hier in einem eigenen Cluster zusammengefasst. Bis auf die Sonderfälle Canarias und Ankara bestehen alle Cluster aus drei oder mehr Einheiten.

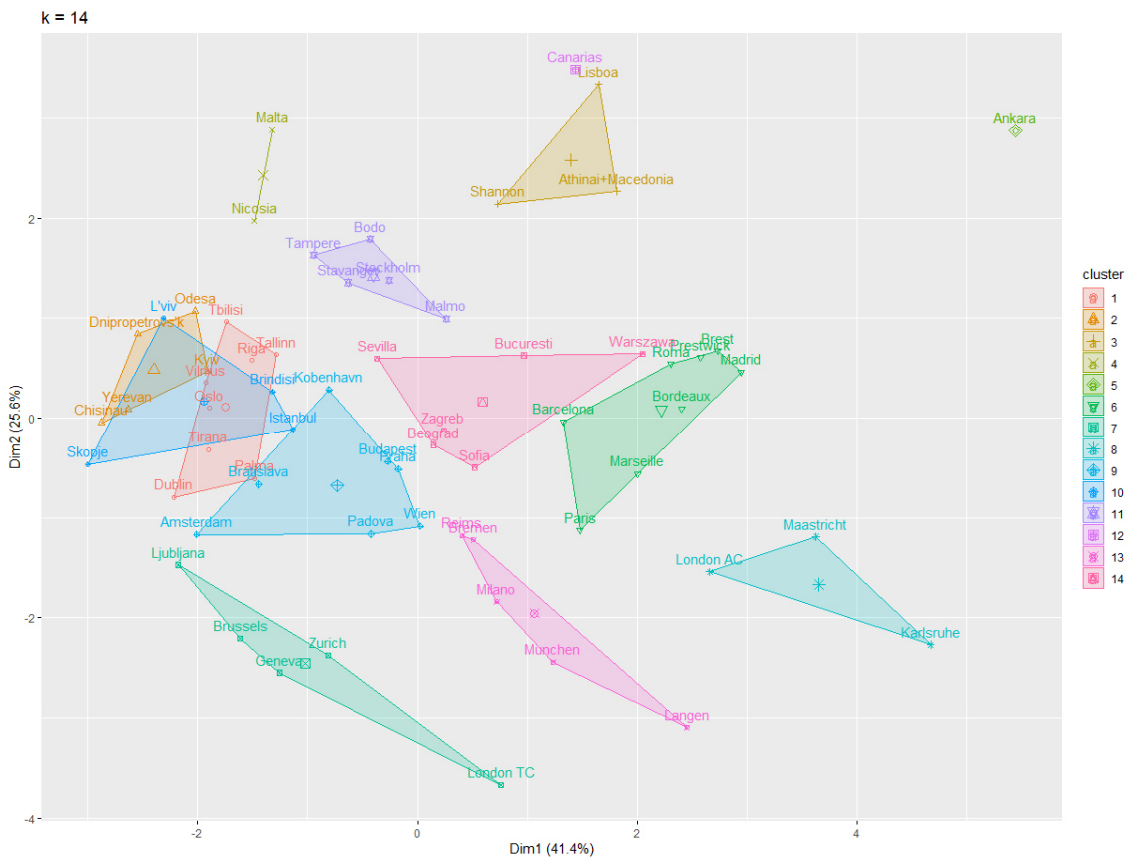


Abbildung 4-5: Aufteilung der ACCs bei 14 Clustern

Der Nachteil des k-means Algorithmus ist u. a., dass er sensitiv auf Ausreißer reagiert und abhängig von zufälligen Startpunkten ist (Müller und Guido 2017; Xu und Wunsch 2009). Dadurch können sich Ergebnisse bei mehrfacher Durchführung ändern. Dies war bei 14 Clusterzentren zu beobachten. Alternativ zum k-means Clustering wurde daher der PAM-Algorithmus angewendet, welcher eine höhere Robustheit gegenüber Ausreißern besitzt (Xu und Wunsch 2009). Dieser unterscheidet sich im Ablauf nur unwesentlich, nimmt aber statt künstlichen Clusterzentren tatsächliche Punkte des Datensatzes (Barbakh et al. 2012). Das heißt, der Algorithmus sucht eine für das Cluster repräsentative Einheit (ACC). Abbildung 4-6 illustriert die Clusterbildung bei k=10. Links sind die Gruppierungen gemäß k-means abgebildet, rechts jene nach PAM-Algorithmus.

Es ist zu sehen, dass sich die Zuordnungen in fast allen Clustern unterscheiden, teilweise jedoch lediglich um eine ACC. Als Sonderfall wird bei PAM lediglich Canarias identifiziert. MUAC, London AC und Karlsruhe bleiben in einem Cluster. Auch die Gruppe, welche u. a. Amsterdam und Brüssel beinhaltet, weist eine konstante Zuordnung auf. Diese Zuordnungen verschieben sich gemäß der gewählten Clusterzahl k. Bei 10 Gruppen ordnet auch PAM die Sonderfälle (hier Canarias, Ankara und London TC) eigenen Gruppen zu.

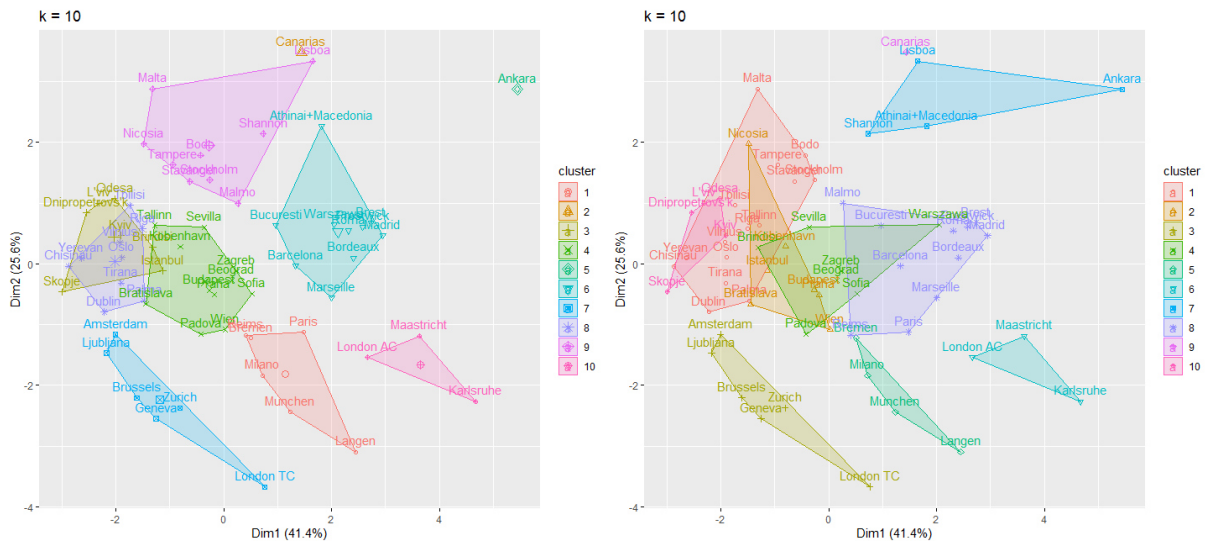


Abbildung 4-6: Unterschiede der Gruppenzuordnung bei k-means und PAM Algorithmus

Aufgrund der Heterogenität der ACCs ist die Median-basierte PAM Methodik vorzuziehen. Auch der Rechenaufwand bleibt bei 63 bzw. 64 Einheiten gering, was den Nachteil dieser Methodik gegenüber k-means verringert. Ein weiteres Kriterium ist die „durchschnittliche Silhouetten-Breite“ (siehe Anhang Abbildung A 16). Im Falle von k-means beträgt sie 0,23 für k=14 und k=10, bzw. 0,21 für k=7. Die Werte für das PAM-Clustering weichen nur geringfügig ab: 0,2 (bei k=7 und k=14) bzw. 0,21 (k=10). Zudem ist bei der PAM-Methode keine Variabilität in den Zuordnungen sichtbar, somit sind die Ergebnisse robust.

Abzugrenzen von den partitionierenden Verfahren ist das hierarchische Clustering. Dies kann agglomerativ (bottom up) oder divisiv (top-down) erfolgen. Das heißt, dass die einzelnen ACCs sukzessive entweder von einem großen Cluster in die einzelnen Einheiten aufgeteilt, oder vice versa von einzelnen ACCs bis zu einem großen Cluster kombiniert werden. Das Ergebnis wird häufig in Dendrogrammen dargestellt, bei denen die Fußzeile die einzelnen ACCs beinhaltet (siehe Abbildung 4-7).

Der Vorteil der hierarchischen Verfahren liegt darin, dass keine ex-ante Annahmen über die Clusteranzahl getroffen werden müssen. Stattdessen wird der Baum auf einer bestimmten Ebene geschnitten und die jeweiligen Zuordnungen als Cluster identifiziert (die roten Boxen in Abbildung 4-7 entsprechen der Gruppierung bei 14 Clustern). Das bedeutet, dass sich die Zuordnungen zu Clustern auch bei unterschiedlicher Anzahl der Cluster nicht ändern – ein Cluster kann lediglich geteilt (agglomerativ) oder zwei Cluster kombiniert (divisiv) werden. Als Ähnlichkeitsmaß wird häufig die ward-Methode genutzt, welche die Varianz innerhalb des Clusters minimiert. Über die Höhe wird die Distanz zwischen den Datenpunkten und/oder Clustern angegeben (Xu und Wunsch 2009).

Abbildung 4-7 illustriert das Ergebnis der agglomerativen, hierarchischen Clusteranalyse für 14 Cluster. Die Zuordnungen unterscheiden sich erwartungsgemäß von den partitionierenden Verfahren. Alternativ filtert das top-down Verfahren eher die bereits durch k-means und PAM Algorithmus identifizierten Ausreißer heraus. Die Zuordnungen der ACCs zu den einzelnen Clustern bleibt jedoch in beiden Ansätzen größtenteils konstant.

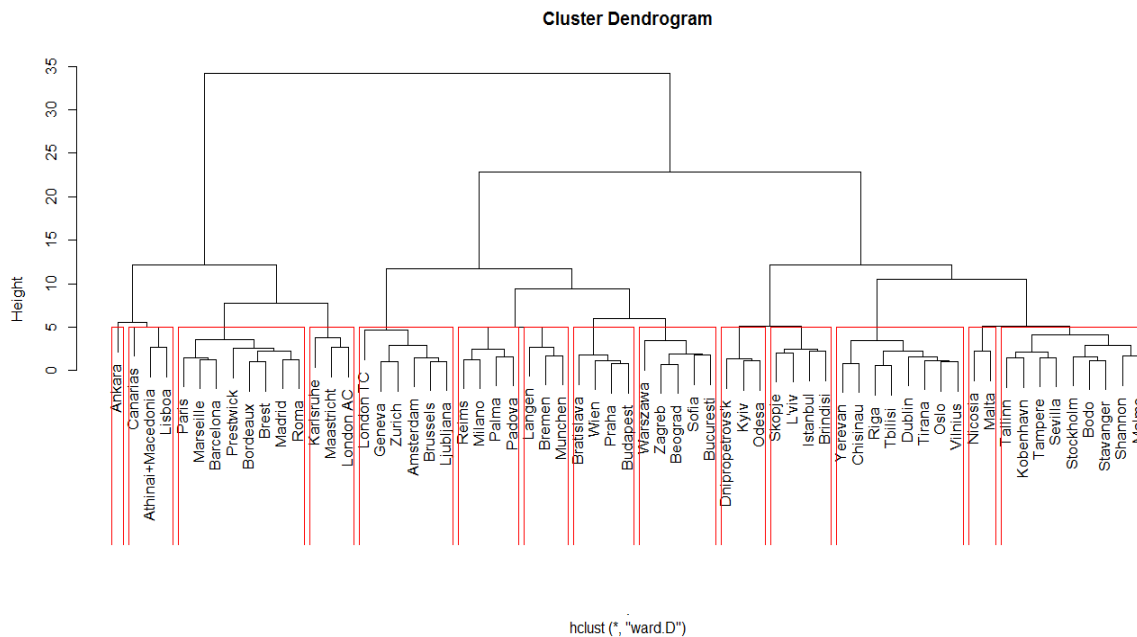
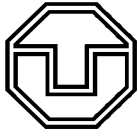


Abbildung 4-7: Ergebnis hierarchische Clusteranalyse

Das Ergebnis umfasst insgesamt 12 Zusammensetzungen der ACCs zu Gruppen, differenziert nach Methodik und Anzahl an Clustern. Jedes Cluster kann durch spezifische Eigenschaften definiert werden, z. B. großer Luftraum oder hohe Komplexität. Um diese spezifischen Eigenschaften zu finden, wurden die Cluster einzeln ausgewertet und die jeweiligen Faktoren der betreffenden ACCs auf Mittelwert und Streuung in Relation zum Gesamtdatensatz untersucht.

Für die entstandenen Cluster werden in einem nächsten Schritt die Zusammenhänge zwischen Faktoren untersucht. Wie bereits diskutiert, sind bei der Berücksichtigung des kompletten Datensatzes meist keine eindeutigen Schlüsse möglich. Nach dem Clustern und Identifizieren der gruppenspezifischen Eigenschaften können jedoch die Zusammenhänge klarer herausgearbeitet werden. Abbildung 4-8 veranschaulicht dies an der Interdependenz zwischen Luftraumgröße und Produktivität. Auf der linken Seite ist dies für den gesamten Datensatz dargestellt, auf der rechten Seite lediglich für die ACCs eines Clusters (Ebene 7 der agglomerativ-hierarchischen Methode). Zudem ist die lineare Trendlinie (rot) eingezeichnet. Diese hat im linken Abschnitt ein Bestimmtheitsmaß von lediglich 0,12, im geclusterten Fall 0,27.

Auf Basis der linken Grafik kann ein negativ parabolischer Zusammenhang vermutet werden. Das bedeutet, dass die Produktivität zunächst mit wachsender Luftraumgröße steigt, jedoch ab einer bestimmten Größe wieder sinkt. In der Tat scheint dies die gelbe Trendlinie zu bestätigen, die zudem ein höheres R^2 aufweist. Jedoch wird diese Trendlinie von einem Extremwert beeinflusst (Canarias). Durch die Eliminierung dieser Einheit erreicht die lineare Trendlinie ebenfalls ein R^2 von 0,22. Unabhängig davon sind für einige Cluster eine negative Interdependenz zwischen Luftraumgröße und Produktivität zu beobachten.

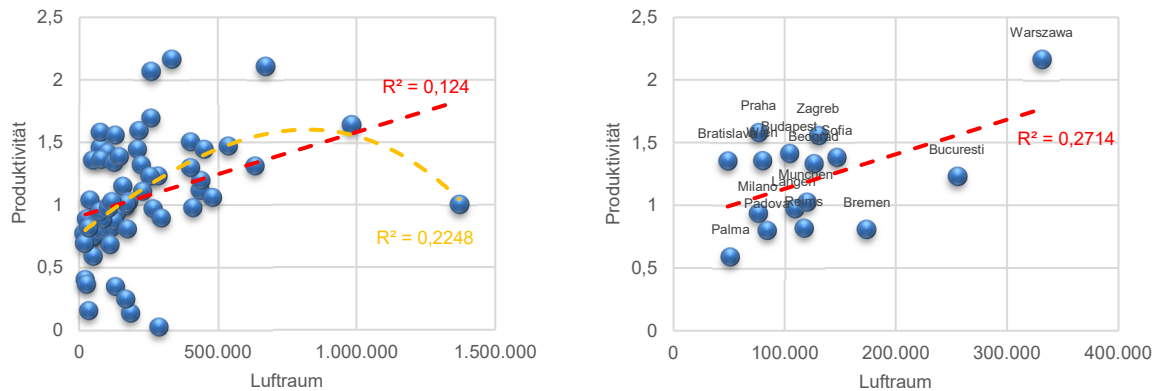
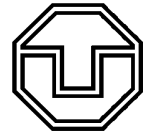


Abbildung 4-8: Zusammenhang zwischen Produktivität und Luftraum vor und nach Clustering

Die Clusteranalyse kann dazu beitragen, Zusammenhänge zwischen den genutzten Faktoren abzubilden. Durch die Anwendung verschiedener Cluster-Algorithmen können zudem Ausreißer identifiziert werden, deren Sonderstatus bei einer weiterführenden Analyse basierend auf einem Clustering berücksichtigt werden sollte (z. B. bei Sensitivitätstests). Die Interdependenzen sind jedoch nicht für alle Cluster sichtbar, insbesondere, wenn andere Faktoren einen wesentlich höheren Beitrag zur Clusterbildung leisteten als die betrachtete Interdependenz.

Ein Problem der Cluster-Analyse stellt hier erneut die geringe Anzahl an Einheiten sowie deren heterogene Charakteristiken dar. Des Weiteren konnten aufgrund der Datenverfügbarkeit wesentliche Eigenschaften, bspw. das Vorhandensein von militärischen Gebieten, nicht berücksichtigt werden. Andere Charakteristiken sind nur sehr vereinfacht approximiert, z. B. die Verkehrsdichte. Schließlich gibt es Unterschiede in den Modellspezifikationen, z. B. die Bestimmung des Maßes der „Ähnlichkeit“ von Einheiten (Distanzmessung).

4.4.2.4 Regressionsanalyse

Wie im vorigen Abschnitt beschrieben, kann der Gesamtdatensatz in mehrere Cluster unterteilt werden. Die Grundgesamtheit umfasst jedoch sehr heterogene Einheiten, sodass das Clustering nur in einzelnen Fällen tiefere Einblicke in Interdependenzen ermöglicht. Ein zweiter Ansatz wird daher über eine Regressionsanalyse erfolgen. Da die Leistungskennzahl auf dieser betrieblichen Ebene eine Produktivitätskennziffer widerspiegelt (nämlich Flugstunden per ATCO-Stunde), wird eine OLS-Regression durchgeführt. Dabei wird, ähnlich den Untersuchungen auf ANSP Ebene, zunächst die Modellgüte maximiert, um eine möglichst genaue Prädiktion zu ermöglichen und weiterhin die Regressionsanalyse stufenweise nach endogenen und exogenen Faktoren aufgebaut.

Ähnlich der Daten auf ANSP Ebene, sind die verwendeten Faktoren auch auf ACC Ebene in ihrer Ausprägung verschieden. Dies kann das Ergebnis einer Regression verzerren. Zur Vermeidung des Effektes stehen drei Anpassungsmethoden zur Verfügung. Erstens ist eine Log-Transformation der Faktoren mit großer Ausprägung (z. B. Luftraumgröße) möglich, wie sie auch auf ANSP Ebene genutzt wurde. Zweitens können die Daten standardisiert werden (z-Transformation). Dadurch weisen die Daten einer Variablen den Mittelwert 0 und die Varianz bzw. Standardabweichung 1 auf. Drittens können die



Daten normiert werden, d. h., die Daten werden derart transformiert, dass sie in einem vorgegebenen Intervall (z. B. [-1;1]) liegen. Die drei Verfahren führen zu einer besseren Vergleichbarkeit der Faktoren untereinander. Standardisierung und Normierung haben zudem den Vorteil, dass die Konstante bzw. Intercept, also der Schnittpunkt mit der y-Achse, in der Regressionsformel vernachlässigt werden kann. Eine detaillierte Beschreibung der Transformationen und deren Anwendung in der Regressionsanalyse bieten u. a. die Publikationen von Haneke et al. (2019), Irizarry (2019) sowie Urban und Mayerl (2018).

Für die Durchführung der Analyse wird primär R genutzt. Das entwickelte Skript vergleicht die Modellgüte für drei Datentransformationen sowie drei verschiedene Modelle:

- Alle Parameter
- Alle Parameter, exklusive Kosten und Größe OPS Room
- Alle Parameter, exklusive Nachfrage und Größe OPS Room

Insgesamt wird demnach die Modellgüte von neun Regressionsmodellen verglichen. Somit kann die Auswirkung der Datentransformation auf die Modellgüte überprüft werden. Die höchste Güte erreichten hier die Modelle mit den normierten Werten (siehe Anhang Tabelle A 12). Innerhalb der Modelle erreicht Modell 2 die geringste Güte. Zusätzlich analysiert das Skript die Regressionsgüte (u. a. Homoskedastizität, Normalverteilung der Residuen und Cook-Graph). Die Graphen sind für die Datentransformationen identisch, unterschieden sich jedoch hinsichtlich der Modelle. Hier weisen jedoch Modell 1 und 3 eine starke Ähnlichkeit auf (siehe als Beispiel Anhang Abbildung A 18).

Nach diesem Robustheitstest wird nun der Einfluss einzelner Parameter geprüft. Hierzu stehen die in Tabelle 4-11 diskutierten Daten zur Verfügung. Zusätzlich wurde mittels einer NEST Auswertung (EUROCONTROL 2018c) zwei Komplexitäts-Proxys hinzugefügt: Der Anteil des Streckenverkehrs¹³⁰ und die durchschnittliche Änderung der Flugflächen¹³¹. Hier ist zu beachten, dass für MUAC ein Wert von 18 % Anteil Streckenverkehr ausgewiesen ist, welcher unrealistisch erscheint. Daher wird ein Anteil von 40 % Streckenverkehr angenommen. Dies begründet sich zum einen durch den hohen Enroute-Anteil (siehe ANSP-Analyse) und den geringen Level-changes pro Flug. Zum anderen muss entgegenwirkend einbezogen werden, dass mehrere Großflughäfen (Frankfurt, London, Amsterdam, Paris, etc.) im Einzugsgebiet des betreffenden Luftraumes sind und somit Flüge ihre zugewiesene Flugfläche erreichen (Departures) oder ihren Sinkflug beginnen (Arrivals). Drittes Kriterium ist ein Vergleich des Wertes für Karlsruhe (ca. 34 %).

Als Parameter werden des Weiteren die FAB-Dummys sowie die Technologie- und Wohlstands-Komponente entsprechend der Regressionsanalyse auf ANSP Ebene berücksichtigt. Die verwendeten unabhängigen Variablen umfassen somit die in Tabelle 4-12 aufgelisteten Faktoren, welche erneut in endogene, teil-exogene und exogene Faktoren unterschieden werden. Da die Faktoren teilweise korrelieren (siehe Anhang Abbildung A 19), wird vor der Anwendung ein VIF-Test durchgeführt.

¹³⁰ Gemeint ist hier jener Verkehr, der weder steigt noch sinkt.

¹³¹ „Average Level Changes“.

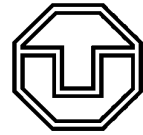


Tabelle 4-12: Effizienzbeeinflussende Faktoren in der Regressionsanalyse (ACC)

	Parameter	Maß	Bedeutung	Einfluss
Endogen	STAFFB	H	ATCO-Stunden pro Sektorstunde ¹³²	-
	TIME (I)	h/a	Durchschnittliche Arbeitszeit pro Lotsen pro Jahr	-
	SIZE (I)	Km ²	Größe des Luftraumes	+
	SECT	Stück	Maximale Anzahl an Sektoren	?
Teillexogen	CRUISE	%	Anteil des Streckenverkehrs	+
	VERT	Anzahl	Proxy für vertikalen Verkehr, beschrieben durch die durchschnittliche Anzahl der Flugflächen-Änderungen pro Flug	-
	DENSITY	LFZ/RE ¹³³	Verkehrsdichte	-
	FAB	0/1	Dummy Variable. Sie ist 1, wenn der übergeordnete ANSP in einem FAB organisiert ist, 0 sonst	+
	TRANS	min	Durchschnittliche Verweildauer im Luftraum	+
	CHANGE	%	Änderungsrate des Verkehrs im Vergleich zum Vorjahr	-
Exog.	COSTS (I)	€	Jährliche Kosten pro Lotsen ¹³⁴	+
	WEALTH (I)	€	BIP/Kopf	+
	RES	Stück	Anzahl der Patente pro 100.000 Einwohner	+

Die Parameter *RES*, *WEALTH*, *FAB*, *DENSITY*, *SIZE* und *TIME* wurden bereits in Abschnitt 4.2.2 beschrieben und die potentielle Wirkung diskutiert. Allerdings unterscheidet sich bei der Luftraumgröße die Zuordnung. Sie wird auf ACC Ebene als endogen bewertet, da die Strukturierung des Luftraumes und die damit verbundene Anzahl und Lage der ACCs dem ANSP obliegt. Ein weiterer Unterschied ist die Berechnung von *DENSITY*. Statt dem PRU Score¹³⁵ wird für die ACC Ebene die Anzahl der Luftfahrzeuge pro Luftraumeinheit bewertet. Aufgrund fehlender Daten können räumliche und zeitliche Konzentrationen des Verkehrs nicht berücksichtigt werden. Daher repräsentiert *DENSITY* einen Durchschnittswert zur Approximation der Verkehrsdichte.

Bei den endogenen Variablen repräsentiert *STAFFB* einen Proxy, welcher die Ressourcen-Puffer hinsichtlich der Fluglotsen ausdrückt. Er berechnet sich aus der Summe der ATCO-Stunden im Verhältnis zu der Summe der Sektorstunden. Da in den meisten Sektoren zwei ATCOs eingesetzt werden, sollte i. d. R. die Variable den Wert „2“ annehmen. Wie Tabelle 4-11 zeigte nimmt diese Variable jedoch einen Maximalwert von 7,21 an. Die Differenz zwischen beiden Werten kann als Puffer für die Ressourcenplanung interpretiert werden. Mit zunehmender Ausprägung dieses Puffers wird eine Verminderung der Produktivität erwartet.

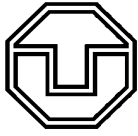
SECT beschreibt, wie umfangreich der Luftraum der ACC untergliedert werden kann. Eine eindeutige Erwartung über die Wirkung ist nicht möglich. Zum einen könnte ein positiver Zusammenhang nachgewiesen werden, da die Kapazität mit zunehmender Ausprägung der Variable flexibler an die Nachfrage angepasst werden kann. Zum anderen könnte eine hohe Sektorisierung auch zu negativen

¹³² Eine Sektorstunde beschreibt die Öffnungszeit eines Sektors.

¹³³ Für die Dichte wird die Anzahl der Luftfahrzeuge durch die Größe des Luftraums der ACC geteilt.

¹³⁴ Im Gegensatz zur ANSP Ebene werden hier aufgrund der anderen Parameter nicht die Kosten pro Lotsenstunde genommen.

¹³⁵ Die Komplexitäts-Scores standen zum Zeitpunkt der Untersuchung nicht zur Verfügung.



Auswirkungen führen, da die Kapazität eines geteilten Sektors aufgrund des Koordinierungsaufwandes nicht genau so hoch ist wie die des ursprünglichen Sektors (Sawtooth-Modell, Standfuss et al. 2018c).

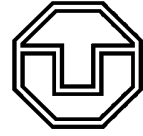
Bei den teilexogenen Faktoren ist, wie bei der Analyse auf ANSP Ebene, eine Unterscheidung in luftraum- und nachfragespezifische Variablen möglich. *CRUISE* beschreibt den Anteil jener Flüge, die sich im Streckenflug befinden und damit keine oder nur geringfügige Änderungen in der Flughöhe aufweisen. Hier wird, analog zur Variable *OVER* auf ANSP Level, ein positiver Zusammenhang zur Leistungsfähigkeit erwartet. Vice versa kann die Anzahl der durchschnittlichen Flugflächenänderungen als Proxy für den Vertikalverkehr (*VERT*) tendenziell produktivitätshemmend angenommen werden. Eine Nutzung des Anteils an Vertikalverkehr ist aufgrund der Singularität nicht möglich. Zudem wird unterstellt, dass Verkehre im Steig- und Sinkflug einen äquivalenten Einfluss auf die Performance haben.

Die Variable *TRANS* beschreibt die durchschnittliche Dauer, die sich ein Luftfahrzeug im Luftraum befindet. Es wird angenommen, dass sich die Produktivität mit zunehmender Ausprägung der Variable erhöht, da eine kurze Verweildauer einen höheren Koordinationsaufwand mit angrenzenden Sektoren verursacht. Die Variable weist zwar eine hohe Korrelation mit der Luftraumgröße auf (Abbildung A 19), dennoch werden beide Variablen getrennt betrachtet, da die Flugrouten im Sektor unterschiedliche Entfernungen implizieren können.

Die relative Änderung des Verkehres im Vergleich zum Vorjahr wird in der Variable *CHANGE* erfasst. Es kann angenommen werden, dass eine hohe Ausprägung der Variable zu einer niedrigeren Produktivität führt. Dies liegt darin begründet, dass auf kurze Frist die Ressourcen nicht beliebig angepasst werden können (z. B. Ausbildung von ATCOs). Eine hohe Ausprägung könnte theoretisch auch zu positiven Auswirkungen führen, insbesondere bei kleineren ACCs, da dadurch die kapazitative Sättigung der Lufträume steigt. Potentielle Sättigungseffekte sind in den Variablen *SECT* und *DENSITY* abgebildet. Dennoch wurde über eine Variable *CHANGE2* überprüft, ob die Richtung der Änderungsrate (positiv = Wachstum, negativ = Rückgang der Verkehrszahlen) die Produktivität beeinflusst.

Der Einfluss der Variable *COSTS* auf die Leistungsfähigkeit entspricht im Wesentlichen jenem auf ANSP Ebene. Es ist jedoch zu beachten, dass hier die jährlichen Kosten pro Lotsen (FTE) verwendet werden. Auch auf ACC Ebene kann ein positiver Einfluss auf die Produktivität erwartet werden, da dadurch ein höherer Anreiz zur sparsamen Nutzung der Ressource gegeben ist. Zudem kann die Ressource Arbeit nicht durch Kapital substituiert werden.

Neben den beschriebenen Variablen wurden weitere Optionen getestet. Als Approximation der Zentralität der ACC innerhalb des europäischen Luftverkehrs wurde die Relevanz bestimmt. Diese Variable (*RELE*) repräsentiert den Verkehrsanteil in dem spezifischen ACC im Verhältnis zur Nachfrage des gesamten Betrachtungsraumes. Des Weiteren kann davon ausgegangen werden, dass eine Störung in einer ACC, welche eine hohe Relevanz aufweist, auch höhere Auswirkungen auf benachbarte Einheiten hat. Die Berücksichtigung dieser Variable in der Regressionsanalyse wurde jedoch verworfen, da die Größe stark nachfrageabhängig ist. Die Nachfrage selbst ist bereits im Produktivitätsmaß enthalten.



Eine weitere Einflussgröße könnten die Kosten für die Luftraumnutzer darstellen. Daher wurde die Berücksichtigung der Variable *UCOSTS* als Proxy für die Bepreisung der Dienstleistungen getestet. Je höher diese Variable ist, desto höher ist auch der Anreiz für Ausweichverkehre. Dieses würde sich langfristig negativ auf die Produktivität auswirken. Von der Nutzung der Variablen innerhalb der Analyse wurde aus mehreren Gründen abgesehen. Erstens sind in den Gesamtkosten auch die Lotsenkosten enthalten, die bereits in der Variable *COSTS* integriert sind. Zweites liegen die Werte nur auf ANSP Ebene vor, d. h. alle ACCs eines ANSP hätten den gleichen Variablenwert. Drittens sind Ausweichverkehre durch eine Vielzahl von weiteren Einflussgrößen abhängig, u. a. von der Luftraumgröße und den Nutzungsgebühren der Nachbareinheiten. Viertens würde diese Variable einen Proxy für die Produktivität selbst darstellen.

Die Regression wird sowohl mit logarithmischer Transformation der Faktoren *TIME*, *SIZE*, *COSTS* und *WEALTH* als auch auf Basis der normierten Werte durchgeführt. Die Normierung betrifft lediglich die kontinuierlichen, unabhängigen Variablen. Dummy-Variablen und die abhängige Variable bleiben davon unberührt. Für die Normierung der Daten stehen mehrere Varianten zur Verfügung, die ebenfalls in R implementiert sind. In der vorliegenden Analyse werden die Daten auf das Intervall [0;1] transformiert. In diesem Fall bleiben die Dummy-Variablen aufgrund der binären Ausprägung von der Transformation unberührt.

Zunächst wird die Modellgüte der Regression maximiert, um so eine möglichst genaue Prädiktion des Produktivitätswertes zu ermöglichen. Die Analyse wird sowohl auf Basis der originären als auch der normierten Werte durchgeführt. Zudem werden verschiedene Kombinationen von Variablen getestet, sodass stark interdependente Faktoren substituiert und die Gesamtmodellgüte verglichen werden kann. Abbruchkriterium ist wiederum ein p-Wert von 0,33. Die Ergebnisse sind in Tabelle 4-13 zusammengefasst¹³⁶. In M1 wird im Grundmodell (vor der Variablenreduktion) lediglich der FAB-Dummy verwendet, in M2 die einzelnen FAB-Zuordnungen. Da der VIF-Test auf Kollinearität für beide Modelle im Ausgangsmodell negativ¹³⁷ ist, werden keine anderen Faktoren substituiert.

Die Ergebnisse zeigen, dass die höchste Modellgüte bei Einbeziehung der einzelnen FABs (M2) erreicht wird. Zwar ist die erklärte Varianz bei Nutzung der originären Daten ca. 1 % höher, jedoch weist das Akaike-Kriterium eine höhere Güte für die normierten Daten aus. Die Variablen *SECT*, *VERT*, *DENSITY*, *UKIRL* und *DKSEFAB* sind in keinem der betrachteten Modelle relevant ($p > 0,33$).

Die Vorzeichen der Koeffizienten sind, mit Ausnahme der Konstanten *INT*, robust. Das heißt, die Interdependenz zwischen den Faktoren und der Produktivität ist modellübergreifend konstant. Die Ausprägungen und Vorzeichen der Koeffizienten spiegeln teilweise die Erwartungen wider, insbesondere die positiven Einflüsse eines hohen Anteils an Streckenverkehr, einer hohen Verweildauer im Luftraum sowie den Lotsenkosten. Des Weiteren ist der negative Einfluss einer hohen Änderungsrate der Nach-

¹³⁶ Die Tabelle enthält lediglich jene Variablen, welche in mindestens einem Modell einen p-Wert $< 0,33$ aufweisen.

¹³⁷ Lediglich der Wert für FABEC war bei der Nutzung originärer Daten signifikant. Diese Kollinearität wird allerdings als Scheinkorrelation gewertet und trat bei Nutzung der normierten Werte nicht auf.



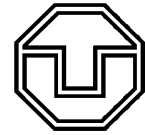
frage auf Basis der normierten Daten erwartungskonform. Auch die Luftraumgröße ist wie erwartet positiv bewertet, jedoch nicht signifikant. Die fehlende Signifikanz ist auch im Falle von *STAFFB* trotz erwartetem negativem Vorzeichen sichtbar. Unerwartet ist dagegen, dass die Arbeitszeit pro Lotsen negativ konnotiert ist. Gleiches gilt für die Vorzeichen der Variablen *WEALTH* und *RES*.

Tabelle 4-13: Regressionsanalyse ACC unter Maximierung der Modellgüte

	Originäre Daten		Normierte Werte	
	M1	M2	M1	M2
INT	-5,361 (1,047)***	-2,183 (1,679)	0,096 (0,083)	0,019 (0,076)
STAFFB	-0,055 (0,034)		-0,089 (0,082)	
TIME (I)		-0,365 (0,214)*		-0,151 (0,087)*
SIZE (I)	0,089 (0,069)			
CRUISE	0,008 (0,002)***	0,006 (0,001)***	0,259 (0,064)***	0,155 (0,058)**
TRANS	0,017 (0,008)**	0,033 (0,004)***	0,517 (0,084)***	0,705 (0,071)***
CHANGE			-0,249 (0,111)**	
FAB	-0,148 (0,141)			
FABEC		-0,128 (0,095)		
FABCE				0,143 (0,055)**
DANUBE		-0,291 (0,174)		
NEFAB		0,285 (0,111)**		0,21 (0,051)***
SWFAB		-0,491 (0,12)***		-0,216 (0,053)***
BALTIC		0,306 (0,165)*		0,265 (0,077)***
BLUEMED		-0,258 (0,095)***		
COSTS (I)	0,479 (0,074)***	0,52 (0,057)***	0,386 (0,083)***	0,554 (0,065)***
WEALTH (I)	-0,05 (0,044)	-0,09 (0,04)**		
RES		-0,004 (0,002)*		-0,279 (0,094)***
Korr. R ²	0,65	0,76	0,60	0,75
Akaike	18,77	-0,97	-69,01	-96,26
N	63	63	63	63

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%***, 99%***

Die Analyse auf Basis der originären Werte zeigt weiterhin, dass die Organisation des ANSP in einem FAB negativ bewertet wird (M1), dies aber abhängig von dem einzelnen FAB ist (M2). Der FAB-Dummy weist jedoch, unabhängig von Modell und Datengrundlage, keine statistische Signifikanz auf. Signifikant positiv wird die Organisation der ACC innerhalb der *BALTIC* FAB, *NEFAB* sowie bei den normierten Daten *FABCE* ausgewiesen. Dagegen ist *SWFAB* in beiden Varianten von M2, *BLUEMED* lediglich bei den originären Daten statistisch signifikant und negativ konnotiert.

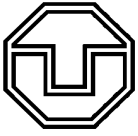


Für Model 2 wird im Folgenden die iterative Regressionsanalyse durchgeführt (Tabelle 4-14). Dabei werden zunächst nur endogene Variablen berücksichtigt (I1). Im zweiten Schritt werden die als teilexogen bewerteten Nachfragevariablen integriert (I2). Schritt drei berücksichtigt zudem die, ebenfalls teilexogene, FAB-Zugehörigkeit (I3). Im letzten Schritt (I4) werden die exogenen Faktoren ergänzt. Der VIF-Test wies in keinem der Iterationen auf ein Multikollinearitätsproblem hin.

Tabelle 4-14: Iterative Regressionsanalyse ACC Ebene, normierte Daten

	I1	I2	I3	I4
INT	0,339 (0,127)***	0,327 (0,153)**	0,045 (0,17)	0,019 (0,149)
STAFFB	-0,078 (0,122)	-0,108 (0,103)	0,039 (0,106)	-0,014 (0,089)
TIME	0,096 (0,175)	0,004 (0,148)	-0,066 (0,146)	-0,094 (0,133)
SIZE	0,316 (0,156)**	0,139 (0,19)	0,039 (0,18)	-0,053 (0,152)
SECT	0,254 (0,149)*	0,071 (0,14)	0,057 (0,138)	0,072 (0,117)
CRUISE		0,14 (0,128)	0,141 (0,129)	0,17 (0,11)
VERT		-0,074 (0,139)	0,025 (0,13)	0,017 (0,109)
DENSITY		0,181 (0,133)	0,098 (0,139)	-0,021 (0,123)
TRANS		0,473 (0,189)**	0,63 (0,182)***	0,689 (0,151)***
CHANGE		-0,449 (0,123)***	-0,188 (0,133)	-0,077 (0,113)
FABEC			0,213 (0,09)**	-0,027 (0,101)
FABCE			0,329 (0,088)***	0,125 (0,096)
DANUBE			0,142 (0,121)	-0,024 (0,107)
UKIRL			0,202 (0,107)*	0,012 (0,105)
NEFAB			0,273 (0,092)***	0,198 (0,094)**
SWFAB			0,116 (0,09)	-0,22 (0,108)**
BALTIC			0,385 (0,12)***	0,248 (0,104)**
BLUEMED			0,094 (0,075)	-0,049 (0,072)
DKSEFAB			0,242 (0,107)**	0,005 (0,117)
COSTS				0,541 (0,136)***
WEALTH				-0,057 (0,184)
RES				-0,239 (0,161)
Korr. R ²	0,13	0,45	0,56	0,70
Akaike	-17,26	-45,57	-53,61	-75,29
N	63	63	63	63

Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***



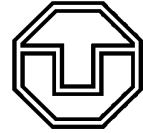
Der iterative Aufbau der Regression führt zu einer sukzessiven Verbesserung der Modellgüte, sowohl hinsichtlich des korrigierten Bestimmtheitsmaßes als auch des Akaike-Kriteriums. Im Gesamtmodell werden 70 % der Varianz aufgeklärt. Die Iterationsschritte I1 und I2 führen zu den erwarteten Einflüssen der Faktoren auf die Produktivität. Unter Einbeziehung von (primär) exogenen Faktoren entsprechen die Vorzeichen der Arbeitszeitvariable TIME und des Proxys für den vertikalen Verkehr VERT jedoch nicht mehr den Erwartungen. Beide Variablen können allerdings auch keine statistische Signifikanz nachweisen.

In Iteration 3 wird nachgewiesen, dass die Organisation des zugehörigen ANSP in einem FAB positiv auf die Produktivität wirkt. Für *FABEC*, *FABCE*, *UKIRL*, *NEFAB*, *BALTIC* und *DKSEFAB* ist diese Interdependenz zudem statistisch signifikant. In I4 dagegen ändert sich das Vorzeichen bei vier von neun FABs. Einen signifikant negativen Einfluss hat gemäß der Auswertung, sofern der korrespondierende ANSP dem SWFAB zugeordnet ist. Im Allgemeinen liefert I4 die am wenigsten plausiblen Ergebnisse. Analog zur ANSP Ebene sind die Variablen RES und WEALTH durch ein negatives Vorzeichen gekennzeichnet, wenngleich dieser Einfluss keine Signifikanz aufweist. Dagegen ist der Einfluss der Lotsenkosten erwartungsgemäß positiv und zudem statistisch signifikant.

Die Regressionsanalyse wurde auch auf Basis der originären Daten durchgeführt (Anhang, Tabelle A 13). Das korrigierte R^2 ist mit 0,72 nur geringfügig höher, das Regressionsmodell weist aber erneut eine geringere Güte hinsichtlich des Akaike-Kriteriums auf. In den ersten drei Iterationsschritten sind die Vorzeichen der Koeffizienten zwischen beiden Datensätzen identisch, unterscheiden sich aber in der Ausprägung und statistischen Signifikanz. Lediglich in der letzten Iteration unterscheiden sich die Vorzeichen bei den Variablen *SIZE* (+), *VERT* (-), *DENSITY* (± 0), *CHANGE* (+), *UKIRL* (-) und *DKSEFAB* (-). Im Falle der ersten beiden Variablen erhöht sich demnach die Plausibilität des Ergebnisses, jedoch ist das positive Vorzeichen bei *CHANGE* nicht erwartungskonform.

Wie vorab diskutiert wurde getestet, inwiefern die Substitution von *CHANGE* (Änderungsrate) mit *CHANGE2* (Wachstumsrate) die Ergebnisse ändert. Dabei war zu beobachten, dass *CHANGE2* in einigen Modellspezifikationen ein positives Vorzeichen aufwies. Dies würde bedeuten, dass die Abweichung der Nachfrage zum Vorjahr (*CHANGE*) generell negativ auf die Performance einwirkt (Tabelle 4-14). Man kann hier jedoch differenzieren, ob diese Änderung auf einem Verkehrszuwachs beruht (positiver Zusammenhang) oder auf einer Abnahme des Verkehrs (negativer Zusammenhang). Statistische Signifikanz konnte für *CHANGE2* jedoch in keinem der Fälle nachgewiesen werden.

Neben der Nutzung der einzelnen Faktoren wurde auch eine Regression auf Basis einer Hauptkomponentenanalyse durchgeführt. Diese führte jedoch weder zu gesonderten Erkenntnissen, noch zu einer Verbesserung der Modellgüte. Sie wird daher nicht gesondert betrachtet.



4.4.2.5 Panel Analyse

Für die Panelanalyse stehen ACC Daten für 62¹³⁸ DMUs über sieben Jahre (2011-2017) zur Verfügung. Sie orientiert sich am Regressionsmodell aus Abschnitt 4.4.2.4. Aufgrund fehlender Daten können jedoch die Variablen *VERT* und *CRUISE* nicht berücksichtigt werden. Zudem müsste für die Nutzung der Variable *CHANGE* das Jahr 2011 aus der Panelanalyse entfernt werden. Da in diesem Falle 62 Beobachtungen nicht berücksichtigt werden können, wurde auf die Verwendung der Änderungsrate des Verkehrs im Regressionsmodell verzichtet. Dies begründet sich auch in der fehlenden Signifikanz der Variable im uniperiodischen Modell.

Die Analyse wird auf Basis der normierten Werte durchgeführt. Gemäß Breusch-Pagan- und Hausman-Test wird eine FEM-Regression gegenüber dem REM oder Pooled OLS bevorzugt¹³⁹. Allerdings lässt der Fixed-Effects Schätzer aufgrund von Kollinearität keine Berücksichtigung der FABs zu. Aus diesem Grund wurden sowohl FEM als auch das REM angewendet.

Die Panel-Analyse wird in fünf Iterationsstufen ausgeführt. In I1 werden lediglich endogene Variablen berücksichtigt. Dieses Modell wird in I2 um die Jahres-Dummys ergänzt. Iteration 3 und 4 fügen die teilexogenen Variablen hinzu. Die beiden Iterationsschritte unterscheiden sich lediglich hinsichtlich der Betrachtung der FABs: Im dritten Schritt wird lediglich der FAB-Dummy integriert, Schritt 4 differenziert die einzelnen FABs. Der vierte Iterationsschritt ist daher nur im Random Effects Modell vorhanden. Iteration 5 umfasst alle Faktoren, inklusive den exogenen Einflüssen.

Tabelle 4-15 fasst die Ergebnisse für das FEM zusammen. Die Modellgüte steigt, mit Ausnahme des Akaike-Kriteriums in I2, mit zunehmender Iterationsstufe an. Zudem sind die Ergebnisse über alle Iterationsschritte hinweg robust, d. h., es gibt zwischen den einzelnen Stufen keine Vorzeichenwechsel in den Koeffizienten.

Sowohl die Konstante *INT* als auch die endogenen Variablen sind in allen Stufen statistisch signifikant. Erwartungskonform hat die Variable *STAFFB* einen negativen Einfluss auf die Produktivität, *SECT* einen positiven. Dagegen entsprechen die negativen Einflüsse von *TIME* und *SIZE* nicht den Erwartungen, bestätigen jedoch die Ergebnisse aus der uniperiodischen Betrachtung (siehe Tabelle 4-14, I4). Der Einfluss der Jahres-Dummys ist gering und, bis auf 2017 sowie teilweise 2016, statistisch nicht signifikant.

Der Einbezug der (teilexogenen) Nachfragedaten erhöhte die Modellgüte des Regressionsmodells und bestätigt ebenfalls die Ergebnisse der iterativen OLS-Regression in Abschnitt 4.4.2.4. Beide Variablen wirken positiv auf die Produktivität ein und sind zudem statistisch signifikant. Die exogenen Faktoren wirken, mit Ausnahme von *WEALTH*, entsprechend den Erwartungen auf die Leistungsfähigkeit ein. Auch die Technologiekomponente (*RES*) ist, im Gegensatz zur uniperiodischen Betrachtung, positiv konnotiert. Signifikant sind jedoch lediglich die Interdependenzen zwischen der Produktivität und den Arbeitskosten der Lotsen (positiv) bzw. dem Wohlstand (negativ). Letztere Beobachtung ist somit

¹³⁸ Die ACCs Tiflis und Simferopol sind nicht enthalten, da die Daten nicht für alle Jahre zur Verfügung stehen.

¹³⁹ Dies ist sowohl bei den originären als auch bei den normierten Daten der Fall.



konsistent über mehrere Produktivitäts-¹⁴⁰ und Regressionsmodelle¹⁴¹ sowie deren Anwendung auf unterschiedlichen operativen Ebenen.

Tabelle 4-15: Iterative Panelanalyse ACC Ebene, Fixed Effects

	<i>I1</i>	<i>I2</i>	<i>I3</i>	<i>I5</i>
INT	0,615 (0,026)***	0,618 (0,026)***	0,432 (0,039)***	0,403 (0,046)***
STAFFB	-0,166 (0,02)***	-0,173 (0,019)***	-0,146 (0,019)***	-0,142 (0,019)***
TIME	-0,242 (0,05)***	-0,238 (0,051)***	-0,235 (0,048)***	-0,224 (0,048)***
SIZE	-0,58 (0,146)***	-0,585 (0,14)***	-0,347 (0,144)**	-0,342 (0,144)**
SECT	0,235 (0,047)***	0,205 (0,045)***	0,086 (0,048)*	0,096 (0,048)**
2012		-0,001 (0,007)	-0,001 (0,006)	-0,004 (0,006)
2013		-0,006 (0,007)	-0,004 (0,007)	-0,006 (0,007)
2014		0 (0,007)	-0,001 (0,006)	-0,002 (0,006)
2015		0,01 (0,007)	0,008 (0,006)	0,003 (0,007)
2016		0,012 (0,007)*	0,012 (0,006)*	0,006 (0,007)
2017		0,03 (0,007)***	0,031 (0,007)***	0,025 (0,007)***
DENSITY			0,421 (0,113)***	0,392 (0,114)***
TRANS			0,361 (0,063)***	0,362 (0,063)***
COSTS				0,103 (0,038)***
WEALTH				-0,18 (0,107)*
RES				0,09 (0,113)
R ² within	0,33	0,39	0,46	0,47
Akaike	-1581,52	-1465,77	-1624,02	-1627,50
N	434	434	434	434

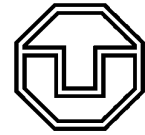
Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***

Tabelle 4-16 stellt die Ergebnisse der REM-Regression dar. Im Wesentlichen bestätigt es die Ergebnisse des Fixed Effects Modells und weist lediglich kleinere Unterschiede hinsichtlich der Ausprägung von Koeffizienten und deren statistischer Signifikanz auf. So ist bspw. die endogene Variable *SIZE* nicht signifikant. Gleiches gilt für *SECT* in den letzten drei Iterationsschritten. Mit Ausnahme des Jahres-Dummys für 2014 sind die Ergebnisse robust hinsichtlich des Vorzeichens.

Die Verwendung eines REM ermöglicht die Integration der FABs in die Regressionsanalyse. Iteration 3 zeigt, dass die Organisation des korrespondierenden ANSP in einem FAB positiv bewertet wird. Dieses Ergebnis ist statistisch signifikant. I4 und I5 verdeutlichen, dass dabei keine Abhängigkeit von

¹⁴⁰ Gemeint sind Produktivitätswerte und DEA-Scores.

¹⁴¹ Gemeint sind die verschiedenen Regressionstypen, z. B. FEM, Trunkated, OLS, usw.

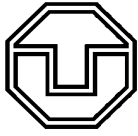


dem FAB besteht: Alle Variablen weisen hier ein positives Vorzeichen auf. In Bezug auf I5 ist dieser Einfluss zudem, mit Ausnahme von FABEC, UKIRL und DKSEFAB, statistisch signifikant.

Tabelle 4-16: Iterative Panelanalyse ACC Ebene, Random Effects

	I1	I2	I3	I4	I5
INT	0,548 (0,032)***	0,553 (0,032)***	0,23 (0,055)***	0,221 (0,052)***	0,208 (0,051)***
STAFFB	-0,162 (0,02)***	-0,169 (0,02)***	-0,142 (0,019)***	-0,137 (0,019)***	-0,132 (0,019)***
TIME	-0,22 (0,05)***	-0,219 (0,05)***	-0,212 (0,047)***	-0,222 (0,047)***	-0,209 (0,047)***
SIZE	-0,052 (0,101)	-0,075 (0,099)	-0,09 (0,099)	-0,066 (0,1)	-0,052 (0,099)
SECT	0,156 (0,043)***	0,133 (0,042)***	0,026 (0,042)	0,045 (0,042)	0,049 (0,042)
2012		-0,001 (0,007)	-0,001 (0,006)	-0,001 (0,006)	-0,004 (0,007)
2013		-0,005 (0,007)	-0,004 (0,007)	-0,005 (0,007)	-0,007 (0,007)
2014		0,001 (0,007)	0 (0,007)	-0,001 (0,007)	-0,003 (0,007)
2015		0,011 (0,007)	0,01 (0,007)	0,009 (0,007)	0,002 (0,007)
2016		0,013 (0,007)*	0,014 (0,007)**	0,013 (0,007)**	0,006 (0,007)
2017		0,03 (0,007)***	0,035 (0,007)***	0,034 (0,007)***	0,027 (0,007)***
DENSITY			0,271 (0,09)***	0,31 (0,092)***	0,284 (0,092)***
TRANS			0,46 (0,058)***	0,467 (0,058)***	0,469 (0,057)***
FAB			0,195 (0,053)***		
FABEC				0,099 (0,06)*	0,079 (0,063)
FABCE				0,306 (0,073)***	0,276 (0,071)***
DANUBE				0,233 (0,107)**	0,213 (0,103)**
UKIRL				0,122 (0,078)	0,111 (0,078)
NEFAB				0,271 (0,071)***	0,295 (0,077)***
SWFAB				0,207 (0,073)***	0,146 (0,074)**
BALTIC				0,397 (0,106)***	0,384 (0,103)***
BLUEMED				0,161 (0,066)**	0,144 (0,064)**
DKSEFAB				0,172 (0,091)*	0,144 (0,091)
COSTS					0,124 (0,038)***
WEALTH					-0,185 (0,096)*
RES					0,029 (0,095)
Akaike	-185,06	-168,26	-349,21	-436,67	-462,07
N	434	434	434	434	434

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%***, 99%***



Für die Panel-Analyse auf ACC Ebene gelten die gleichen Einschränkungen wie bei der Analyse auf ANSP Ebene. Zwar stehen mit 63 bzw. 62 DMUs fast die doppelte Anzahl an Beobachtungen zur Verfügung, jedoch fehlen aufgrund von Datenverfügbarkeit wesentliche Einflussgrößen, die bspw. Komplexität, Militär und Wetterphänomene repräsentieren.

Die mangelnde Interpretierbarkeit einiger Faktoren kann darin begründet liegen, dass andere wesentliche Einflüsse oder signifikante Unterschiede zwischen den ACCs, z. B. auf institutioneller Ebene, nicht quantitativ erfasst werden können. Diese werden mittels einer qualitativen Untersuchung im nachfolgenden Abschnitt aufgeführt. Zudem werden Daten auf Sektorebene ausgewertet.

4.4.3 Qualitative Analyse ACC

4.4.3.1 Stakeholder Konsultationen

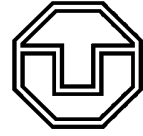
Um die betrieblichen Praktiken der ACCs zu erfassen, wurde ein Fragebogen entwickelt, welcher Aspekte der Kapazitäts- und Personalplanung sowie der verwendeten Systeme und Tools adressierte. Dadurch sollten wesentliche Unterschiede zwischen den ACCs identifiziert und potentielle best-practices abgeleitet werden. Die Fragen wurden im Rahmen von Konsultationen in den ACCs mit Experten des operativen Ablaufes diskutiert und durch Informationen sowie Beobachtungen im Kontrollraum ergänzt. Die Ergebnisse dieser Studie wurden in FABEC (2019) veröffentlicht.

Bei Erstellung des Fragebogens sowie dessen Analyse muss berücksichtigt werden, dass insbesondere die Kapazitätsplanung verschiedene zeitliche Horizonte umfassen kann. Daher ist eine Differenzierung in strategische und taktische sowie prä-taktische Maßnahmen notwendig. Ergänzend wurden die durch NEST berechneten Daten auf Sektorebene validiert (siehe Abschnitt 4.4.3.5).

Zunächst fand im Jahr 2016 ein Vergleich zwischen den beiden ACCs statt, welche ausschließlich den oberen Luftraum kontrollieren: Maastricht-UAC und der UAC Karlsruhe. Dadurch sollte eine betriebliche Vergleichbarkeit der Verkehrscharakteristiken sichergestellt und die Unterschiede in den Leistungs- und Kapazitätsparametern erklärt werden. In den Jahren 2017 bis 2019 wurden die verbleibenden ACCs des FABEC-Raumes in die Analyse einbezogen. Diese umfassen: Karlsruhe, Maastricht UAC, Brüssel, Langen, München, Amsterdam Bremen, Genf, Zürich, Brest, Marseille, Reims, Bordeaux und Paris.

Der Fragebogen beinhaltet 30 Fragen, welche in fünf Bereiche zugeordnet werden: Personal, Luftraum, ATFM, Nachfrage und Betrieb. Die Ergebnisse werden in zwei Gruppen aufgeteilt. Zum einen werden strategische Kapazitätstreiber und -hemmer untersucht, welche es ermöglichen, eine bestimmte angestrebte Sektorkapazität zu erreichen. Zum anderen sind die taktischen und prä-taktischen Kapazitätstreiber und –inhibitoren, die es dem ACC ermöglichen, den Luftraumnutzern am Tag des Betriebs die nachgefragte Kapazität bereitzustellen, adressiert. Die Kapazität kann entweder durch die Einflüge pro Stunde gemessen werden (Entries per hour) oder durch die Anzahl der Luftfahrzeuge, welche sich gleichzeitig im Luftraum befinden (Occupancy).

Nach Abschluss dieser Datenaufnahme zu den betrieblichen Praktiken wurden die Ergebnisse gegenübergestellt, um die Heterogenität darzustellen sowie die daraus resultierenden Auswirkungen



auf die Produktivität und Kapazität abzuleiten. Im Folgenden werden die wesentlichen Unterschiede in Methoden und Tools präsentiert. Inhalt und Struktur orientieren sich an FABEC (2019).

4.4.3.2 Strategischer Planungshorizont

Auf Sektorebene nutzen ACCs unterschiedliche Methoden zur Bestimmung des Kapazitätsbedarfs. Dies beinhaltet u.a. eine regelmäßige Überprüfung, ob der deklarierte Kapazitätswert (Capacity Default Value, CDV) den (sich stetig ändernden) Verkehrscharakteristiken entspricht. Bei Bedarf wird der CDV angepasst. Hierfür nutzen ACCs sowohl CAPAN¹⁴², eine Methodik zur Bewertung der Sektorenkapazität (Cook 2016; EUROCONTROL 2016c), als auch Echt- und Schnellzeit-Simulationen. Brüssel nutzt neben CAPAN auch BADA¹⁴³-Daten (EUROCONTROL 2020b; Nuic et al. 2010), Amsterdam lediglich ein Workload-Model. Das Monitoring erfolgt nach Einflügen pro Stunde bzw. Occupancy Count und/oder Überlastungsberichten der Lotsen. Die französischen ACCs nutzen zudem das Tool SALTO (DSNA 2018).

Ergänzend wird in einigen ACCs unter Mithilfe von operativen Experten der Kapazitätswert angepasst. Im ACC Bremen werden beispielsweise für jede Sektorfamilie (EBG) Arbeitsgruppen gebildet, um operative Optimierungen oder Umgestaltungen zu diskutieren. Auch mögliche EBG-übergreifende Initiativen werden diskutiert. Zusätzlich gibt es eine Expertengruppe, die für eine strategische Planung des gesamten Luftraums der Kontrollzentrale verantwortlich ist. Hier könnte u. a. die dynamische Sektorisierung (Standfuss et al. 2018d) einen Beitrag leisten und die Methodik anhand realer Untersuchungsräume angepasst und angewendet werden. Arbeitsgruppen werden auch in den ACCs Paris, Brest und Marseille eingesetzt. In Bezug auf die Kapazitätsüberwachung ist zu beachten, dass das Monitoring nach Occupancy Counts in Sektoren, die einen hohen Anteil an Abflügen aufweisen, aufgrund von häufigen und kurzfristigen Schwankungen nicht zuverlässig ist.

Eine wesentliche Unterstützung bei der Nutzung der Luftraumkapazität bieten ATS-Systeme. Die ACCs der DFS und LVNL sind mit dem Indra iTEC-System ausgestattet, das in Zusammenarbeit mit ENAIRE und NATS entwickelt wurde. Maastricht betreibt ein speziell auf die individuellen Bedürfnisse angepasstes Indra-System, Belgocontrol analog dazu nutzt ein Thales-System. Die ATS-Systeme unterscheiden sich teilweise im Funktionsumfang (z. B. What-If-Probing und Konflikterkennung) und Anpassungsmöglichkeiten (z. B. Elektronische Kontrollstreifen, Anpassbares Interface).

In den Konsultationen wurde deutlich, dass die Konflikterkennung unter Umständen Probleme verursacht. Die Systeme zur mittelfristigen Konflikterkennung berücksichtigen keine Änderung der Routings. Dies führt im unteren Luftraum häufig zu Fehlalarmen.

¹⁴² ATC Capacity Analyser Tool.

¹⁴³ Base of Aircraft Data.



4.4.3.3 Personalplanung und -einsatz

Um die Kapazität langfristig zu erhöhen oder Personalfluktuationen¹⁴⁴ auszugleichen, ist die Ausbildung neuer Lotsen notwendig. Da das vollständige Schulungsprogramm (je nach ACC und sektorenbezogener Komplexität) 2,5 bis 5 Jahre dauern kann, ist eine frühzeitige Planung notwendig. Die niederländische LVNL bereitet hierfür einen 10-Jahres-Plan vor, wofür auch die Demografie genau analysiert wird. Das ACC Zürich überprüft zweimal jährlich die langfristige Personalplanung. Die DFS führt viermal im Jahr strategische Besprechungen zur Personalbesetzung für alle ACCs durch, wobei eine standardisierte Methodik zur Festlegung der erforderlichen Personalstärke verwendet wird. In Frankreich wird die strategische Planung auf nationaler Ebene verwaltet. Die Zahl der Mitarbeiter wird vom Verkehrsministerium festgelegt. Die Zahl der erforderlichen ATCOs wird dann jährlich angepasst. Diese Anpassungen berücksichtigen die Verkehrsentwicklung, Personal im Ruhestand sowie eine statistische Prognose des „nicht verfügbaren“ Personals (z. B. aufgrund von Schwangerschaft).

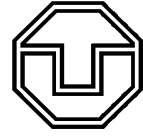
Generell sollten nach Expertenmeinung bei der Erstellung eines langfristigen Personalplans die folgenden Faktoren berücksichtigt werden:

- Luftraummerkmale: Im oberen Luftraum hängt die Personalplanung stärker von der Verkehrsentwicklung ab als im unteren Luftraum
- Prognostizierte Verkehrsbelastung, die sich auf die erforderliche Sektorisierung und Personalausstattung auswirkt
- Arbeitszeiten sowie Urlaubsanforderungen und –ansprüche
- Natürliche Personal-Fluktuation, u. a. Anzahl der Controller, die in den Ruhestand gehen oder den Job wechseln möchten
- Demografie der Besetzung, z. B. wegen Elternzeit oder Krankheitsquote

Die langfristige Personalplanung stellt die ACCs vor unterschiedliche Herausforderungen. Dazu gehören Unsicherheiten in Bezug auf prognostizierten und tatsächlichen Verkehr sowie die damit notwendige Personalpuffer, siehe auch Hergert und Thießen (2014) sowie Standfuss und Whittome (2019b). Allerdings ist die optimale Personalpuffergröße nicht klar definiert. Die Kontrollzentralen verwenden oft einen historischen Puffer, der jedoch nicht betrieblich validiert wird. Die Schweizer ACCs Genf und Zürich sehen eine große Herausforderung in den hohen Ausfallraten bei der Lizenzierung der Lotsen (über 50 %). In der Vergangenheit verzeichnete Maastricht ebenfalls hohe Abbruchquoten, konnte diese jedoch senken. Die Lotsenausbildung für die FIR Bremen wies eine Ausfallquote von lediglich 20 % auf. In Frankreich spielen Attraktivitätsfaktoren eine Rolle, da Lotsen bestimmte Standorte bevorzugen. Zusätzlich wird in der Öffentlichkeit zunehmend die Erhöhung der ATM-Kapazität als Konflikt zum Umweltschutz gesehen.

Die kurzfristige Dienstplanung ist insbesondere von der Größe des Controller-Pools sowie dessen Einsatz-Flexibilität abhängig. Jeder Lotse ist für eine bestimmte Anzahl an Sektoren einsetzbar bzw. lizenziert. Je größer die Anzahl der lizenzierten Sektoren ist, desto flexibler kann ein ATCO eingesetzt

¹⁴⁴ Z. B. wegen Rente.



werden. Jedoch erhöht dies die Ausbildungsdauer und verkompliziert die Aufrechterhaltung der Lizenzen. In den vergangenen Jahren ist ein Trend zum „Cross-Training“ sichtbar, d. h. dass Lotsen für verschiedene Sektorfamilien bzw. EBGn ausgebildet werden. Insgesamt sind jedoch starke Unterschiede zwischen den ACCs identifiziert worden. Im Brüsseler ACC (bestehend aus zwei EBGn) waren zum Zeitpunkt der Studie 30 ATCOs „cross-trained“. Alle neuen ATCOs werden ebenfalls für beide Sektorfamilien ausgebildet.

In Amsterdam können alle Fluglotsen für alle Sektoren eingesetzt werden. Es wurde jedoch festgestellt, dass das nationale Arbeitsrecht hier limitierender Faktor ist. Das ACC München nutzt allgemein große Sektorfamilien. Jede EBG ist hier mit ca. 80 ATCOs besetzt, die jeweils für zwei bis drei EBGn einsetzbar sind. Im ACC Genf werden ebenfalls alle Lotsen für alle Sektoren ausgebildet. Dagegen sind in Bremen Lotsen innerhalb einer Sektorfamilie nicht für alle Sektoren einsetzbar, was die Personaleinsatzplanung erschwert. In Zukunft wird die Luftraumstruktur angepasst, um die Lizenzierung und damit die Planungsprozesse zu vereinfachen.

Da Lotsen in Schichten arbeiten, ist auch hier die Anpassungsfähigkeit von der Flexibilität abhängig. Dies hängt unter anderem von Anreizmöglichkeiten ab, z. B. für Überstunden oder Teilzeit-Arbeit. In Brüssel sind Schichten zeitlich festgelegt, die Anzahl der Lotsen ist jedoch flexibel. Es werden Stand-by-Schichten verwendet, um Verkehrsschwankungen abzufedern. Die Stand-by Schichten werden, genau wie im ACC Amsterdam, extra vergütet. Im Münchner ACC sind rund 30 % der Lotsen Teilzeitbeschäftigte. Die Bereitschaft für zusätzliche Schichten ist sehr hoch. In Langen werden Standby-Schichten aufgrund des Tarifvertrags nicht genutzt. Daher wird für die meisten vor Ort vorhandenen Sektorfamilien ein Puffer-ATCO eingesetzt. Genf und Zürich bieten unbezahlten Urlaub an, wenn keine Fluglotsen erforderlich sind, und eine 200%ige Überstundenzahlung, wenn eine zusätzliche Schicht geleistet wird.

In den ACCs werden verschiedene Tools genutzt, um die Personaleinsatzplanung zu unterstützen. Diese dienen zur Analyse der Verkehrsnachfrage (z. B. NEST, CHMI¹⁴⁵ oder SALTO) und unterstützen bei der Dienstplangenerierung (u. a. BIOMAN und ePEP). Sie unterscheiden sich in der Funktionalität und basieren auf verschiedenen Plattformen (z. B. Excel). Es konnte nicht eindeutig geklärt werden, welches Tool die beste Funktionalität aufweist.

4.4.3.4 Kurzfristige Maßnahmen und Herausforderungen

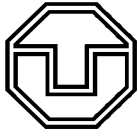
Die Möglichkeiten, auf kurzfristige Änderungen des Verkehrs zu reagieren, sind von ACC zu ACC unterschiedlich. Sie hängen maßgeblich von den verwendeten Flow Management Tools¹⁴⁶, der Sektorisierung¹⁴⁷ und den ATFM-Maßnahmen¹⁴⁸ ab. ATFM-Tools verbessern die Kapazitätsbereitstellung signifikant. Da die Tools größtenteils an die Bedürfnisse des betreffenden ANSP bzw. der spezifischen

¹⁴⁵ Collaboration Human Machine Interface.

¹⁴⁶ Zur Unterstützung der Abschätzung des Traffic Flows und der daraus resultierenden Arbeitslast.

¹⁴⁷ Gemeint ist die Flexibilität durch die Möglichkeit, Sektoren zu öffnen oder zu schließen.

¹⁴⁸ Regulatorische Maßnahmen werden ACC-übergreifend als „Last resort“ bezeichnet.



ACC angepasst sind, kann jedoch keines der in Anhang Tabelle A 14 dargestellten Tools als „best case“ angesehen werden.

Neben den verwendeten Tools trägt auch die Möglichkeit zur Öffnung und Schließung von Sektoren zur Flexibilität bei. Die Sektoren der ACCs Brüssel, Genf und Zürich müssen bspw. für mindestens 20 Minuten, in den ACCs Maastricht, Langen und München für 30 Minuten geöffnet sein. Dagegen beträgt die minimale Öffnungszeit für Amsterdam ACC lediglich 5 Minuten. In Bremen kann jeder Sektor zu jedem Zeitpunkt und für jeden beliebigen Zeitraum geöffnet werden. In Karlsruhe können Sektoren kurzfristig für mindestens eine Stunde geöffnet werden, dies variiert jedoch innerhalb der EBGn. Üblicherweise werden Sektoren alle 15¹⁴⁹ oder 30 Minuten geöffnet. Ein Sektor-Split führt allerdings auch immer zu einer Verringerung der maximal möglichen Kapazität (Standfuss et al. 2018c).

Alternativ zu einer Teilung des Sektors kann der Verkehr reguliert¹⁵⁰ werden. Die Regulierung eines Sektors wird von allen ACCs allerdings lediglich als „letztes mögliches Mittel“ gesehen, z. B. wenn zu wenig Personal zur Verfügung steht um einen weiteren Sektor zu öffnen. Hier wird eine engere Kooperation mit den Airlines angestrebt, um herauszufiltern, welche Flüge gesteuert werden sollen. Der Schwellenwert richtet sich entweder nach dem Occupancy oder dem Entry Count, in wenigen Fällen werden beide Indikatoren verwendet.

In Karlsruhe wurde ein Watch-Value-Verfahren implementiert. Neben dem angegebenen Kapazitätswert (CDV) hat jeder Sektor einen "Watch-Value low" (niedriger als CDV) und "Watch-Value high" (höher als CDV). Innerhalb dieses Rahmens kann der Supervisor oder die Flow Management Position (FMP) entscheiden, ob der Sektor geöffnet bzw. geschlossen werden oder eine Regulierung stattfinden soll. Schlüsselfaktoren hierfür sind u. a. Wetterbedingungen, Verkehrsnachfrage und -mix sowie die Fluktuation des Verkehrs.

Eine wesentliche Einschränkung der kurzfristigen Planbarkeit wurde von den operativen Experten im unvorhergesehenen Verkehr in Form von „Intrudern“¹⁵¹ oder „Runaways“¹⁵² gesehen. Diese erschweren eine effiziente Ressourcenplanung zur Bereitstellung der Kapazität. Dies kann durch Änderung der Flugpläne, Verspätungen oder Verkehrssteuerungsmaßnahmen der Upstream-Sektoren¹⁵³ bedingt sein. Innerhalb eines Zeithorizontes von zwei bis drei Stunden ist die Prognose etwa 80 % genau (Brüssel und Bremen).

Taktische „directs“ können in diesem Zusammenhang die Effizienz des Sektors erhöhen und so in Überlast-Situationen auch eine Entlastung schaffen. Durch diese Abkürzungen wird ein Teil der Flugstrecke (und damit Treibstoff und Emissionen) eingespart. Diese kann von den Lotsen angewiesen oder von den Piloten angefordert werden. Es ist jedoch auch möglich, dass die Kapazitätsplanung in den

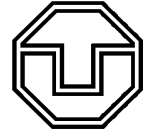
¹⁴⁹ Z. B. 10:00, 10:15, 10:30, etc.

¹⁵⁰ Gemeint ist das Anwenden einer Steuerungsmaßnahme.

¹⁵¹ Ein ungeplanter Flug möchte den Sektor durchfliegen.

¹⁵² Ein geplanter Flug durchquert nicht oder nicht zur geplanten Zeit in den Sektor.

¹⁵³ Upstream-Sektoren sind jene Sektoren, welche das LFZ bereits durchflogen hat.



Downstream-Sektoren¹⁵⁴ dadurch gestört wird. Hinsichtlich der Erzielung eines effizienten Gleichgewichts zwischen tactical directs und der Abfolge des Flugplans wurde keine Konsistenz zwischen den ACCs gefunden. Es wird jedoch als wichtig angesehen, die Fluggesellschaften für die Problematik zu sensibilisieren, sodass die eingereichten Flugpläne denjenigen entsprechen, die die Fluggesellschaften tatsächlich fliegen.¹⁵⁵ Darüber hinaus wurden umfassende Informationen zum Operational Air Traffic (OAT) als entscheidend für eine effiziente Ressourcenplanung hervorgehoben.

Auch der Einfluss des Wetters wurde in den Analysen adressiert. Wetterereignisse waren im Jahr 2017 für 23,2 % der europäischen ATFM-Verspätungen verantwortlich. Die ANSPs und ACCs streben eine breitere Zusammenarbeit auf globaler (ICAO, WMO) und regionaler Ebene (MetAlliance) an. Amsterdam ACC verfügt über ein System, welches das aktuelle Wetter aufzeichnet und die weitere Entwicklung vorhersagt. Integriert sind der Flughafen Amsterdam Schiphol und das Kapazitätsmanagement. Zusätzlich verfügt Amsterdam ACC über eine MET-Position innerhalb des Kontrollraums.

4.4.3.5 Sektorkapazität

Neben der qualitativen Analyse wurde eine Datenbank mit betrieblichen Daten auf Sektorebene erstellt und innerhalb der Konsultationen validiert. Dazu wurde NEST verwendet. Da dieses Tool jedoch nicht die tatsächlichen Sektoröffnungsschemen verwendet¹⁵⁶, musste hier von einer signifikanten Abweichung zwischen ausgegebenen und tatsächlichen Daten ausgegangen werden. Daher war es notwendig, die tatsächlichen Sektoröffnungszeiten direkt von den ANSPs und ACCs anzufordern. Dadurch konnten mittels NEST verschiedene operative Parameter ausgewertet werden. Die Daten umfassen einen vierwöchigen Zeitraum im Sommer 2017. Ziel der Analyse ist es, die Unterschiede in der Sektorkapazität und Produktivität zwischen den ACCs zu validieren und die Faktoren zu verstehen, die zu diesen Unterschieden beitragen.

Zunächst werden die ACCs nach operativen Gesichtspunkten gegenübergestellt. Brest und Marseille kontrollieren den größten Luftraum innerhalb FABECs, sowohl hinsichtlich Fläche als auch Volumen. Dem gegenüber ist der Luftraum der ACC Brüssel am kleinsten. Die Nachfrage (gemessen in Flugstunden) ist wiederum in Karlsruhe und Maastricht am höchsten (siehe Abbildung 4-9) und zwischen 2012 und 2017 stetig gewachsen. Zudem gehören, gemäß PRU-Berechnungen, acht FABEC ACCs zu den 12 ACCs mit der höchsten Komplexität.

Tabelle 4-17 zeigt die deskriptiven Statistiken für eine Auswahl von Parametern¹⁵⁷. Sie verdeutlicht die Heterogenität der betrachteten Einheiten. Jedoch ist zu beachten, dass die Daten ein bereits gesteuertes System repräsentieren (unterschiedlicher Grad der Sektorisierung). Die Streuung der Zahlen ist für einige Parameter wie Luftraumgröße oder Öffnungszeit sehr hoch. Ein hoher Unterschied zwischen dem Maximum und dem 95 % -Quantil weist auf Extremwerte hin, wie dies bei CDV der Fall ist. Es kann davon ausgegangen werden, dass ein CDV von 121 einen Messfehler darstellt.

¹⁵⁴ Downstream-Sektoren sind jene Sektoren, welches das LFZ noch durchfliegen wird.

¹⁵⁵ Abweichungen „planned“, „initial“ und „actual“.

¹⁵⁶ NEST nutzt die Sektoröffnungsschemen aus den geplanten Konfigurationen.

¹⁵⁷ Exklusive Französischer ACCs.

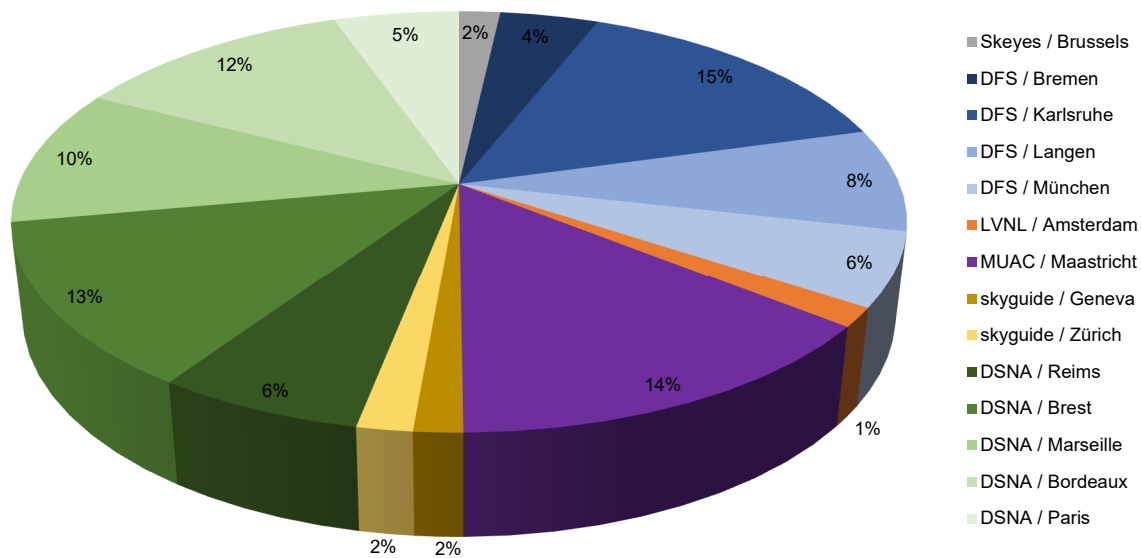
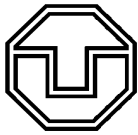


Abbildung 4-9: Anteil der ACCs an der Nachfrage innerhalb FABECs

Neben der Heterogenität besteht eine der entscheidenden Eigenschaften der Daten darin, dass sie durch ein vorab „gesteuertes“ System erfasst wurden. Das heißt, der Datensatz umfasst sowohl Elementarsektoren als auch den Verbund aller Sektoren einer EBG. Man kann dies als „Grad der Sektorisierung“ verstehen. Dessen Ausprägung steigt mit der Nachfrage und der damit verbundenen Teilung von Sektoren. Diese „Zersplitterung“ ist jedoch nicht homogen, sodass bspw. ein dreimaliges Teilen in ACC A vergleichbar ist mit einem einmaligen Split in ACC B, z. B. hinsichtlich der Luftraumgröße. Daher wird vermutet, dass eine zweidimensionale Analyse¹⁵⁸ nicht aussagekräftig ist. Dies wird auch an der Streuung von Werten deutlich, wie Abbildung 4-10 zeigt.

Kapazität kann, wie bereits beschrieben, u. a. in Einflügen pro Stunde gemessen werden. Abbildung 4-10 zeigt die durchschnittlichen Verteilungen der Einflüge pro Stunde, die in einzelnen Sektoren innerhalb der ACCs stattfanden¹⁵⁹. Um Extremwerte auszuschließen, wurde für die Analysen hauptsächlich das 95. Perzentil genutzt. Das Diagramm beinhaltet zudem den CDV, der als deklarierte Sektorkapazität verwendet wird. Die Abbildung zeigt, dass die höchsten Kapazitäten im oberen Luftraum erzielt werden. Bordeaux weist die höchste Streuung auf, was auf eine hohe Variabilität in der Nachfrage hinweist. Ähnliche Beobachtungen sind auch für die Kapazitätsbewertung nach Occupancy möglich. Im oberen Luftraum ist die Zahl der gleichzeitig im Luftraum befindlichen Flugzeuge deutlich höher als im unteren Luftraum. Hier weist insbesondere MUAC hohe Werte auf. Im unteren Luftraum sind die Differenzen der Occupancy counts relativ gering (siehe Anhang, Abbildung A 20).

¹⁵⁸ Interdependenz zwischen zwei Faktoren, z. B. Kapazität und Komplexität.

¹⁵⁹ Gewichtet nach Dauer der Sektoröffnung.

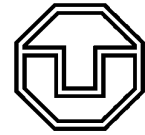


Tabelle 4-17: Deskriptive Statistik Sektordaten

	Min	Mittelwert	Median	95 % Q	Max	Stabw.
Fläche (km ²)	1.823	20.160	15.423	45.222	120.552	15.878
Volumen (km ² x 100 feet)	249.983	2.855.095	1.857.241	9.308.362	19.451.884	3.137.024
Höhenbänder	20	166	163	315	450	86
Öffnungszeit (h)	34	220	199	436	672	141
Anzahl Sektor-Öffnungen	2	53	49	102	239	29
CDV	16	45	44	62	121	14
Ø Nachfrage pro Tag	19	248	177	641	1.429	207
Ø Occupancy	0,3	4,4	3,8	9,5	15,8	2,6
Ø Entry Counts [20min]	0,6	10,0	9,3	18,6	22,7	4,5
Ø Entry Counts [60min]	1,3	29,9	27,7	57,9	69,0	14,2
ATCO Hours on Board	69	433	333	887	1.619	297
No. of level changes	2	307	159	901	4.402	423
Distanz (km)	29.188	592.717	423.597	1.571.267	4.677.043	613.692
Gesamtflugzeit (h)	79	826	602	2.121	6.582	829
Ø Geschwindigkeit (km/h)	285	711	779	851	890	148
Anteil steigender Flüge	5 %	35 %	36 %	56 %	83 %	15 %
Anteil sinkender Flüge	7 %	36 %	35 %	62 %	88 %	17 %
Anteil Streckenflüge	0 %	29 %	17 %	71 %	78 %	25 %
Anteil LWT ¹⁶⁰ LFZ	0 %	4 %	2 %	13 %	39 %	5 %
Anteil MWT ¹⁶¹ LFZ	60 %	84 %	85 %	94 %	96 %	8 %
Anteil HWT ¹⁶² LFZ	0 %	12 %	11 %	31 %	40 %	8 %
Lotsenproduktivität ¹⁶³	0,14	1,99	1,63	4,42	7,28	1,21

Wie bereits in der qualitativen Analyse diskutiert, hat die Komplexität des Verkehrs eine entscheidende Bedeutung für die Kapazität sowie den Zeitpunkt von Regulierungsmaßnahmen. Für die Untersuchungen stand jedoch lediglich eine Metrik auf ANSP und ACC Ebene zur Verfügung (EUROCONTROL 2006, 2008a). Von den Experten wird vor allem der Vertikalverkehr als komplexitätsfördernd beschrieben. Im Umkehrschluss würde das bedeuten, dass sich der Anteil des Streckenverkehrs¹⁶⁴ positiv auf die Kapazität auswirkt. Abbildung 4-11 zeigt, dass dies nicht nachgewiesen werden kann. Gleiches gilt für den Zusammenhang zwischen Streckenverkehr und Occupancy.

Aufgrund der Problematik des „gesteuerten Systems“ wurde als alternative abhängige Variable die Verkehrsdichte verwendet. Zudem wurden basierend auf den verfügbaren Daten die in Anhang Tabelle A 15 dargestellten Metriken als grobe Approximation der Sektor-Komplexität entwickelt. Ein Zusammenhang konnte jedoch in keinem der Fälle nachgewiesen werden. Wesentliche Bestandteile, z. B. die Anzahl der Kreuzungspunkte oder der Anteil militärischen Verkehrs, sind als Komplexitätsfaktoren aufgrund fehlender Datenverfügbarkeit nicht berücksichtigt.

¹⁶⁰ Low wake turbulence.

¹⁶¹ Medium wake turbulence.

¹⁶² High wake turbulence.

¹⁶³ Die Produktivität entspricht den kontrollierten Flugstunden pro Lotsenstunde.

¹⁶⁴ „share of cruising flights“.

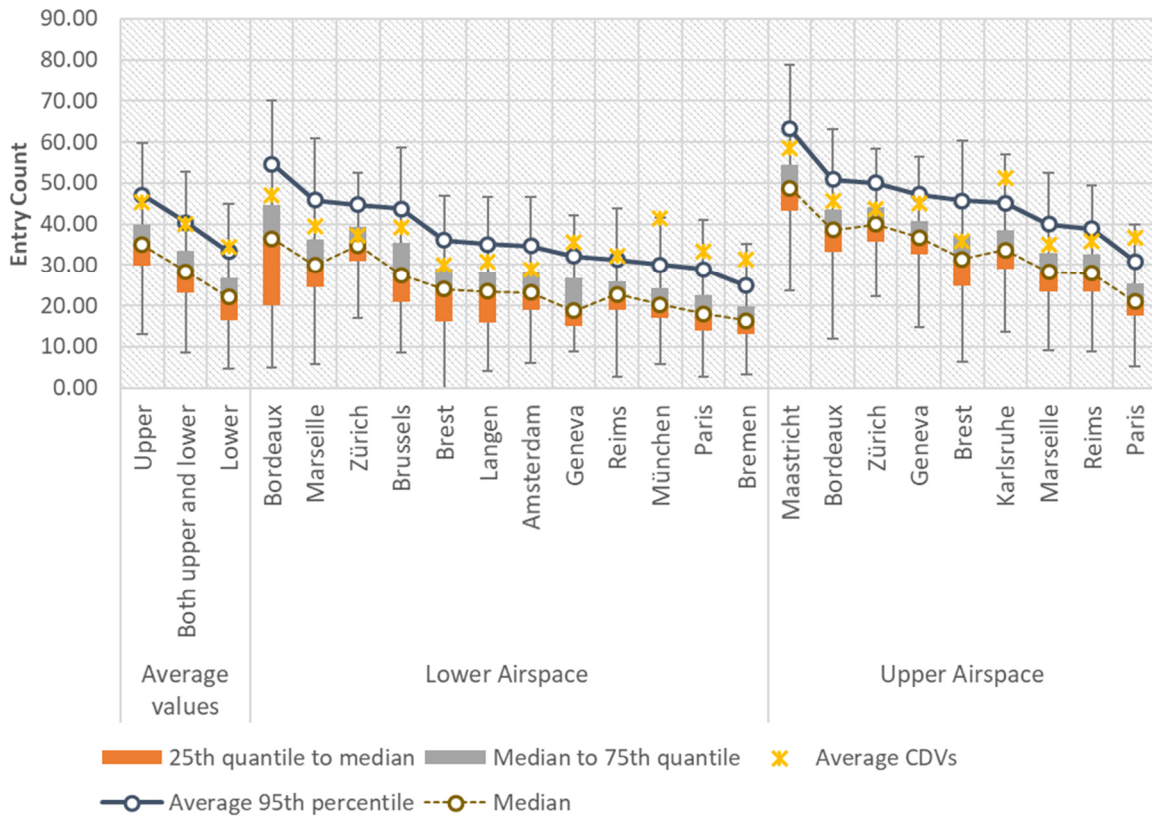
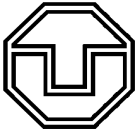


Abbildung 4-10: Kapazität einzelner ACCs (Entry Counts)
(FABEC 2019)

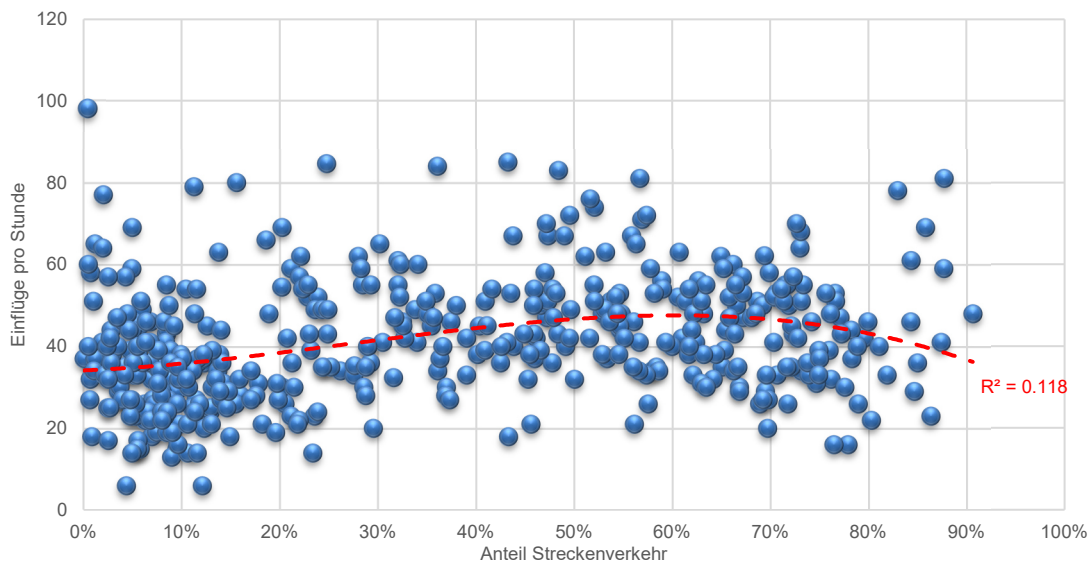
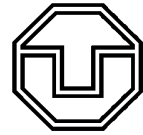


Abbildung 4-11: Interdependenz zwischen Streckenverkehr und Kapazität

Die Abhängigkeit weiterer Parameter wurde getestet, jedoch konnten auch in diesen Fällen keine statistisch signifikanten Zusammenhänge nachgewiesen werden. Dabei spielte es keine Rolle, ob die Sektoren beider Lufträume oder nur jene der oberen Lufträume betrachtet wurden. Auch eine vorab durchgeführte Cluster-Analyse lieferte keine Verbesserung hinsichtlich der Interpretierbarkeit der Ergebnisse. Ein Grund dafür könnte die Nutzung des zweidimensionalen Analyseverfahrens sein, welches



lediglich die Zusammenhänge zwischen zwei Parametern bestimmt. Dabei werden mögliche Effekte anderer Faktoren nicht abgebildet.

Alternativ könnte eine Regressionsanalyse durchgeführt werden. Da die Sektoren aber keine DMU darstellen (sondern lediglich das übergeordnete ACC) ist die Nutzung der Produktivität als abhängige Variable nicht sinnvoll. Über eine Regression kann jedoch der Einfluss der Faktoren auf die Kapazität des Sektors analysiert werden. Dabei muss der unterschiedliche Grad an Sektorisierung berücksichtigt werden. Daher sind skalenabhängige Größen wie „Occupancy“ oder „Entries per hour“ nicht als abhängige Variable verwendbar.

Eine skalenunabhängige Größe stellt die Verkehrsdichte dar. Diese könnte als abhängige Variable der Regression dienen. Obwohl mehrere Optionen getestet wurden, war eine Regressionsanalyse auf Basis der vorliegenden Daten jedoch nicht möglich. Dies lag insbesondere an der geringen Anzahl an abhängigen Variablen (und dadurch resultierend OVB) sowie der hohen Korrelation zwischen den einzelnen Faktoren.

4.5 Größenvorteile im Air Traffic Management

4.5.1 Hintergrund

Historisch bedingt ist der europäische Luftraum durch einen hohen Grad an Fragmentierung gekennzeichnet. Im Vergleich mit dem amerikanischen Luftraum wird die europäische Struktur als ineffizient und kostenintensiv beschrieben (EUROCONTROL und FAA 2019a, 2019b). Die Regressionsanalysen auf ANSP Ebene bestätigen, dass der Luftraum einen wesentlichen Einfluss auf Produktivität und Effizienz hat (vgl. Abschnitt 4.2).

Eine Lösungsoption stellt die Organisation des Luftverkehrs in Functional Airspace Blocks dar (vgl. Abbildung 2-11). Dieser Ansatz aus dem SES I Paket umfasst, aus praktischer Perspektive, die Konsolidierung mehrerer Flugsicherungsdienstleister zu einem gemeinsamen Luftraum. Hier sollen u. a. durch die sukzessive Anpassung der Systeme die Interoperabilität gewährleistet und dadurch Koordinationsaufwände reduziert werden. Die zugrundeliegende Annahme dabei ist, dass die Flugsicherungsdienstleister unter steigenden Skalenerträgen operieren. Das bedeutet, dass eine Vergrößerung der operativen Einheiten zu einer Steigerung von Produktivität und Effizienz führen würde.

Da sich der Verkehr nicht gleichmäßig über Europa verteilt, stellt eine Harmonisierung von Systemen und Verfahren eine große Herausforderung dar. Hard- und Software sind je nach ANSP, teilweise je nach ACC, unterschiedlich und an individuelle Bedürfnisse auf Basis der Verkehrscharakteristika angepasst (vgl. Abschnitt 4.4). Dies gilt insbesondere für die Air Traffic Service Systeme zur Konflikterkennung und für die Koordinierung mit anderen Sektoren (FABEC 2019).

In der europäischen „Core Area“, d. h. der Region mit der höchsten Verkehrsdichte, sind sieben ANSPs im FABEC organisiert. Es erscheint zunächst plausibel, dass hier eine harmonisierte Abwicklung des Verkehrs Effizienzvorteile bringen könnte. Andererseits sind mit der deutschen (DFS) und der französischen Flugsicherung (DSNA) zwei Unternehmen diesem FAB zugeordnet, welche - gemessen in



Flügen oder Flugstunden - innerhalb Europas ohnehin zu den größten Dienstleistern gehören. Hinzu kommt, dass die Schweiz einen der komplexesten Lufträume aufweist (EUROCONTROL 2020i).

Es ist bislang unklar, ob die vermuteten Größenvorteile tatsächlich realisierbar sind. Ziel dieser Analyse ist daher, die Skalanelastizität der europäischen Flugsicherungen zu vergleichen und die potentiellen Vorteile einer operativen Vergrößerung zu untersuchen. Des Weiteren werden Einflussfaktoren und räumliche Abhängigkeiten identifiziert. Ergebnisse dieser Untersuchung wurden in Standfuss et al. (2017b) und Standfuss et al. (2019a) publiziert.

4.5.2 Analyse der Skalenerträge und Skaleneffizienz

Die Auswertung der Skaleneffekte ermöglicht einen tieferen Einblick, welche ANSPs von einer operativen Konsolidierung der Prozesse profitieren könnten und welche nicht. Steigende Skalenerträge implizieren einen solchen Vorteil. ANSPs mit konstanten Skalenerträgen wirtschaften bereits im Betriebsoptimum. Für eine weitere Vergrößerung sind Größenvorteile nicht prognostizierbar. Werden fallende Skalenerträge ausgewiesen, so ist das Betriebsoptimum bereits überschritten. Eine Konsolidierung kann hier zu Größennachteilen führen.

Die Analyse der Größenvorteile wird mittels einer DEA unter variablen Skalenerträgen (siehe auch Abschnitt 3.4.2.2) auf Basis der PRU Daten von 2016 durchgeführt. Modelle und Implikationen entsprechen jenen des Abschnitts 3.4. In einem ersten Schritt wird die Skaleneffizienz aller DMUs, respektive ANSPs, bestimmt. Eine geringe Skaleneffizienz ist hier gleichzusetzen mit einer nicht optimalen Betriebsgröße. Über eine Analyse der Skaleneffekte wird in einem zweiten Schritt ermittelt, ob die skalenineffizienten Einheiten aus betrieblicher Sicht „zu klein“ sind, also durch Konsolidierung Größenvorteile realisieren könnten, oder die optimale Betriebsgröße bereits überschritten wurde. Im dritten Schritt wird über die räumliche Autokorrelation bestimmt, wie groß die Effizienzinterdependenz zwischen den einzelnen ANSPs ist. Zudem werden mögliche exogene Einflussfaktoren untersucht.

Um auf Basis der Modelle potentielle Größenvorteile ableiten zu können, wird ein Skalenindikator (Scale Indicator, SI) eingeführt. Dieser beschreibt modellübergreifend, wie hoch die Wahrscheinlichkeit eines ökonomischen Vorteiles durch Skaleneffekte ist. Können für eine Einheit in einem Modell steigende Skalenerträge nachgewiesen werden, entspricht dies einem Wert von +1, bei fallenden Skalenerträgen -1 und bei konstanten Skalenerträgen 0. Diese Werte werden, ohne weitere Gewichtung, über alle vier Modelle aufsummiert. Der Indikator kann zudem als Robustheitstest interpretiert werden.

Die Ergebnisse variieren gemäß Tabelle 4-18 zwischen den Modellen, insbesondere bei der Anzahl der Flugsicherungen mit konstanten und fallenden Skalenerträgen. Die Anzahl der ANSPs mit steigenden Skalenerträgen weist eine relativ geringe Streuung auf. Nach Modell 1 (M1) arbeiten überdurchschnittlich viele ANSPs skalenineffizient. Dies liegt in der Methodik und der begrenzten Anzahl der Einheiten begründet. In Modell 2 wurden die Output Parameter daher zu einer Größe zusammengefasst. Dadurch sinkt auch die Anzahl der als effizient ausgewiesenen Einheiten in allen weiteren Modellen, da hier eine Dimension eliminiert wird.

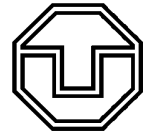


Tabelle 4-18: Anzahl der DMUs mit steigenden, konstanten oder fallenden SKE

	M1	M2	M2A	M2B
IRS	13	14	20	17
CRS	14	3	8	2
DRS	10	20	8	18

Modell 2A berücksichtigt die Heterogenität der ANSPs, indem MUAC aus der Analyse ausgeklammert wurde. Dadurch veränderte sich Verlauf und Form der Randproduktionsfunktion. Es gelten nun mehr Einheiten als skaleneffizient als in Model 2. Auch kann nach diesem Modell eine höhere Anzahl an Flugsicherungen Größenvorteile realisieren. Dagegen sinkt die Anzahl der ANSPs mit Größennachteilen von 20 auf acht. Auch dies kann auf die veränderte Produktionsfunktion zurückgeführt werden. Model 2B unterscheidet sich hinsichtlich der Ergebnisse nur geringfügig von Model 2. Intuitiv ist dies nachvollziehbar, da hier lediglich die Berechnung der CFH und CIU für jeden ANSP individuell vorgenommen wurden.

Aus den Ergebnissen in Tabelle 4-18 kann abgeleitet werden, dass auch jene der ANSPs (teilweise) zwischen den Modellen variieren müssen. Tabelle 4-19 veranschaulicht diese Abweichungen anhand einiger Beispiele. Konsistent sind die Ergebnisse der Skalenerträge bspw. für Alcontrol (steigend), DSNA (fallend) und EANS (konstant). Dagegen sind die Ergebnisse für Finnland abhängig vom gewählten Modell, bei denen jeweils 2x fallende und 2x steigende Skalenerträge ausgewiesen werden. Dies ist dadurch begründet, dass die Skaleneffizienz für ANS Finnland sehr hoch ist. Dadurch kommt es zu einem kippenden Effekt¹⁶⁵, welcher auch für andere ANSPs mit modellinkonsistenten Skalenerträgen erkennbar ist. Die Ergebnisse für alle ANSPs sind in Tabelle A 16 dargestellt.

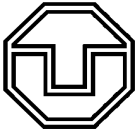
Tabelle 4-19: Skalenerträge ausgewählter ANSPs

ANSP	M1	M2	M2A	M2B	SI
Alcontrol	irs	irs	irs	irs	4
DSNA	drs	drs	drs	drs	-4
EANS	crs	crs	crs	crs	0
ANS Finnland	irs	drs	irs	drs	0

Abbildung 4-12 illustriert die Skalenindikatoren für jedes Land Europas.¹⁶⁶ Zu beachten ist, dass bei den Skalenindikatoren keine Gewichtung nach Modell erfolgte. Die Abbildung verdeutlicht, dass für die Mehrheit der ANSPs keine Größenvorteile vorausgesagt werden können. Dies liegt entweder an negativen Skalenindikatoren (-4 bis -1), oder einer methodenübergreifenden Streuung der Ergebnisse. Dazu werden einige ANSPs bereits als skaleneffizient ausgewiesen.

¹⁶⁵ Aufgrund der hohen Skaleneffizienz „kippen“ die Skaleneffekte von IRS auf DRS, bzw. vice versa, da der ihr zugehörige Benchmark die Produktionsfunktion beeinflusst.

¹⁶⁶ Aus Gründen der Übersichtlichkeit wird jeder ANSP in den zugehörigen Landesgrenzen dargestellt, auch wenn es in der Realität Abweichungen geben kann.



Mit einer hohen Wahrscheinlichkeit können Größenvorteile lediglich die Dienstleister in den Benelux-Staaten, die in FABCE koordinierten ANSPs von Österreich, Tschechien, der Slowakei, Slowenien und Ungarn sowie Anbieter in Teilen des Baltikums und der Balkanstaaten realisieren. Außerdem arbeiten ANSPs in Moldawien und Armenien modellübergreifend unter steigenden Skalenerträgen.

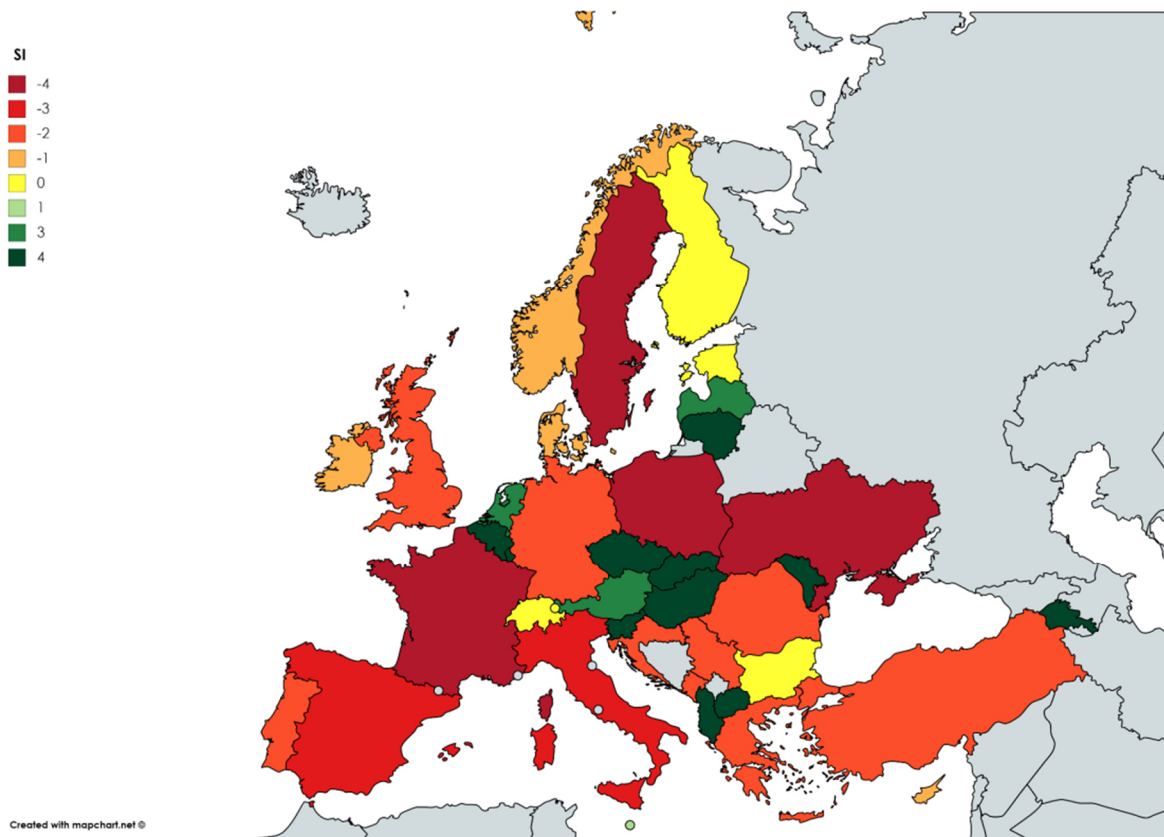
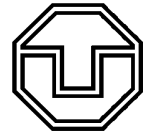


Abbildung 4-12: Skalenindikator für die europäischen Flugsicherungen

Die Skalenindikatoren der ANSPs entsprechen größtenteils den Erwartungen. Insbesondere die großen ANSPs in Frankreich, Italien oder Spanien weisen fallende Skalenerträge auf. Kleinere Flugsicherungen dagegen, etwa Albanien, Armenien oder Slowenien, würden von einer Konsolidierung profitieren. Allerdings gibt es Ausnahmen. Für einige verhältnismäßig kleine ANSPs, etwa Estland, Malta und Zypern, werden mehrheitlich keine zunehmenden Skalenerträge berechnet. Malta und Zypern könnten jedoch gemäß Modell 2B Größenvorteile realisieren. Die Flugsicherungen von Polen und Schweden arbeiten unter abnehmenden Skalenerträgen, obwohl ihre Outputs im mittleren Bereich angesiedelt sind. Beide ANSPs weisen über alle Modelle hinweg fallende Skalenerträge auf.

Für 13 der 27 ANSPs können Größenvorteile nachgewiesen werden ($SI \geq 3$), das heißt, dass diese ANSPs in mindestens drei von vier Modellen steigende Skaleneffekte aufwiesen. Drei von ihnen wurden lediglich in Modell 1 als skaleneffizient eingeordnet. Insbesondere Flugsicherungen mit sehr kleiner Unternehmensgröße, bspw. Belgocontrol und LVNL (Niederlande), können von einem Konzept wie den FABs profitieren. Allerdings ist mittels der angewendeten Methodik keine Ableitung möglich, wie groß ein FAB im Optimum sein sollte. Beide ANSPs werden derzeit im FABEC zusammen mit der



DFS, Skyguide, DSNA, MUAC und ANA LUX¹⁶⁷ organisiert. Wie in Abbildung 4-12 dargestellt, können für DFS und DSNA keine Größenvorteile nachgewiesen werden. Das bedeutet, dass bei der Konsolidierung aller Lufträume der sieben ANSPs ebenfalls Größennachteile möglich sind.

Größennachteile konnten für sechs ANSPs eindeutig nachgewiesen werden ($SI \leq -3$), insbesondere für Spanien, Italien und Frankreich. Für 18 ANSPs konnte keine eindeutige Aussage getroffen werden (SI zwischen -2 und 2). Dies liegt zum einen daran, dass sich die Ergebnisse modellübergreifend unterscheiden, zum anderen arbeiten einige Flugsicherungen bereits im Betriebsoptimum, sodass keine Größenvor- oder -nachteile entstehen.

Eine weitere Analyse bestätigt die Notwendigkeit, die europäische Heterogenität zu berücksichtigen, hier insbesondere am Beispiel der Sonderstellung von MUAC. Modell 2A zeigt, dass dieser ANSP als Benchmark die Produktionsfunktion erheblich verschiebt und dadurch die Ergebnisse stark beeinflusst. Abbildung 4-13 stellt die potentiellen Größenvorteile für Modell 2A grafisch dar. Zusätzlich ist anhand der Farbabstufung erkennbar, ob ein ANSP wesentliche oder lediglich kleine Größenvor- oder -nachteile hat. Die Grenze liegt bei einer Skaleneffizienz von 0,9 für Diseconomies of Scale, (DoS) bzw. 0,8 für Economies of Scale (EoS).

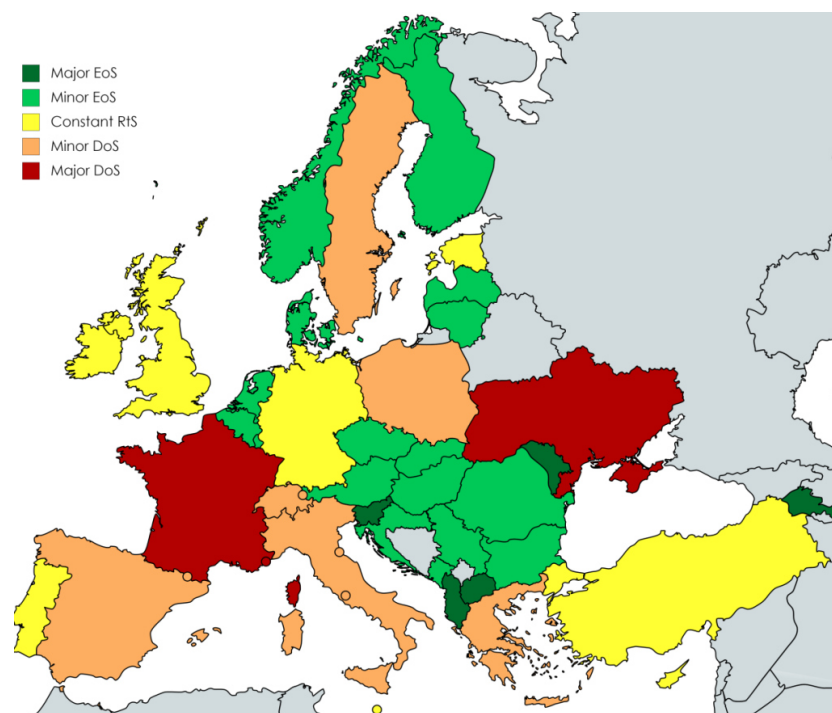
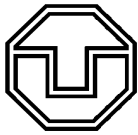


Abbildung 4-13: Skaleneffekte nach Model 2A

Die Ergebnisse der Skalenerträge nach Modell 2A sind bis auf wenige Ausnahmen plausibel. Entsprechend den Erwartungen sind jene Länder mit geringer Nachfrage oder kleiner Luftraumgröße durch steigende Skalenerträge gekennzeichnet. Dies betrifft vor allem die ost- und südosteuropäischen Staaten sowie drei der vier skandinavischen Staaten. Länder mit hoher Nachfrage und/oder Komplexität

¹⁶⁷ ANA LUX wird in den PRU Daten nicht ausgewiesen. Der Luftraum wird maßgeblich von Belgocontrol bzw. MUAC kontrolliert.



weisen Größennachteile auf, insbesondere Frankreich, Italien und Spanien. Eine Ausnahme bildet hier die Ukraine. Für den korrespondierenden ANSP werden modellübergreifend, so auch in 2A, Größennachteile ausgewiesen. Zwar ist der Luftraum verhältnismäßig groß, in Anbetracht der niedrigen Nachfrage waren sinkende Skalenerträge jedoch nicht zu erwarten. Ausgehend von Modell 2A sind im FAB Kontext lediglich für FABCE und NOFAB Größenvorteile prognostizierbar. Allerdings können keine Aussagen getroffen werden, ob die Allokation der ANSPs zu dem jeweiligen FAB optimal ist.

4.5.3 Strukturelle Einflussfaktoren

Die Vorteile einer Konsolidierung sind von mehreren endogenen und exogenen Faktoren abhängig. Exogene Faktoren, beispielsweise die Komplexität des Luftraumes, werden bei einer DEA nicht explizit berücksichtigt. Sowohl Stellungnahmen der ANSPs als auch die EUROCONTROL-Berichte (ACE und PRR) legen jedoch nahe, dass die Performance der Flugsicherung stark von der Komplexität des Luftraumes abhängig ist.

Zur Berechnung der Komplexität wird durch die PRU für jeden ANSP ein Complexity Score ermittelt. Dieser setzt sich zusammen aus der relativen Verkehrsdichte (Luftfahrzeuge pro Luftraumeinheit) und einem Strukturindex, welcher die potentiellen Konflikte zwischen zwei Luftfahrzeugen ausdrückt (EUROCONTROL 2008a). Ein weiterer exogener Faktor ist die Größe des Luftraumes. Die Grenzen des Luftraumes richten sich derzeit maßgeblich nach jenen des staatlichen Hoheitsgebietes. Das FAB-Konzept kann prinzipiell als Konsolidierung individueller Lufträume gesehen werden.

Im Folgenden soll untersucht werden, ob ein Zusammenhang zwischen den Faktoren „Komplexität“ bzw. „Größe des Luftraumes“ und dem Abstand vom Skalenoptymum (bzw. der optimalen Betriebsgröße), wie er in den obigen Analysen ermittelt wurde, existiert. Hierfür wurde die Abweichung vom Skalenoptymum, respektive der durch die DEA berechneten optimalen Betriebsgröße, exemplarisch für Model 2B bestimmt. Dadurch wird berücksichtigt, dass eine Skaleneffizienz durch steigende oder fallende Skalenerträge verursacht werden kann. Flugsicherungsdienstleister, für die steigende Skalenerträge ermittelt wurden (z. B. Albanien), sind durch ein negatives Vorzeichen gekennzeichnet. Die DMUs auf der Abszisse (Wert = 0) wurden durch die DEA als skaleneffizient bewertet. Je weiter eine Einheit von der Abszisse entfernt ist, desto höher ist die berechnete Skaleneffizienz.

Gemäß der Argumentation von EUROCONTROL nimmt die Performance mit zunehmender Komplexität ab. Diese These greift Abbildung 4-14 auf und stellt die Komplexität der Abweichung vom Skalenoptymum gegenüber. Für die Gesamtheit aller ANSPs ist zu konstatieren, dass die Datenpunkte eine hohe Streuung aufweisen, der Korrelationskoeffizient beträgt 0,26. Ein direkter Zusammenhang zwischen Skaleneffizienz und Komplexität kann daher für die Gesamtheit aller ANSPs nicht nachgewiesen werden.

Ergänzende Erkenntnisse kann jedoch eine Analyse von Teilgruppen bieten. Zunächst werden nur die ANSPs betrachtet, für die im DEA Modell steigende Skalenerträge berechnet wurden (untere Hälfte in c). Diese ANSPs sind maßgeblich durch einen kleinen Luftraum gekennzeichnet. Auch bei dieser Gruppe ist eine hohe Streuung ($r = 0,27$) nachweisbar. Jedoch sinkt tendenziell die Abweichung

zur Abszisse (und damit die Skaleneffizienz der ANSP) mit zunehmender Komplexität. Ein wesentlicher Faktor des Complexity Scores ist die „Adjusted Density“, welche eine Näherung für die Zahl der Luftfahrzeuge pro Luftraumeinheit darstellt (EUROCONTROL 2016b, 2020). Dadurch werden Lufträume mit einer hohen Verkehrsdichte (z. B. Schweiz) als verhältnismäßig komplex ausgewiesen.

Vice versa sind ANSPs mit weniger Verkehr durch eine geringe Komplexität gekennzeichnet. Genau diese ANSPs (kleiner Luftraum, wenig Verkehr)¹⁶⁸ würden durch einen staatenübergreifenden Zusammenschluss der Lufträume profitieren und könnten dadurch die Effizienz steigern. Bei ANSPs mit hoher Komplexität ist die Abweichung zum Betriebsoptimum geringer. Eine grenzübergreifende Anpassung der Luftraumstruktur durch eine Konsolidierung der ANSPs von Belgien und den Niederlanden würde bspw. nur geringfügig zu (operativen) Effizienzvorteilen führen, da diese Lufträume bereits an ihrer Kapazitätsgrenze arbeiten und somit die Sektorisierung bereits an die Nachfrage angepasst ist. Dies legt die Vermutung nahe, dass allgemein mit steigender Komplexität, respektive steigender Verkehrsdichte, der Produktivitätsgewinn durch eine Konsolidierung der Lufträume sinkt.

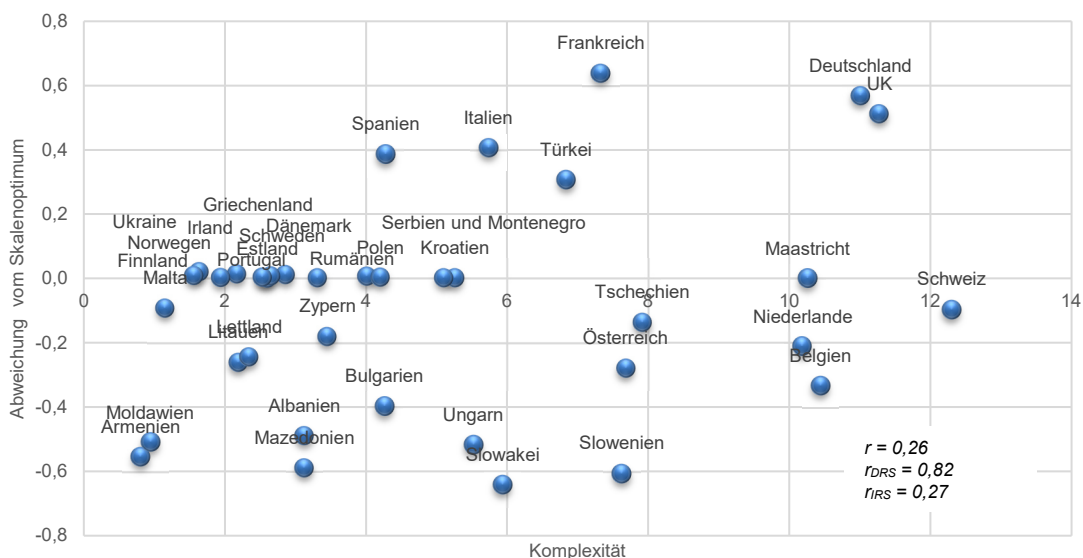


Abbildung 4-14: Zusammenhang zwischen Komplexität und Skaleneffizienz (Modell 2B)

Bei ANSPs, für die fallende Skalenerträge ausgewiesen werden (‘obere‘ Hälfte in Abbildung 4-14), ist die Streuung der Datenpunkte geringer und der Korrelationskoeffizient mit 0,82 größer als bei den ANSPs mit zunehmenden Skalenerträgen. ANSPs mit deutlicher Abweichung von der optimalen Betriebsgröße sind durch eine hohe Komplexität gekennzeichnet (z. B. Deutschland und Großbritannien). Dabei ist auch zu berücksichtigen, dass sich in den größeren Ländern der Verkehr nicht gleichmäßig im Luftraum verteilt. Beispielsweise ist in Deutschland die Verkehrsdichte im Süden höher als im Norden. Bis auf Italien und Spanien weisen alle ANSPs mit einer geringen Komplexität (<6) eine hohe Skaleneffizienz auf.

¹⁶⁸ Z. B. Mazedonien und Albanien.

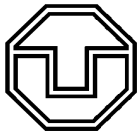


Abbildung 4-15 zeigt, dass mittels hyperbolischer Regression ein funktionaler Zusammenhang zwischen Luftraumgröße und Abstand zum Skalenoptimum hergestellt werden kann. Aus der Grafik ist ablesbar, dass zunächst die Effizienz mit einer Vergrößerung des Luftraums steigt (steigende Skalenerträge).

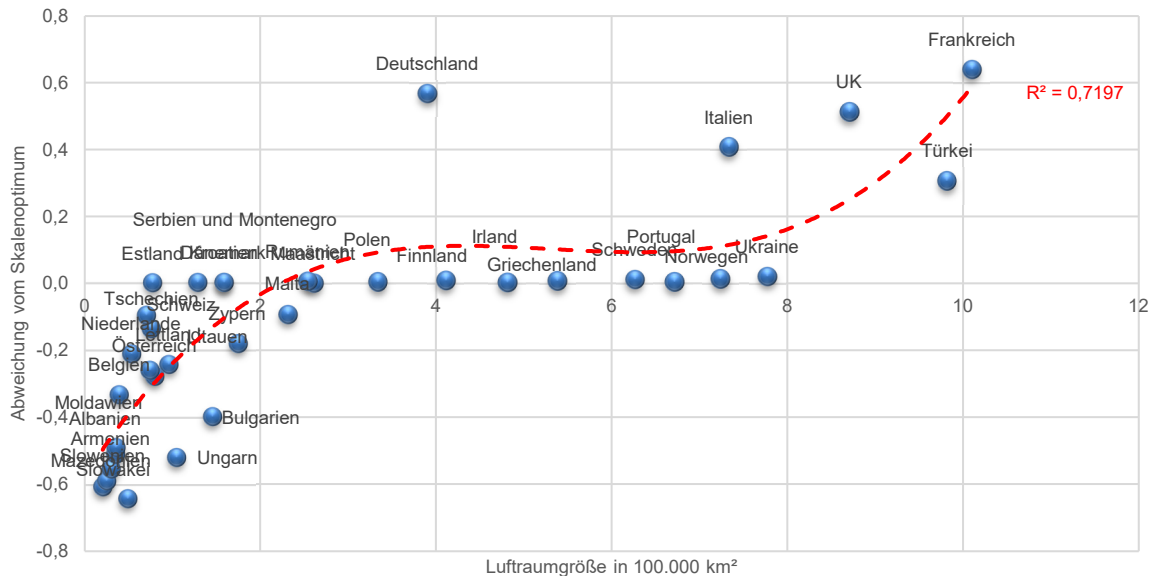


Abbildung 4-15: Zusammenhang zwischen Luftraum und Skaleneffizienz (Modell 2B)¹⁶⁹

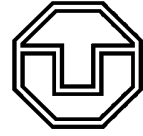
Zwischen Luftraumgrößen von 400.000 und 800.000 km² ist eine optimale Größe erreicht. Zwar verläuft die Regressionslinie parallel zur Abszisse (Wert von $\approx 0,1$ statt 0,0 - dadurch wäre der Luftraum bereits zu groß), dies kann jedoch durch die höhere Streuung bzw. Abweichung bei den DMUs mit fallenden Skalenerträgen (z. B. Deutschland) erklärt werden. Ab einem Wert von 800.000 km² sinkt die Skaleneffizienz aufgrund fallender Skalenerträge.

4.6 Sonstige Einflussfaktoren

4.6.1 Prognosequalität

In den Abschnitten 4.2 bis 4.5 wurde eine Ursachenanalyse auf verschiedenen operativen Ebenen sowie über mehrere Zeithorizonte durchgeführt. Primäres Ziel war die Quantifizierung der effizienzbeeinflussenden Effekte. Obwohl diese Analysen ein breites Spektrum an Faktoren abdecken, konnten einige Aspekte nicht integriert werden. Dies liegt zum einen an fehlenden Daten, zum anderen daran, dass Themen Gegenstand derzeitiger oder zukünftiger Untersuchungen sind und somit noch kein endgültiges Ergebnis vorliegt. In diesem Abschnitt werden einige dieser Faktoren adressiert und diskutiert, inwieweit sie die Leistungsfähigkeit einer Flugsicherung beeinflussen. Dabei werden sowohl statistische Methoden angewandt (vorliegender Abschnitt) als auch Ergebnisse auf Basis komparativer Analysen mit dem amerikanischen Luftraum präsentiert (Abschnitte 4.6.2 bis 4.6.4).

¹⁶⁹ Spanien repräsentiert einen Extremwert und wurde daher nicht abgebildet.



Die Ressourcenplanung von Flugsicherungsorganisationen umfasst eine taktische und strategische Ebene (FABEC 2019), siehe auch Abschnitt 4.4. Dabei besteht der generelle Kompromiss zwischen den Kosten der Kapazitätsbereitstellung und den Kosten unzureichender Kapazität¹⁷⁰ (Standfuss et al. 2018c). Die folgenden Abschnitte fokussieren die Ressourcenplanung und die Abweichungen aufgrund von Prognosefehlern. Ansatz und Ergebnisse basieren auf Standfuss und Whittome (2019b).

Der Begriff „Ressource“ bezieht sich in diesem Abschnitt auf die Anzahl der ATCOs bzw. ATCO-Stunden. Dieser Bedarf wird durch die prognostizierte Verkehrsentwicklung und die dadurch benötigte Anzahl an Sektoren sowie deren Öffnungszeiten bestimmt. Einbezogen werden zusätzlich Faktoren, z. B. Fluktuationsrate, „On-the-job“-Auszubildende und Elternzeit. Eine Erhöhung der (geplanten) Kapazität ist jedoch durch institutionelle (z. B. Gewerkschaftsvereinbarungen) und technische Einschränkungen (z. B. Anzahl der Sektoren oder ATCO-Arbeitspositionen) begrenzt. In der mittelfristigen Planung (bis zu 12 Monate) sind zusätzlich Anpassungen von operativen Verfahren¹⁷¹ möglich. Die Personalplanung ist vor allem durch die jährliche Gesamtarbeitszeit begrenzt.

Um die Kapazität langfristig anzupassen bzw. zu erweitern, kann ein ANSP zusätzliches Personal einstellen. Hinsichtlich der Fluglotsen erfolgt dies entweder "ab-initio" (Neueinsteiger) oder als "ready-entry" (hat bereits als Fluglotse gearbeitet). Allerdings ist sowohl die Anzahl der Ready-Entries als auch die Möglichkeit, das Personal in Theorie und ‚on-the-job‘ zu schulen, begrenzt. Darüber hinaus gibt es einen Trade-off in der Ausbildung hinsichtlich der Lizensierung. Werden Fluglotsen für eine große Anzahl von Sektoren ausgebildet, so erhöht das die Einsatz-Flexibilität. Bei einer Ausbildung für eine geringere Anzahl von Sektoren werden Ausbildungszeit und -kosten verringert (vgl. Abschnitt 4.4.3.3). Die Bestimmung eines optimalen Personalstocks, insbesondere hinsichtlich der Anzahl der beschäftigten Fluglotsen, ist zudem durch volatile Verkehrsströme durch Unsicherheiten gekennzeichnet (Standfuss et al. 2018c, 2019b), was wiederum einen effizienten Einsatz von Ressourcenpuffern voraussetzt. Durch diese Komplexität der Ressourcenplanung sind valide Prognosen notwendig, um eine effiziente „Balance“¹⁷² von angebotener und nachgefragter Kapazität zu erreichen.

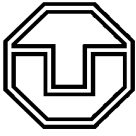
STATFOR veröffentlicht zweimal jährlich eine 7-Jahres-Prognose (EUROCONTROL 2018d) sowie etwa alle zwei Jahre eine 20-Jahres-Prognose (EUROCONTROL 2019b). Diese Analysen berücksichtigen mehrere Szenarien, die als Konfidenzintervall interpretiert werden können. Darüber hinaus wird die räumliche Verteilung geschätzt, d. h. eine ANSP-individuelle Wachstumsrate berechnet.

Abbildung 4-16 zeigt verschiedene Szenarien für die prognostizierten Flugbewegungen im Jahr 2050. In einem Maximal-Szenario geht STATFOR von einem jährlichen Wachstum von durchschnittlich 2,7 % aus. Das heißt, dass im Jahr 2050 etwa 27 Mio. Flugbewegungen erwartet werden. Des Weiteren

¹⁷⁰ Z. B. Verspätungskosten.

¹⁷¹ Z. B. Anflüge, Holdings, etc.

¹⁷² Im operativen Kontext „Demand Capacity Balancing“ (DCB) genannt.



zeigt die Abbildung eine hohe Unsicherheit. Zwischen dem Maximal-Szenario und dem Minimal-Szenario, welches von lediglich 1,1 % jährlichem Verkehrswachstum ausgeht, besteht ein Nachfrage-Unterschied von etwa 16 Mio. Flügen (EUROCONTROL 2013a).

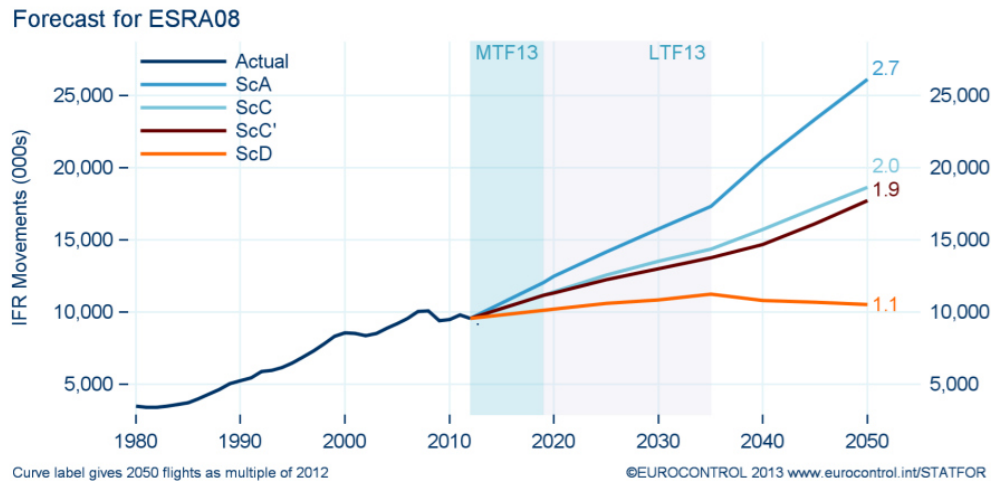


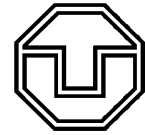
Abbildung 4-16: Verkehrsprognose 2050
(EUROCONTROL 2013a)

Da sich diese Unsicherheiten direkt auf die Ressourcenplanung auswirken, sollen für die prognostizierten Flüge (und deren Konfidenzintervalle) die für die Nachfrage notwendigen ATCOs (gemessen in FTE) approximiert werden. Dafür werden zunächst die Flüge in Flugstunden umgerechnet, da diese aggregiert werden können. Im Jahr 2017 überwachten europäische Fluglotsen 16,2 Mio. Flugstunden. Unter Berücksichtigung der Gesamtzahl der Flüge betrug die durchschnittliche Länge eines Fluges in Europa 1,54 Stunden. Unter der Annahme, dass sich dieser Wert bis 2050 nicht wesentlich verändert, beträgt die Unsicherheit in der STATFOR-Prognose etwa 24 Mio. Flugstunden. Die Jahresarbeitsleistung eines Fluglotsen betrug 2017 durchschnittlich 897 Flugstunden. Eine Unsicherheit von 24 Mio. Flugstunden pro Jahr bedeuten demnach eine Spanne von 27.000 ATCOs bzw. FTEs.

In der 7-Jahres-Prognose sind die Konfidenzintervalle enger als in der langfristigen Perspektive. Allerdings sind auch die Möglichkeiten, die Ressourcen an die Nachfrage anzupassen, begrenzt. STATFOR prognostizierte für die DFS für das Jahr 2017 zwischen 3,46 und 3,95 Mio. Flüge im 6-Jahres Horizont und zwischen 3,02 und 3,35 Mio. Flüge im 3-Jahres-Horizont. Dies bedeutet, dass das Konfidenzintervall von 490.000 Flügen auf 330.000 Flüge gesunken ist. Auf Basis der PRU Daten entspricht dies rund 238.000 bzw. 160.000 Flugstunden. Im Durchschnitt werden von den DFS-ATCOs 780 Flugstunden pro Jahr abgefertigt. Das Konfidenzintervall beträgt somit 305 bzw. 205 ATCOs¹⁷³.

Um die Qualität von Prognosen zu messen, stehen mehrere Metriken zur Verfügung. Die Verzerrung kann dabei systematische und/oder zufällige Fehler aufweisen. Die gebräuchlichsten Indikatoren sind der "absolute" und der "relative Prognose Fehler". Dabei wird die Abweichung B durch die Differenz

¹⁷³ Die Werte repräsentieren lediglich eine grobe Approximation. Der Umrechnungswert kann zwischen den einzelnen ACCs, z. B. zwischen Karlsruhe und Bremen, variieren, da die tatsächlichen Flugstunden pro ATCO von Verkehrsmerkmalen, z. B. der Komplexität, abhängen.



(4-8) oder dem Verhältnis (4-9) zwischen prognostizierten (M_P) und tatsächlichen Verkehrsaufkommen (M_A) gebildet. Bei aggregierten Daten für n Einheiten werden analog die Formeln (4-10) und (4-11) verwendet. Da die ANSPs ein "homogenes" Produkt anbieten, aber heterogen in der Größe sind, wird in diesem Fall der ‚Mean Absolute Percentage Error‘ als die zu bevorzugende Metrik angesehen (Vlinen et al. 2017).

$$\text{Absolute Forecast Bias (AFB)} \quad B = \sum (M_P - M_A) \quad (4-8)$$

$$\text{Relative Forecast Bias (RFB)} \quad B = \frac{\sum M_P}{\sum M_A} \quad (4-9)$$

$$\text{Mean Absolute Deviation (MAD)} \quad B = \frac{1}{n} \sum |M_P - M_A| \quad (4-10)$$

$$\text{Mean Absolute Percentage Error (MAPE)} \quad B = \frac{1}{n} \sum \left| \frac{M_P - M_A}{M_A} \right| \quad (4-11)$$

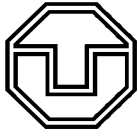
Im Folgenden wird die Prognosegüte der 7-Jahre-STATFOR-Vorhersagen analysiert. Dafür wurde die Basisprognose (ohne Konfidenzintervall) mit dem tatsächlichen Bedarf verglichen. Darüber hinaus umfasst die Analyse drei Zeiträume (1 Jahr, 3 Jahre, 6 Jahre), die als kurz-, mittel- und langfristig bezeichnet werden¹⁷⁴. Um Konsistenz zu erreichen, werden für die Analyse nur Daten verwendet, die in den Februar-Berichten veröffentlicht wurden. Dabei ist zu beachten, dass der 'Relative Forecast Bias' die Genauigkeit angibt (z. B. 0,99 bedeutet 99 % Genauigkeit).

Tabelle 4-20 zeigt die Ergebnisse des prognostizierten Verkehrsaufkommens für die deutsche Flugsicherung am Beispiel der Jahre 2015 bis 2017. Ergänzend sind in Anhang Tabelle A 17 weitere Jahre dargestellt. Erwartungsgemäß sind langfristige Prognosen ungenauer als mittel- oder kurzfristige Vorhersagen. Es gibt nur eine Ausnahme: Der MAPE für 2011 ist in der langfristigen Perspektive niedriger (und damit präziser) als in der mittelfristigen. Insgesamt wurde die Nachfrage in den meisten Fällen überschätzt, insbesondere lang- und mittelfristig. In der Folge wird auch der Bedarf an Ressourcen überschätzt, was zu einer sinkenden Kosteneffizienz führt.

Tabelle 4-20: Prognosequalität der STATFOR Vorhersagen für die DFS, 2015-2017

Bezugsjahr	Prognosejahr	M_P	M_A	AFB	RFB	MAD	MAPE
2017	2016	3.226.000	3.259.000	-33.000	0,99	33.000	0,01
2017	2014	3.192.000	3.259.000	-67.000	0,98	67.000	0,02
2017	2011	3.638.000	3.259.000	379.000	1,12	379.000	0,12
2016	2015	3.142.000	3.146.000	-4.000	1,00	4.000	0,00
2016	2013	3.233.000	3.146.000	87.000	1,03	87.000	0,03
2016	2010	3.551.000	3.146.000	405.000	1,13	405.000	0,13
2015	2014	3.056.000	3.080.000	-24.000	0,99	24.000	0,01
2015	2012	3.255.000	3.080.000	175.000	1,06	175.000	0,06
2015	2009	3.605.000	3.080.000	525.000	1,17	525.000	0,17

¹⁷⁴ Die Begriffe sind abzugrenzen von den „kurz-, mittel- und langfristigen“ Prognosen von STATFOR.



Eine ganzheitlichere Sicht der Prognosequalität kann durch die Berücksichtigung mehrerer Unternehmen oder mehrerer Jahre ermöglicht werden. Da für die Analyse lediglich Daten der DFS zur Verfügung standen, wurden diese zeitlich aggregiert. Ziel ist es zu differenzieren, welche Jahre eine relativ große oder kleine Abweichung zwischen prognostiziertem und tatsächlichem Verkehr aufweisen.

Abbildung 4-17 stellt die MAD- und MAPE-Werte für die Jahre 2008 bis 2017 dar. Dabei ist zu beachten, dass aus Konsistenzgründen nur die Vorhersagen der Februar-Berichte berücksichtigt wurden. Für jedes Prognose-Jahr wurden MAPE und MAD auf der Grundlage von sieben Vorhersagen berechnet, z. B. die Vorhersagen von 2011-2017 für das Prognose-Jahr 2017. Hier betrug der mittlere absolute prozentuale Fehler 3%. Dieses Maß ist jedoch skalenunabhängig, d. h. es berücksichtigt die Größe der DMU. Wie aus der mittleren durchschnittlichen Abweichung hervorgeht, betrug die mittlere Verzerrung ca. 100.000 Flüge. Dies ist ein praktikablerer Indikator, um den Einfluss der Verzerrung auf die Ressourcenplanung zu untersuchen.

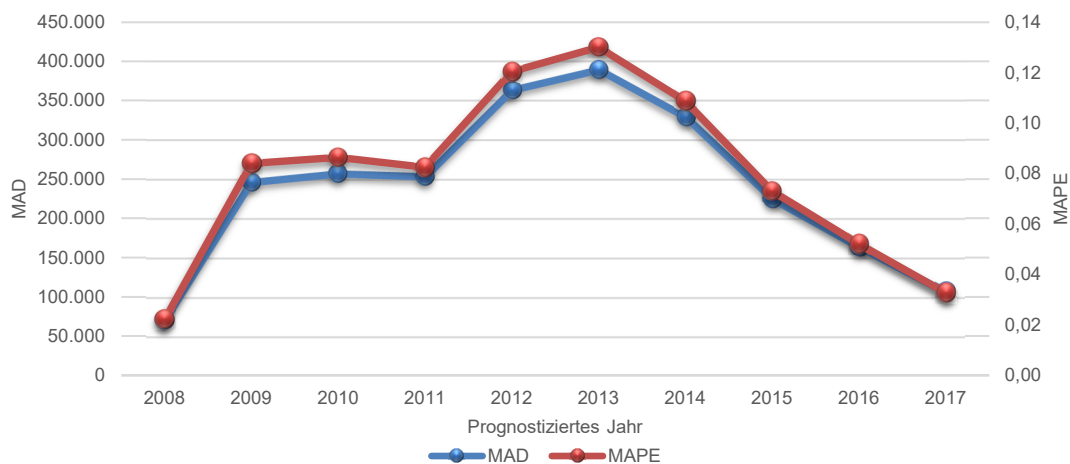
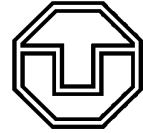


Abbildung 4-17: STATFOR Prognosefehler nach Jahr für die DFS

Abbildung 4-17 zeigt, dass die Genauigkeit der Vorhersagen für die Jahre nach 2013 stetig gestiegen ist. Die beste Genauigkeit wird jedoch für die Prognose der Nachfrage im Jahr 2008 nachgewiesen. Eine signifikante Veränderung ist im Vergleich zu 2009 sichtbar, als sich der MAPE-Score vervierfacht hat. Die höchsten Werte für MAPE und MAD und damit die geringste Prognosegüte betrifft das Jahr 2013. Dies wird durch die starken Verzerrungen der Vorhersagen in den Jahren 2007 und 2008 verursacht: Die beginnende und in den Vorhersagen nicht einberechnete Finanzkrise führte hier zu einer signifikanten Überschätzung der Verkehrsnachfrage.

Eine zweite Möglichkeit, die Genauigkeit von Vorhersagen zu beurteilen, ist die Auswertung einzelner Berichte. Die entsprechende Frage ist, wie gut eine bestimmte Prognose mit den tatsächlichen Verkehrszahlen der einbezogenen Jahre übereinstimmt. Dabei nahm die Prognosegenauigkeit der spezifischen Berichte in den ersten sieben Jahren kontinuierlich ab (siehe Anhang, Abbildung A 21). Die maximale durchschnittliche Abweichung betrug 600.000 Flüge für den Bericht 2008. Der hohe Wert ist auf erhebliche Verzerrungen für die Vorhersagen für 2012-2014 zurückzuführen, die zu Abweichungen von bis zu 947.000 Flügen zwischen der prognostizierten und der tatsächlichen Verkehrsnachfrage



fürten. Die Berichte 2009 und 2010 zeichneten sich durch eine höhere Genauigkeit aus; die Werte für MAD und MAPE stiegen jedoch 2011 wieder an.

Fehlerhafte Abschätzungen des zu erwartenden Verkehrs haben einen direkten Einfluss auf die erforderlichen Ressourcen und damit auf die Kosten. Tabelle 4-20 zeigt, dass die Zahl der Flüge im Jahr 2017 in der Siebenjahresperspektive um 379.000 Flüge überschätzt und in der Dreijahresperspektive um 67.000 Flüge unterschätzt wurde. Nach den Daten der PRU bedeuten die Verzerrungen des Berichts eine Ressourcen-Fehlabschätzung von 136 bzw. 42 ATCOs. Da PRU die ATCO-Kosten veröffentlicht, ist es möglich, die entsprechenden Gesamtkosten abzuleiten. Im Jahr 2017 beliefen sich die Kosten auf 217.490€ pro ATCO, was bedeutet, dass die Abweichung zwischen prognostiziertem und tatsächlichem Verkehr ca. 30 Mio. € bzw. 9 Mio. € entspricht. Kosten für die Ausbildung und etwaige Overheadkosten sind darin nicht berücksichtigt. Geht man von einer ATCO-Schulungszeit von fünf Jahren aus, so ergab sich die größte Abweichung zwischen dem tatsächlichen und dem prognostizierten Bedarf im Jahr 2013 (875.000 Flüge). Die Überschätzung würde somit zu einer Überbesetzung von 545 ATCOs (31 % der tatsächlichen FTE im Jahr 2013) führen. Die verursachten Kosten betragen über 101 Mio. €. Die durchschnittliche Verzerrung in der Fünf-Jahres-Perspektive beträgt 378.900 Flüge oder 236 ATCOs.

Auch die Kosten sind lediglich als grobe Approximation zu verstehen. Erstens wird die Ressourcen- und Personaleinsatzplanung dadurch erschwert, dass jedes ATCO nur für die Überwachung bestimmter Sektoren lizenziert ist. Folglich könnten die Auswirkungen der verzerrten Nachfrageprognose noch stärker sein als berechnet. Zweitens berücksichtigen die Kosten lediglich jene für Fluglotsen. Eine umfangreiche Kapazitätserweiterung könnte jedoch auch Investitionen in anderes Personal sowie andere Systeme oder Verfahren erfordern. Drittens werden bei der Approximation Größeneffekte ausgeklammert. Die Interdependenz zwischen Nachfrage und ATCOs ist jedoch möglicherweise nicht linear. Viertens beschränkt sich die Untersuchung auf Daten der DFS.

Auf Basis dieses Abschnittes und der korrespondierenden FABEC Studie (Standfuss und Whittome 2019b) wurde die Methodik auf weitere ANSPs bzw. FABs angewendet (FABEC und BLUE MED FAB 2019). Diese Untersuchung war jedoch FAB-intern und kein Bestandteil der Dissertation. Da für die vorliegende Arbeit lediglich Daten der DFS zur Verfügung standen, konnte der Aspekt der Prognosequalität nicht in der Regressionsanalyse berücksichtigt werden. Es ist Gegenstand derzeitiger Untersuchungen, welche Auswirkungen eine fehlerhafte Prognose auf die Produktivität bzw. Effizienz der Flugsicherungen hat. Zudem beschäftigen sich Studien mit alternativen Prognoseverfahren, u. a. Gelhausen et al. (2018).

4.6.2 Anteil des General Aviation Verkehrs

In offiziellen Benchmarking Reports bildet der Vergleich mit der US-amerikanischen Flugsicherung einen wesentlichen Schwerpunkt. Auf Basis operativer (EUROCONTROL und FAA 2019a) und finanzieller (EUROCONTROL und FAA 2019b) Kennzahlen werden beide Lufträume miteinander verglichen (siehe auch Abschnitt 3.2.1.5). Dabei ist zu konstatieren, dass die amerikanische FAA höhere Produktivitäts- und Kosteneffizienzwerte aufweist. Dieser Performance-Gap wurde in Zusammenarbeit



mit FABEC untersucht, da sich Nachfragestrukturen, Finanzierung der Dienstleistungen sowie Arbeitsweisen teilweise elementar unterscheiden. Ziel der Untersuchung war es, den Einfluss der operativen (Standfuss und Whittome 2018, 2019a) und finanziellen Heterogenität (Whittome und Standfuss 2018) zwischen beiden Lufträumen auf die Produktivitäts- und Kosteneffizienzwerte zu bestimmen. Der vorliegende Abschnitt fasst diese Untersuchungen zusammen.

Die Nachfrage sowie deren Charakteristika unterscheiden sich zwischen amerikanischem und europäischem Luftraum signifikant. Bereits FABEC (2014) stellte die proklamierten Effizienzvorsprünge anhand einiger operativer Differenzen und der daraus fehlenden Vergleichbarkeit in Frage. Elementarer Bestandteil war der hohe Unterschied des Anteils an General Aviation (GA) Flügen nach Instrumentenflugregeln. Daher wurde untersucht, ob die Höhe des IFR-GA-Verkehrs sowie dessen Verteilung im Luftraum einen Einfluss auf die Leistungskennziffern haben. Da keine ANSP- oder ACC-individuellen Daten über den GA Verkehr in Europa zur Verfügung stehen, war eine Berücksichtigung in der Regressionsanalyse nicht möglich.

Die grundlegenden Thesen waren:

1. In Europa nutzt IFR GA primär jene Lufträume, die bereits vom kommerziellen Verkehr nachgefragt werden (z. B. Fluggesellschaften).
2. In den USA gibt es primär Punkt-zu-Punkt-Verbindungen in Gegenden mit geringem Verkehr
3. In den USA nutzen IFR GA Flugzeuge, beispielsweise Business Jets, häufig sekundäre Flughäfen, wodurch der Verkehr einfacher separiert werden kann

Um diese Thesen zu überprüfen, wurden drei Ansätze verfolgt. Erstens fand ein Vergleich der am häufigsten frequentierten IFR GA-Flughäfen in den USA und Europa hinsichtlich der Dichte (Entfernung zwischen den Flughäfen) und deren räumliche Verteilung statt. Zweitens wurden Dichtediagramme von Europa und den USA sowohl horizontal als auch vertikal ausgewertet. In Europa war es zudem möglich, Unterschiede zwischen geplanten und tatsächlichen Trajektorien zu überprüfen (Vorhersagbarkeit der benötigten Kapazität). Drittens ermöglichte eine Fuzzy-Clustering-Analyse die Identifikation von Verkehrsclustern für beide Verkehrstypen. Dafür wurden DDR2¹⁷⁵-Daten (EUROCONTROL 2017b) analysiert und die Flugbahnen in äquidistante Punkte getrennt. Durch das Fuzzy-Clustering konnte eine vordefinierte Anzahl von Clustern abgeleitet und auf die europäische Karte projiziert werden (zur Näheren Beschreibung des Vorgehens, siehe Abschnitt 5.2.2.2). Für den Luftverkehr in den USA standen keine äquivalenten Flugdaten zur Verfügung.

Zunächst wurde die Lage der 25 meistfrequentierten IFR GA Flughäfen in Europa und den USA verglichen. Abbildung 4-18 zeigt, dass sich in Europa 20 von 25 Flughäfen in einem Umkreis von 1900 km (blauer Kreis) befinden. Offensichtlich nutzt IFR GA größtenteils Hauptflughäfen wie Zürich und Madrid, jedoch kaum Sekundärflughäfen¹⁷⁶.

¹⁷⁵ Demand Data Repository der EUROCONTROL.

¹⁷⁶ Z. B. Egelsbach (EDFE) für Frankfurt (EDDF).

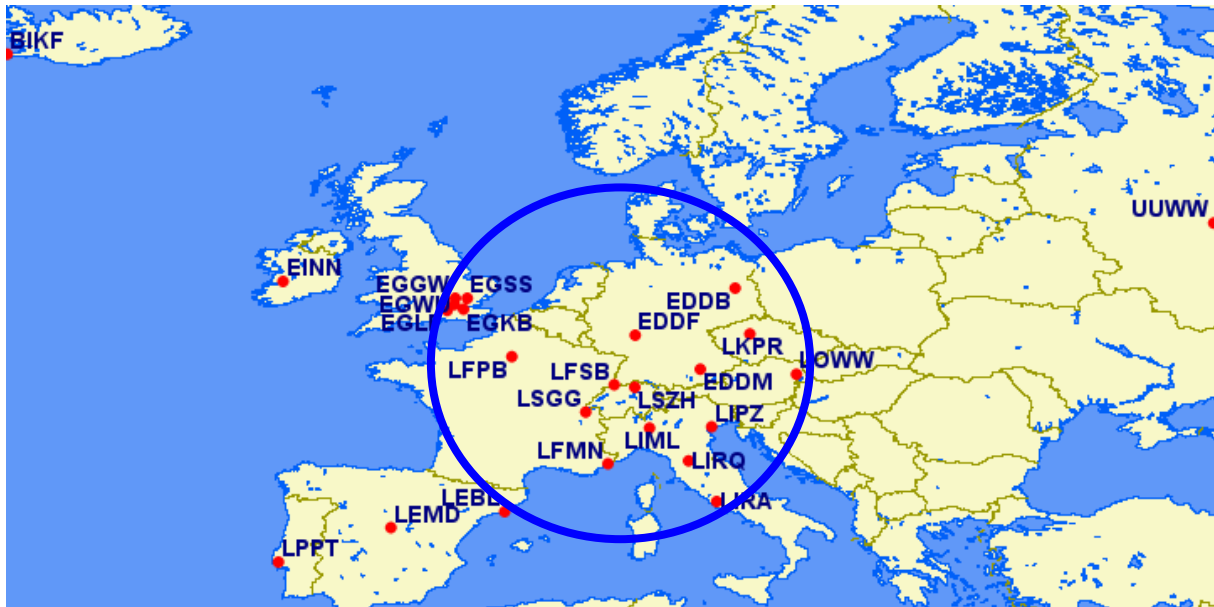
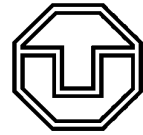


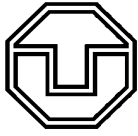
Abbildung 4-18: Die 25 meistfrequentierten GA Flughäfen in Europa

Bei Verwendung des gleichen Radius im US-Luftraum (orangefarbener Kreis) befinden sich maximal sieben US-Flughäfen im Umkreis von 1900 km (Abbildung 4-19). Darüber hinaus handelt es sich bei den meisten Flughäfen um sekundäre Flughäfen wie KFXE für Ford Lauderdale, KPWK für Chicago, KVNY für Los Angeles und KTEB für New York und Newark. In der Folge ist die Dichte an Flughäfen in Europa signifikant größer als in den USA.



Abbildung 4-19: Die 25 meistfrequentierten GA Flughäfen in den USA

Eine genauere Betrachtung der Verkehrsströme und Trajektorien ermöglichte eine Datenauswertung über NEST. Das Tool ermöglicht die Erstellung von Dichteplots für Europa, wie sie in Abbildung 4-20 dargestellt sind. Je nach Anzahl der Flüge pro Tag werden die zugehörigen Streckenabschnitte in einem Bereich von grün (geringes Verkehrsaufkommen) bis rot (hohes Verkehrsaufkommen) eingefärbt. Da IFR GA nur etwa 2-3% des gesamten IFR-Verkehrs ausmacht, wurde die Skala an diesen Verkehr



angepasst. Die Dichtediagramme zeigen die Trajektorien eines Beispieltages von IFR GA (rechts) und dem kommerziellen IFR-Verkehr (links). Dabei handelt es sich um die geplanten Trajektorien. Für die tatsächlichen Trajektorien, siehe Anhang, Abbildung A 22.

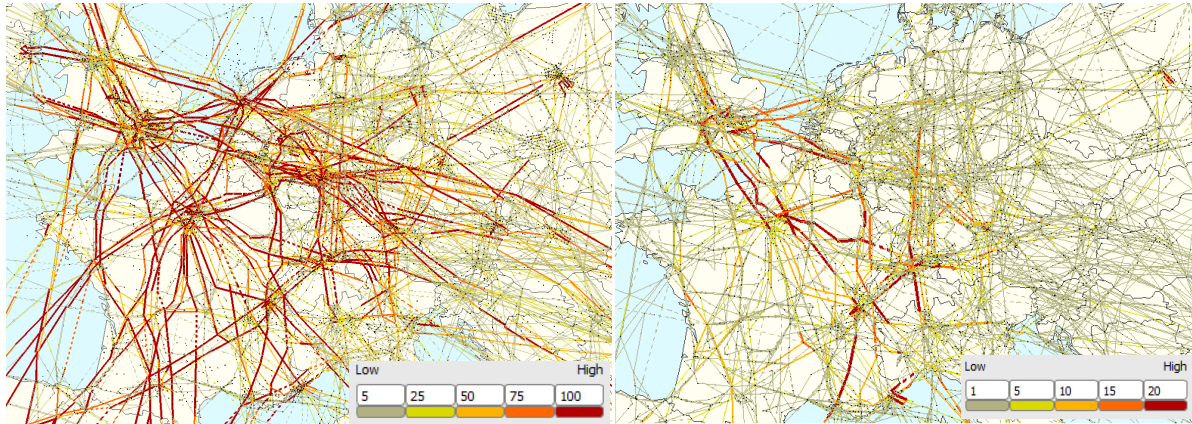


Abbildung 4-20: Density Plots für IFR und IFR GA, geplante Trajektorien

Die Dichte-Plots zeigen, dass IFR und IFR GA prinzipiell die gleichen Routen verwenden, was zu einem zusätzlichen Arbeitsaufwand für die Lotsen führt. IFR-GA-Verkehr ist hauptsächlich im europäischen Zentrum, insbesondere in den Lufträumen von FABEC und Großbritannien, aktiv. Dadurch kann davon ausgegangen werden, dass die IFR-GA Nachfrage einen zusätzlichen Aufwand für die Lotsen bedeutet. Darüber hinaus weist der IFR-GA-Verkehr eine höhere Streuung zwischen geplanten und tatsächlichen Trajektorien auf als der kommerzielle IFR-Verkehr, wodurch die Vorhersagbarkeit des Verkehrs geringer ist (siehe Anhang, Abbildung A 22).

Die Plots bestätigen die These, dass der IFR-GA-Verkehr hauptsächlich die Lufträume im europäischen Zentrum nutzt. Sie bieten jedoch nur eine horizontale Ansicht. Da Businessjets meist höhere Flugflächen als kommerzielle Linienverkehre nutzen ist es möglich, dass diese beiden Verkehrsströme nur zu Beginn und Ende eines City Pairs mit anderem Verkehr interagieren. Allerdings trifft dies nur auf Jets zu. Flugzeuge mit Propellertriebwerken erreichen diese Flughöhen nicht.

Um die europäische Situation mit den Verkehrsströmen in den USA zu vergleichen, stellte die FAA Trajektorien-Plots für IFR GA Flüge und sonstige IFR Flüge zur Verfügung¹⁷⁷. Die Plots umfassen einen 16-Tage-Zeitraum. Die Verkehrsströme in den USA weisen eine hohe Dichte in der südwestlichen Region und im östlichen Teil des Landes auf. Die Trajektorien-Endpunkte bestätigen, dass IFR GA häufig sekundäre Flughäfen nutzt. Der kommerzielle Verkehr nutzt dagegen primär die großen Drehkreuze. Dies ermöglicht eine Trennung der Verkehrsströme für Terminal- und Streckenverkehre.

Auch diese Diagramme bieten zunächst nur eine horizontale Ansicht. Tiefere Einblicke kann die Analyse vertikaler Interdependenzen bieten. Zu diesem Zweck wurde das City-Pair Fort Lauderdale und Atlanta untersucht, einschließlich Primär- und Sekundärflughäfen. Anhand der vertikal separierten DichtepLOTS¹⁷⁸ wird gezeigt, dass der kommerzielle IFR-Verkehr vom IFR GA größtenteils vertikal getrennt

¹⁷⁷ Aufgrund der Sensibilität der Daten durften diese Plots nicht veröffentlicht werden.

¹⁷⁸ Bereitgestellt wurden Density Plots der Flugflächen 0-180, 180-220, 220-260, etc., bis 420-460.

ist. Das bedeutet, dass der IFR-GA-Verkehr nicht den Taskload des Lotsen erhöht, sondern zu einer effizienteren Luftraumauslastung führt.

Zur Validierung der Ergebnisse wurde eine Fuzzy-Clustering-Analyse verwendet, um die Bereiche mit dem höchsten Verkehrsaufkommen zu identifizieren. Da keine Daten für den US-Luftraum verfügbar sind, wurde die Methodik nur auf den europäischen Luftraum angewendet. Die räumliche Abgrenzung basiert auf der im Vergleichsbericht verwendeten Definition und dient dazu, die in diesem Bereich abfliegenden und/oder ankommenden Flüge aus den DDR-Flugdaten zu filtern. Für das Clustering wurde das Tool RouGe des Deutschen Zentrums für Luft- und Raumfahrt verwendet.

Abbildung 4-33 zeigt das Verkehrs-Clustering für IFR GA in Europa. Die blaue Linie steht für die räumliche Abgrenzung, die orangefarbenen Punkte repräsentieren die äquidistanten Punkte der Flugbahnen, die aus DDR2-Daten berechnet werden. Je geringer der Abstand zwischen den Clusterpunkten ist, desto höher ist die Verkehrsdichte. Da sowohl die äquidistanten Punkte als auch die Clusterpunkte (grün) im Kernbereich Europas eine höhere Dichte erreichen, bestätigt die Clusteranalyse die Ergebnisse der Dichtediagramme und der räumlichen Analyse. Die Untersuchung wurde auf Basis von 25 und 50 Clusterpunkten durchgeführt.

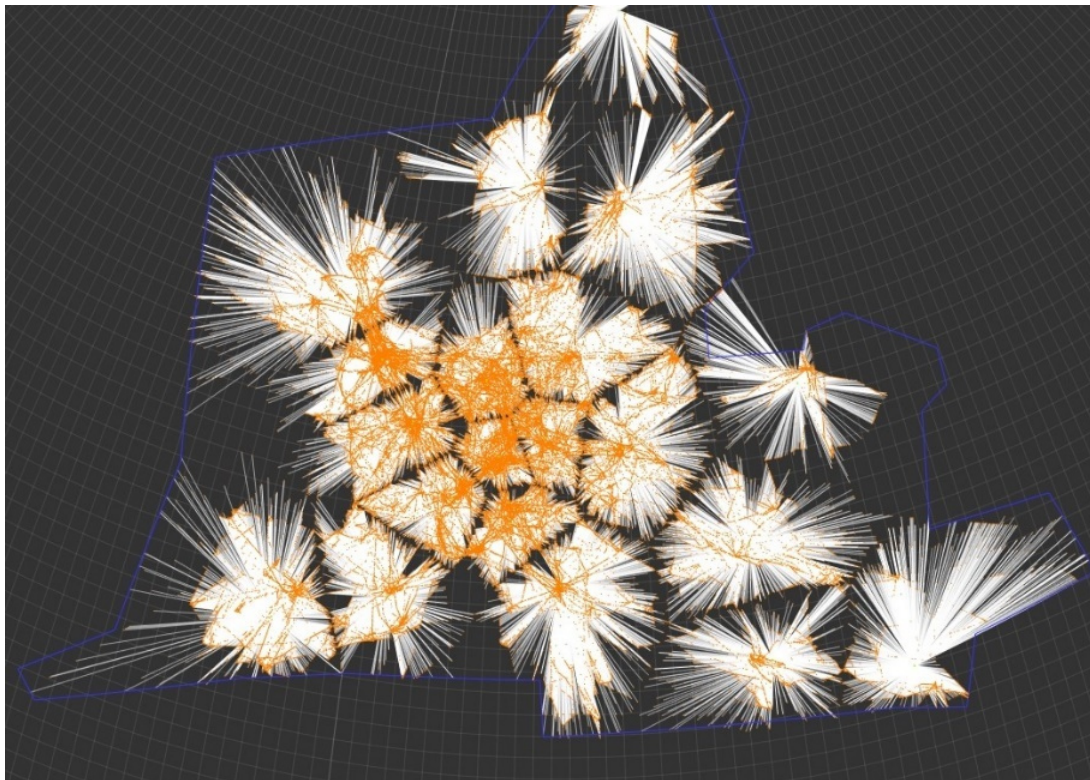


Abbildung 4-21: Fuzzy Clustering des europäischen GA Luftverkehrs

Auf Basis der Untersuchungen konnte nachgewiesen werden, dass ein Teil der Performance-Lücke zwischen USA und Europa auf den Anteil des IFR GA Verkehrs zurückzuführen ist. Dies ist dadurch begründet, dass in Europa die räumliche Distanz zwischen den meistfrequentierten IFR GA Flughäfen im Vergleich zu den USA gering ist. Die parallele Nutzung von Flughäfen und Flugrouten durch IFR und IFR GA führt zu einem Mehraufwand in Europa. Dagegen nutzt IFR GA in den USA



häufig Sekundärairports und die Trajektorien sind horizontal und vertikal getrennt. Dadurch repräsentiert der hohe IFR GA Anteil eine bessere Auslastung der ohnehin vorhandenen Kapazität. Die Verkehrsnachfrage kann als exogen angesehen werden und ist somit nicht von einem ANSP beeinflussbar. Eine Approximation zeigt, dass der IFR GA Anteil für ca. sechs bis acht Prozentpunkte der Performance-Differenz, welche im offiziellen Benchmarking-Report (EUROCONTROL und FAA 2019a) veröffentlicht wurde, verantwortlich ist (Standfuss und Whittome 2019a).

4.6.3 Einfachbesetzung der Sektoren

Im vorigen Abschnitt wurde diskutiert, dass exogene Effekte der Verkehrsnachfrage einen Einfluss auf die Performance haben können. In einem weiteren Schritt sollen zwei endogene Faktoren untersucht werden. Auch hier kann auf Basis des EU-US Vergleichs die Wirkung der Heterogenität bestimmt werden. Ein grundlegender Unterschied zwischen dem US-amerikanischen und europäischen ANS-Betriebskonzept besteht in der Anwendung von ‚Single ATCO Sector Operations‘ (SSO)¹⁷⁹. Das heißt, dass nur ein Lotse für einen Sektor verantwortlich ist. Innerhalb des FAA-Gebiets ist diese Ein-Mann-Besetzung Standard. In Europa werden Sektoren überwiegend von zwei ATCOs betreut (Dual ATCO Sector Operations, DSO), jedoch wurden SSO bei einigen Flugsicherungszentren (z. B. Belgocontrol) bereits implementiert. Des Weiteren werden sie bei einigen ANSPs zur taktischen Personaleinsatzplanung angewendet¹⁸⁰, sofern bestimmte Verkehrskriterien erfüllt sind.

Die Übertragung eines SSO Konzeptes würde einen hohen Modellierungsaufwand bedeuten, da aufgrund der Arbeitslast des Lotsen der Luftraum in kleinere Sektoren aufgeteilt werden müsste. Stattdessen wird impliziert, dass dieses Betriebskonzept lediglich in Zeiten mit geringer Verkehrsnachfrage Anwendung findet. Für die Approximation des Einsparpotentials werden folgende Modellannahmen getroffen, differenziert nach Tag- und Nachtschichten:

- Nacht
 - SSO ist während der Nachtschichten (22:00 bis 06:00 Uhr) prinzipiell anwendbar
 - In dieser Zeit sind nur 2¹⁸¹ statt 3 Lotsen notwendig
 - Es gibt nur einen Sektor pro EBG
- Tag
 - SSO ist nicht während der Tagschicht (06:00 bis 22:00 Uhr) anwendbar.
 - Der Tag besteht aus zwei Schichten (z. B. Früh- und Spätschicht)
 - Im Durchschnitt besteht eine EBG aus 4 Sektoren
 - Diese Sektoren werden in 60 % der Zeit zusammengelegt

Basierend auf dieser Annahme kann das erforderliche Personal berechnet werden:

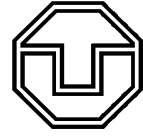
- $40\% \cdot 4 \text{ Sektoren} \cdot 2,5 \text{ ATCOs}^{182} = 4 \text{ ATCOs}$
- $60\% \cdot 2 \text{ Sektoren} \cdot 2,5 \text{ ATCOs} = 3 \text{ ATCOs}$

¹⁷⁹ Bezeichnungen teilweise abweichend je nach Institution. SSO wird u. a. auch als Single Person Operation (SPO) oder seltener Single Man Operations (SMO) bezeichnet.

¹⁸⁰ Z. B. im Krankheitsfall.

¹⁸¹ Aufgrund von Pausen müssen 2 ATCOs eingeplant werden.

¹⁸² Die Approximation über 2,5 statt 2 ATCOs liegt in Pausenregelungen begründet.



Für die ganze Tagschicht beträgt der modellierte Bedarf $2 \times (4 + 3) = 14$ ATCOs. Hinzu kommen die Lotsen, die für die Nachtschicht eingesetzt werden:

- Personal SSO = 14 ATCOs am Tag, 2 ATCOs in der Nacht = 16 ATCOs
- Personal DSO = 14 ATCOs am Tag, 3 ATCOs in der Nacht = 17 ATCOs

In der Folge würde die Einführung von SSO Einsparungen von $1/17 = 5,9\%$ der FTEs bedeuten.

Abhängig von der operativen Größe der europäischen Flugsicherungsdienstleister würden die europaweiten Einsparungen durch die Einführung von SSO 884 ATCOs (0,51 %) oder 1,1 Mio. ATCO-Stunden (0,49 %) betragen. ANSPs mit mehr als 1 Mio. ATCO-Stunden pro Jahr sind dabei mit den vollen 5,9 % berücksichtigt worden. Bei ANSPs ab 450.000 ATCO-Stunden reduzierte sich die potentielle Einsparung um 1 %. Bei allen restlichen Flugsicherungsorganisationen wurde eine Reduktion von 3,9 % angenommen. Diese Schwellenwerte haben keinen operativen Hintergrund, wurden jedoch vorab mit den Experten von FABEC diskutiert. Wenn in der ANSP SSO bereits implementiert wurde, wurden keine Einsparungen einberechnet. Eine Auflistung der einzelnen ANSPs sowie der Einsparpotentiale befindet sich in Anhang Tabelle A 18.

Der reduzierte Bedarf an Ressourcen führt zu einer Erhöhung der europäischen Leistungsindikatoren. Die Performance-Lücke verringert sich für die 2019 veröffentlichten Daten um 2-3 Prozentpunkte (Standfuss und Whittome 2019a). Dies entspricht dem Performance Gap, welcher im vorherigen Bericht berechnet wurde (Standfuss und Whittome 2018). Auf Basis dieses Berichtes wurde die durch die Implementierung von SSO bedingte Kosteneinsparung bestimmt¹⁸³. Dafür wurden zunächst die ATCO-Kosten je ANSP im Falle des bisherigen Betriebskonzeptes und bei SSO bestimmt. Die Differenz zwischen den Kosten wurde über 38 ANSPs aufsummiert. Eine Implementierung von SSO würde demnach im europäischen Kontext 124 Mio. EUR pro Jahr einsparen. Die Kosten pro Flugstunde verringern sich um 8€, die Kosten pro Flug um 12€. Auch diese Werte sind als Approximation zu sehen.

4.6.4 Jährliche Arbeitszeit pro Lotsen

Aufgrund verschiedener landesspezifischer Determinanten, z. B. Arbeits- und Tarifverträge, politische Rahmenbedingungen oder der Einfluss von Gewerkschaften, unterscheiden sich die jährlichen Lotsenarbeitszeiten innerhalb Europas. Abbildung 4-22 zeigt, dass die Zahlen zwischen 938 (DFS) und 1949 Stunden pro Lotsen (DCAC Cyprus) variieren. Die amerikanische FAA (grün dargestellt) ist durch die dritthöchste Lotsenarbeitszeit gekennzeichnet.

Um den Einfluss auf die Performance zu testen, wird die durchschnittliche Jahresarbeitszeit eines amerikanischen Lotsen (1814 Stunden¹⁸⁴) als Benchmark genutzt und die Arbeitszeiten der europäischen Lotsen angepasst. Basierend darauf kann der Bedarf an ATCOs (FTE) errechnet werden. Für Zypern und Malta wird die (im Vergleich zu den USA höhere) Jahresarbeitszeit konstant gehalten, da andernfalls ein „negatives“ Einsparpotential impliziert würde und dies im Rahmen der Effizienz nicht zielführend wäre.

¹⁸³ Die Kosten pro Lotsen bzw. pro Lotsenstunde ist durch die PRU Daten ableitbar.

¹⁸⁴ Dieser Wert ist aus den Vergleichsberichten ableitbar.

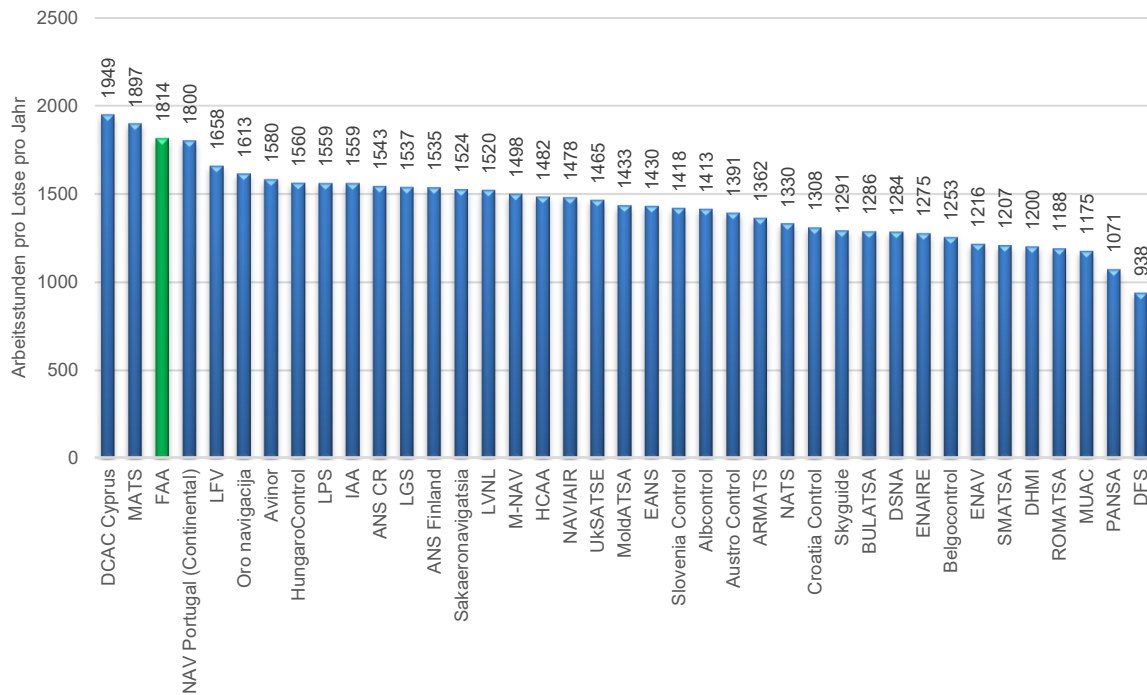


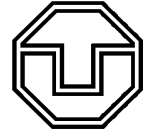
Abbildung 4-22: Jährliche Arbeitsstunden je ATCO

Die Erhöhung der Arbeitsstunden je ATCO führt zu einer Minderung des Ressourcenbedarfs. Bei der modellierten Fixierung auf 1814 Stunden je ATCO sinkt der paneuropäische Bedarf an Fluglotsen von 18.003 um 28,6% auf 12.849 FTEs. Im Vergleich zu den USA umfasst die daraus folgende Verbesserung der Performance etwa 19 Prozentpunkte. Die Ressourceneinsparungen wirken sich zudem direkt auf die Kosten aus, welche um 790 Mio. € sinken (Standfuss und Whittome 2019a). Diese Überlegungen sind jedoch primär von hypothetischer Natur. Wie bereits in Abschnitt 4.2.2 diskutiert ist die Arbeitszeit der Lotsen teilexogen, bedingt durch den Einfluss der rechtlichen Rahmen und der Gewerkschaften. Dennoch ist das offenbarte Einsparpotential bzw. der korrespondierende Beitrag zur Performance-Differenz zwischen USA und Europa signifikant höher als jener der SSO.

4.7 Zwischenfazit

In den Abschnitten 4.2.3 und 4.2.4 wurden verschiedene Regressionsmodelle getestet, welche einen funktionalen Zusammenhang zwischen abhängiger Variable (Produktivitäts- oder Effizienzscores) und mehreren unabhängigen Variablen herstellen. Die Ergebnisse sind größtenteils plausibel und decken sich mit den Annahmen in Tabelle 4-2.

Abhängig von der verwendeten Leistungsmetrik werden in der Literatur verschiedene Methoden empfohlen. Bei nicht-beschränkter abhängiger Variable kann eine OLS-Regression durchgeführt werden, ansonsten werden Tobit-Modelle oder eine Trunkated Regression empfohlen. Eine generelle Entscheidung zwischen Tobit- oder Trunkated Modell ist abhängig von der Spezifikation (Auswahl der zu berücksichtigenden Faktoren). In den meisten Fällen wurde für das Tobit-Modell eine höhere Modellgüte berechnet. Insbesondere aufgrund der hohen Ähnlichkeit der Ergebnisse bietet die Anwendung einer



OLS-Regression¹⁸⁵ Vorteile gegenüber der Nutzung von Tobit-Modellen. Insgesamt weichen beide Ergebnisse nur in Ausnahmefällen voneinander ab, sodass die Effekte tendenziell mit beiden Methoden bestimmt werden können.

Im Sinne einer möglichst genauen Prädiktion der abhängigen Variablen wurden mittels Maximierung der Modellgüte die signifikanten Faktoren herausgefiltert. Hier ist zu konstatieren, dass die Produktivität gemäß PRU genauer geschätzt werden kann als die DEA-Scores. Verstärkt wird dieser Effekt dadurch, dass je nach DEA-Modell die signifikanten Variablen variieren. Modellunabhängig waren lediglich die Variablen NONA, AIRP, STAT, GINI und SI. Das heißt, dass die Signifikanz der Variablen von der ökonomischen Modellierung in der First Stage DEA sowie den betrachteten Einheiten (inklusive oder exklusive MUAC) abhängen.

Aufgrund der wenigen Beobachtungen sind die Ergebnisse primär als Tendenz zu sehen. Eine Erweiterung des Datensatzes kann u. a. durch die Beobachtungen mehrerer Jahre erfolgen, welche in Form einer Panel-Analyse ausgewertet werden (siehe Abschnitt 4.3). Des Weiteren fehlen (datenbedingt) wesentliche Faktoren, bspw. Menge und Verteilung des GA-Verkehrs oder Anzahl, Lage und Nutzungshäufigkeit militärischer Gebiete. Zu berücksichtigen ist auch, dass eine Analyse der ANSP Ebene lediglich Durchschnittswerte beinhaltet. Insofern muss bspw. kritisch hinterfragt werden, inwieweit ein einzelner Komplexitätswert für den gesamten Luftraum aussagekräftig ist. Eine disaggregierte Analyse kann hier helfen, weitere Effekte zu identifizieren und zu quantifizieren (siehe Abschnitt 4.4).

Innerhalb der Ursachenanalyse wurde die Vermutung bestätigt, dass ein Benchmarking auf ANSP Ebene nicht ausreicht, um eine holistische Untersuchung der Wirkungsprozesse zu ermöglichen. Daher wurde die Analyse auf Subebenen erweitert. Auf diesen spielen operative Kennziffern wie Kapazität und Humanressourcen eine wichtige Rolle. Dies hat auch Auswirkungen auf die Benchmarking-Methodik, die im Falle von ACCs insbesondere auf Indexzahlen beruht. Dies ist jedoch auch durch die Datenverfügbarkeit begründet: Auf disaggregierter Ebene stehen zwar mehr Einheiten für ein Benchmarking zur Verfügung, jedoch sind für diese Einheiten weniger Faktoren erfasst.

Die qualitative Analyse hat gezeigt, dass auch auf disaggregierter Ebene eine signifikante Heterogenität besteht, sowohl auf betrieblicher als auch auf institutioneller Ebene. Dies erklärt u. a. die in Abschnitt 4.4.2 diskutierten, teilweise kontraintuitiven Ergebnisse. Ein wesentliches Problem stellt hier die Datenverfügbarkeit dar. Um alle Einflussgrößen auf Produktivität bzw. Kapazität abzuschätzen, müssen alle relevanten Faktoren berücksichtigt werden. Eine statistische signifikante Beziehung zwischen der von den ACCs erreichten Kapazität und den (verfügbaren) Einflussfaktoren konnte, auch auf Sektorebene, nicht nachgewiesen werden.

Wesentlicher Nutzen der Analyse war die Sensibilisierung auf die Unterschiede zwischen den ACCs, u. a. hinsichtlich der verwendeten Systeme und deren Funktionalität. Die ACCs haben dadurch die Möglichkeit, unterstützende Systeme in der nächsten Modernisierung den eigenen Bedürfnissen anzupassen. Es scheint jedoch sowohl aufgrund der unterschiedlichen Nachfragecharakteristiken und

¹⁸⁵ Wenngleich hier die diskutierten methodischen Einschränkungen gelten.



Umwelteinflüssen als auch aus ökonomischen Gesichtspunkten zweifelhaft, ob alle ACCs auf ein einheitliches System umgestellt werden sollten.

Die Untersuchung deckte zudem einige, ACC-übergreifende, offene Fragen auf, welche im weiteren Verlauf von Studien getestet werden sollten. Hier ist u. a. die richtige bzw. optimale Größe von Ressourcen-Puffern zu nennen. Dies könnte über den Ansatz eines „Level-of-Service“ Konzeptes betrachtet werden, welches die marginalen Kosten von Ressourcen und Delay gegenüberstellt. Die offenen Fragen betreffen zudem einzelne Daten und die Interpretation von Extremwerten.

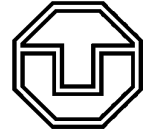
Die Analyse der Skaleneffekte zeigt, dass modellunabhängig für etwa die Hälfte der ANSPs Größenvorteile nachgewiesen werden können. Die FABs bieten jedoch Potentiale zur Optimierung sowohl hinsichtlich der Zusammensetzung als auch der Struktur. Dadurch wird vor allem die operative Implementierung erschwert. Die Luftraumgrenzen der FABs werden größtenteils durch die politischen Grenzen der zugehörigen ANSPs determiniert. Als sinnvoller kann eine Allokation gemäß den Verkehrsflüssen gesehen werden (Standfuss et al. 2018d).

Ein weiteres Problem stellt die unterschiedliche Anzahl der in einem FAB organisierten ANSPs dar. Die bereits vollständig implementierten FABs DK-SWE und UK-Ireland bestehen aus lediglich zwei ANSPs, während bspw. FABEC insgesamt sieben Flugsicherungsdienstleister umfasst. Da sich Systeme und Tools wesentlich zwischen den ANSPs (bzw. ACCs) unterscheiden (siehe Abschnitt 4.4.3), erhöht sich auch der Koordinationsaufwand bezüglich der Systemabstimmungen zwischen den einzelnen ANSPs. Dadurch wird die in der Interoperabilitätsverordnung geforderte Harmonisierung der Systeme erschwert.

Verkehrscharakteristiken wie Menge und Fluss sowie deren Volatilität unterscheiden sich je nach Flugsicherung, sodass die Systeme an diese Determinanten angepasst und bspw. durch individuelle Tools unterstützt werden müssen. Alternativ wäre eine Konsolidierung von Einheiten auf ACC Ebene bzw. Einheiten mit geringer Heterogenität hinsichtlich Verfahren und Systemen denkbar.¹⁸⁶ Da diese Einheiten meist äquivalente Anforderungen haben, ist hier die Anpassung von Systemen effektiver. Zuletzt können auch sprachliche, kulturelle, rechtliche und operative Unterschiede zusätzlich eine Herausforderung in der Umsetzung darstellen. Davon ableitend bilden die betrachteten Modelle einen Ansatz zur Überprüfung der Größenvorteile, die Ergebnisse sind jedoch abhängig von den methodischen Einschränkungen sowie der Datengrundlage.

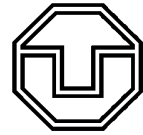
Die Analyse der Skaleneffizienz und Skaleneffekte legt nahe, dass die Konsolidierung von Lufträumen zu FABs nicht zwangsläufig zu ökonomischen Vorteilen führen muss. Dabei wurden lediglich für jeden ANSP individuell mögliche Vorteile analysiert, nicht jedoch, ob diese in den derzeit konstruierten FABs umgesetzt werden können. Die Streuung der Ergebnisse legt nahe, dass die Modelle sensitiv auf die gewählten Inputs und Outputs sowie die Anzahl der DMUs reagieren. Diese fehlende Robustheit begrenzt die generelle Aussagekraft der Analyse. Dennoch können für einige ANSPs modellübergreifend relativ eindeutige Aussagen getroffen werden. Flugsicherungsunternehmen mit einem SI von -4

¹⁸⁶ Bspw. die Konsolidierung von ACCs, die lediglich den oberen Luftraum betreuen.



oder -3 sind bereits in ihrer jetzigen Betriebsgröße oberhalb des Optimums, unabhängig von ihrer Zuordnung zu einem FAB.

Neben den bisher beschriebenen und diskutierten Einflussfaktoren können weitere Effekte zu Effizienz und Produktivität beitragen, u. a. die Prognosequalität als Basis der Ressourcenplanung, operative Konzepte sowie die Finanzierung der ANS Dienstleistungen. Diese wurden in Zusammenarbeit mit den operativen Experten von DFS bzw. FABEC modelliert und analysiert. Eine Zusammenfassung der Ergebnisse liefert Abschnitt 4.6. Dieser ist als Ergänzung der qualitativen und quantitativen Analysen intendiert.



5 Zusammenfassung und Handlungsempfehlungen

5.1 Ergebnisse der Dissertation

5.1.1 Überblick

In der vorliegenden Dissertation wurden bislang im Wesentlichen zwei der drei Forschungsfragen adressiert: Wie lässt sich die Leistungsfähigkeit von Flugsicherungsunternehmen bestimmen und welche Faktoren beeinflussen die Effizienz. Das vorliegende Kapitel fasst die Ergebnisse zusammen und diskutiert zukünftige Herausforderungen sowie alternative Betriebskonzepte (Abschnitt 5.2). Abschnitt 5.3 überführt die Erkenntnisse in konkrete Handlungsempfehlungen zur Steigerung der Leistungsfähigkeit im europäischen Luftverkehrsmanagement.

Die Analysen basierten auf dem ökonomischen Modell, welches in Abbildung 3-15 dargestellt wurde. Neben relevanten Ressourcen (Inputs) und Dienstleistungen (Outputs) wurden endogene und exogene Einflussfaktoren untersucht. Eine klare Differenzierung in endogene und exogene Effekte ist jedoch schwierig: Wie bereits in Kapitel 3.5 diskutiert, sind einige der betrachteten Faktoren teil-exogen.

Die Ergebnisse der Dissertation können in fünf Schwerpunkte untergliedert werden, wie in Abbildung 5-1 dargestellt ist. Das erste Cluster (Leistungsbewertung) umfasst die erste Forschungsfrage (vgl. Abschnitt 1.2) sowie deren Ergebnisse. Cluster 2 bis 5 adressieren primär die Ursachenanalyse und damit die zweite Forschungsfrage. Die Elemente des Modells und damit auch die Cluster der Ergebnisse stellen ein interagierendes System dar. Daher sind die abgebildeten Cluster nicht trennscharf. Beispielsweise sind die Elemente der Cluster Luftraumstruktur und Betrieb auch abhängig von politischen Vorgaben. Des Weiteren können die untersuchten Faktoren je nach operativer Ebene unterschiedlich wirken.

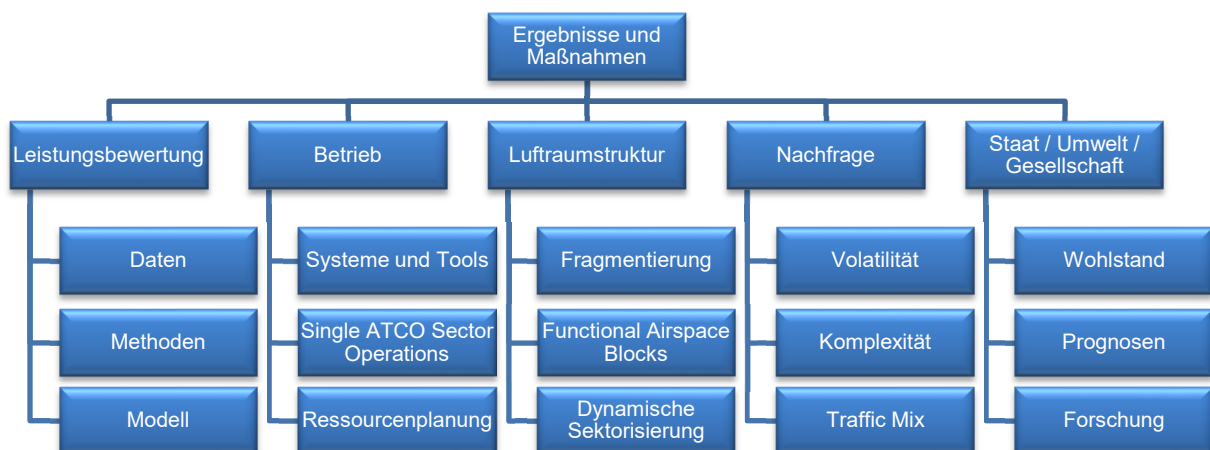


Abbildung 5-1: Analyse-Cluster und deren Elemente



5.1.2 Modellierung und Leistungsbewertung

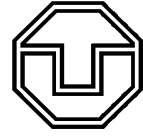
Eine zentrale Aufgabe bestand in der Überführung des ökonomischen Modells eines ANSP in eine geeignete Benchmarking-Methodik zur Messung der Performance. Je nach Betriebsebene sind nur ausgewählte Determinanten des Grundmodells (Abbildung 3-15) relevant. Dadurch unterscheidet sich die Leistungsbewertung zwischen ANSP und ACC sowohl hinsichtlich der angewendeten Methodik als auch bezüglich der verwendeten Inputs und Outputs. Auch die Umweltfaktoren können variieren. Das ökonomische Modell hilft, die einzelnen Faktoren nach Relevanz für die jeweilige Betriebsebene zu bewerten. Die operative Expertise der FABEC PMG sowie die für FABEC (2019) durchgeführten Feldstudien lieferten einen wesentlichen Beitrag dazu.

Ein wesentlicher Schritt der Untersuchung bestand in der Analyse der verfügbaren Daten in Zusammenarbeit mit der FABEC PMG PBTF. Diese wurden auf mehreren operativen Ebenen hinsichtlich Erfassungsmethodik und Qualität überprüft. Die Validierung ist insbesondere deshalb notwendig, da die DEA Methodik sensitiv auf Extremwerte reagiert. Eine mangelhafte Datenqualität kann daher zu verzerrten Ergebnissen führen. Da offizielle Berichte und Studien primär PRU Daten verwenden, fokussierte auch die Datenanalyse diese Quelle. Dabei konnte festgestellt werden, dass es innerhalb FABECs z. T. unterschiedliche Verfahren zur Akquirierung bzw. Verrechnung gibt. Dies ist insbesondere bei den Ressourcen bzw. der dafür entstehenden Kosten der Fall. Es ist anzunehmen, dass diese Inkonsistenzen auch in bzw. zu anderen ANSPs bzw. FABs existieren. Eine höhere Qualität können ANSP-individuelle Daten¹⁸⁷ aufweisen, wie die disaggregierte Analyse in Abschnitt 4.4 gezeigt hat. Diese können zur Validierung der öffentlich zugänglichen Quellen genutzt werden.

Für das Benchmarking wurden mehrere Methoden verglichen und die jeweiligen Vor- und Nachteile herausgearbeitet. Da stochastisch-parametrische Methoden umfangreiche a priori Annahmen voraussetzen, führt eine Fehlschätzung zu verzerrten Ergebnissen. Da diese Effekte in vielen Publikationen (vgl. Abschnitt 3.2.2) zu beobachten waren, wurde innerhalb der Dissertation der deterministische Ansatz der Data Envelopment Analysis verfolgt. Es konnte nachgewiesen werden, dass die DEA auf die im ATM relevanten Inputs und Outputs angewendet werden kann. Die Robustheit konnte durch die Formulierung mehrerer Modelle, die sich in den verwendeten Inputs und Outputs sowie der berücksichtigten ANSPs unterschieden, nachgewiesen werden. Die Auswahl der Inputs und Outputs wurde durch die im vorigen Absatz diskutierte Datenanalyse unterstützt. Erweiterungen des DEA Grundmodells führten, mit Ausnahme der Supereffizienz-DEA, nicht zu einer Steigerung der Ergebnisqualität oder zu ergänzenden Erkenntnissen.

Auf disaggregierter Ebene umfassen die PRU Daten nur einige ausgewählte Werte für ACCs. Als alternative Datenquelle wurde im Rahmen der Dissertation die Software NEST von EUROCONTROL verwendet. Die zugrundeliegenden Daten ermöglichen eine Analyse der Sektoren und stellen daher die höchstmögliche Granularität für das Performance Benchmarking dar. Innerhalb der Feldstudien wurden diese Daten validiert. Bei der Verwendung von NEST-Daten ist jedoch zu berücksichtigen, dass

¹⁸⁷ Z. B. die Daten des Leistungsinformationszentrums der DFS.



innerhalb des Tools keine Aktualisierung hinsichtlich der tatsächlichen Sektor-Konfigurationen stattfindet. Die aktualisierten Trajektorien¹⁸⁸ werden den geplanten Sektoren-Schemata zugeordnet. Daraus folgt, dass in den Daten Sektoren enthalten sein können, die im realen Betrieb nicht oder in einer anderen Zeitspanne geöffnet wurden. Die tatsächlichen Konfigurationen sind nicht veröffentlicht, wurden im FABEC Rahmen jedoch auf Anfrage durch die ANSPs bereitgestellt. Des Weiteren sind Sektoren per se ungeeignet für ein Benchmarking, da sie keine DMUs darstellen. Allerdings ermöglicht die Aggregation ausgewählter Datenpunkte auch die Analyse auf ACC Ebene und kann dadurch die PRU-Daten ergänzen.

5.1.3 Ursachenanalyse

5.1.3.1 Methodik

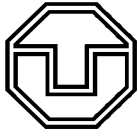
Eine wesentliche Motivation des Dissertationsvorhabens repräsentierte die Identifikation von „Stellschrauben“, welche es einem ANSP ermöglichen, die eigene Leistungsfähigkeit zu verbessern. Auf Basis der qualitativen und quantitativen Analysen sowie des zugrundeliegenden ökonomischen Modells wurden mehrere endogene, exogene und teilexogene Faktoren untersucht. Die „Stellschrauben“ der ANSP sind dabei durch die endogenen Faktoren abgebildet. Zusätzlich können auf Basis der teil-exogenen Aspekte Handlungsempfehlungen, z. B. gegenüber politischen Entscheidungsträgern, abgeleitet werden.

Für die Ursachenanalyse wurde eine Vielzahl (potentieller) Einflussfaktoren berücksichtigt, die auf Basis der Expertenbefragungen sowie durch die Zusammenarbeit innerhalb der FABEC PMG PBTF identifiziert wurden. Der Einfluss dieser Faktoren wurde mittels verschiedener Regressionsmethoden getestet. Auch hier war die Suche nach einer geeigneten Methodik sowie die Überprüfung ihrer Anwendbarkeit ein wesentlicher Bestandteil der Untersuchungen. Während für Produktivitätskennwerte eine OLS-Regression anwendbar ist, sind Effizienzwerte zwischen 0 und 1 beschränkt. Daher wurde getestet, ob Ergebnisse zwischen OLS-, Tobit- und Trunkated Regression abweichen.

Die Koeffizienten und statistischen Signifikanzen unterscheiden sich primär aufgrund des Benchmarking Modells. Innerhalb eines Modells weichen die Koeffizienten zwischen den einzelnen Regressionsansätzen nur in geringem Ausmaß ab. Insbesondere zwischen OLS- und Tobit Regression sind nur marginale Differenzen beobachtbar. Bei den Modellgüte-optimierten Regressionen sind die statistisch signifikanten Koeffizienten teilweise unterschiedlich und haben daher auch unterschiedliche Ausprägungen der Koeffizienten. Die iterativen Regressionen bestätigten diese Beobachtungen.

Für Panelmodelle wurden Pooled Regression, Fixed- und Random-Effects Modelle angewendet. Anhand von statistischen Tests (z. B. Hausman-Test) konnte die jeweils geeignete Methode identifiziert werden. Dabei wurde in den meisten Fällen der FEM-Schätzer bevorzugt. Dennoch führten nicht alle Ansätze zu plausiblen Ergebnissen. Dies war sowohl auf ANSP- als auch ACC Ebene zu beobachten (siehe dazu die folgenden Abschnitte). Eine qualitative Analyse (insbesondere auf disaggregierter

¹⁸⁸ „actual trajectories“ vs. „planned trajectories“ oder „initial trajectories“.



Ebene) ist sinnvoll, da nicht alle effizienzbeeinflussenden Faktoren quantifizierbar sind. Dennoch können Arbeitsweisen, Planungsstrategien oder Systemunterstützungen einen effizienten Betriebsablauf fördern.

Die folgenden Abschnitte orientieren sich in ihrer Struktur an den in Abbildung 5-1 dargestellten Ergebnis-Clustern. Sie fassen den Einfluss einzelner Faktoren auf die Leistungsfähigkeit der DMUs zusammen. Dabei ist zu berücksichtigen, dass nicht alle wesentlichen Einflussfaktoren quantifiziert werden konnten und damit in der Regressionsanalyse fehlen. Des Weiteren konnten einige Faktoren lediglich als Dummy-Variable integriert werden. Weist die Ausprägung dieser Variable eine hohe Varianz auf, so sinkt die Aussagekraft der Dummy-Variable.

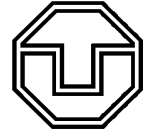
5.1.3.2 Betrieb

Die in Abschnitt 4.4 diskutierten Feldstudien ermöglichten einen umfangreichen Einblick in Betriebs- und Planungsmechanismen einzelner ACCs. Hier konnten mittels des entwickelten Fragebogens signifikante Unterschiede identifiziert werden, u. a. hinsichtlich Ressourcen-Puffern oder der verwendeten Systeme und Tools. FABEC-ANSPs nutzen verschiedene ATS-Systeme, die sich auch in Funktionalitäten unterscheiden. Hier könnte ein Problem hinsichtlich der Interoperabilitäts-Klausel der SES-Statuten entstehen (als Beispiel sei der Unterschied zwischen vektoren- und trajektorienbasierter Systeme genannt). Diese Faktoren sind jedoch nicht quantifizierbar, sodass der Effekt auf die Produktivität oder Effizienz nicht mittels Regression analysiert werden konnte.

Ein wesentliches Ergebnis von FABEC (2019) war zudem eine Beschreibung jener Faktoren, die in weiteren Studien untersucht werden sollten. Dies betrifft unter anderem die Lizenzierung der Lotsen (Erhöhung der Flexibilität versus Ausbildungskosten) sowie die Implementierung von Single ATCO Sector Operations. Die Analysen von Standfuss und Whittome (2019a) bestätigen, dass SSO in signifikanten Vorteilen resultieren können. In der Modellierung wurde dieses Konzept lediglich in Nachtschichten angewendet. Dennoch resultierte eine Approximation in einem jährlichen Einsparpotential von 124 Mio. €. Dieses Konzept würde jedoch eine neue Strukturierung des Luftraumes erfordern, da die Sektoren für SSO verkleinert werden müssten. Zudem stellen politische oder rechtliche Restriktionen, wie im Falle von skyguide¹⁸⁹, wesentliche Inhibitoren dar. Dennoch ist eine ANSP-individuelle Überprüfung empfehlenswert, ob SSO zumindest zu bestimmten Tageszeiten und/oder bei spezifischen Verkehrsaufkommen anwendbar ist.

Gemäß den durchgeführten Regressionsanalysen hat die jährliche Arbeitszeit des Lotsen einen Einfluss auf die Leistungsfähigkeit. Dabei unterscheidet sich das Ergebnis jedoch zwischen uni- und multiperiodischen Ansätzen. Abbildung 4-22 zeigte, dass sich die durchschnittlichen jährlichen Arbeitszeiten in Europa stark unterscheiden. Eine hohe Arbeitszeit wirkt sich positiv auf ATCO-Produktivität und DEA-Score aus, jedoch negativ auf das TFP-Wachstum. Diese Ergebnisse sind ambivalent. Da durch höhere Jahresarbeitszeiten Ressourcen gespart oder anders eingesetzt werden können, ist ein

¹⁸⁹ Aufgrund des Flugzeugunglücks von Überlingen wäre hier bspw. eine Einmann-Besetzung (SSO) nicht umsetzbar.



positiver Zusammenhang zu erwarten. Die in Abschnitt 4.6.4 gerechnete Approximation ergab ein Einsparpotential von 790 Mio. €. (wenngleich Allokationseffekte innerhalb der ANSPs vernachlässigt wurden). Allerdings können lange Arbeitszeiten auch zu einer Ermüdung der Lotsen und damit zu einer geringeren individuellen Arbeitsleistung führen. Dies würde ein negatives Vorzeichen erklären. In diesem Fall gäbe es, ähnlich der Betriebsgröße (siehe Abschnitt 4.5), einen Wendepunkt, an dem die Performance mit der Arbeitszeit sinkt.

Ein eng verknüpfter Effekt ist auf ACC Ebene zu beobachten. Hier wurde das Verhältnis von ATCO-Stunden zu Sektorstunden als Proxy für Ressourcenpuffer deklariert. Die Regressionsanalyse zeigt, dass sich dieser Faktor negativ auf die Produktivität auswirkt. Der Zusammenhang entspricht den Erwartungen. Dennoch sind diese Puffer notwendig, um die Nachfragevolatilität abzufedern (siehe auch Abschnitt 5.1.3.4). Diese Ressourcenpuffer können dazu führen, dass mehr ATCOs (FTEs) eingesetzt werden, diese jedoch aufgrund der Nachfrageschwankungen weniger Zeit „on board“ verbringen. Dadurch reduziert sich die jährliche Arbeitszeit je ATCO.

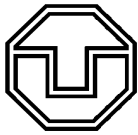
Innerhalb der Regressionsanalyse wurde weiterhin der Anteil an Nicht-ATCOs an den Personalressourcen adressiert. Da die Nicht-ATCOs keinen unmittelbaren Beitrag zur Wertschöpfung liefern, wäre ein negativer Zusammenhang plausibel. Dies wird, bis auf wenige Ausnahmen in einzelnen Iterationsschritten, durch die Regressionsmodelle bestätigt, sowohl auf Basis der Produktivität gemäß PRU als auch der DEA-Scores. Des Weiteren wirkt sich ein interner Wetterdienst („Internal MET Service“) positiv auf die Leistungsfähigkeit aus. Komplizierter ist jedoch der Einfluss der Eigentumsform. Hier wurde primär in „joint stock company“ und „state owned“ unterschieden, da eine reine Unterscheidung nach staatlich und privatisiert nicht möglich ist. Die Ergebnisse zeigen, dass der Faktor „staatlich“ eher zu einer Ineffizienz führt, als die Organisation in einer GmbH. Die Entscheidung über die Unternehmensform kann jedoch als teil-exogen und politisch beeinflusst betrachtet werden.

5.1.3.3 Luftraum

Sowohl in bisherigen Publikationen und offiziellen Berichten als auch innerhalb der vorliegenden Dissertation wurde der Luftraum als wesentlicher Faktor von Produktivität und Effizienz identifiziert. Die Ergebnisse der Regressionsanalyse zeigen, dass die Leistungsfähigkeit mit der Größe des Luftraumes steigt. Dies gilt sowohl für die ANSP- als auch für die ACC-Ebene. Auch „Oceanic Airspaces“ wirken positiv auf die Leistungsindikatoren ein. Jedoch wurden bei der Regression keine Skaleneffekte berücksichtigt. Zudem wurde ein negativer Zusammenhang zwischen Performance und der maximalen Anzahl an Sektoren nachgewiesen. Daher spielt nicht nur die Größe des Luftraumes eine übergeordnete Rolle, sondern auch dessen Fragmentierung. Um dieses Thema zu adressieren, wurde die derzeit verfolgte Strategie der Functional Airspace Blocks¹⁹⁰ adressiert. Alternativ wird zudem eine Strukturierung durch die dynamische Sektorisierung¹⁹¹ in Abschnitt 5.2.2 diskutiert.

¹⁹⁰ Makroskopischer Ansatz.

¹⁹¹ Mikroskopischer Ansatz.

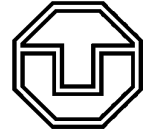


Mittels einer Analyse der Skaleneffizienz wurde überprüft, inwieweit sich Größenvorteile im europäischen Flugsicherungssektor realisieren lassen. Dies implizierte, dass die Vorteile durch die Konsolidierung von Lufträumen entstehen. Dies erfolgte unter Berücksichtigung der Arbeitsweise und den betrieblichen Einheiten eines ANSP sowie der Heterogenität der angebotenen Dienstleistungen. Die Ergebnisse deuten darauf hin, dass Größenvorteile nicht für alle Einheiten realisierbar sind, sondern ein Wendepunkt existiert, an dem eine weitere Vergrößerung der operativen Einheit zu Größennachteilen führt.

Größenvorteile konnten für 13 der 27 ANSPs eindeutig nachgewiesen werden ($SI \geq 3$), das heißt, dass diese ANSPs in mindestens drei von vier Modellen steigende Skaleneffekte aufwiesen. Drei von ihnen wurden lediglich in Modell 1 als skaleneffizient eingeordnet. Insbesondere Flugsicherungen mit sehr kleiner Unternehmensgröße, bspw. Belgocontrol und LVNL (Niederlande), können von einem Konzept wie den FABs profitieren. Die angewendete Methodik ermöglicht jedoch keine Ableitung, wie groß ein FAB im Optimum sein sollte. Größennachteile konnten für sechs ANSPs eindeutig nachgewiesen werden ($SI \leq -3$), insbesondere für Spanien, Italien und Frankreich. Für 18 ANSPs konnte keine eindeutige Aussage getroffen werden (SI zwischen -2 und 2). Dies liegt zum einen daran, dass sich die Ergebnisse modellübergreifend unterscheiden, zum anderen arbeiten einige Flugsicherungen bereits im Betriebsoptimum, sodass keine Größenvor- oder -nachteile bestätigt werden können.

Des Weiteren wurde der Zusammenhang von Skaleneffizienz, differenziert nach steigenden und fallenden Skalenerträgen, mit Luftraumgröße und Verkehrskomplexität untersucht. Es konnte nachgewiesen werden, dass die Effizienz mit steigender Luftraumgröße bis zu einem Wendepunkt zunimmt. Diese hyperbolische Approximation erreicht ein Bestimmtheitsmaß von 72 %. Demnach sind ANSPs mit einer Luftraumgröße unter 300.000 km² primär von steigenden Skalenerträgen gekennzeichnet, während ein Luftraum von über 800.000 km² ausnahmslos zu fallenden Skalenerträgen führte. Innerhalb dieses Bereiches sind die ANSPs, bis auf wenige Ausnahmen (z. B. Deutschland und Italien), skaleneffizient oder nur marginal skaleneffizient.

Größenvor- und -nachteile können neben der Luftraumgröße auch von Menge und Konzentration des Verkehrs abhängen. Bereits in der Regressionsanalyse wurden auf ANSP- und ACC Ebene überwiegend positive Zusammenhänge zwischen der Dichte des Verkehrs und dem Leistungsindikator ausgewiesen (siehe auch Abschnitt 5.1.3.4). Ein direkter Zusammenhang zwischen Skaleneffizienz und Komplexität ist jedoch nicht nachweisbar. Differenziert man die Analyse nach ANSPs mit fallenden und steigenden Skalenerträgen, so ist für die erste Kategorie ein negativer Zusammenhang zwischen Komplexität und Skaleneffizienz abzuleiten. Zudem ist der Korrelationskoeffizient mit 0,82 relativ hoch. Das heißt, je höher die Komplexität, desto höher ist die Skaleneffizienz (im Falle von fallenden Skalenerträgen). Für steigende Skalenerträge ist keine signifikante Korrelation zu beobachten. Bei ANSPs mit einer hohen Verkehrsdichte (z. B. Belgocontrol) ist der Abstand zum Skalenoptimum geringer. Infolgedessen sind die potentiellen Vorteile einer Konsolidierung (z. B. Belgien-Niederlande) tendenziell geringer als für ANSPs mit geringer Verkehrsdichte (z. B. Albanien-Mazedonien).



Die FABs bieten Potentiale zur Optimierung, sowohl hinsichtlich der Zusammensetzung als auch der Struktur. Dies bestätigen die Regressionsanalysen, in denen für die FABs kein robuster positiver oder negativer Zusammenhang ausgewiesen wird. In den Panelanalysen ist der generelle¹⁹² Zusammenhang überwiegend negativ konnotiert, Vorzeichen und Ausprägung des Koeffizienten ist jedoch abhängig von dem individuellen FAB. Diese Ergebnisse werden von der Regressionsanalyse auf ACC Ebene bestätigt. Daraus lassen sich einige potentielle Einflussfaktoren für eine erfolgreiche Implementierung des Konzeptes ableiten.

Die Luftraumgrenzen der FABs werden größtenteils durch die politischen Grenzen der zugehörigen ANSPs determiniert. Als sinnvoller kann eine Allokation gemäß Verkehrsflüssen gesehen werden. Auch die Anzahl der in einem FAB organisierten ANSPs unterscheidet sich. Verkehrscharakteristiken wie Menge und Fluss sowie deren Volatilität unterscheiden sich je nach Flugsicherung, sodass die Systeme an diese Umwelteinflüsse angepasst und bspw. durch individuelle Tools unterstützt werden müssen. Alternativ wäre eine Konsolidierung von Einheiten auf ACC Ebene bzw. Einheiten mit geringer Heterogenität hinsichtlich Verfahren und Systemen denkbar.¹⁹³ Da diese Einheiten meist äquivalente Anforderungen haben, ist hier die Anpassung von Systemen effektiver. Zuletzt können auch sprachliche, kulturelle, rechtliche und operative Unterschiede zusätzlich eine Herausforderung in der Umsetzung eines FAB darstellen.

Die Analyse zeigt, dass der makroskopische Ansatz einer Umstrukturierung kritisch gesehen werden muss. Um die Ziele des Flight Path 2050¹⁹⁴ zu erreichen, ist eine alternative Aufteilung des europäischen Luftraumes jedoch empfehlenswert, insbesondere um prognostizierte Zuwächse in der Verkehrsnachfrage sicher und effizient bewältigen zu können. Starre Sektorengrenzen können hier als Einschränkung gesehen werden, da die Verkehrsströme an diese Struktur angepasst werden. Effizienter wäre u. U. eine dynamische Anpassung der Sektorengrenzen an die aktuelle Verkehrsnachfrage sowie deren Schwankungen. Ein potentieller mikroskopischer Ansatz zur Anpassung der Luftraumstruktur stellt das Flight Centered Air Traffic Management dar (siehe Abschnitt 5.2.2).

5.1.3.4 Nachfrage

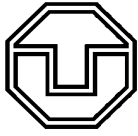
Bereits im ökonomischen Modell wurde dargestellt, dass die Verkehrscharakteristiken einen Einfluss auf die Leistungsfähigkeit eines ANSP nehmen. Menge, Verteilung und Flüsse haben eine direkte Auswirkung auf Interaktionen zwischen den Flugzeugen und damit auf den Taskload eines Lotsen (EUROCONTROL 2006; Gerdes et al. 2016). Der Taskload begrenzt wiederum die Kapazität des Sektors und damit den möglichen Output. Dies führt zu hohen Produktivitätswerten für Flugsicherungsdienstleister, welche hauptsächlich (oder ausschließlich) den oberen Luftraum kontrollieren.

In der Folge entwickelte PRU eine Metrik, um die durch die Luftraumnutzer verursachte Komplexität zu bewerten. Die Bewertung enthält vier Parameter: Angepasste Dichte, horizontale Interaktionen,

¹⁹² FAB ja/nein.

¹⁹³ Bspw. die Konsolidierung von ACCs, die lediglich den oberen Luftraum betreuen.

¹⁹⁴ U. a. Stärkung der Wettbewerbsfähigkeit und Reduktion der CO₂-Emissionen.



Geschwindigkeitsinteraktion und vertikale Interaktionen. Der Score wird u. a. verwendet, um Produktivitätsindikatoren und die Verkehrskomplexität zu vergleichen. Darüber hinaus wurde der Complexity Score zur Festlegung der Leistungsziele für die Referenzperiode 3 genutzt (European Commission 2018).

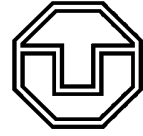
Der Einfluss der Komplexität auf die Leistungskennwerte wurde mittels Regressionsanalyse überprüft. Bereits in Vorab-Studien konnte kein Zusammenhang zwischen dem Complexity Score und der Produktivität nachgewiesen werden (Standfuss 2020). Diese Untersuchung fand sowohl auf ANSP als auch auf ACC Ebene statt. Außerdem wurden Extremwerte eliminiert und lediglich Enroute Services fokussiert. Keine der Anpassungen führte zu einer statistisch signifikanten Interdependenz. Diese Erkenntnisse wurden mittels Regressionsanalyse größtenteils bestätigt, wenngleich einige Modelle eine positive Interdependenz schätzen.

In einem weiteren Schritt wurden daher die Komplexitäts-Komponenten einzeln betrachtet. Auch hier konnten keine robusten Einflüsse nachgewiesen werden. Die „Adjusted Density“ war in den meisten Modellen statistisch signifikant und positiv konnotiert. Der Horizontal- und Vertikal-Score erreichte in den wenigsten Modellen bzw. deren Spezifikationen (Iterationsschritten) statistische Signifikanz und war zudem meist positiv. Lediglich der Koeffizient des Speed Scores war in einigen Modellen statistisch signifikant und negativ. Dies würde zu dem Ergebnis führen, dass Komplexität keinen nachweisbaren Einfluss auf die Leistungsfähigkeit besitzt. Tatsächlich legt dies jedoch die Vermutung nahe, dass das Komplexitätsmaß der PRU per se einige Schwächen aufweist (Standfuss und Rosenow 2020).

Erstens repräsentiert die Verkehrsdichte einen wesentlichen Anteil des Gesamtscores (EUROCONTROL 2020). So wird bspw. aus dem Performance Review Report deutlich, dass die Komplexität parallel zur Verkehrsnachfrage steigt (EUROCONTROL 2019h). Zweitens ist nach Angaben der ATCOs militärischer Verkehr ein Hauptfaktor für Komplexität. Dieser beinhaltet in erster Linie Anzahl und Dauer von TRAs. Nach Ansicht von Betriebsexperten ist zudem die Vorhersehbarkeit der militärischen Nachfrage gering, was die negativen Auswirkungen auf die Bereitstellung von Kapazitäten noch verstärkt. Dennoch ist der militärische Verkehr nicht im Komplexitätsmaß integriert.

Die Komplexität hängt drittens von den Wetterbedingungen wie Wind oder Sichtweite ab. Einerseits führt gutes Wetter zu einem höheren Anteil des VFR- bzw. GA-Verkehrs, was die Komplexität zusätzlich erhöhen kann. Andererseits können Gebiete mit schlechtem Wetter zu Umleitungen (abweichende Trajektorien) führen, was ebenfalls zu einer hohen Komplexität beitragen kann (Rosenow et al. 2017b). Viertens basiert die Berechnung der Scores auf ex-ante-Simulationen und der daraus abgeleiteten potentiellen Interaktionen. Eine Analyse der tatsächlichen Interaktionen sowie die Berücksichtigung der diskutierten Inhibitoren könnten zu einem zielführenderen Score führen. Fünftens wird primär der Enroute-Verkehr adressiert. Der PRU Complexity Score ist daher für Terminal Dienstleistungen sowie Gate-to-Gate Betrachtungen ungeeignet.

Für die Erfassung und Bewertung der Nachfrage existieren alternative Ansätze (Diaconu et al. 2014; Djokic 2014; Salaün et al. 2010; Schelbert et al. 2012; Sridhar et al. 1998). Diese konnten innerhalb der Dissertation jedoch aufgrund der Datenverfügbarkeit nicht berücksichtigt werden. Zum einen



sind diese Metriken primär für Enroute-Verkehr konzipiert, während für die Benchmarking-Modelle eine Gate-to-Gate-Perspektive bevorzugt wurde. Zum anderen fehlen die für die Berechnung notwendigen Inputdaten. Es wird dennoch empfohlen, ein alternatives Maß für die Komplexität innerhalb des ANSP-Benchmarkings zu entwickeln.

Mit der Komplexität gekoppelt ist der Anteil an Überflügen und Inlandsflügen. Überflüge finden meist im oberen Luftraum statt, wo auch der Vertikalverkehr geringer ist. Dieser spielt wiederum bei Inlandsflügen eine Rolle. Das wird durch die Regressionsanalyse bestätigt. In den meisten Modellen sind Überflüge erwartungsgemäß positiv konnotiert, Inlandsflüge negativ. Statistische Signifikanz wird insbesondere bei der Tobit Regression und den Panel-Modellen erreicht.

Eine weitere Einflussgröße repräsentieren kurzfristige, unplanmäßige Änderungen des Verkehrs (unanticipated Traffic), entweder durch zusätzliche Flüge („Intruder“) oder dem Wegfallen geplanter Flüge („Runaways“), z. B. durch Routenänderungen. Diese Abweichungen können auf mikroskopischer Ebene Mehrbelastungen oder Mehrkosten verursachen. Der Einfluss dieser Abweichungen wurde daher auf Sektorebene durch FABEC untersucht (FABEC 2018b, 2018c). Auch langfristige Schwankungen (z. B. Saisonalität) steigern die Vorhaltekosten für Ressourcen. EUROCONTROL berechnet diese Abweichung durch das Verhältnis von peak-load zu average-load. In Zusammenarbeit mit FABEC wurden alternative Metriken auf Anwendbarkeit untersucht, welche die Schwankungen aller Zeitabschnitte berücksichtigt (Standfuss et al. 2019b).

Innerhalb der Dissertation wurden verschiedene Maße für die saisonale Volatilität auf Anwendbarkeit getestet. Dabei sind Streumaße (Standardabweichung oder Variationskoeffizient) von Metriken der Ungleichverteilung (GINI oder HHI) zu unterscheiden. Aufgrund der Skalendependenz führt die Anwendung der Streumaße zu signifikant unterschiedlichen Ergebnissen. Dagegen sind GINI und HHI durch eine hohe Korrelation gekennzeichnet. Auf Basis der Regressionsanalyse konnte festgestellt werden, dass eine Ungleichverteilung des Verkehrs negativ auf die Performance wirkt. Dieses Ergebnis ist modellübergreifend robust und weist statistische Signifikanz auf.

Ein weiterer potentieller Einflussfaktor ist der Traffic Mix. Er kann als wesentlicher Faktor eines holistischen Komplexitätsmaßes angesehen werden. Innerhalb einer FABEC-Studie wurde zudem der IFR GA Verkehr untersucht. Hier bestehen zwischen den USA und Europa wesentliche Unterschiede in der Nachfragestruktur. Aufgrund fehlender Datenverfügbarkeit konnte diese Kenngröße nicht in die Regression einbezogen werden. Jedoch war es möglich, etwa drei Prozentpunkte des in offiziellen Berichten publizierten Leistungsunterschiedes zu erklären.

Der letzte Punkt adressiert den Einfluss der Prognosequalität auf die Ressourcenplanung. Im Allgemeinen sind langfristige Vorhersagen verzerrter als kurzfristige, da die beeinflussenden Ereignisse nicht umfassend vorhersagbar sind (zufälliger Fehler). Dies ist bspw. an der aktuellen Corona-Pandemie zu beobachten. Die prognostizierten BIP-Zuwächse und abgeleiteten Verkehrszahlen rechnen diese Pandemie nicht ein, sodass die Nachfrage überschätzt wird.



In Standfuss und Whittome (2019b) wurden verschiedene Metriken zur Bestimmung der Prognosequalität angewendet. Diese Metriken berechnen in erster Linie absolute Werte, was zu einem Verzerrungsindikator führt, der nicht berücksichtigt, ob die Nachfrage über- oder unterschätzt wurde. Im Zusammenhang mit den Flugsicherungsdiensten könnte dies jedoch notwendig sein, da die Kostenstrukturen unterschiedlich sind. Einerseits kann eine Unterschätzung der Nachfrage zu einem Mangel an Ressourcen führen. Wenn die Nachfrage die Kapazität übersteigt, resultiert dies in Verspätungen und dadurch in Kosten für die Luftraumnutzer. Andererseits wird im Falle einer Überschätzung die Kosteneffizienz des Flugsicherungsdienstes abnehmen, da das Verhältnis zwischen Output und der dafür eingesetzten Ressourcen nicht optimal ist¹⁹⁵. Das Hauptziel wäre die Minimierung der Gesamtkosten, die sich aus den Betriebskosten (ANSPs) und den Verspätungskosten (Luftraumnutzer) zusammensetzen.

5.1.3.5 Umweltfaktoren

Neben den betrieblichen und verkehrlichen Einflussfaktoren wurden innerhalb der Dissertation auch weitere, exogene Effekte untersucht. Die bisher diskutierten Analysecluster gemäß Abbildung 5-1 weisen dabei oft Interdependenzen zu politischen, geographischen oder volkswirtschaftlichen Implikationen auf. So ist bspw. die Schaffung der FABs sowie der Allokation der individuellen ANSPs politisch, aber auch geographisch geprägt. Staffelungsabstände und dafür notwendige Maßnahmen sind ebenfalls institutionell vorgegeben.

Andere untersuchte Effekte betreffen den Wohlstand des Landes und, damit verbunden, das Lohnniveau in den jeweiligen Staaten. Dies hat wiederum einen direkten Einfluss auf Ressourcenkosten. Es wurde außerdem angenommen, dass der Stand der Technologie die Performance beeinflusst. In den Regressionsanalysen konnte aber weder der Proxy für Forschung noch jener für Wohlstand diesen Zusammenhang nachweisen. Dies kann darin begründet liegen, dass in der Leistungskennzahl maßgeblich operative Faktoren verwendet wurden. Auf ACC Ebene konnte allerdings gezeigt werden, dass die Kosten pro Lotsen sich positiv auf die Leistungsfähigkeit auswirken. Mit steigenden Kosten pro Lotsen wächst auch der Anreiz für einen ANSP, Lotsen effizient einzusetzen.

Auch die Finanzierung der europäischen Flugsicherungsdienste unterliegt einer Anreizregulierung und basiert auf Gebühren. Dadurch unterscheidet sich die Finanzierung vom amerikanischen System, welches primär steuerfinanziert ist. Beide Systeme haben Vor- und Nachteile. In der Dissertation wurde gezeigt, dass das Gebührensystem bspw. Umwege induzieren kann. Außerdem konnte innerhalb einer FABEC Studie nachgewiesen werden, dass den europäischen ANSPs bei der Finanzierung nach amerikanischem System mehr finanzielle Mittel zur Verfügung stünden (Whittome und Standfuss 2018). Wenngleich diese Studie lediglich eine grobe Approximation darstellt, so ist der Unterschied zwischen beiden Finanzierungsoptionen signifikant hoch. Es kann daher angenommen werden, dass implizite Mess- und Modellierungsfehler das grundsätzliche Ergebnis nicht beeinflussen.

¹⁹⁵ Die Kosten der Ressourcen wurden mittels eines implizierten, linearen Zusammenhangs zwischen Nachfrage und Lotsenbedarf approximiert.

5.2 Ausblick

5.2.1 Neue Marktteilnehmer

In zukünftigen Luftverkehrskonzepten werden zunehmend neue Marktteilnehmer zu berücksichtigen sein. Dazu zählen kommerzielle Raumtransporte, unbemannte Transportflugzeuge, Drohnen oder Individualverkehr (z. B. Lufttaxis). Die erwartete, steigende Nachfrage nach kommerziellen Raumtransporten erfordert ein interoperables Luft- und Raumfahrtmanagement (Kaltenhäuser et al. 2017). Eine nahtlose und sichere Integration in den zivilen Luftraum ist von entscheidender Bedeutung (siehe Abbildung 2-11). Dabei müssen verschiedene Konzepte für künftige Raumfahrtoperationen, von Raketenstartoperationen bis hin zu suborbitalen Flügen mit Betrieb in Flughafenumgebung in das ATM integriert werden. Dafür ist es notwendig, dass das operationelle Luftraumkonzept zeitlich und räumlich flexibel ist, um die Flugbahnänderungen des Raumfahrzeuges zu berücksichtigen.

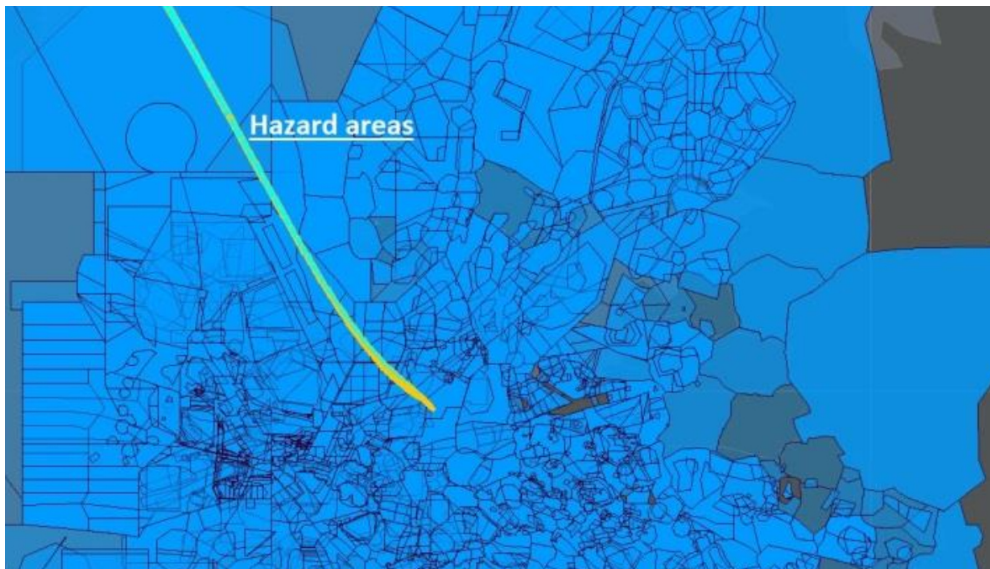
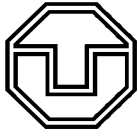


Abbildung 5-2: Integration von Weltraumflügen in den europäischen Luftraum
(Luchkova et al. 2016)

Verschiedene Studien beschäftigen sich auch mit der Luftraumbewirtschaftung bei steigendem luftgebundenen Individualverkehr (Sunil et al. 2015). Die Verkehrssystemtheorie unterscheidet im bodengebundenen Verkehr drei Phasen: Freie Strömung, gebundene Strömung (synchronisierter Verkehr) und Stau (Kerner 1998). Jeder Zustand erfordert unterschiedliche Arten von Lösungen hinsichtlich des Verkehrsmanagements. Auf den Luftverkehr übertragen repräsentiert eine freie Strömung die Flüge im oberen Luftraum, während die synchronisierte bzw. überlastete Phase charakteristisch für den Terminalbereich und die Landebahn ist (Schultz et al. 2017). Aufgrund der steigenden Verkehrsnachfrage durch private Luftfahrzeuge wird erwartet, dass ein zukünftiges Stadtverkehrsmanagement in erster Linie die synchronisierten und überlasteten Phasen bewältigen muss. Dies setzt eine spezifische Anpassung des ATM voraus. Verschiedene Arten von Managementkonzepten wurden bereits innerhalb des Projektes Metropolis getestet (Sunil et al. 2015).



Auch im Enroute-Bereich werden neue Teilnehmer (z. B. Drohnen) sowie alternative Transportkonzepte (z. B. unbemannte Frachtmaschinen) den zukünftigen Luftraum prägen. Diese neuen Marktteilnehmer erfordern eine Integration in das Luftverkehrsmanagement (EUROCONTROL 2018f; Temme 2016). Es ist anzunehmen, dass die Kapazität derzeitiger Strukturen nicht ausreicht, um die wachsende Nachfrage zu befriedigen. Eine Alternative stellt das Flight-Centered ATM dar. Es sieht vor, dass nicht mehr der Verkehrsfluss einer quasi-fixen Struktur folgt, sondern sich die Struktur dynamisch an die Nachfrage der Luftraumnutzer anpasst. Zwei wesentliche Konzepte umfassen den Ansatz einer dynamischen Sektorisierung (Gerdes et al. 2018; Standfuss et al. 2018d) oder ein sektorloses ATM (Birkmeier und Korn 2014; Korn et al. 2010). Inwiefern diese Konzepte dazu beitragen, die zukünftigen Anforderungen zu erfüllen, wird im folgenden Abschnitt anhand der dynamischen Sektorisierung diskutiert.

5.2.2 Dynamische Luftraumstrukturen

5.2.2.1 Hintergrund

Die zukünftige Luftraumstruktur Europas sieht nach den derzeitigen Plänen eine Organisation in FABs vor. In Abschnitt 4.5 konnte jedoch gezeigt werden, dass dieser Ansatz nicht für alle Flugsicherungen zu operativen Vorteilen führt. Um die ehrgeizigen Ziele des FP2050 sowie den Verkehrsprognosen Rechnung zu tragen, kann neben dem makroskopischen Ansatz der FABs auch ein mikroskopischer Ansatz als notwendig erachtet werden. Auch signifikante Verkehrsrückgänge, wie aktuell durch die Corona-Pandemie zu beobachten sind, sollten bei der Entwicklung einer zukünftigen Luftraumstruktur bzw. von alternativen Betriebskonzepten berücksichtigt werden.

Ein wesentliches Konzept des FC-ATM stellt neben der sektorlosen Flugführung (Birkmeier und Korn 2014; Korn et al. 2010) die dynamische Sektorisierung (Gerdes et al. 2016) dar. Während heute der Verkehr maßgeblich den vorgegebenen Luftraumstrukturen folgt, wird im Ansatz der dynamischen Sektorisierung diese Struktur dem Verkehr angepasst. Ziel ist eine alternative Strukturierung nach funktionalen Gesichtspunkten, welche eine nachfrageorientierte Zielfunktion (Aufkommen und Flow) berücksichtigt (Standfuss et al. 2016b, 2018d). Dynamische Konzepte sollen es zudem ermöglichen, die in Abschnitt 5.2.1 diskutierte Nachfrage zukünftiger Marktteilnehmer in das ATM zu integrieren (Schultz et al. 2018a). Erste Ansätze für die Implementierung dynamischer Sektorstrukturen wurden u. a. von Delahaye et al. (1998;), Sergeeva et al. (2015) sowie Sherali und Hill (2011) vorgeschlagen. Die folgenden Abschnitte beschreiben das entwickelte Konzept und die sich daraus ergebenden operativen und finanziellen Vorteile.

5.2.2.2 Ansatz

Bei einer (automatischen) Strukturierung des Luftraumes müssen verschiedene Aspekte berücksichtigt werden:

- Die Form der Einheiten sollte konvex sein und keine spitzen Winkel aufweisen.
- Es ist vorteilhaft, die Grenzen des Sektors parallel zu den Hauptverkehrsströmen zu platzieren, da dies den Koordinationsaufwand reduziert.

- Die Grenzen der Sektoren sollten zudem nicht nahe von Konfliktpunkten (z. B. kreuzende Hauptverkehrsströme) platziert werden.
- Die Größe und Form der Sektoren ist zudem an die Komplexität des Luftraumes und den damit einhergehenden Taskload des/der Lotsen anzupassen.

Der innerhalb des Dissertationsvorhabens fokussierte Schwerpunkt liegt auf der ökonomischen Bewertung dieses Ansatzes. Dennoch soll die Funktionsweise übersichtshalber erläutert werden. Die Erzeugung, Anpassung und Optimierung der Sektorgrenzen erfolgt gemäß dem Schema in Abbildung 5-3. Zu Beginn des Prozesses erfolgt die Leistungsbewertung des derzeitigen Luftraumsystems. Daraus ergeben sich Verbesserungspotentiale für die Luftraumstruktur. Diese Struktur wird dann in drei Schritten generiert und optimiert (Gedes et al. 2018; Standfuss et al. 2016b, 2018d).

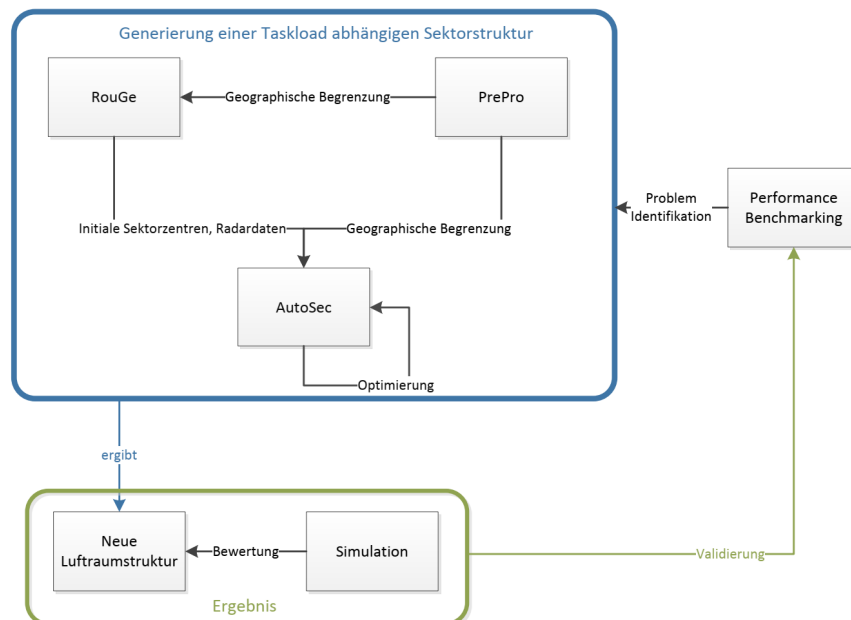
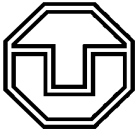


Abbildung 5-3: Methodischer Ansatz der dynamischen Luftraumoptimierung

Mit Hilfe des Tools „PrePro“ (Pre Processor) werden die benötigten Luftraumdaten aus den DDR2 Daten der EUROCONTROL generiert. Hierfür wird zunächst die Begrenzung des zu untersuchenden Bereiches¹⁹⁶ definiert (Hull). Mittels „RouGe“ (Route Generator) werden Radardaten gefiltert, basierend auf dem generierten Hull. Das heißt, die Trajektorien der im Untersuchungsraum durchgeführten Flüge werden ermittelt, basierend auf Start- und Zielflughafen sowie den Wegpunkten inklusive zugehöriger Zeiten (Abbildung 5-4, links). Diese Daten werden in das Programm „AutoSec“ (Automatic Sectorization) überführt. RouGe unterteilt zudem die Trajektorien in äquidistante Abschnitte. Jeder Abschnitt jeder Trajektorie wird mittels eines Punktes repräsentiert (Abbildung 5-4, Mitte). Aus diesen Punkten wird durch ein Fuzzy-Clustering-Verfahren eine vorher definierte Anzahl an Verkehrsclustern gebildet (Temme 2002). In diesem Ansatz wird das Fuzzy-c-means-Clusteringverfahren verwendet, bei dem

¹⁹⁶ Z. B. Deutschland, Europa o. ä.



eine Gruppe von Daten durch ihr Zentrum¹⁹⁷ (Schwerpunkt) repräsentiert wird (Krishnapuram und Keller 1993). Diese Zentren (Abbildung 5-4, rechts) werden als Initialisierung für ein Voronoi Diagramm verwendet.

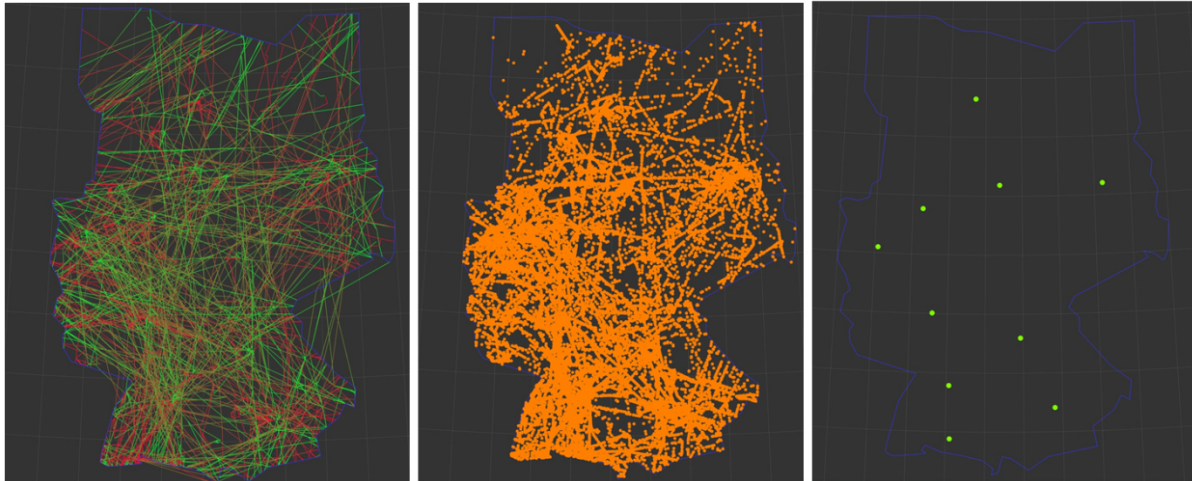


Abbildung 5-4: Von der Trajektorie zum Verkehrs-Cluster

Das Voronoi-Diagramm ermöglicht die Unterteilung einer Fläche bzw. eines Raumes in eine bestimmte Anzahl an Zellen (Berg et al. 2008). Diese Anzahl entspricht jener der durch das Fuzzy Clustering identifizierten Zentren. Die dadurch entstehenden Zellen werden durch Knoten und Kanten bestimmt. Kanten werden von allen Punkten erzeugt, die zu zwei Zentren gehören. Knoten/Ecken sind durch die Punkte definiert, welche drei verschiedenen Zentren zugeordnet sind (Abbildung 5-5).

In der Grafik repräsentieren die roten Punkte die Zentren, die blauen Linien die zugehörigen Knoten und Kanten. Die konvexe, geographische Umrandung (Hull) wird durch das grüne Viereck gebildet. Diese Konvexität kann bei Lufträumen nicht garantiert werden. Daher wurde die Berechnung der Sektorgrenzen derart erweitert, dass das Voronoi-Diagramm mittels der „line-segment-intersecting“-Methode an beliebige Grenzen angepasst werden kann (Gerdes et al. 2016).

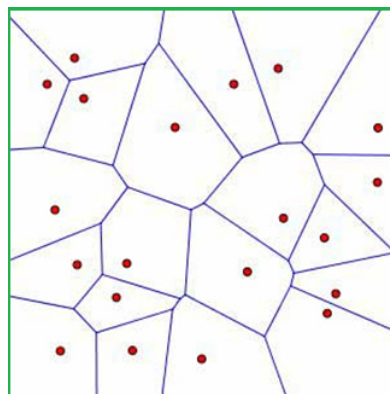
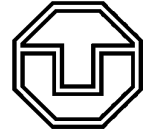


Abbildung 5-5: Beispiel für ein Voronoi Diagramm¹⁹⁸

¹⁹⁷ Gewichtet mittels Zugehörigkeitsgrad.

¹⁹⁸ Angelehnt an www.ams.org.



Die neue Grundstruktur des Luftraumes berücksichtigt in der aktuellen Aufteilung lediglich die Verkehrsnachfrage auf Basis der Trajektorien. Diese Struktur wird in einem nächsten Schritt über evolutionäre Algorithmen (Gerdes et al. 2004; Michalewicz 1996) hinsichtlich zusätzlicher Zielfunktionen optimiert, welche die zu Beginn diskutierten Anforderungen an die Sektoren widerspiegeln. Dadurch wird bspw. mittels geeigneter Taskload-Modelle (Gerdes et al. 2016) gewährleistet, dass die Arbeitslasten benachbarter Sektoren ähnlich hoch sind. Steigt in einem Sektor die Arbeitslast an, verschiebt sich die Grenze des Luftraumes, sodass der betroffene Sektor kleiner wird. Steigt die Arbeitslast über einen Grenzwert hinaus weiter an, so können die Sektoren automatisch neu strukturiert werden – bspw. durch die Öffnung eines zusätzlichen Sektors. Zudem passt sich die Struktur zeitlichen Schwankungen an und reagiert somit dynamisch auf die Volatilität des Verkehrs (Standfuss et al. 2018d).

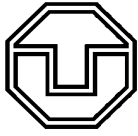
5.2.2.3 Betriebliche und ökonomische Vorteile

Im Allgemeinen ermöglicht die dynamische Sektorisierung ein flexibles Luftraummanagement, vor allem im Falle exogener Ereignisse, und vermeidet oder reduziert externe Effekte. Die wesentlichen betrieblichen Vorteile sind:

- Verbesserte Kapazitätsauslastung, wodurch Steuerungs-Maßnahmen und damit Umwege, Emissionen und Verspätungen oder Flugausfälle reduziert werden.
- Grundlegende Maßnahmen zur Kapazitätsanpassung bleiben gültig (z. B. Teilung und Zusammenlegen von Sektoren), wodurch die Lotsen-Akzeptanz des Konzepts gewährleistet ist. Dies repräsentiert einen wesentlichen Vorteil gegenüber dem sektorlosen ATM.
- Verringerung negativer externer Effekte, z. B. von Streiks.
- Verringerung der Folgen negativer exogener Effekte, z. B. Unwettern oder militärischen Aktivitäten.
- Dank flexiblerer Personalplanung können Flugsicherungsunternehmen das prognostizierte Verkehrswachstum mit der gleichen Menge¹⁹⁹ an Personalressourcen managen (besseres Demand Capacity Balancing).

Das Konzept der dynamischen Sektorisierung ist primär für den oberen Luftraum anwendbar, da hier im Wesentlichen Streckenverkehr stattfindet und die Komplexität gering ist. Somit könnte der gesamte obere europäische Luftraum neu strukturiert und die Sektoren unabhängig von Landesgrenzen allokiert werden. Dies hätte wesentliche operative und ökonomische Vorteile. Die dynamische Sektorisierung vermeidet Sektor-Teilungen bei steigender Nachfrage. Dadurch können kleinere Schwankungen in der Nachfrage abgedeckt werden. Dies ist insbesondere in hochfrequentierten Lufträumen ein wesentlicher Vorteil, da dadurch ein höheres Verkehrsaufkommen bei gleichem Ressourcenverbrauch abgewickelt werden kann. In schwach frequentierten Lufträumen kann die Kapazität wiederum effizienter genutzt werden, da durch die landesgrenzenübergreifende Optimierung Größenvorteile realisiert werden können. In beiden Lufträumen ermöglicht dies einen flexibleren und besser planbaren Personaleinsatz (Standfuss et al. 2018a).

¹⁹⁹ Oder verhältnismäßig geringerem Zuwachs.



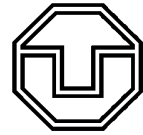
Für beide Luftraumtypen (oberer und unterer) schafft die Implementierung der dynamischen Sektorisierung ein System, welches robust gegenüber Unsicherheiten ist, die durch (unvorhersehbare) kurzfristige Verkehrsentwicklungen verursacht werden. Da die Struktur dem Fluss folgt, führen Verschiebungen in den Verkehrsmustern zu einer Verschiebung der Sektoren. Infolgedessen können Puffer für Personal reduziert werden, wodurch Zusatz- und Bereitschaft-Schichten und folglich Kosten vermieden werden. Gemäß Abschnitt 4.5 können Größenvorteile vor allem in den südosteuropäischen Ländern erwartet werden. Im europäischen Kerngebiet ist eher mit einer effizienteren Kapazitätsnutzung und damit der Reduktion von Verspätungen und Flugausfällen zu rechnen. Dies führt zu einer Verminderung der Kosten für Fluggesellschaften.

Es ist jedoch derzeit nicht möglich, den wirtschaftlichen Nutzen durch eine anpassungsfähige Sektorisierung en detail zu quantifizieren. Die Möglichkeit, die Gesamtzahl der Sektoren zu reduzieren, hängt von der Menge und der Komplexität des Verkehrs ab. Bis dato steht kein validierter Vergleich zur Verfügung, wie viele Sektoren (und damit ATCO-Stunden) tatsächlich eingespart werden können. Geht man von einer Reduktion um 5% aus, so könnten pro Jahr in Europa 130 Mio. € gespart, oder die Ressourcen zur Kapazitätserweiterung eingesetzt werden. Die tatsächliche Kapazitätssteigerung ist jedoch nicht abschätzbar, da der Paradigmenwechsel die Flug-Trajektorien beeinflussen könnte.

Vor allem die externen Kosten repräsentieren sehr hohe Einsparpotentiale, u. a. im Falle von Streiks. Im März 2018 verursachte der Streik der französischen DSNA innerhalb von drei Tagen (35 Stunden) eine Verspätung von über 6.600 Stunden (Flight Service Bureau 2018). Basierend auf einer Studie der University of Westminster verursacht jede Minute der Verspätung Kosten von 81€ (Cook und Tanner 2011), was zu Gesamtverspätungskosten von 32,4 Mio. € führt, oder umgerechnet 925.000€ pro Stunde. Diese Kosten wären teilweise vermeidbar, da ein anderer lizenziertes ANSP den oberen Luftraum teilweise oder vollständig betreuen könnte.

Im unteren Luftraum ist das Konzept aufgrund des hohen Spezialisierungsgrades und der erhöhten Komplexität nicht uneingeschränkt anwendbar. Daher wurde mit der DFS Niederlassung Bremen ein alternatives Konzept entwickelt. Dieses sieht vor, dass alle Sektoren einer EBG zusammengelegt werden (wie nachts bereits teilweise existent). Über einen Algorithmus können, ähnlich der sektorlosen Flugführung, Flüge den Executive-ATCOs zugewiesen werden. Dies reduziert nicht nur den Taskload, sondern erhöht gleichzeitig die Kapazität, da die nicht mehr benötigten Planner-ATCOs als Executive-ATCOs zur Verfügung stehen.

Um die wirtschaftlichen Vorteile einer dynamischen Sektorisierung in niedrigeren Lufträumen abzuleiten, ist es notwendig, die Anzahl der Sektoren und EBGn zu erfassen. Aus den NEST-Daten (EUROCONTROL 2018c) kann die Anzahl der Elementarsektoren berechnet werden. Allerdings sind auf europäischer Ebene keine Daten verfügbar, wie diese Sektoren zu EBGn zugeordnet werden. Auf der Grundlage von FABEC-Daten waren diese Informationen für Frankreich, Deutschland, Belgien, Luxemburg, die Schweiz und die Niederlande jedoch verfügbar. Die entsprechenden Flugsicherungsdienste betreuen einen Luftraum von insgesamt 1,7 Mio. km², welcher in 38 EBGn mit rund 300 Elementar-Sektoren gegliedert ist.



In diese Analyse werden keine Anflugsektoren integriert. Darüber hinaus ist der ACC Karlsruhe aufgrund der ausschließlichen Erbringung von Dienstleistungen im oberen Luftraum ausgeschlossen. Daher enthält der Datensatz 12 ACCs mit 251 elementaren Sektoren in 27 EBGn. Es wurde außerdem berücksichtigt, dass alle französischen und schweizerischen ACCs auch Dienstleistungen für obere Lufträume anbieten. Da die Komplexität der oberen Lufträume geringer ist und die Sektoren im Vergleich zu den unteren Lufträumen größer ausgelegt werden können, wurde modellhaft angenommen, dass mindestens 75 % der elementaren Sektoren dem unteren Luftraum zugeordnet sind. Darüber hinaus werden Elementar-Sektoren hauptsächlich in Zeiten hoher Verkehrsnachfrage verwendet. In der Modellierung wird davon ausgegangen, dass während der Nacht (22 bis 6 Uhr) die Sektoren ohnehin kombiniert sind und damit der EBG entsprechen. Zusätzlich wird angenommen, dass die Sektoren in 50 % der Zeit zusammengelegt werden.

Der Bedarf für einen Planner-ATCO wird wie folgt geschätzt:

- 0 in der Nachtzeit
- 1 für Tagschicht bei geringem Verkehr (50 %)
- 2 für Tagschicht bei hohem Verkehr (50 %)

Darüber hinaus wird angenommen, dass die Anzahl und Verteilung der Flüge sich nicht wesentlich verändert. Daraus folgt, dass die Anzahl der ausführenden ATCOs der Anzahl der elementaren oder kombinierten Sektoren entspricht.

Simuliert man die benötigten ATCO Stunden für den Ansatz im unteren Luftraum, dann sinken die ATCO-Stunden um 38 % auf 3.416 Stunden. Im Jahresvergleich bedeutet dies eine Einsparung von 773.800 ATCO-Stunden allein bei den in FABEC betreuten Lufträumen. Abbildung 5-6 verdeutlicht die jährlichen Kosteneinsparungen der Flugsicherungsdienste in den unteren Lufträumen durch die Implementierung einer dynamischen Sektorisierung. Insbesondere die französische DSNA und die deutsche DFS könnten hierdurch Vorteile erzielen. Die Gesamtkostenersparnis im Bereich FABEC beläuft sich auf 99 Mio. € pro Jahr.

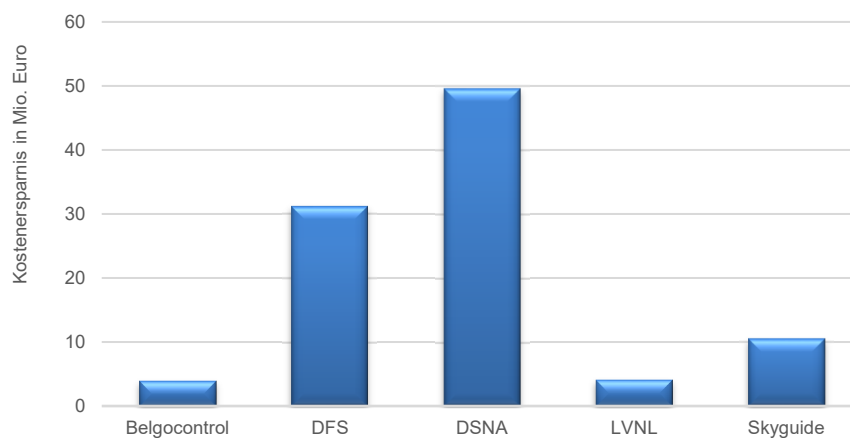


Abbildung 5-6: Kosteneinsparung durch dynamische Sektorisierung im unteren Luftraum



5.2.2.4 Schlussfolgerungen

Die dynamische Sektorisierung ist ein innovatives Konzept zur Strukturierung oder Umstrukturierung von Lufträumen. Durch die Definition von betrieblichen und wirtschaftlichen Zielfunktionen sowie sektorspezifischen Einschränkungen kann der Algorithmus auf verschiedene Regionen und Verkehrscharakteristiken angewendet werden. Sie bietet dadurch eine Alternative zur sektorlosen Flugführung und kann bspw. als Zwischentechnologie fungieren.

Der dynamische Sektorisierungsansatz bietet eine Lösung für Verkehrssituationen, die eine mittel- und kurzfristige Luftraumstrukturanpassung erfordern. Hier inbegriffen sind u. a. die charakteristischen morgendlichen und abendlichen Peaks bei interkontinentalen Verbindungen, Militäraktionen (Islami et al. 2017), Vulkanasche-Eruptionen (Luchkova et al. 2015), Konvektionszonen (Rosenow et al. 2018a), Verhinderung von Kondensstreifen (Rosenow et al. 2017a), Berücksichtigung kommerzieller Raumfahrtoperationen (Kaltenhäuser et al. 2017) sowie die Integration neuer Marktteilnehmer (Schultz et al. 2018b; Temme 2016). Darüber hinaus ergeben sich im Rahmen eines ökologisch effizienten Flugbetriebs neue Anforderungen an die effiziente Nutzung des Luftraums.

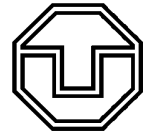
Auf Basis dynamischer Sektorgrenzen wurden zwei differenzierte Ansätze für den oberen und unteren Luftraum erarbeitet. Dies ist aufgrund der verschiedenen Nachfragecharakteristiken notwendig. Für beide Konzepte wurde der betriebliche Nutzen herausgestellt und finanzielle Einsparpotentiale approximiert. Dadurch konnte nachgewiesen werden, dass eine dynamische Sektorisierung wesentliche Vorteile gegenüber dem makroskopischen Ansatz aufweist. Sie kann sowohl die allgemein gültigen als auch die individuell erforderlichen Eigenschaften in eine Optimierung der Luftraumstruktur einbeziehen.

Der Ansatz basiert jedoch auf einer Reihe von Annahmen, die politischer bzw. rechtlicher Anpassungen bedürfen. Insbesondere die transnationale Verschiebung von Sektorengrenzen wird aufgrund der geltenden Lufthoheit eine Herausforderung darstellen. Allerdings sind hierfür bereits Ansätze in den Statuten zum Single-European-Sky manifestiert. Demnach darf jeder lizenzierte ANSP jeden europäischen Luftraum betreuen. Zur Umsetzung des Betriebskonzeptes ist zudem ein neuer Ansatz hinsichtlich der Finanzierung notwendig.

5.2.3 Weitere Forschungsaspekte

Abgeleitet aus den Erkenntnissen, welche in Abschnitt 5.1 zusammengefasst wurden, sowie den zukünftigen Herausforderungen (Abschnitt 5.2.1) und Betriebskonzepten (Abschnitt 5.2.2), können ergänzende Forschungsfragen formuliert werden. Ein wesentlicher Aspekt wird in der Flexibilisierung der Luftraumstrukturen gesehen. Auf makroskopischer Ebene sollte überprüft werden, ob die aktuelle FAB-Zusammensetzung verbessert werden kann. Konsolidierungen von Lufträumen mit ähnlichen Aufgaben bzw. Charakteristiken (z. B. oberer Luftraum) könnten einen höheren Nutzen bieten. Auch Zusammen-schlüsse auf ACC statt ANSP Ebene könnten Gegenstand zukünftiger Forschungen sein.

Die dynamische Sektorisierung stellt ebenfalls eine potentielle Verbesserung der Luftraumstruktur dar. Sie kann als Übergangstechnologie zu einem sektorlosen Luftverkehrsmanagement gesehen werden. Allerdings sind für dieses Konzept viele technische, ökonomische und rechtliche Fragen offen.



Als Beispiel sei hier ein alternatives Finanzierungskonzept genannt. Dies wird insbesondere dann nötig, wenn sich Sektorengrenzen über Staatsgrenzen hinweg verschieben. Des Weiteren ist eine Berechnung des ökonomischen Nutzens des Betriebskonzeptes im oberen Luftraum offen. Durch eine Simulation kann die Anzahl der notwendigen Sektorstunden (und damit Lotsenstunden) berechnet werden. Diese können dann mit dem tatsächlichen derzeitigen Ressourcenbedarf verglichen werden. Ergänzt wird dieser Nutzen durch die Reduktion von Verspätungen und Umwegen und damit der Vermeidung von externen Kosten.

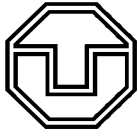
Der bestmögliche Einsatz der vorhandenen Ressourcen setzt unter anderem präzise Prognosen des Verkehrsaufkommens voraus. Es konnte innerhalb von Vorstudien jedoch nachgewiesen werden, dass die vorhandenen Prognosen (STATFOR) ein verhältnismäßig großes Konfidenzintervall nutzen. Dies erschwert eine effiziente Personalplanung signifikant (siehe Abschnitt 4.6.1). Insbesondere in stark frequentierten Lufträumen kann es in Folge dessen zu Verspätungen kommen. Gemäß des Flightpath 2050 sollte im Durchschnitt der Schwellenwert von 0,5 Verspätungsminuten pro Flug nicht überschritten werden. Es wäre daher zu untersuchen, wie hoch das Konfidenzintervall maximal sein dürfte, um diese Zielvorgabe zu erreichen. Dieses Intervall kann als notwendiger Ressourcenpuffer interpretiert werden.

Ein weiterer Aspekt sind die Mess- bzw. Berechnungsverfahren (Datenerhebung). So konnte bspw. mittels Regressionsanalyse nur in Ausnahmefällen ein Einfluss der Komplexität auf die Performance nachgewiesen werden. Dieses unerwartete Ergebnis legt den Schluss nahe, dass die Metrik der PRU möglicherweise unvollständig ist oder methodische Schwächen aufweist. In einer Vorstudie konnte nachgewiesen werden, dass wesentliche Elemente, welche den Taskload eines Lotsen erhöhen, nicht im Komplexitätsmaß der PRU integriert sind (Standfuss 2020). Ein wesentlicher Forschungsaspekt kann daher in der Modellierung und/oder Anwendung einer alternativen Metrik gesehen werden (Standfuss und Rosenow 2020). Auch für die Messung des Outputs mittels Composite Flight Hours könnte eine Gewichtung auf Basis betrieblicher Daten wesentliche Vorteile bringen.

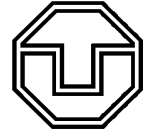
5.3 Handlungsempfehlungen

Basierend auf den in der Dissertation gewonnenen Erkenntnissen können die im folgenden zusammengefassten Handlungsempfehlungen formuliert werden. Diese sind differenziert nach institutionellen, akademischen, betrieblichen und politischen Aspekten, wenngleich auch hier eine eindeutige Zuordnung nicht immer möglich ist. Eine Gewichtung der einzelnen Empfehlungen gestaltet sich schwierig, da die einzelnen Punkte oft verknüpft sind und/oder einander bedingen. Dennoch kann die Reihenfolge der Unterpunkte als eine grobe Gewichtung der einzelnen Aspekte interpretiert werden.

1. Institutionelle Handlungsempfehlungen für EUROCONTROL sowie die publizierten Performance Berichte:
 - 1.1. Europaweit verbindliche Vorgaben für eine konsistente Erfassung der PRU Daten bzw. der Verrechnung von Komponenten.
 - 1.2. Eindeutige Allokation einzelner (Kosten-) Komponenten zu Enroute- und Terminal-Services.



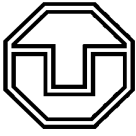
- 1.3. Keine Zulassung von Schätzdaten.
 - 1.4. Bereitstellung der tatsächlichen Sektor-Konfigurationen und -Öffnungszeiten für NEST.
 - 1.5. Überarbeitung des aggregierten Outputmaßes „Composite Flight Hours“. Gewichtung nach individuellen operativen Merkmalen (z. B. Komplexität).
 - 1.6. Entwicklung eines neuen Komplexitätsmaßes. Umstellung von Plan- auf Ist-Daten sowie Einbeziehung bisher unberücksichtigter Faktoren (u. a. Militär und Wetter). Bereits publizierte, alternative Ansätze sollten auf Möglichkeiten der Implementierung überprüft werden.
 - 1.7. Nutzung von multidimensionalen Analysemethoden für das ANSP-Benchmarking. Auf diesem operativen Level ist ein Indexziffer-Ansatz nicht zielführend. Auch für die Ursachenforschung reichen zweidimensionale Grafiken meist nicht aus, da Quereffekte nicht abgebildet werden.
 - 1.8. Für den EU-US Vergleich: Berücksichtigung von operativen und finanziellen Unterschieden in den Performance-Indikatoren bzw. der Ergebnispräsentation.
2. Empfehlungen für akademische Studien
- 2.1. Erstellung geeigneter ökonomischer Modelle je nach betrachteter operativer Ebene. Dies betrifft Inputs und Outputs, allerdings auch die Orientierung und die unterstellten Skaleneffekte.
 - 2.2. Umfangreiche Recherche der Datenqualität und -erhebung. Insbesondere bei modellierten Indikatoren (Komplexität, CFH) sollte eine Überprüfung auf Anwendbarkeit stattfinden.
 - 2.3. Bei der Überführung in ein Benchmarking Modell sollten monetäre Faktoren vermieden werden, insbesondere solange die Punkte 1.1.-1.3. nicht erfüllt sind.
 - 2.4. Anwendung von Panelmodellen oder anderen multiperiodischen Analysen nur für Daten ab dem Jahr 2008.
 - 2.5. Die DEA ist anwendbar, sollte aber nicht mehr als vier Faktoren parallel betrachten. Sonderformen führen nicht zu einem Zusatznutzen, sondern verzerren teilweise die Ergebnisse. Allerdings kann die Supereffizienz-DEA Extremwerte von DMUs identifizieren.
 - 2.6. Die verwendete Methodik sollte vergleichsweise auch auf einen Datensatz ohne MUAC angewendet werden.
 - 2.7. Eine Leistungsbewertung auf ANSP Ebene kann nicht alle Effekte vollständig erklären und macht u. U. eine disaggregierte Analyse notwendig.
 - 2.8. Qualitative Analysen können helfen, Unterschiede in der Leistungsfähigkeit zu erklären.
 - 2.9. Untersuchung alternativer Luftraumstrukturierungen bei Voll-Implementierung von Single ATCO Sector Operations (siehe auch 3.1.).



- 2.10. Weiterentwicklung des Konzeptes der dynamischen Sektorisierung. Test der tatsächlich benötigten Sektoren und Vergleich mit den derzeit existierenden (je nach Verkehrsszenario). Ableitung technischer Voraussetzungen (z. B. Display Anzeigen für Lotsen) sowie Entwicklung eines Finanzierungskonzeptes.
 - 2.11. Entwicklung eines Level-of-Service Konzeptes zur Reduzierung der Gesamtkosten, bestehend aus Betriebskosten des ANSP und Verspätungskosten der Luftrumnutzer. Ableitung notwendiger Ressourcenpuffer.
3. Betriebliche Handlungsempfehlungen
- 3.1. Sicherstellung der Interoperabilität zwischen den ATS Systemen. Die Funktionalitäten dieser Systeme sollten zwischen den ANSPs abgestimmt und bei einer Erneuerung (Upgrade) angepasst werden.
 - 3.2. Test auf Möglichkeiten der Implementierung von Single ATCO Sector Operations in ausgewählten Verkehrssituationen (z. B. Nachts).
 - 3.3. Eruiierung einer „optimalen“ jährlichen Arbeitszeit pro Lotse. Untersuchung der signifikanten Differenzen in den Jahresarbeitszeiten sowie der beeinflussenden Effekte (z. B. Gewerkschaften, Puffer, etc.).
 - 3.4. Untersuchung und Definition (kosten)optimaler Ressourcenpuffer, um auf volatilen Verkehr sowie Fehlprognosen reagieren zu können (siehe auch Punkt 2.11.).
 - 3.5. Überprüfung der Lizenzierung von Lotsen für zusätzliche Sektoren. Dies erhöht die Flexibilität der Personaleinsatzplanung.
 - 3.6. Flexibilisierung der Arbeitsschichten (Start- und Endzeit) und ggf. Teamzusammensetzungen.
 - 3.7. Einhaltung des Flugplanes. Vermeidung von Änderungen in der Flugroute oder der Gewährung von Directs, da dies in Downstream-Sektoren zu Problemen führen kann und die Vorhersagbarkeit einschränkt.
 - 3.8. Überprüfung, ob ein internes Wetterbüro wesentliche Vorteile bringt.
 - 3.9. Sofern möglich, Senkung des Anteils an Nicht-ATCOs am Gesamtpersonal.
4. Handlungsempfehlungen für politische Institutionen und Entscheidungsträger
- 4.1. Überprüfung der Functional Airspace Blocks auf Allokation²⁰⁰ und Allokationsebene²⁰¹.
 - 4.2. Entwicklung und Implementierung alternativer Finanzierungskonzepte, welche Anreize zu Flugplanabweichungen bzw. Umwegen vermeiden.

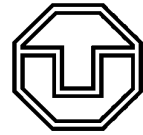
²⁰⁰ Wie viele ANSPs pro FAB und welche.

²⁰¹ ANSP oder ACC Ebene.



- 4.3. Überwachung der Prognosequalität und deren Auswirkungen auf Delay und Kosteneffizienz der ANSPs.
- 4.4. Schaffung rechtlicher Voraussetzungen multilateraler ANSPs zur operativen Konsolidierung hinsichtlich FABs oder der Ermöglichung einer landesgrenzenübergreifenden, dynamischen Sektorisierung bzw. sektorlosen Flugführung.
- 4.5. Überprüfung, ob und bei welchen ANSPs eine (Teil-)Privatisierung möglich ist.

Die Dissertation liefert durch die Untersuchungen und den daraus abgeleiteten Handlungsempfehlungen einen Beitrag, das zukünftige Luftverkehrsmanagement effizienter zu gestalten. Der holistische Ansatz umfasst das Benchmarking sowie die Ursachenforschung auf mehreren operativen Ebenen. Die Komplexität des Systems, dessen Verknüpfungen zu Politik und Umweltfaktoren sowie die Datenverfügbarkeit stellen eine Herausforderung dar. Dadurch war es nicht möglich, die Aspekte vollständig zu quantifizieren. Die operative Expertise von FABEC trug jedoch dazu bei, dass die wesentlichen Faktoren identifiziert und untersucht wurden. Infolgedessen konnten klare Tendenzen hinsichtlich des Einflusses auf die Leistungsfähigkeit der ANSPs abgeleitet werden. Der holistische Ansatz ermöglicht sowohl eine Weiterentwicklung durch die Ergänzung von (nichteuropäischen) ANSPs als auch die Inkludierung zusätzlicher oder alternativer Daten der jeweiligen operativen Ebene.



Anhang

A ATM

A1 Luftraumklassen

In Abbildung A 1 ist die Luftraumstruktur in Deutschland dargestellt. In einigen anderen Ländern sind zudem die Lufträume A und B existent. Die Buchstaben geben an, ob ein Luftraum kontrolliert (A, B, C, D und E) oder unkontrolliert (F und G) ist und welche spezifischen Charakteristika er aufweist (siehe Tabelle A 1). Eine weitere Einteilung ist die Differenzierung in oberen und unteren Luftraum. In weiten Teilen Europas ist die Grenze bei Flugfläche 245²⁰². Der Klassifizierung nach entspricht das der Luftraumklasse C (Charly) (DFS 2015, 2019; European Commission 2012a).

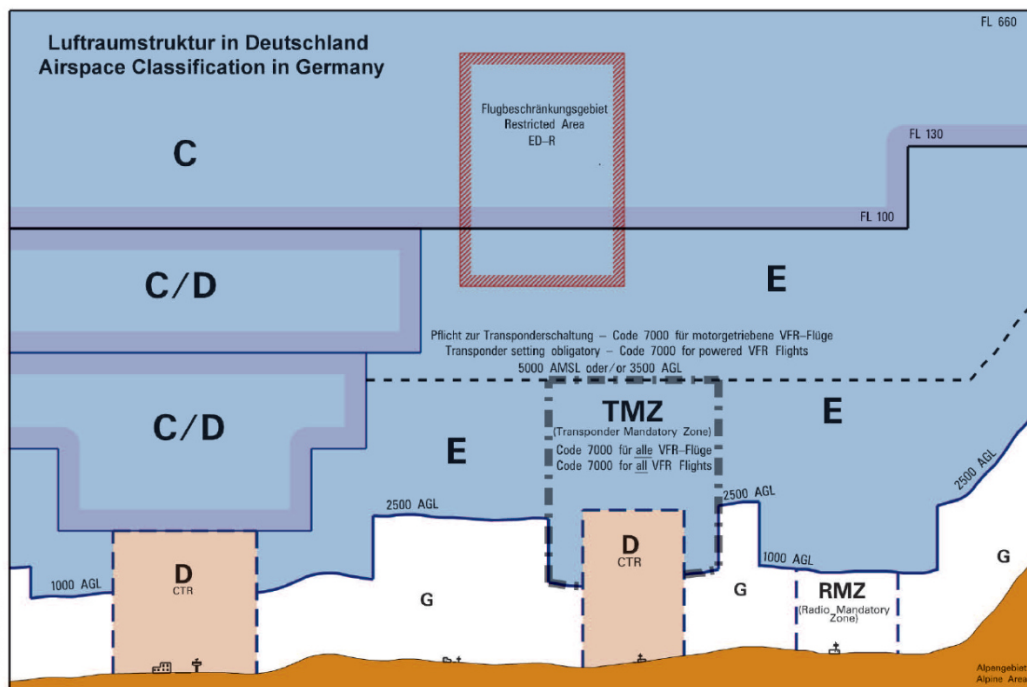


Abbildung A 1: Luftraumklassen gemäß DFS (2015)

Die Unterscheidung dient vor allem der verkehrlichen Charakterisierung. Im oberen Luftraum findet primär Streckenverkehr statt, während der untere Luftraum einen signifikanten Anteil an Vertikalverkehr²⁰³ aufweist, wobei die Geschwindigkeitsunterschiede zwischen den Luftfahrzeugen höher sind. Die Komplexität des unteren Luftraumes wird daher oft höher bewertet als die des oberen Luftraumes. Sie ist außerdem abhängig vom Design der Sektoren und vom Traffic Mix (Majumdar und Ochieng 2007). Zudem ist unter Flugfläche 100 eine Maximalgeschwindigkeit von 250kt erlaubt.

²⁰² Das entspricht 24.500 Fuß.

²⁰³ Steigender und sinkender Flugverkehr.

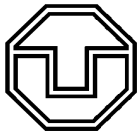


Tabelle A 1: Luftraumklassen und deren Charakteristiken

Luftraum ²⁰⁴	IFR	VFR	Freigabe	Staffelung	FIS
Alfa	Ja	Nein	Ja	Ja	Nein
Bravo	Ja	Ja	Ja	IFR ⇄ IFR IFR ⇄ VFR VFR ⇄ IFR VFR ⇄ VFR	Nein
Charly	Ja	Ja ²⁰⁵	Ja	IFR ⇄ IFR IFR ⇄ VFR	VFR ⇄ VFR
Delta	Ja	Ja	Ja	IFR ⇄ IFR	IFR ⇄ VFR VFR ⇄ IFR VFR ⇄ VFR
Echo	Ja	Ja	Nur für IFR	IFR ⇄ IFR	IFR ⇄ VFR VFR ⇄ VFR
Foxtrott	Max 1	Ja	Nein	Nein	IFR ⇄ VFR VFR ⇄ IFR VFR ⇄ VFR
Golf	Ja	Ja	Nein	Nein	IFR ⇄ IFR IFR ⇄ VFR VFR ⇄ IFR VFR ⇄ VFR

A2 EURCONTROL Raum

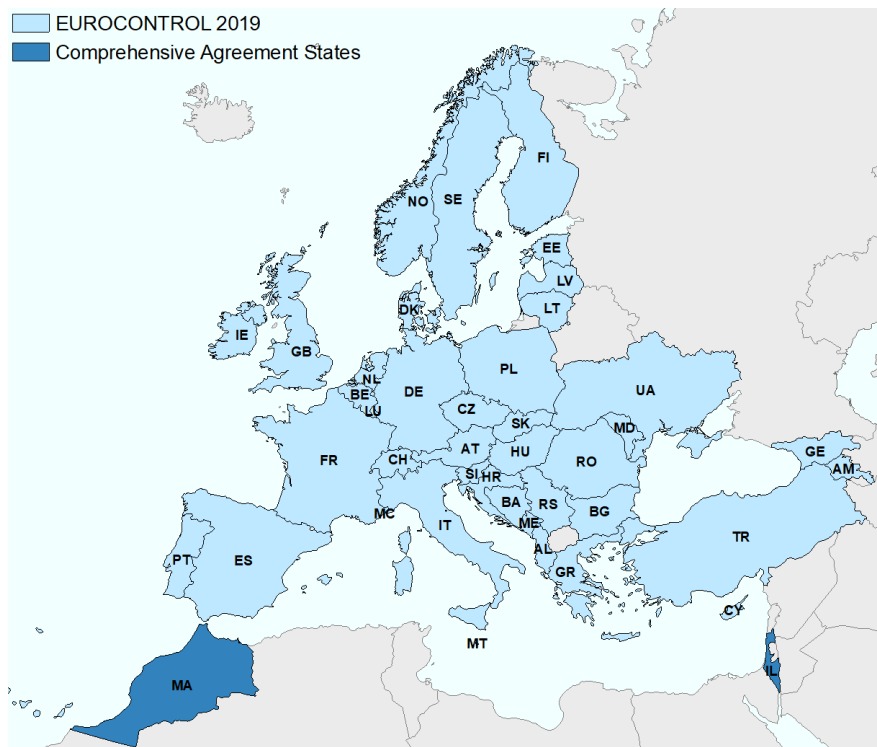
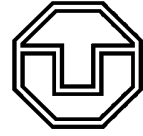


Abbildung A 2: EUROCONTROL Staaten
 (EUROCONTROL 2020d)

²⁰⁴ Die Lufträume Alpha, Bravo und Foxtrott sind in Deutschland nicht existent.

²⁰⁵ In Deutschland nur als CVFR.



A3 EUROCONTROL Gremien

Unter der Schirmherrschaft der EUROCONTROL begannen im Jahr 1998 die europäischen Bestrebungen einer Leistungsbewertung im Flugsicherungsbereich. Auf Basis der Strategie der Europäischen Zivilluftfahrtkonferenz (European Civil Aviation Conference, ECAC), eine regionale Unterorganisation der ICAO, sollte damit ein unabhängiges und transparentes Bewertungssystem geschaffen werden. Weitere Motivationen waren eine gegenseitige Rechenschaftspflicht, die Schaffung von Grundlagen für Investitionsanalysen sowie das Anheimstellen von Leitlinien als Grundlage für staatliche Regulierungen (EUROCONTROL 2016f).

Im Zuge dieser Bestrebungen wurde die Performance Review Commission gegründet. Sie ist eine ständige und von EUROCONTROL unabhängige Kommission, welche mit der Bereitstellung von Informationen hinsichtlich der ATM-Performance beauftragt ist. Diese publiziert sie innerhalb der jährlichen Performance Review Reports²⁰⁶ sowie der ATM-Cost-Effectiveness Reports²⁰⁷ (vgl. Abschnitt 3.2.1) (EUROCONTROL 2007, 2018c).

Die PRC wird in ihren Tätigkeiten von der Performance Review Unit unterstützt. Die Mitgliedsstaaten übermitteln hierfür Daten. Dabei gibt es hinsichtlich Umfang und Erfassungsmethodik Vorgaben und Richtlinien. Auf Basis von umfangreichen Datenanalysen erarbeitet die PRU Empfehlungen zur Verbesserung der Leistungsfähigkeit (EUROCONTROL 2016e).

PRU und PRC werden seit September 2010 als Performance Review Body des Single European Sky zusammengefasst. Neben dem Monitoring und Benchmarking setzt die PRB auch Zielvorgaben für die Verbesserung von Key Performance Indikatoren in den Bereichen Kosteneffizienz, Umwelt, Kapazität und Sicherheit.²⁰⁸ Zudem übernimmt sie die Beratungsfunktion für die Europäische Kommission hinsichtlich der Bewertung und Umsetzung der Funktionalen Luftraumblöcke (European Commission 2010c, 2013, 2014a).

A4 Performance Benchmarking Civil Air Navigation Services Organisation

Die Civil Air Navigation Services Organisation (CANSO) ist eine internationale Organisation, welche die Interessen der Flugsicherungsdienstleister vertritt. Sie kooperiert mit anderen Institutionen wie der ICAO oder der IATA. Die Tätigkeitsfelder umfassen u. a. Wirtschaftlichkeit und Effizienz. In diesem Rahmen werden die „Global ANS Performance Reports“ veröffentlicht.

Diese Studien weisen einen bedeutend geringeren Umfang auf als die Performance Reports der EUROCONTROL. Ein Vorteil dieser Studien ist jedoch die Anzahl der betrachteten Gebiete: Als weltweit agierende Organisation beschränkt sich CANSO nicht auf den europäischen Luftraum, sondern inte-

²⁰⁶ Kurz: PRR.

²⁰⁷ Kurz: ACE-Report.

²⁰⁸ Vgl. Commission Regulation (EU) No 691/2010 sowie Commission Implementing Regulation (EU) No 390/2013.



griert Flugsicherungsdienstleister aus allen Regionen. Allerdings variieren die involvierten Flugsicherungen zwischen den einzelnen Studien. Die nachfolgenden Angaben beziehen sich auf den Global ANS Performance Report 2015 (CANSO 2015).

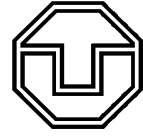
Die Daten für die Benchmarking-Aktivitäten werden direkt von den CANSO-Mitgliedstaaten bereitgestellt. Der Bericht beinhaltet die Leistungsbewertung von 28 Flugsicherungen sowie 5-12 KPIs²⁰⁹. Diese umfassen sowohl betriebliche als auch finanzielle Indikatoren, u. a.:

- Kosten pro IFR Flugstunde
- Kosten pro Lotsenstunde
- Lotsenproduktivität
- Umsatz pro Flugstunde
- Anteil Kapitalkosten an Gesamtkosten

Im Vergleich zum Jahr 2013 verzeichneten 79% der teilnehmenden Flugsicherungen zunehmende Verkehrszahlen. Den höchsten Anteil an IFR-Flugstunden wiesen die FAA (USA), Nav Canada und JANS (Japan) sowie AAI (Indien) auf. Die indische Flugsicherung sieht sich zudem den geringsten Kosten je IFR-Flugstunde gegenüber. Insgesamt konnten 56% der ANSPs ihre Kosten pro IFR-Flugstunde reduzieren.

Ähnlich den ACE-Reports veröffentlicht auch CANSO sogenannte Fact-Sheets der betrachteten Flugsicherungen. Die Angaben umfassen betriebliche Determinanten, die Anzahl an Einrichtungen sowie personelle Kennzahlen. Außerdem werden weiterführende Angaben, bspw. zu Investitionen oder zu den Besitzverhältnissen, bereitgestellt.

²⁰⁹ Die Reports werden in eine Zusammenfassung, einen ANSP-spezifischen sowie einen industriespezifischen Bericht unterteilt.



B Effizienzanalyse

B1 Charnes-Cooper Transformation

Das beschriebene Maximierungsproblem stellt ein lineares Quotientenprogramm dar. Gemäß der Zielfunktion wird dabei nur der Nenner minimiert, nur der Zähler maximiert, oder beides gleichzeitig optimiert. Vorteilhafter wäre es, wenn Zähler oder Nenner konstant gesetzt werden könnte. Eine Möglichkeit zur Lösung des Problems besteht in der Charnes-Cooper Transformation. Dabei werden Zähler und Nenner der Zielfunktion und Nebenbedingungen durch die aggregierten Inputs des Unternehmens i dividiert. Dadurch wird der Nenner auf 1 normiert. Diese Normierung wird als neue Nebenbedingung eingeführt. Dies führt dazu, dass die (modifizierten) Aggregationsgewichte ebenfalls in der Summe 1 ergeben. Zur mathematischen Herleitung in verschiedenen Schreibweisen siehe (Cantner et al. 2007).

B2 Kosten pro Lotsenstunde (on duty)

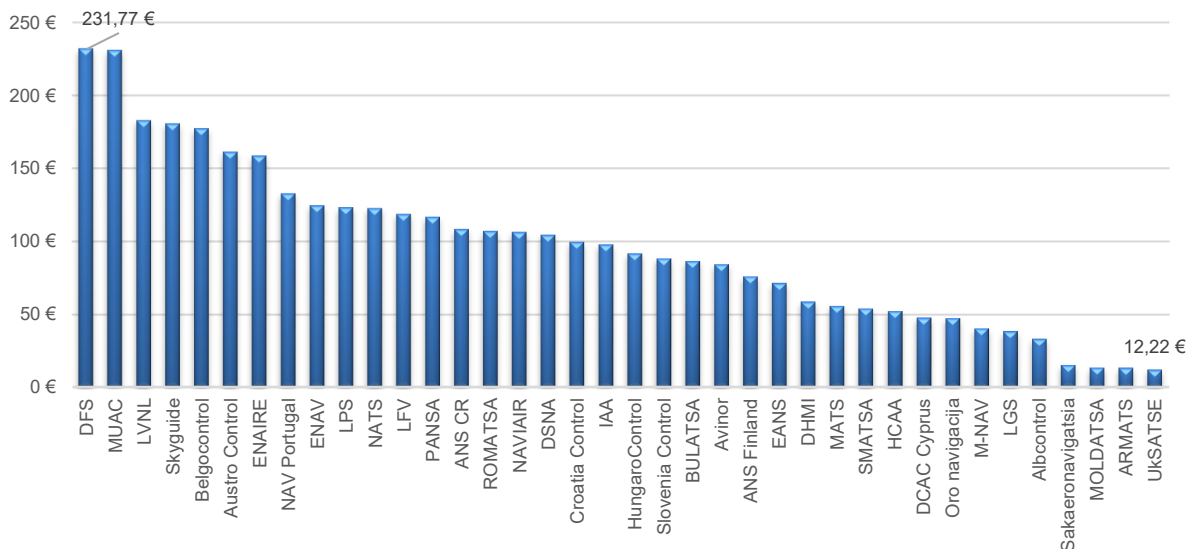


Abbildung A 3: Arbeitskosten pro Lotsenstunde in Europa (2017)
(EUROCONTROL 2020h)

B3 Ablauf-Schema Monte Carlo Simulation zur Sensitivitätsanalyse Gewichtung CFH

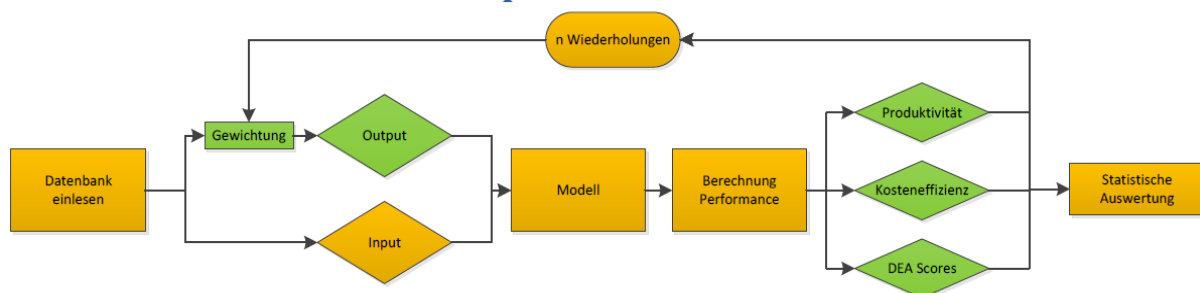
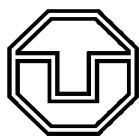


Abbildung A 4: Schema Monte Carlo Simulation zur Berechnung der Composite Flight Hours



B4 Ergebnisse der DEA

Tabelle A 2: DEA-Ergebnisse nach CRS, 2016, Modellvergleich

ANSP	Land	M1	M2	M2A	M2B
Albcontrol	Albanien	0,433	0,404	0,454	0,403
ANS CR	Tschechien	0,718	0,501	0,865	0,529
ARMATS	Armenien	0,626	0,373	0,373	0,476
Austro Control	Österreich	0,779	0,451	0,785	0,421
Avinor	Norwegen	1,000	0,609	0,734	0,879
BULATSA	Bulgarien	0,707	0,667	0,783	0,675
Croatia Control	Kroatien	0,679	0,638	0,722	0,638
DCAC Cyprus	Zypern	0,936	0,878	1,000	0,868
DFS	Deutschland	1,000	0,540	1,000	0,497
DHMI	Türkei	0,879	0,648	0,872	0,654
DSNA	Frankreich	0,598	0,418	0,656	0,427
EANS	Estland	1,000	1,000	1,000	0,976
ENAIRES	Spanien	0,995	0,733	0,964	0,746
ENAV	Italien	0,884	0,513	0,787	0,489
ANS Finland	Finnland	1,000	0,557	0,615	0,571
HCAA	Griechenland	0,759	0,640	0,756	0,607
HungaroControl	Ungarn	0,651	0,525	0,850	0,562
IAA	Irland	0,954	0,707	1,000	0,711
LFV	Schweden	0,973	0,743	0,843	0,677
LGS	Lettland	0,977	0,811	0,867	0,845
LPS	Slowakei	0,483	0,389	0,672	0,404
LVNL	Niederlande	1,000	0,371	0,647	0,287
MATS	Malta	0,971	0,948	0,984	1,000
M-NAV	Mazedonien	0,463	0,430	0,430	0,410
MoldATSA	Moldawien	0,891	0,461	0,461	0,615
MUAC	International	1,000	1,000		1,000
NATS	UK	1,000	0,560	1,000	0,545
NAV Portugal	Portugal	1,000	0,838	1,000	0,842
NAVIAIR	Dänemark	1,000	0,577	0,848	0,556
Oro navigacija	Litauen	0,575	0,458	0,505	0,455
PANSA	Polen	0,838	0,725	0,845	0,713
ROMATSA	Rumänien	0,518	0,461	0,617	0,489
Sakaeronavigatsia	Georgien	0,559	0,535	0,569	0,642
Belgocontrol	Belgien	0,918	0,335	0,579	0,279
Skyguide	Schweiz	0,931	0,497	0,859	0,563
Slovenia Control	Slowenien	0,425	0,395	0,443	0,384
SMATSA	Serbien	0,709	0,679	0,774	0,785
UkSATSE	Ukraine	0,523	0,411	0,422	0,431
Arithmetisches Mittel		0,799	0,590	0,745	0,607

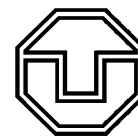


Tabelle A 3: DEA-Ergebnisse nach VRS, 2016, Modellvergleich

ANSP	Land	M1	M2	M2A	M2B
Albcontrol	Albanien	0,985	0,985	1,000	0,985
ANS CR	Tschechien	0,780	0,606	0,891	0,631
ARMATS	Armenien	1,000	1,000	1,000	1,000
Austro Control	Österreich	0,837	0,509	0,835	0,498
Avinor	Norwegen	1,000	0,621	0,768	0,983
BULATSA	Bulgarien	0,710	0,670	0,792	0,677
Croatia Control	Kroatien	0,695	0,644	0,740	0,643
DCAC Cyprus	Zypern	0,974	0,905	1,000	0,904
DFS	Deutschland	1,000	1,000	1,000	0,967
DHMI	Türkei	1,000	1,000	1,000	1,000
DSNA	Frankreich	1,000	1,000	1,000	1,000
EANS	Estland	1,000	1,000	1,000	1,000
ENAIRES	Spanien	1,000	1,000	1,000	1,000
ENAV	Italien	0,886	0,748	0,798	0,721
ANS Finland	Finnland	1,000	0,566	0,634	0,577
HCAA	Griechenland	0,794	0,647	0,799	0,611
HungaroControl	Ungarn	0,859	0,748	1,000	0,748
IAA	Irland	0,962	0,710	1,000	0,713
LFV	Schweden	0,995	0,759	0,927	0,686
LGS	Lettland	1,000	0,959	1,000	0,959
LPS	Slowakei	0,743	0,738	0,817	0,759
LVNL	Niederlande	1,000	0,495	0,732	0,623
MATS	Malta	0,994	0,982	1,000	1,000
M-NAV	Mazedonien	0,996	0,996	1,000	0,996
MoldATSA	Moldawien	1,000	1,000	1,000	1,000
MUAC	International	1,000	1,000		1,000
NATS	UK	1,000	1,000	1,000	1,000
NAV Portugal	Portugal	1,000	0,844	1,000	0,847
NAVIAIR	Dänemark	1,000	0,581	0,878	0,568
Oro navigacija	Litauen	0,729	0,665	0,667	0,934
PANSA	Polen	0,873	0,736	0,889	0,720
ROMATSA	Rumänien	0,530	0,464	0,619	0,493
Sakaeronavigatsia	Georgien	0,962	0,952	1,000	0,681
Belgocontrol	Belgien	0,979	0,550	0,685	0,714
Skyguide	Schweiz	0,961	0,529	0,863	0,584
Slovenia Control	Slowenien	0,938	0,938	1,000	0,938
SMATSA	Serbien	0,729	0,686	0,794	0,798
UkSATSE	Ukraine	0,527	0,423	0,474	0,442
Arithmetisches Mittel		0,906	0,780	0,881	0,800



Tabelle A 4: DEA Ergebnisse 2007-2016, heterogener Datensatz, VRS, Modell 2

ANSP	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Albcontrol	0,977	0,991	0,985	0,982	1,000	0,978	0,991	1,000	0,952	0,985
ANS CR	0,539	0,561	0,650	0,616	0,576	0,541	0,538	0,543	0,588	0,606
ARMATS			1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Austro Control	0,620	0,599	0,672	0,594	0,529	0,518	0,487	0,506	0,518	0,509
Avinor	0,532	0,554	0,690	0,681	0,619	0,638	0,527	0,711	0,482	0,621
BULATSA	0,523	0,575	0,613	0,613	0,572	0,592	0,578	0,681	0,594	0,670
Croatia Control	0,563	0,613	0,707	0,732	0,681	0,694	0,585	0,621	0,522	0,644
DCAC Cyprus	0,876	0,895	0,889	0,890	0,920	0,918	0,918	0,915	0,878	0,905
DFS	1,000	0,996	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
DHMI	0,818	0,912	1,000	1,000	0,978	1,000	1,000	1,000	1,000	1,000
DSNA	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
EANS	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
ENAIRE	1,000	1,000	0,943	1,000	1,000	0,999	1,000	1,000	1,000	1,000
ENAV	0,834	0,751	0,730	0,797	0,763	0,898	0,795	0,728	0,773	0,748
ANS Finland	0,686	0,717	0,729	0,713	0,677	0,624	0,518	0,541	0,528	0,566
HCAA	0,574	0,616	0,709	0,814	0,731	0,762	0,676	0,691	0,596	0,647
HungaroControl	0,788	0,789	0,806	0,797	0,756	0,741	0,749	0,756	0,731	0,748
IAA	0,709	0,707	0,824	0,712	0,662	0,705	0,695	0,718	0,653	0,710
LFV	0,909	0,913	0,910	0,716	0,817	0,652	0,656	0,702	0,745	0,759
LGS	0,945	0,948	0,960	0,957	0,957	0,960	0,960	0,962	0,962	0,959
LPS	0,616	0,619	0,623	0,631	0,680	0,645	0,673	0,768	0,723	0,738
LVNL	0,562	0,526	0,555	0,556	0,581	0,528	0,526	0,519	0,506	0,495
MATS	0,963	0,963	0,971	0,968	1,000	0,980	0,989	0,993	0,981	0,982
M-NAV	0,985	0,976	0,990	0,995	0,995	1,000	0,998	1,000	1,000	0,996
MoldATSA	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
MUAC	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
NATS	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
NAV Portugal	0,537	0,544	0,587	0,719	0,710	0,777	0,787	0,830	0,747	0,844
NAVIAIR	0,634	0,550	0,630	0,631	0,624	0,614	0,620	0,662	0,593	0,581
Oro navigacija	0,653	0,655	0,654	0,656	0,676	0,656	0,655	0,657	0,655	0,665
PANSA	0,878	0,588	0,798	0,793	0,752	0,774	0,798	0,598	0,628	0,736
ROMATSA	0,491	0,484	0,516	0,536	0,498	0,518	0,452	0,520	0,485	0,464
Sakaeronavigatsia									0,939	0,952
Belgocontrol	0,526	0,529	0,527	0,536	0,536	0,532	0,538	0,536	0,540	0,550
Skyguide	0,634	0,625	0,602	0,600	0,573	0,548	0,527	0,538	0,527	0,529
Slovenia Control	0,947	0,940	0,952	0,946	0,940	0,944	0,940	0,940	0,937	0,938
SMATSA	0,644	0,642	0,810	0,833	0,683	0,679	0,706	0,672	0,634	0,686
UkSATSE	0,463	0,420	0,496	0,574	0,483	0,419	0,379	0,374	0,357	0,423
Arithm. Mittel	0,762	0,756	0,798	0,800	0,783	0,779	0,764	0,775	0,757	0,780

(Aus Gründen der Übersichtlichkeit werden nur die Jahre 2007 bis 2016 dargestellt)

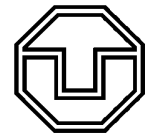
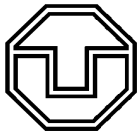


Tabelle A 5: DEA Ergebnisse 2007-2016, homogener Datensatz, VRS, Modell 2

ANSP	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Albcontrol	0,977	0,991	0,985	0,982	1,000	0,978	0,991	1,000	0,952	0,985
ANS CR	0,539	0,561	0,650	0,616	0,576	0,541	0,538	0,543	0,588	0,606
Austro Control	0,620	0,599	0,672	0,594	0,529	0,518	0,487	0,506	0,518	0,509
Avinor	0,532	0,554	0,690	0,681	0,619	0,638	0,527	0,711	0,482	0,621
Belgocontrol	0,526	0,529	0,527	0,536	0,536	0,532	0,538	0,536	0,540	0,550
BULATSA	0,523	0,575	0,613	0,613	0,573	0,592	0,579	0,681	0,594	0,670
Croatia Control	0,563	0,613	0,707	0,732	0,681	0,694	0,585	0,621	0,522	0,644
DCAC Cyprus	0,876	0,895	0,892	0,895	0,921	0,919	0,919	0,915	0,889	0,909
DFS	1,000	0,996	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
DHMI	0,818	0,912	1,000	1,000	0,978	1,000	1,000	1,000	1,000	1,000
DSNA	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
EANS	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
ENAIRE	1,000	1,000	0,943	1,000	1,000	0,999	1,000	1,000	1,000	1,000
ENAV	0,834	0,751	0,730	0,797	0,763	0,898	0,795	0,728	0,773	0,748
ANS Finland	0,686	0,717	0,729	0,713	0,677	0,624	0,518	0,541	0,528	0,566
HCAA	0,574	0,616	0,709	0,814	0,731	0,762	0,676	0,691	0,596	0,647
HungaroControl	0,788	0,789	0,808	0,802	0,759	0,743	0,751	0,763	0,731	0,748
IAA	0,709	0,707	0,824	0,712	0,662	0,705	0,695	0,718	0,653	0,710
LFV	0,909	0,913	0,910	0,716	0,817	0,652	0,656	0,702	0,745	0,759
LGS	0,945	0,948	0,962	0,962	0,959	0,961	0,960	0,968	0,964	0,966
LPS	0,616	0,619	0,623	0,631	0,680	0,645	0,673	0,768	0,723	0,738
LVNL	0,562	0,526	0,555	0,556	0,581	0,528	0,526	0,519	0,506	0,495
MATS	0,963	0,963	0,972	0,970	1,000	0,983	0,989	0,994	0,984	0,984
M-NAV	0,985	0,976	0,992	1,000	1,000	1,000	0,999	1,000	1,000	1,000
MoldATSA	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
MUAC	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
NATS	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
NAV Portugal	0,537	0,544	0,587	0,719	0,710	0,777	0,787	0,830	0,747	0,844
NAVIAIR	0,634	0,550	0,630	0,631	0,624	0,614	0,620	0,662	0,593	0,581
Oro navigacija	0,653	0,655	0,655	0,657	0,676	0,656	0,655	0,657	0,655	0,665
PANSA	0,878	0,588	0,798	0,793	0,752	0,774	0,798	0,598	0,628	0,736
ROMATSA	0,491	0,484	0,516	0,536	0,498	0,518	0,452	0,520	0,485	0,464
Skyguide	0,634	0,625	0,602	0,600	0,573	0,548	0,527	0,538	0,527	0,529
Slovenia Control	0,947	0,940	0,954	0,951	0,940	0,946	0,944	0,947	0,942	0,939
SMATSA	0,644	0,642	0,810	0,833	0,683	0,679	0,706	0,672	0,634	0,686
UkSATSE	0,463	0,420	0,496	0,574	0,483	0,419	0,379	0,374	0,357	0,423
Arithm. Mittel	0,762	0,756	0,793	0,795	0,777	0,773	0,758	0,770	0,746	0,770

(Aus Gründen der Übersichtlichkeit werden nur die Jahre 2007 bis 2016 dargestellt)



C Ursachenanalyse

C1 PCA

Abbildung A 5 zeigt das Ergebnis einer Regression auf Basis der Hauptkomponenten. Die Modellgüte ist mit einem R^2 von 0,03 sehr gering. Hierbei spielte es keine Rolle, welches Regressionsmodell angewendet wurde. Signifikant ist die Komplexitätskomponente und jene für Luftraumgröße.

```
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.66168   0.11483   5.762 5.26e-06 ***
PC1           0.06448   0.03929   1.641  0.1133
PC2           0.07410   0.03608   2.054  0.0506 .
PC3          -0.11180   0.05908  -1.892  0.0701 .
PC4           0.01175   0.13332   0.088  0.9305
PC5          -0.05165   0.05646  -0.915  0.3690
PC6           0.01541   0.05550   0.278  0.7835
PC7          -0.03420   0.10844  -0.315  0.7551
PC8           0.03988   0.05623   0.709  0.4848
PC9          -0.09307   0.09698  -0.960  0.3464
PC10         -0.14927   0.16178  -0.923  0.3650
PC11          0.06605   0.06523   1.013  0.3210
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1918 on 25 degrees of freedom
Multiple R-squared:  0.3295,    Adjusted R-squared:  0.03446
F-statistic: 1.117 on 11 and 25 DF,  p-value: 0.3893
```

Abbildung A 5: Regressionsanalyse mit Hauptkomponenten

C2 Logarithmische Transformation der DEA Werte

Bsp. Modell 2B: Histogramm der DEA-Scores (links) und deren logarithmische Transformation (rechts).

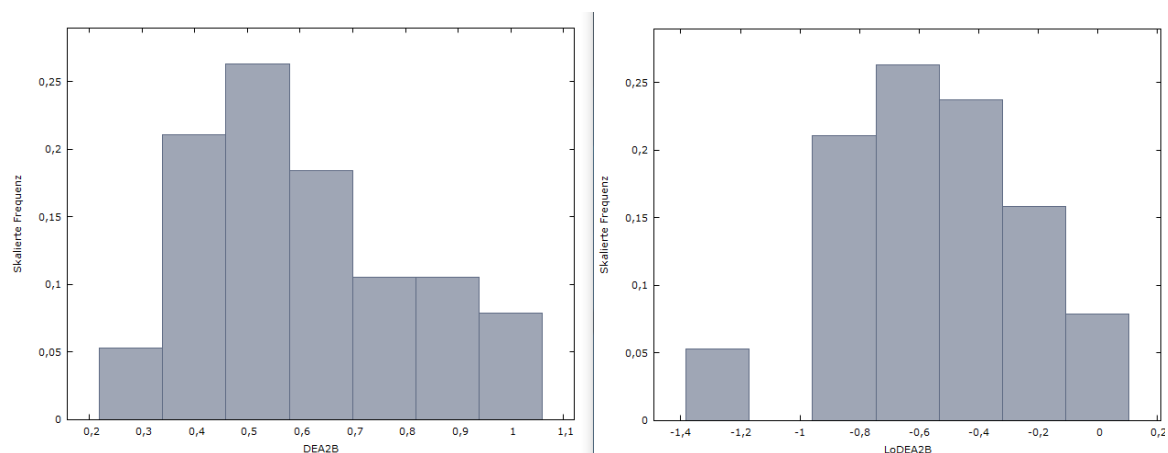
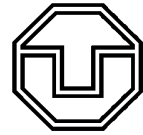


Abbildung A 6: Logarithmische Transformation DEA Werte



C3 Korrelationsmatrix der Faktoren

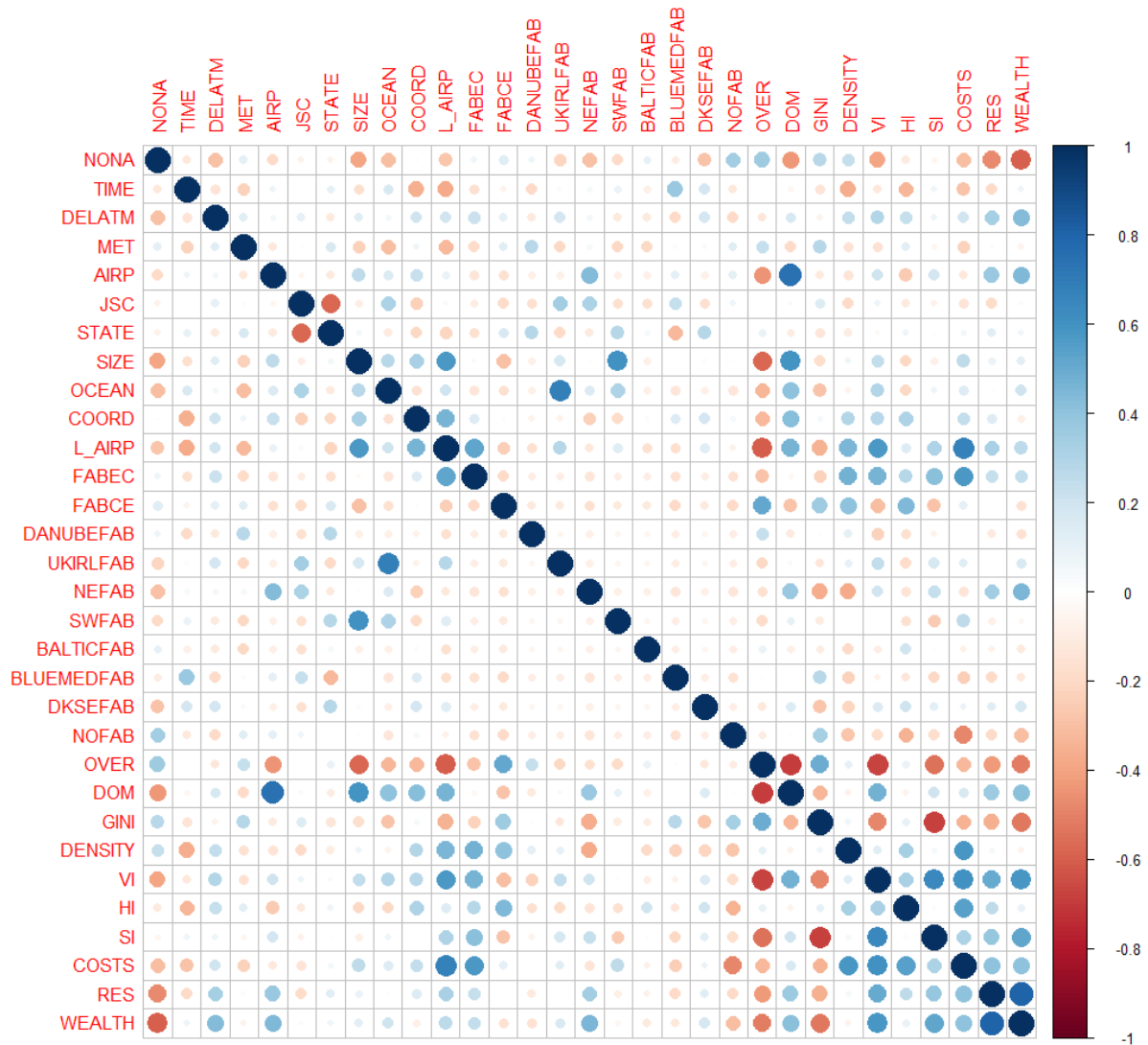


Abbildung A 7: Korrelationsmatrix der quantitativen Einflussfaktoren, ANSP Ebene



C4 Ergebnis Regression

Tabelle A 6: Vergleich Ergebnisse Trunkated Regression

	M1	M2	M2A	M2B
INT	1,494 (1,737)	-0,647 (1,729)	-1,361 (1,553)	4,028 (1,755)**
NONA	-1,121 (0,391)***	-0,41 (0,266)	-0,949 (0,388)**	-0,506 (0,248)**
TIME (I)	0,463 (0,44)	0,281 (0,356)	0,725 (0,35)**	-0,731 (0,369)**
DELATM	-0,006 (0,078)	-0,006 (0,054)	-0,083 (0,053)	0,045 (0,053)
MET	0,096 (0,049)**	0,004 (0,042)	0,037 (0,038)	0,016 (0,041)
AIRP	0,384 (0,193)**	0,182 (0,121)	0,435 (0,125)***	0,165 (0,118)
JSC	0,034 (0,075)	0,017 (0,057)	0,047 (0,065)	-0,072 (0,06)
STATE	-0,108 (0,068)	-0,102 (0,053)*	-0,115 (0,065)*	-0,136 (0,052)***
SIZE (I)	-0,094 (0,136)	0,168 (0,103)	0,145 (0,102)	-0,019 (0,102)
OCEAN	-0,068 (0,223)	-0,023 (0,075)	-0,196 (0,225)	0,037 (0,074)
COORD	-0,016 (0,02)	0,002 (0,016)	0,001 (0,016)	-0,008 (0,015)
L_AIRP	-0,101 (0,094)	-0,036 (0,051)	-0,082 (0,066)	-0,08 (0,048)*
NOFAB	0,061 (0,093)	0,014 (0,085)	0,1 (0,084)	-0,009 (0,083)
OVER	-0,836 (0,384)**	0,186 (0,189)	0,112 (0,184)	0,126 (0,187)
DOM	-0,864 (1,027)	-0,726 (0,456)	-1,16 (0,785)	0,128 (0,47)
GINI	-3,728 (1,004)***	-1,888 (0,729)***	-3,878 (0,927)***	-2,715 (0,692)***
DENSITY	0,021 (0,017)	-0,006 (0,01)	0,019 (0,011)*	-0,007 (0,01)
VI	0,093 (0,729)	0,062 (0,437)	0,006 (0,583)	0,424 (0,424)
HI	0,384 (0,438)	0,012 (0,357)	0,286 (0,365)	-0,297 (0,315)
SI	-1,015 (0,789)	-0,553 (0,516)	-1,137 (0,693)	-1,245 (0,524)**
COSTS	0,001 (0,001)	0 (0,001)	0,002 (0,001)*	0 (0,001)
RES	-0,01 (0,003)***	-0,002 (0,002)	-0,007 (0,002)***	-0,004 (0,001)***
Log-Lik.	37,18	35,85	37,39	37,11
N	38	38	37	38

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%** , 99%***

C5 Panel-Analyse

Prinzipiell sind ANSPs mit einer konstanten Änderungsrate der TFP von denen zu unterscheiden, die entweder stark fluktuierende Raten oder einen oder mehrere Peaks aufweisen. Dazu wird zunächst die Differenz (Spanne) zwischen maximaler und minimaler Änderungsrate untersucht. Ist zusätzlich die Standardabweichung hoch, kann auf eine ausgeprägte Fluktuation in den Wachstumsraten geschlossen werden. In Tabelle A 7 sind einige ANSPs inklusive der deskriptiven Statistik für Modell 2A dargestellt, sortiert nach der Spanne.

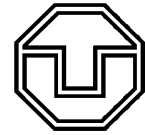


Tabelle A 7: Deskriptive Statistik Panel-Daten ANSP Ebene, Modell 2A

ANSP	Land	Min	GM	Max	Stabw	Span
Albcontrol	Albanien	0,700	0,995	1,314	0,153	0,614
Avinor	Norwegen	0,707	0,976	1,268	0,161	0,561
HCAA	Griechenland	0,620	0,963	1,158	0,146	0,538
LVNL	Niederlande	0,789	0,975	1,187	0,090	0,398
ENAIRE	Spanien	0,852	1,011	1,223	0,098	0,371
NATS	UK	0,920	0,999	1,229	0,080	0,309
UkSATSE	Ukraine	0,899	1,003	1,178	0,094	0,279
DFS	Deutschland	0,867	1,003	1,133	0,068	0,266
MATS	Malta	0,956	1,064	1,216	0,094	0,260
ENAV	Italien	0,901	1,015	1,095	0,061	0,194
ANS Finland	Finnland	0,898	0,988	1,084	0,063	0,186
MUAC	Maastricht	0,915	1,024	1,098	0,053	0,183
BULATSA	Bulgarien	0,995	1,042	1,155	0,051	0,160
Belgocontrol	Belgien	0,961	1,012	1,089	0,035	0,128
DSNA	Frankreich	0,915	0,987	1,036	0,035	0,121
Skyguide	Schweiz	0,941	0,994	1,042	0,029	0,101

Für Enroute Services können vier Modelle aufgestellt werden, wobei je einmal ein Input und ein Output substituiert werden. Aufgrund des zur Verfügung stehenden Output Parameters für Terminal Services werden lediglich zwei Modelle gegenübergestellt, welche sich in einem Inputfaktor unterscheiden (Lotsenstunden versus Anzahl der Lotsen).

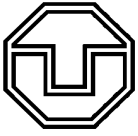
Tabelle A 8: Panel-Analyse, Enroute-Modelle

	3A	3B	4A	4B
Inputs	ACC ATCO-Hours Number of ACCs		Number of ACC ATCOs Number of ACCs	
Outputs	IFR Flight Hours	IFR Flights	IFR Flight Hours	IFR Flights

Tabelle A 9: Panel-Analyse, Terminal-Modelle

	5A	5B
Inputs	Tower ATCO-Hours Number of Towers	Tower ATCOs Number of Towers
Outputs	IFR Airport Movements	

Bei der Betrachtung der Malmquist-Indizes für die Enroute Services (Abbildung A 8) weist die jährliche Verlaufskurve der TFP-Änderungsrate (rot) starke Ähnlichkeiten zu jener der Gate-to-Gate Perspektive auf. In den meisten Jahren ist ein Wirtschaftswachstum zu verzeichnen. Im Jahr 2009 wird einen Verlust von über 5 % ausgewiesen, was den Wert der Gate-to-Gate Betrachtung übersteigt. Die Begründung kann ebenfalls in der beginnenden Eurokrise gesehen werden. Parallel zur aggregierten Betrachtung ist auch hier das Jahr 2012 durch eine Rezession gekennzeichnet. Zudem ist der Effekt der Komponenten (technische Effizienz und technischer Fortschritt) äquivalent. Dagegen ist ein wesentlicher Unterschied im TFP des Jahres 2006 zu sehen, welches hier ein Wachstum von ca. 4 % ausweist



(-1 % bei Gate-to-Gate). Das legt die Vermutung nahe, dass diese Rezession durch die Terminal Services bedingt ist.

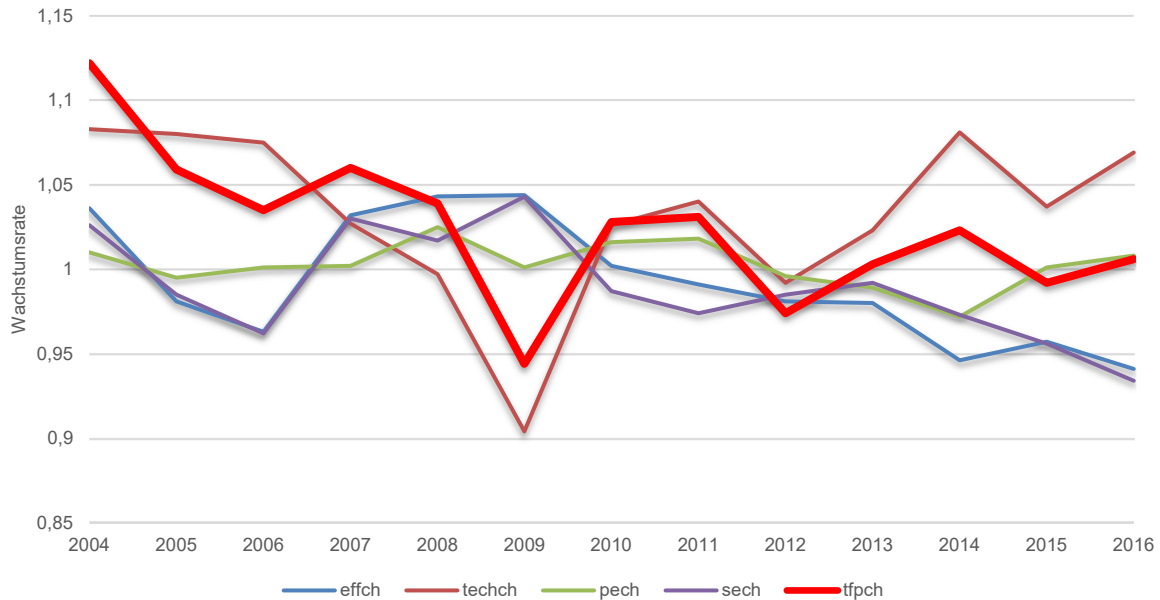


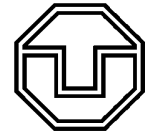
Abbildung A 8: Entwicklung Malmquist Indizes, Enroute Services

Die modellbedingte Abweichung der TFP-Kurvenverläufe ist sehr gering. Es gibt jedoch signifikante Abweichungen zwischen den Kurven der Komponenten. Insbesondere der technische Wandel schwankt in einer jährlichen Betrachtung zwischen den Modellen. Die Unterschiede werden hauptsächlich durch den im Modell gewählten Output bestimmt, da die Ergebnisse für Modell 3A den Werten 4A ähneln, analog 3B und 4B. Dies wird durch den europaweiten Durchschnitt der Indizes unter Berücksichtigung der gesamten Zeitspanne bestätigt (Tabelle 4-18). Es zeigt sich, dass die TFP-Änderung nicht signifikant zwischen den Modellen schwankt, sondern ein durchschnittliches jährliches Wachstum von 2,3 bis 3,0 % ausweist. Betrachtet man Flüge als Output, so ist eine geringe positive Änderung der technischen Effizienz sichtbar (1,2 % bzw. 1,5 %). Sie stagniert (Modell 1A) oder nimmt sogar ab (Modell 2A), wenn man stattdessen die kontrollierten Flugstunden betrachtet.

Tabelle A 10: Panel-Analyse, Ergebnisse für Enroute

	effch	techch	pech	sech	tfpch
M3A	0,999	1,027	1,004	0,994	1,026
M3B	1,012	1,017	1,005	1,008	1,030
M4A	0,991	1,032	1,003	0,989	1,023
M4B	1,015	1,012	1,006	1,010	1,027

Abbildung A 9 zeigt die europäischen ANSPs und ihre Produktivitätsänderung zwischen 2003 und 2016 beispielhaft für Modell 3A. Die „Big Player“, wie die deutsche DFS, die spanische ENAIRE oder die französische DSNA, erzielen ein geringes Wachstum von 0,1 % bis 0,3 %. Das größte Wachstum wurde durch den bulgarischen ANSP erzielt. Dieser Effekt ist direkt mit dem größten TFP-Verlust



gekoppelt, welcher durch die ukrainische UkSATSE repräsentiert wird. Insbesondere in Bulgarien stiegen u. a. durch die Ukraine-Krise die kontrollierten Flugstunden im betrachteten Zeitraum um mehr als das Doppelte. Darüber hinaus erzielt Rumänien ein jährliches Wachstum von 4,5 %. Die steigende Produktivität könnte hier darauf hinweisen, dass im Falle von Rumänien und Bulgarien Größenvorteile realisiert werden konnten (vgl. Abschnitt 4.5) und/oder eine höhere Auslastung erreicht wurde.

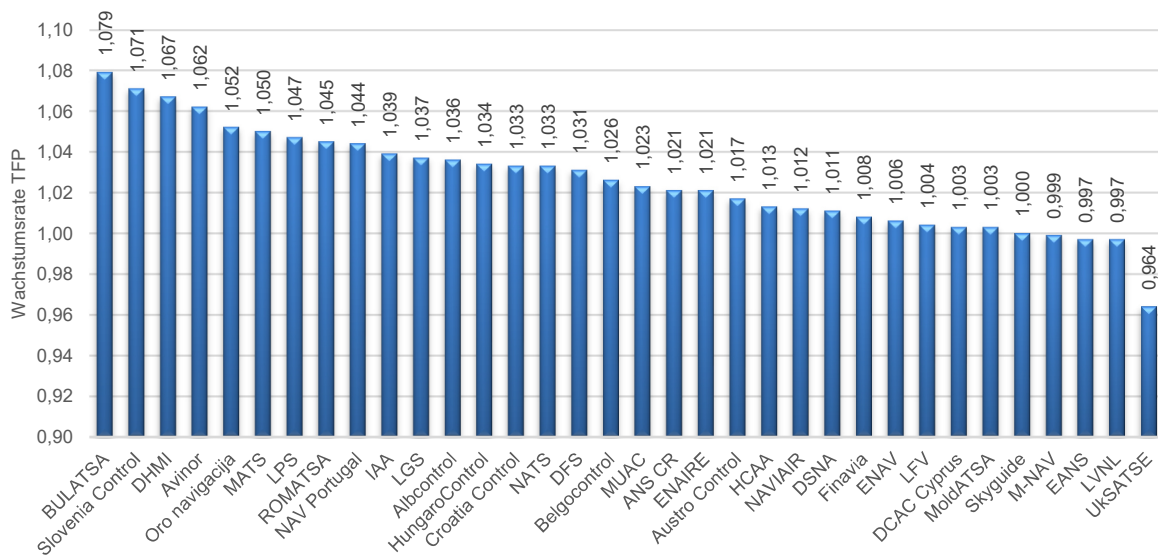


Abbildung A 9: Malmquist Indizes europäischer Flugsicherungen, Enroute, Modell 3A

Allgemein weisen südosteuropäische Flugsicherungsdienste ein relativ hohes jährliches Wachstum auf, das jenes der Gate-to-Gate Betrachtung übersteigt. Eine mögliche Erklärung ist, dass diese Flugsicherungsdienste auch primär Enroute Services anbieten (90 % in Bulgarien und 75 % in Rumänien). Zum Vergleich stellen Abbildung A 10 und Abbildung A 11 die Ergebnisse für die Terminal Services gemäß Modell 5B dar.

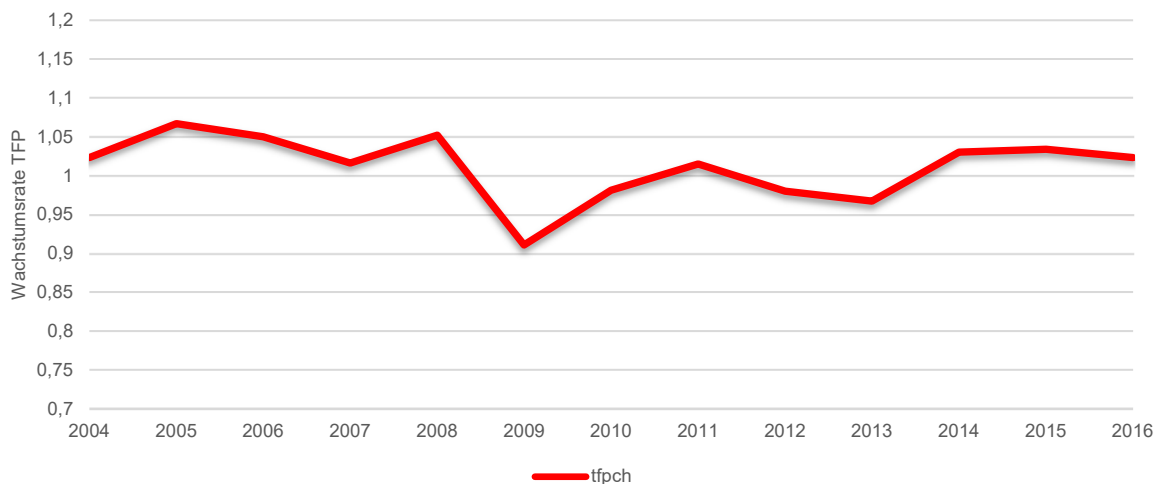


Abbildung A 10: Entwicklung Malmquist Indizes, Terminal Services, Modell 5B

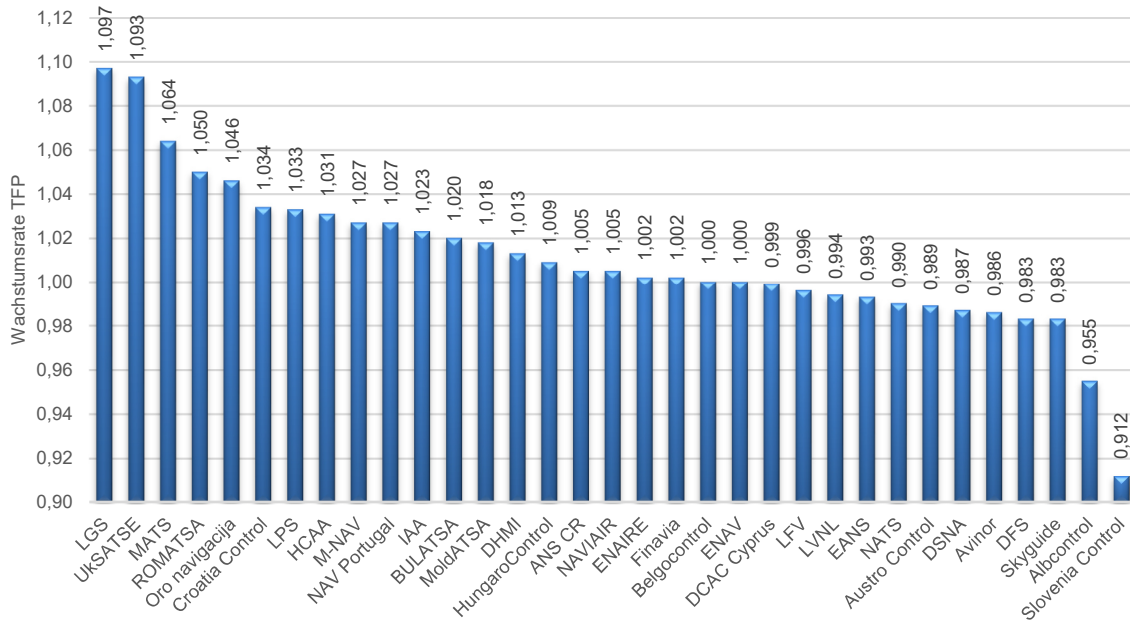
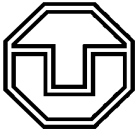
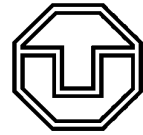


Abbildung A 11: Malmquist Indizes europäischer Flugsicherungen, Terminal, Modell 5B



C6 Panel-Regression

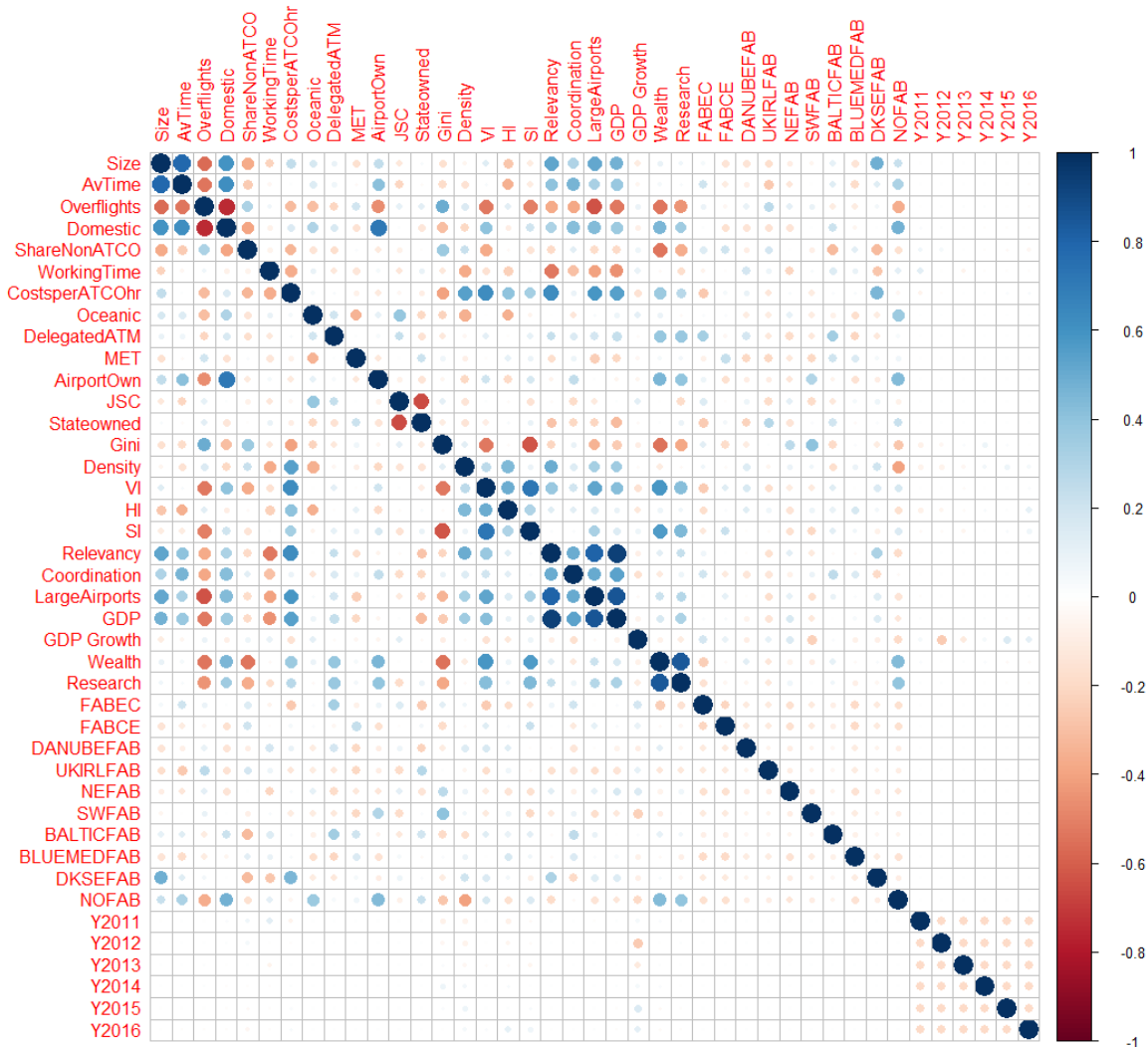


Abbildung A 12: Korrelationsmatrix Einflussfaktoren ANSP, Panel-Analyse

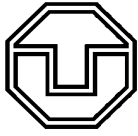
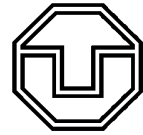


Tabelle A 11: Regression basierend auf Panel-Daten (Modell 1)

	Fixed Effects					Random Effects				
	I1	I2	I3	I4	I5	I1	I2	I3	I4	I5
INT	0,786 (0,287)***	0,791 (0,294)***	0,758 (0,42)*	0,875 (0,587)	1,389 (0,698)**	0,726 (0,269)***	0,723 (0,275)***	0,727 (0,396)*	0,689 (0,543)	1,3 (0,626)**
NONA	0,002 (0,055)	0,002 (0,056)	0,021 (0,08)	0,118 (0,105)	0,02 (0,133)	-0,11 (0,052)	-0,01 (0,053)	0,016 (0,073)	0,093 (0,092)	-0,023 (0,117)
TIME (I)	0,031 (0,039)	0,031 (0,039)	0,014 (0,049)	-0,001 (0,065)	-0,023 (0,073)	0,04 (0,036)	0,041 (0,037)	0,02 (0,047)	0,016 (0,06)	-0,008 (0,066)
DELATM	0,016 (0,016)	0,016 (0,016)	0,01 (0,017)	0,014 (0,018)	0,023 (0,022)	0,009 (0,015)	0,009 (0,015)	0,006 (0,016)	0,011 (0,016)	0,023 (0,02)
MET	0,001 (0,013)	0,001 (0,014)	0,002 (0,017)	0,009 (0,018)	0,013 (0,02)	0,006 (0,012)	0,006 (0,013)	0,009 (0,015)	0,012 (0,016)	0,017 (0,017)
AIRP	0,001 (0,019)	0 (0,019)	-0,015 (0,023)	-0,008 (0,051)	0,009 (0,059)	-0,003 (0,018)	-0,004 (0,019)	-0,016 (0,023)	-0,01 (0,046)	0,006 (0,054)
JSC	-0,008 (0,018)	-0,008 (0,019)	-0,014 (0,025)	-0,027 (0,036)	-0,041 (0,038)	0 (0,017)	0,001 (0,017)	-0,014 (0,023)	-0,027 (0,033)	-0,05 (0,034)
STATE	0,01 (0,018)	0,01 (0,018)	0,008 (0,022)	0 (0,029)	-0,003 (0,029)	0,007 (0,017)	0,007 (0,017)	0 (0,021)	-0,006 (0,026)	-0,016 (0,026)
Y2012		0,004 (0,02)	0,003 (0,023)	0,01 (0,024)	0,003 (0,027)		0,006 (0,02)	0,008 (0,022)	0,011 (0,023)	0,003 (0,025)
Y2013		0,002 (0,019)	0,002 (0,019)	0,002 (0,019)	0,003 (0,02)		0,002 (0,019)	0,002 (0,019)	0,002 (0,019)	0,004 (0,019)
Y2014		-0,009 (0,02)	-0,009 (0,022)	0 (0,024)	-0,007 (0,027)		-0,007 (0,02)	-0,004 (0,022)	0,001 (0,023)	-0,01 (0,025)
Y2015		-0,014 (0,019)	-0,014 (0,019)	-0,012 (0,019)	-0,01 (0,02)		-0,015 (0,019)	-0,014 (0,019)	-0,012 (0,019)	-0,008 (0,019)
Y2016		-0,007 (0,02)	-0,008 (0,022)	0 (0,024)	-0,009 (0,027)		-0,005 (0,02)	-0,003 (0,022)	0,001 (0,023)	-0,012 (0,025)
SIZE (I)			0,011 (0,008)	0,013 (0,012)	0,004 (0,014)			0,011 (0,008)	0,016 (0,011)	0,006 (0,013)
OCEAN			0,03 (0,026)	0,031 (0,029)	0,024 (0,032)			0,037 (0,024)	0,041 (0,025)	0,026 (0,028)
COORD			0,003 (0,005)	-0,002 (0,006)	-0,002 (0,006)			0,001 (0,004)	-0,004 (0,005)	-0,003 (0,005)
L_AIRP			-0,015 (0,012)	-0,024 (0,023)	-0,024 (0,025)			-0,015 (0,011)	-0,02 (0,021)	-0,023 (0,022)
NOFAB			-0,017 (0,026)	-0,044 (0,033)	-0,027 (0,042)			-0,021 (0,025)	-0,041 (0,03)	-0,017 (0,038)
OVER				-0,071 (0,058)	-0,107 (0,064)*				-0,062 (0,055)	-0,1 (0,058)*
DOM				-0,094 (0,202)	-0,112 (0,224)				-0,087 (0,179)	-0,097 (0,196)
GINI				-0,032 (0,223)	-0,176 (0,27)				0,101 (0,207)	-0,019 (0,252)
DENSITY				-0,004 (0,004)	-0,003 (0,004)				-0,004 (0,003)	-0,004 (0,004)
VI				0,237 (0,182)	0,285 (0,188)				0,151 (0,163)	0,242 (0,168)
HI				0,073 (0,099)	0,092 (0,107)				0,114 (0,089)	0,127 (0,093)
SI				-0,264 (0,161)	-0,359 (0,175)**				-0,155 (0,148)	-0,278 (0,16)*
COSTS					0 (0)					0 (0)
GDP_G					0,003 (0,003)					0,006 (0,002)**
WEALTH (I)					-0,014 (0,021)					-0,02 (0,018)
RES					0 (0,001)					0 (0,001)
R ² within	0,02	0,03	0,08	0,08	0,08	-434,61	-426,01	-424,25	-416,46	-419,79
Akaike	-418,56	-410,27	-409,27	-403,87	-399,15	204	204	204	204	204
N	204	204	204	204	204	204	204	204	204	204

Standardfehler in Klammern. Signifikanzlevel 90 %*, 95 %**, 99 %***



C7 Regressions- und Clusteranalysen auf ACC Ebene

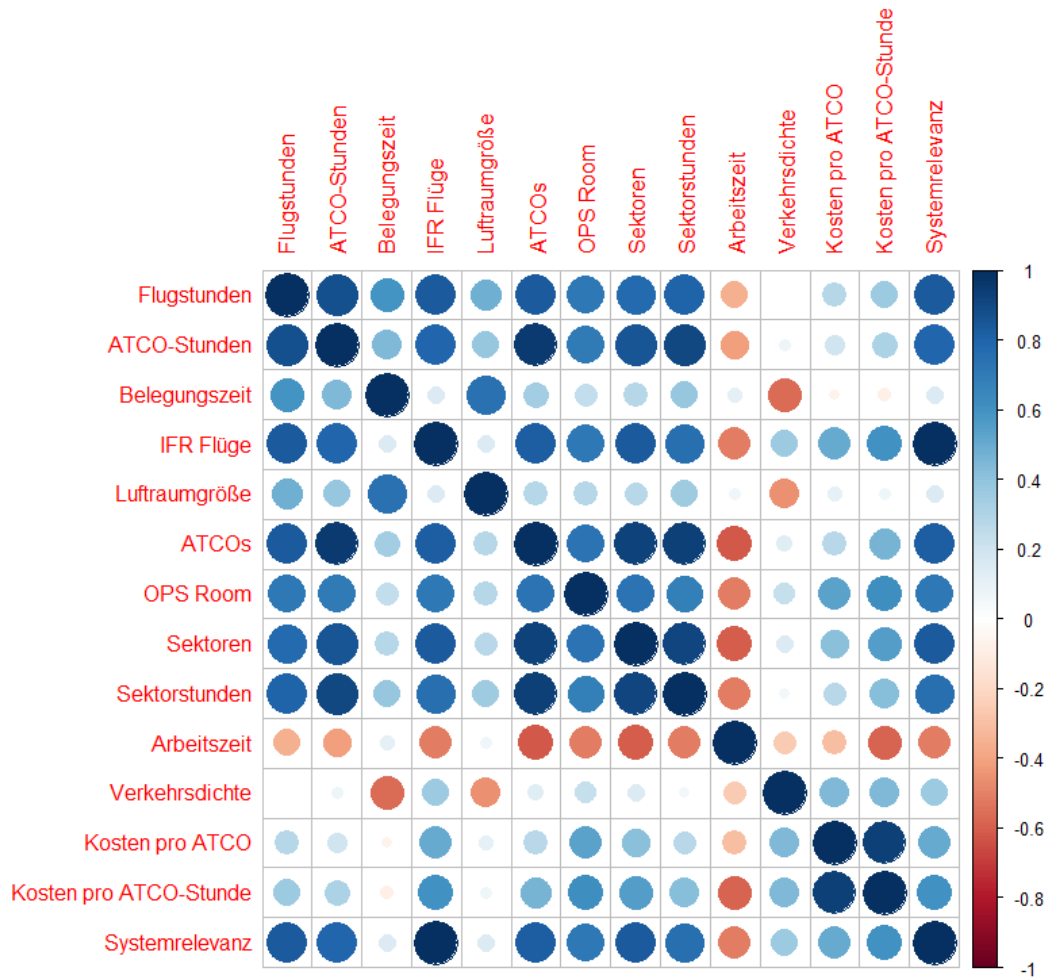


Abbildung A 13: Korrelationsmatrix potentieller Einflussfaktoren auf ACC Ebene

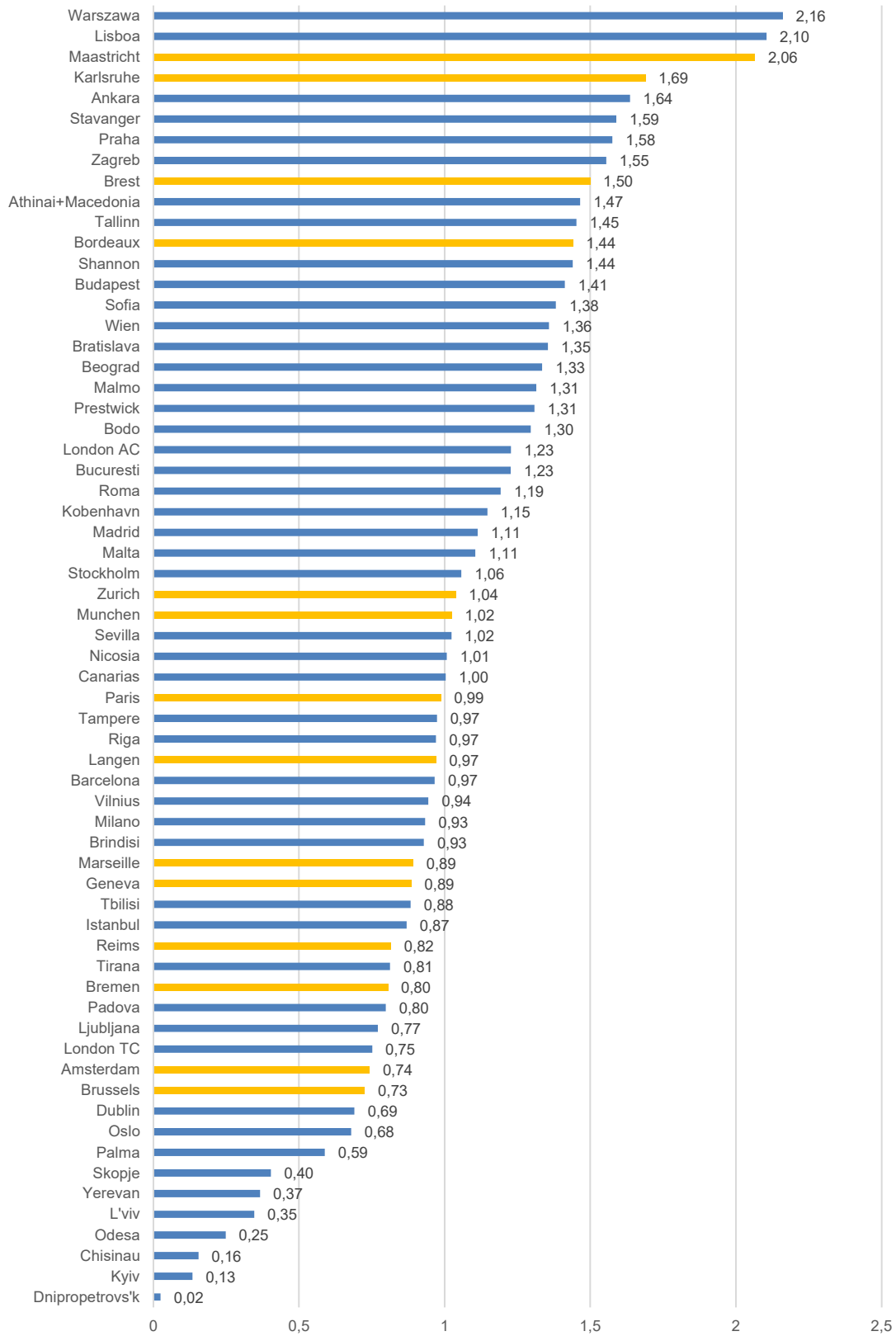


Abbildung A 14: ATCO-Produktivität europäischer ACCs

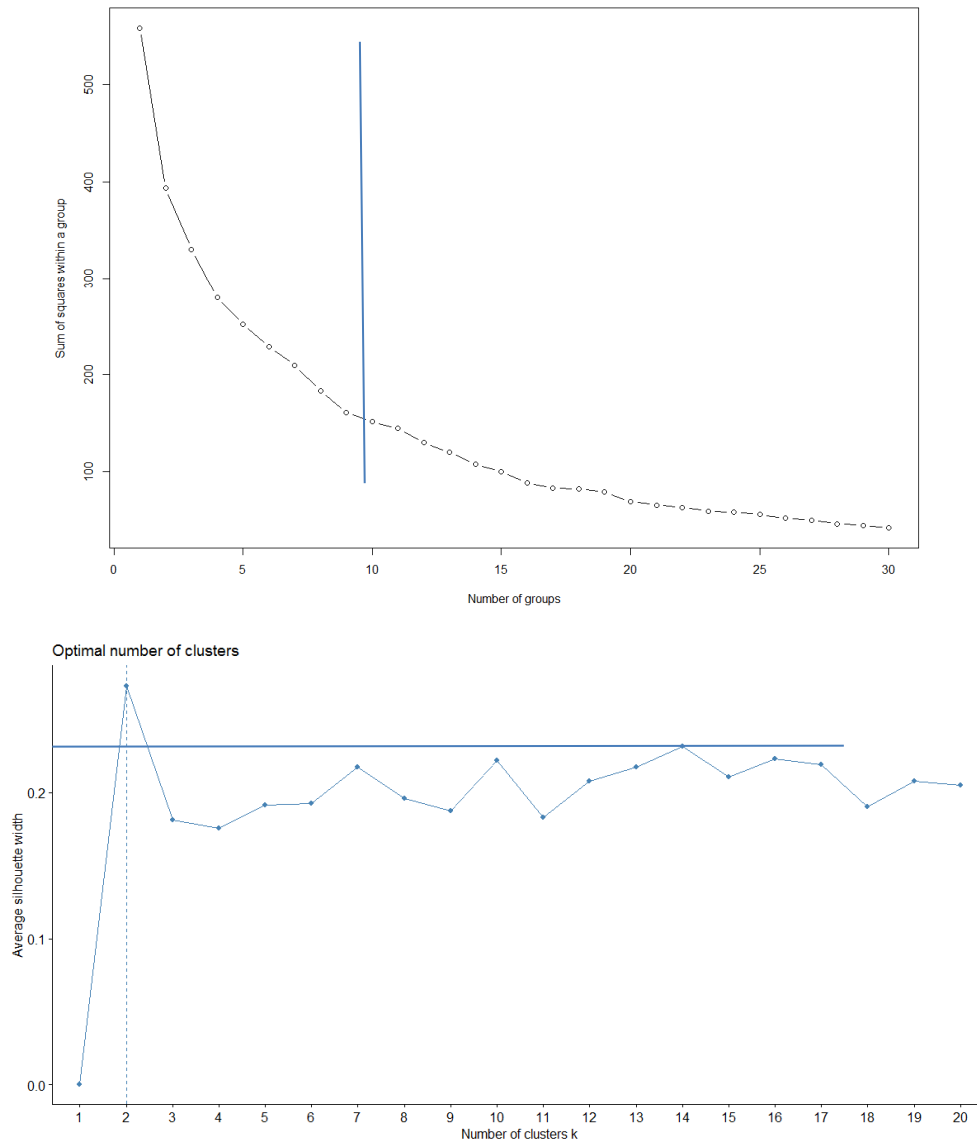
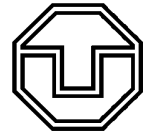


Abbildung A 15: Grafische Bestimmung der Cluster-Anzahl

Der Silhouetten Koeffizient berechnet die durchschnittliche Unähnlichkeit zwischen der Beobachtung i und allen anderen Punkten innerhalb desselben Clusters. Im optimalen Fall sollten alle Werte $S_i > 0$ sein.

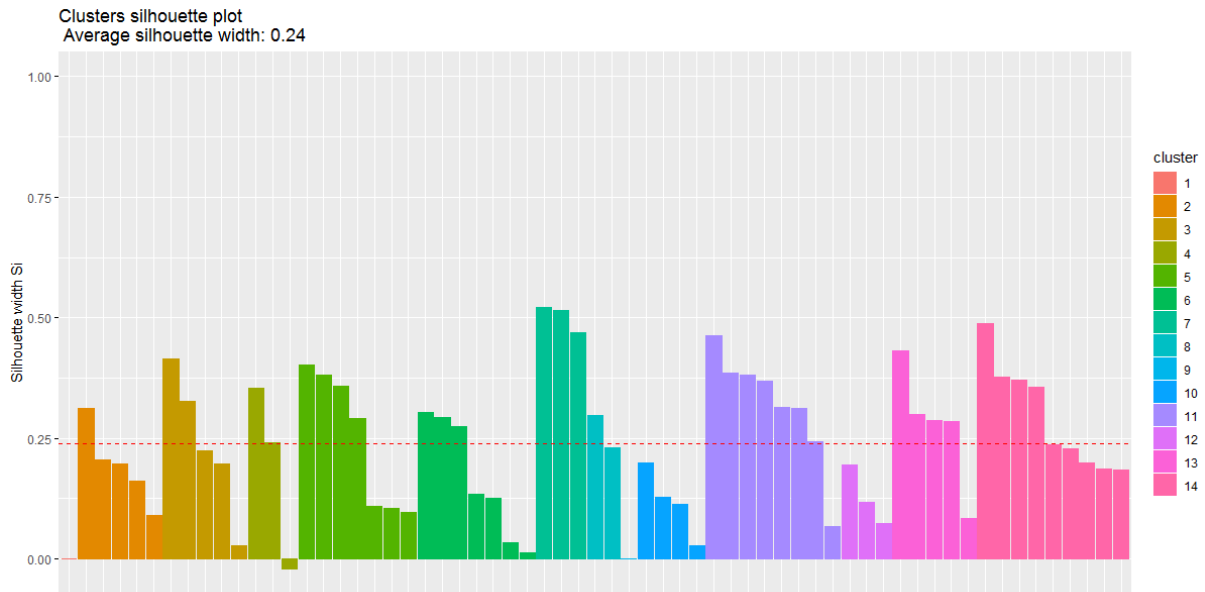
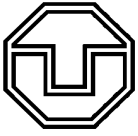


Abbildung A 16: Silhouetten-Diagramm zur Überprüfung der Cluster-Anzahl

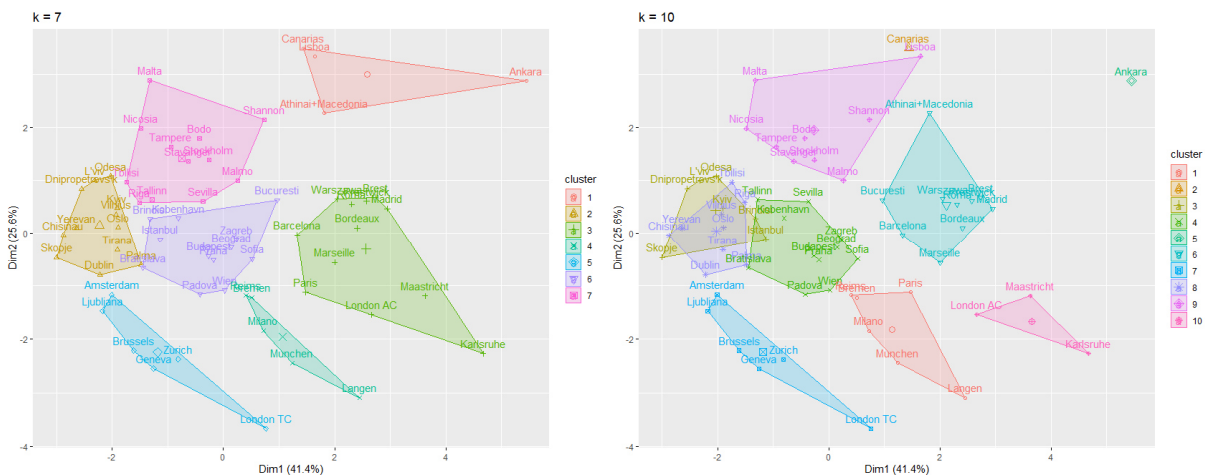


Abbildung A 17: Ergebnisse k-means Clustering

Tabelle A 12: Modellgüte Regression, ACC Ebene

	R1	R2	R3	RS1	RS2	RS3	RN1	RN2	RN3
Akaike	27,856	39,938	24,051	128,369	140,450	124,563	-67,779	-55,697	-71,584
Bayes	57,860	59,226	49,768	158,373	159,738	150,281	-37,775	-36,409	-45,867
R ²	0,707	0,585	0,706	0,707	0,585	0,706	0,706	0,585	0,706
Adjusted R ²	0,637	0,531	0,65	0,637	0,532	0,65	0,65	0,531	0,65
F-Statistik	10,07	11,05	12,51	10,07	11,05	12,51	12,51	11,05	12,51
Std-Fehler	0,271	0,308	0,266	0,602	0,684	0,592	0,124	0,1443	0,124

Für eine detaillierte Beschreibung des Akaike- und Bayes-Kriteriums siehe Burnham und Anderson (2004).

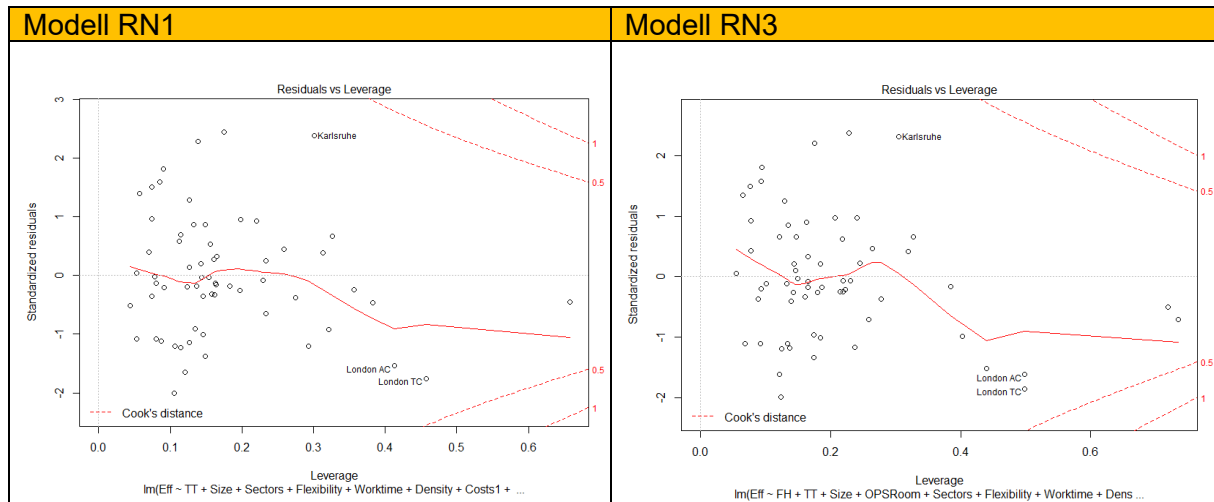
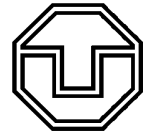


Abbildung A 18: Regressionsgüte ACC am Beispiel Cook-Graph

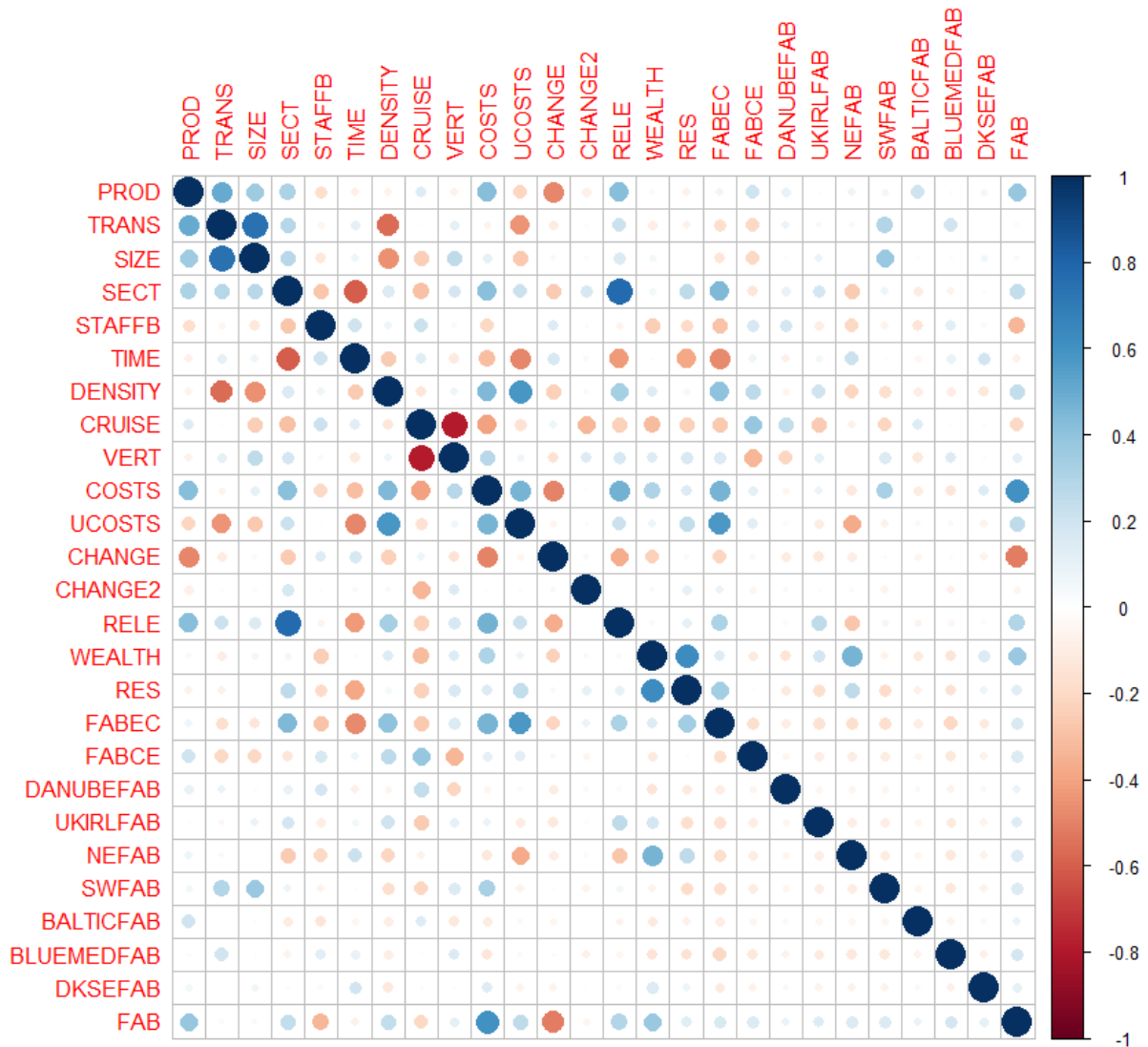


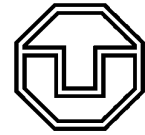
Abbildung A 19: Korrelationsmatrix Parameter der Regression auf ACC Ebene



Tabelle A 13: Iterative Regressionsanalyse ACC Ebene, originäre Daten

	<i>I1</i>	<i>I2</i>	<i>I3</i>	<i>I4</i>
INT	-1,157 (2,805)	-1,182 (2,622)	0,597 (2,727)	-3,081 (2,828)
STAFFB	-0,036 (0,049)	-0,065 (0,042)	0,008 (0,047)	-0,028 (0,039)
TIME	0,013 (0,406)	-0,152 (0,334)	-0,226 (0,343)	-0,285 (0,312)
SIZE	0,183 (0,063)***	0,264 (0,113)**	0,075 (0,118)	0,026 (0,097)
SECT	0,007 (0,01)	-0,006 (0,009)	0 (0,009)	0,001 (0,007)
CRUISE		0,007 (0,004)*	0,005 (0,004)	0,005 (0,003)
VERT		-0,003 (0,013)	0,003 (0,013)	-0,001 (0,01)
DENSITY		0,034 (0,014)**	0,014 (0,014)	0 (0,012)
TRANS		0,014 (0,01)	0,03 (0,011)***	0,03 (0,009)***
CHANGE		-2,236 (0,717)***	-1,064 (0,783)	0,169 (0,677)
FABEC			0,427 (0,196)**	-0,191 (0,225)
FABCE			0,678 (0,192)***	0,048 (0,229)
DANUBE			0,292 (0,258)	-0,278 (0,24)
UKIRL			0,417 (0,232)*	-0,128 (0,232)
NEFAB			0,553 (0,203)***	0,216 (0,219)
SWFAB			0,246 (0,19)	-0,564 (0,237)**
BALTIC			0,775 (0,263)***	0,261 (0,241)
BLUEMED			0,2 (0,158)	-0,278 (0,169)
DKSEFAB			0,488 (0,237)**	-0,119 (0,27)
COSTS				0,531 (0,123)***
WEALTH				-0,08 (0,049)
RES				-0,004 (0,003)*
Korr. R ²	0,18	0,49	0,56	0,72
Akaike	70,20	44,49	41,25	14,30
N	63	63	63	63

Standardfehler in Klammern. Signifikanzlevel 90%*, 95%** , 99%***



C8 Ergänzende Ergebnisse ACC Capacity Benchmarking

Tabelle A 14: Bewertung von in FABEC genutzter ATFM-Tools

Tool	Vorteile	Nachteile
iFMP	Ermöglicht die Integration interner Daten und erhöht somit die Datenqualität. Enthält Szenarien. Erhöht die Effizienz der Sektor-Konfigurationsplanung.	Geringe Genauigkeit für den Verkehr von Flughäfen unterhalb des Verantwortungsbereichs. Berechnete Komplexitätsangaben werden noch nicht zur Entscheidungsfindung herangezogen.
CHMI	Etabliertes Standardtool, das seit mehreren Jahren in ganz Europa eingesetzt wird. Nützlich für die Verwaltung der Flughafenkapazität.	Keine Berücksichtigung von Operational Air Traffic in der Vorhersage und Berechnung. Keine Berechnung der Komplexität oder Arbeitslast.
AirMagic	Simulationsbasiert. Berechnungen zur Komplexität und Arbeitslast. OAT-Verkehr wird berücksichtigt.	
Internal workload model	Individuell entworfen für den Luftraum. Berücksichtigt A-CDM. Harmonisiertes Modell, das von der strategischen bis zur taktischen Phase verwendet wird.	Berücksichtigt nicht die genutzten Landebahn (AMS) oder die militärischen Lufträume.
CRYSTAL	Erweiterte Berechnung des Occupancy Counts. Informationen werden benutzerfreundlich dargestellt. Häufigeres Informationsupdate als in CHMI. Darstellung von unerwartetem Verkehr.	Verwendet keine FDPS-Trajektorien (basierend auf OLDI-Nachrichten).
NOP	Nutzt nur eine Datenquelle.	
n-CONNECT	Sollte bis 2021 alle von PCP IR geforderten Funktionalitäten abdecken (z. B. kurzfristige ATFCM-Maßnahmen).	Noch kein näher spezifizierter Funktionsumfang.

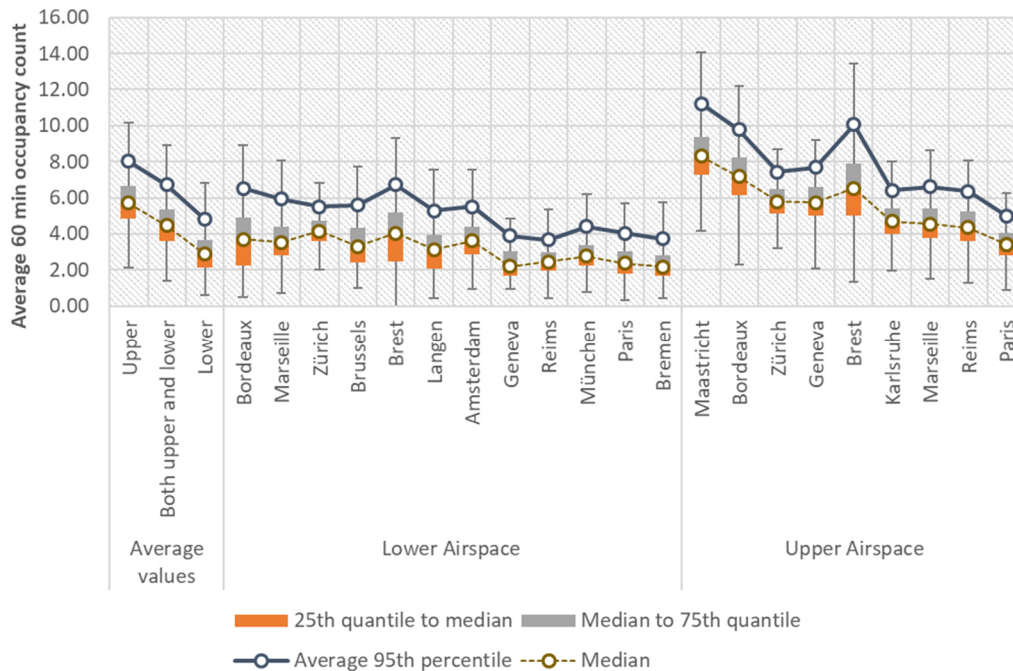


Abbildung A 20: Anzahl der gleichzeitig in einem Sektor befindlichen Luftfahrzeuge (FABEC 2019)

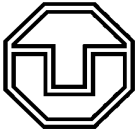
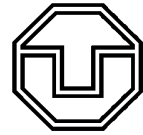


Tabelle A 15: Ansätze zur Berechnung der Komplexität auf Sektorebene

		A	B	C	D
Komponente	Verkehrsdichte	$= \frac{95\% \text{ Occupancy}}{\text{Volume}}$	$= \frac{\text{Avg. Occupancy}}{\text{Volume}}$		
	Heterogenität des Verkehrs	= aircraft type variance ²¹⁰	= flight type variance ²¹¹		
	Horizontale Konflikte	= 1 – Cruising flights		= 1 – Cruising flights	
	Vertikale Konflikte	= Average level changes per covered Flightlevel			
Gewichtung		Keine	Keine	70%/15%/15%	Keine

²¹⁰ Varianz zwischen Low, medium und high wake turbulence LFZ.

²¹¹ Varianz zwischen steigenden, sinkenden und im Streckenflug befindlichen Flugzeugen.



C9 Skalenerträge europäischer Flugsicherungen, 2016

Tabelle A 16: Skalenerträge, ANSP Ebene, 2016

ANSP	M1	M2	M2A	M2B	SI
Albcontrol	irs	irs	irs	irs	4
ANS CR	irs	irs	irs	irs	4
ARMATS	irs	irs	irs	irs	4
Austro Control	crs	irs	irs	irs	3
Avinor	crs	drs	irs	drs	-1
Belgocontrol	irs	irs	irs	irs	4
BULATSA	drs	drs	irs	irs	0
Croatia Control	drs	drs	irs	drs	-2
DCAC Cyprus	drs	drs	crs	irs	-1
DFS	crs	drs	crs	drs	-2
DHMI	crs	drs	crs	drs	-2
DSNA	drs	drs	drs	drs	-4
EANS	crs	crs	crs	crs	0
ENAIRE	crs	drs	drs	drs	-3
ENAV	crs	drs	drs	drs	-3
ANS Finland	irs	drs	irs	drs	0
HCAA	irs	drs	drs	drs	-2
HungaroControl	irs	irs	irs	irs	4
IAA	irs	drs	crs	drs	-1
LFV	drs	drs	drs	drs	-4
LGS	crs	irs	irs	irs	3
LPS	irs	irs	irs	irs	4
LVNL	crs	irs	irs	irs	3
MATS	crs	crs	crs	irs	1
M-NAV	irs	irs	irs	irs	4
MoldATSA	irs	irs	irs	irs	4
MUAC	crs	crs		crs	0
NATS	crs	drs	crs	drs	-2
NAV Portugal	crs	drs	crs	drs	-2
NAVIAIR	crs	drs	irs	drs	-1
Oro navigacija	irs	irs	irs	irs	4
PANSA	drs	drs	drs	drs	-4
ROMATSA	drs	drs	irs	drs	-2
Skyguide	drs	irs	drs	irs	0
Slovenia Control	irs	irs	irs	irs	4
SMATSA	drs	drs	irs	drs	-2
UkSATSE	drs	drs	drs	drs	-4



C10 Prognosequalität

Tabelle A 17: Prognosequalität STATFOR basierend auf DFS-Daten

Bezugsjahr	Prognosejahr	M _P	M _A	AFB	RFB	MAD	MAPE
2017	2016	3.226.000	3.259.000	-33.000	0,99	33.000	0,01
2017	2014	3.192.000	3.259.000	-67.000	0,98	67.000	0,02
2017	2011	3.638.000	3.259.000	379.000	1,12	379.000	0,12
2016	2015	3.142.000	3.146.000	-4.000	1,00	4.000	0,00
2016	2013	3.233.000	3.146.000	87.000	1,03	87.000	0,03
2016	2010	3.551.000	3.146.000	405.000	1,13	405.000	0,13
2015	2014	3.056.000	3.080.000	-24.000	0,99	24.000	0,01
2015	2012	3.255.000	3.080.000	175.000	1,06	175.000	0,06
2015	2009	3.605.000	3.080.000	525.000	1,17	525.000	0,17
2014	2013	3.051.000	3.030.000	21.000	1,01	21.000	0,01
2014	2011	3.397.000	3.030.000	367.000	1,12	367.000	0,12
2014	2008	3.977.000	3.030.000	947.000	1,31	947.000	0,31
2013	2012	3.082.000	2.990.000	92.000	1,03	92.000	0,03
2013	2010	3.257.000	2.990.000	267.000	1,09	267.000	0,09
2013	2007	3.740.000	2.990.000	750.000	1,25	750.000	0,25
2012	2011	3.240.000	3.018.000	222.000	1,07	222.000	0,07
2012	2009	3.281.000	3.018.000	263.000	1,09	263.000	0,09
2012	2006	3.562.000	3.018.000	544.000	1,18	544.000	0,18
2011	2010	3.060.000	3.078.000	-18.000	0,99	18.000	0,01
2011	2008	3.588.000	3.078.000	510.000	1,17	510.000	0,17
2011	2005	3.394.000	3.078.000	316.000	1,10	316.000	0,10

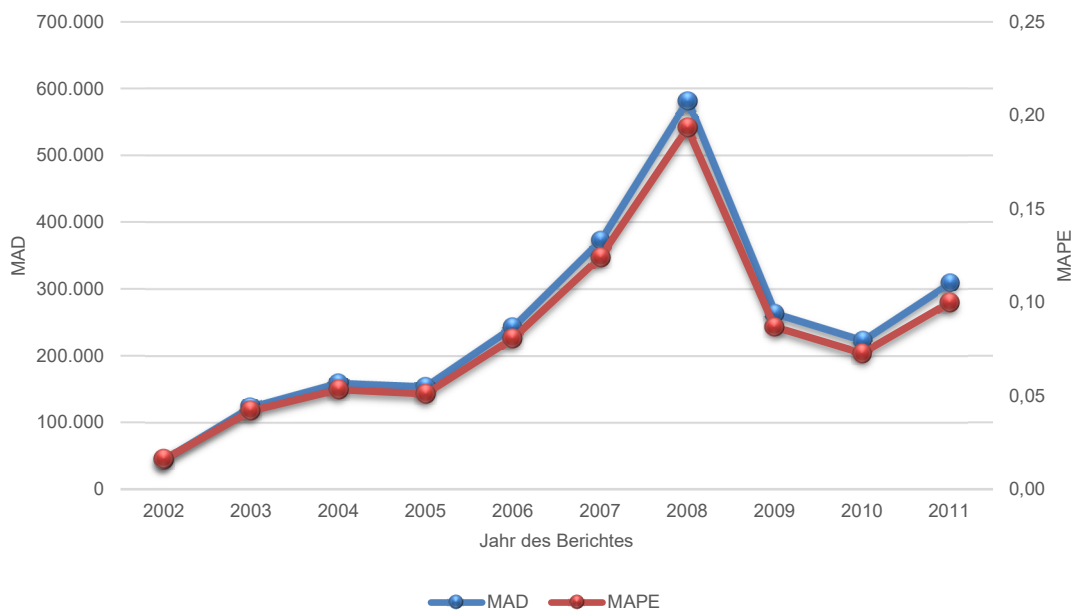
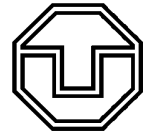
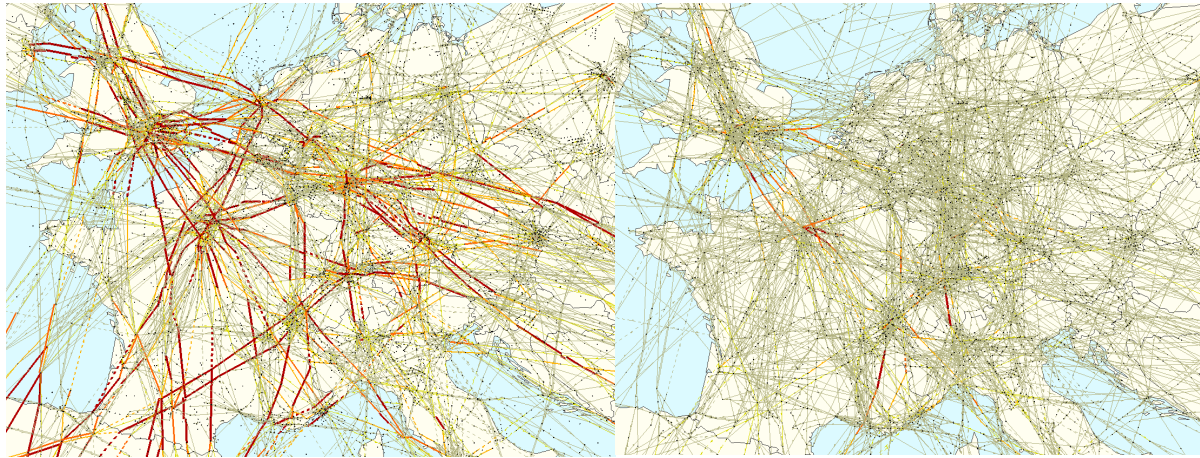


Abbildung A 21: Prognosequalität STATFOR



C11 Weitere Einflussfaktoren



IFR Trajektorien

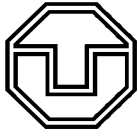
IFR GAT Trajektorien

Abbildung A 22: Density Plots für IFR und IFR-GAT, tatsächliche Trajektorien

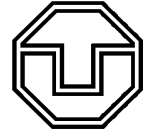
Tabelle A 18: Potentielle Einsparungen durch Single Sector Operations

ANSP	SSO?	Ein- sparung	ATCO-Hrs	ATCOs	SSO ATCOs	SSO ATCO- Hrs
Albcontrol	Nein	3,9%	75.453	49	47	72.524
ANS CR	Nein	3,9%	301.269	192	185	289.573
ARMATS	Nein	3,9%	111.628	82	79	107.294
Austro Control	Nein	3,9%	408.924	291	280	393.048
Avinor	Nein	4,9%	634.709	407	387	603.720
Belgocontrol	Ja	0%	305.320	232	232	305.320
BULATSA	Nein	3,9%	318.548	248	238	306.180
Croatia Control	Nein	3,9%	336.294	234	225	323.238
DCAC Cyprus	Nein	3,9%	172.084	86	83	165.403
DFS	Nein	5,9%	1.751.189	1.777	1.672	1.648.178
DHMI	Nein	5,9%	1.446.656	1.120	1.054	1.361.559
DSNA	Nein	5,9%	3.572.088	2.782	2.618	3.361.965
EANS	Nein	3,9%	83.680	52	50	80.431
Enaire	Nein	5,9%	2.047.833	1.779	1.674	1.927.372
ENAV	Teilw. ²¹²	5,0%	1.825.164	1.414	1.343	1.733.906
ANS Finland	Nein	3,9%	280.001	183	176	269.130
HCAA	Nein	4,9%	844.192	496	472	802.976
HungaroControl	Ja	0%	270.205	173	173	270.205
IAA	Ja	0%	311.322	204	204	311.322
LFV	Nein	4,9%	776.199	470	447	738.302
LGS	Nein	3,9%	128.400	93	89	123.415
LPS	Nein	3,9%	122.634	82	79	117.873
LVNL	Nein	3,9%	336.859	178	171	323.781

²¹² ENAV nutzt SSO lediglich für die Anflugkontrolle. Daher würde ein Einsparpotential von 5,0 unterstellt.



ANSP	SSO?	Ein- sparung	ATCO-Hrs	ATCOs	SSO ATCOS	SSO ATCO- Hrs
MATS	Nein	3,9%	107.782	54	52	103.598
M-NAV	Nein	3,9%	85.663	65	62	82.337
MoldATSA	Nein	3,9%	105.134	73	70	101.052
MUAC	Nein	3,9%	299.908	268	258	288.264
NATS	Nein	5,9%	1.726.166	1.415	1.332	1.624.627
NAV Portugal	Ja	0%	400.532	220	220	400.532
NAVIAIR	Nein	3,9%	302.240	208	200	290.506
Oro navigacija	Nein	3,9%	135.861	87	84	130.586
PANSA	Nein	4,9%	550.403	479	456	523.530
ROMATSA	Nein	4,9%	547.015	448	426	520.308
Skyguide	Nein	4,9%	450.185	362	344	428.205
Slovenia Control	Nein	3,9%	130.170	91	88	125.116
SMATSA	Nein	3,9%	319.104	277	266	306.715
UkSATSE	Nein	5,9%	1.065.672	842	792	1.002.985
Sum			22.686.485	17.513	16.629	21.565.078

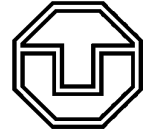


Literaturverzeichnis

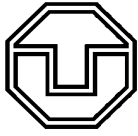
- Aberle, G. (2014): „Transportwirtschaft: Einzelwirtschaftliche und gesamtwirtschaftliche Grundlagen“. 4. Auflage, München, Wien: Oldenbourg Verlag.
- ACARE (2011): „Flightpath 2050 - Europe's Vision for Aviation“. Luxemburg: European Commission.
- ACI, CANSO, IATA, ICAO, ICCAIA (2019): „Aviation Benefits Report 2019“. Industry High Level Group.
- Adler, N., Delhay, E., Kivel, A., Proost, S. (2017): „Motivating Air Navigation Service Provider Performance“. Leuven: KU Leuven Discussion Papers.
- Ahmada, M. F., Ishtiaq, M., Hamid, K., Khurram, M. U., Nawaz, A. (2017): „Data Envelopment Analysis and Tobit Analysis for Firm Efficiency in Perspective of Working Capital Management in Manufacturing Sector of Pakistan“. In: *International Journal of Economics and Financial Issues*, Vol. 7, pp. 706–713.
- Aigner, D., Knox-Lovell, C. A., Schmidt, P. (1977): „Formulation and estimation of stochastic frontier production function models“. In: *Journal of Econometrics*, Vol. 6 (1), pp. 21–37.
- Albers, S., Klapper, D., Konradt, U., Walter, A., Wolf, J. (2009): „Methodik der empirischen Forschung“. 3. Auflage, Wiesbaden: Gabler Verlag.
- Albrecht, L., Klumpp, M., Keuschen, T. (2012): „DEA-Effizienzvergleich deutscher Verkehrsflughäfen in den Bereichen Passage und Fracht“. In: *ild Schriftenreihe Logistikforschung*, Vol. 28.
- Albright, J. (2015): „Saying Yes to Portable Electronic Devices“. In: *Aviation Week Network, Business & Commercial Aviation*,.
- Andersen, P., Petersen, N. C. (1993): „A Procedure for Ranking Efficient Units in Data Envelopment Analysis“. In: *Management Science*, Vol. 39, pp. 1261–1264.
- Arnaldo, R. M., Comendador, V. F. G., Barrangan, R., Pérez, L. (2014): „European Air Navigation Service Providers' Efficiency Evaluation Through Data Envelopment Analysis (DEA)“. St. Petersburg: International Council of Aeronautical Sciences Conference (ICAS).
- ATAG (2005): „The economic and social benefits of air transport“. Geneva.
- ATAG (2014): „Aviation Benefits Beyond Borders“. Geneva.
- ATAG (2018): „Aviation Benefits Beyond Borders“. Geneva.
- Auer, B., Rottmann, H. (2015): „Statistik und Ökonometrie für Wirtschaftswissenschaftler - Eine anwendungsorientierte Einführung“. 3. Auflage, Wiesbaden: Gabler.
- Bacher, J., Pöge, A., Wenzig, K. (2011): „Clusteranalyse: Anwendungsorientierte Einführung in Klassifikationsverfahren“. 3. Auflage, München: Oldenbourg Verlag.
- Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2018): „Multivariate Analysemethoden: Eine anwendungsorientierte Einführung“. 15. Auflage, Berlin, Heidelberg: Springer.
- Banker, R. D. (1986): „Stochastic data envelopment analysis“. Pittsburgh: Carnegie-Mellon University.
- Banker, R. D., Charnes, A., Cooper, W. W. (1984): „Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis“. In: *Management Science*, Vol. 30, pp. 1078–1092.



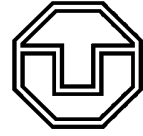
- Banker, R. D., Natarajan, R. (2008): „Evaluating Contextual Variables Affecting Productivity Using Data Envelopment Analysis“. In: *Operations Research*, Vol. 56, pp. 48–58.
- Barbakh, W. A., Wu, Y., Fyfe, C. (2012): „Non-Standard Parameter Adaptation for Exploratory Data Analysis“. Berlin, Heidelberg: Springer.
- Bartlett, M. S. (1951): „The effect of standardization on a Chi-square approximation in factor analysis“. In: *Biometrika*, Vol. 38, pp. 337–344.
- Battese, G. E., Corra, G. S. (1977): „Estimation of a Production Frontier Model With Application to the Pastoral Zone off Eastern Australia“. In: *Australian Journal of Agricultural Economics*, Vol. 21 (3), pp. 169–179.
- Bem, A., Daszyńska-Żygadło, K., Hajdíkóvá, T., Juhás, P. (2020): „Finance and Sustainability“. In: *Proceedings from the Finance and Sustainability Conference*, Cham: Springer Nature Switzerland.
- Berg, M. de, Cheong, O., Kevald, M. van, Pvermars, M. (2008): „Computational Geometry - Algorithms and Applications“. Berlin, Heidelberg: Springer.
- Biebig, P., Althof, W., Wagener, N. (2017): „Seeverkehrswirtschaft: Kompendium“. 4. Auflage, München: Oldenbourg Verlag.
- Bielecki, A. (2011): „Efficient Frontier Analysis“. In: *Zusätzliche Beiträge zu Methodik der empirischen Forschung*, Wiesbaden: Gabler.
- Bilotkach, V., Gitto, S., Jovanović, R., Mueller, J., Pels, E. (2015): „Cost-efficiency benchmarking of European air navigation service providers“. In: *Transportation Research Part A: Policy and Practice*, Vol. 77, pp. 50–60.
- Birkmeier, B., Korn, B. (2014): „Five transition strategies for sectorless ATM“. Colorado Springs: Digital Avionics Systems Conference (DASC).
- Blondiau, T., Delhaye, E., Adler, N., Proost, S. (2016): „Productivity Measurement of Air Traffic Management in Europe“. COMPAIR Project.
- Bogetoft, P. (2013): „Performance Benchmarking: Measuring and Managing Performance“. New York: Springer Science & Business Media.
- Bogetoft, P., Otto, L. (2011): „Benchmarking with DEA, SFA, and R“. New York: Springer Science & Business Media.
- Bortz, J. (2013): „Statistik für Sozialwissenschaftler“. 3. Auflage, Berlin, Heidelberg: Springer.
- Brunnberg, D. (2018): „Kapitalmarktkommunikation und Finanzanalysten“. Wiesbaden: Springer-Verlag.
- Buis, M. L. (2010): „Analyzing Proportions“. Tübingen: Institut für Soziologie, Eberhard Karls Universität.
- Burnham, K. P., Anderson, D. R. (2004): „Multimodel Inference: Understanding AIC and BIC in Model Selection“. In: *Sociological Methods & Research*, Vol. 33 (2), pp. 261–304.
- Button, K. (2010): „Transport Economics“. 3. Auflage, Cheltenham, Northampton: Edward Elgar Publishing.
- Button, K., Neiva, R. (2013): „Single European Sky and the functional airspace blocks: Will they improve economic efficiency?“. In: *Journal of Air Transport Management*, Vol. 33, pp. 73–80.



- Button, K., Neiva, R. (2014): „Economic Efficiency of European Air Traffic Control Systems“. In: *Journal of Transport Economics and Policy*, Vol. 49, pp. 626–639.
- Buxbaum, J., Standfuss, T. (2014): „Data Envelopment Analysis - Vorschlag einer neuen Methodik zum Performancevergleich der Flugsicherungen“. In: *Innovation im Fokus*, Vol. 1, pp. 11–16.
- CANSO (2015): „Global ANS Performance Report 2015“. Amsterdam.
- Cantner, U., Krüger, J., Hanusch, H. (2007): „Produktivitäts- und Effizienzanalyse - Der nicht parametrische Ansatz“. Berlin: Springer-Verlag.
- Carree, M. A. (2002): „Technological inefficiency and the skewness of the error component in stochastic frontier analysis“. In: *Economics Letters*, Vol. 77, pp. 101–107.
- Carta, A., Steel, M. F. J. (2012): „Modelling multi-output stochastic frontiers using copulas“. In: *Computational Statistics & Data Analysis*, Vol. 56 (11), pp. 3757–3773.
- Cento, A. (2009): „The Airline Industry: Challenges in the 21st Century“. Heidelberg: Physica-Verlag.
- Cerny, B. A., Kaiser, H. F. (1977): „A Study Of A Measure Of Sampling Adequacy For Factor-Analytic Correlation Matrices“. In: *Multivariate Behavioral Research*, Vol. 12, pp. 43–47.
- Charnes, A., Cooper, W. W., Rhodes, E. (1978): „Measuring the efficiency of decision making units“. In: *European Journal of Operational Research*, Vol. 2 (5), pp. 429–444.
- Chen, Y., Du, J. (2015): „Super-Efficiency in Data Envelopment Analysis“. In: *Data Envelopment Analysis: A Handbook of Models and Methods*, New York: Springer.
- Clark, T. S., Linzer, D. A. (2015): „Should I Use Fixed or Random Effects?“. In: *Political Science Research and Methods*, Vol. 3 (2), pp. 399–408.
- Coelli, T. (2008): „A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program“. Armindale: University of New England.
- Coelli, T. J., Rao, P. D. S., O'Donnell, C. J., Battese, G. E. (2005): „An Introduction to Efficiency and Productivity Analysis“. New York: Springer.
- Collier, T., Johnson, A. L., Ruggiero, J. (2011): „Technical efficiency estimation with multiple inputs and multiple outputs using regression analysis“. In: *European Journal of Operational Research*, Vol. 208, pp. 153–160.
- Cook, A. (2016): „European Air Traffic Management: Principles, Practice and Research“. London: Routledge.
- Cook, A., Tanner, G. (2011): „European airline delay cost reference values“. London: University of Westminster, Department of Transport Studies. Published by EUROCONTROL Performance Review Unit (Brussels).
- Cooper, W. W., Seiford, L. M., Zhu, J. (2011): „Data Envelopment Analysis - History, Models and Interpretations“. In: *Handbook on Data Envelopment Analysis*, New York, Dordrecht, Heidelberg, London: Springer.
- Ćujić, M., Jovanović, M., Savić, G., Jakšić, M. L. (2015): „Measuring the Efficiency of Air Navigation Services System by Using DEA Method“. In: *International Journal for Traffic and Transport Engineering*, Vol. 5 (1), pp. 36–44.



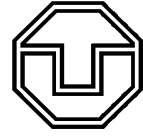
- Daduna, J. R., Voß, S. (2000): „Informationsmanagement im Verkehr“. Berlin, Heidelberg: Springer.
- Daly, K. J. (2008): „Financial volatility: Issues and measuring techniques“. In: *Physica A: Statistical Mechanics and its Applications*, Vol. 387, pp. 2378–2393.
- Delahaye, D., Schoenauer, M., Alliot, J. M. (1998): „Airspace sectoring by evolutionary algorithms“. Anchorage: IEEE International Congress on Evolutionary Computation (CEC).
- Deltuvaitė, V. (2018): „Main Determinants of Volatility in Air Traffic and Its Impact on ANSPs' Performance“. Warsaw: FABEC Workshop on Volatility in air traffic and its impact on ATM Performance.
- Deutscher Bundestag (2009): „Grundgesetz für die Bundesrepublik Deutschland (GG), Artikel 87d“.
- DFS (2015): „Aeronautical Information Publication VFR“. Langen.
- DFS (2019): „Luftraumstruktur/Sichtflregeln in der Bundesrepublik Deutschland“. Langen: www.dfs.de/dfs_homepage/de/Flugsicherung/Luftraum/luftraumstruktur_092016.pdf (10.04.2019).
- Diaconu, A. G., Stanciu, V., Pleter, O. T. (2014): „Air traffic complexity metric for en-route and terminal areas“. In: *Scientific Bulletin University Politehnica of Bucharest*, Vol. 76, pp. 13–24.
- Djokic, J. (2014): „Investigation into Air Traffic Complexity as a Driver of a Controller's Workload“. Technische Universität Dresden.
- DSNA (2018): „ATFCM and ATC - Filling the gap through extended CDM concept“. Paris: <https://www.ecologique-solidaire.gouv.fr/sites/default/files/Dynamic-ATFCM.pdf> (17.02.2020).
- Eibner, W. (2010): „Understanding International Trade: Theory & Policy / Anwendungsorientierte Außenwirtschaft: Theorie & Politik“. Berlin: Walter de Gruyter.
- Embrechts, P., Lindskog, F., McNeil, A. (2001): „Modelling Dependence with Copulas and Applications to Risk Management“. ETHZ Zurich: Department of Mathematics.
- EUROCONTROL (2006): „Complexity Metrics for ANSP Benchmarking Analysis“. Brussels: ACE Working Group on Complexity.
- EUROCONTROL (2007): „Terms of Reference and Rules of Procedure“. Brüssel: Performance Review Commission.
- EUROCONTROL (2008a): „Airspace Complexity For Regulation Purposes“. EEC Note No. 13/2008.
- EUROCONTROL (2008b): „ATCO in OPS“. In: *Specification for Economic Information Disclosure*, Brussels.
- EUROCONTROL (2010): „SES II Proposed EU-wide Performance Targets for the period 2012-2014“. Brüssel: Performance Review Commission.
- EUROCONTROL (2011): „SES II Performance Scheme - Assessment of National / FAB Performance Plans with Performance Targets for the period 2012-2014“. Brussels: Performance Review Body.
- EUROCONTROL (2012): „Specification for Economic Information Disclosure“. Brussels: Performance Review Unit and Performance Review Commission.
- EUROCONTROL (2013a): „Challenges in Growth - European Air Traffic in 2050“. Brussels: Network Manager.



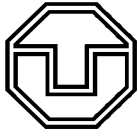
- EUROCONTROL (2013b): „SES Performance Scheme Reference Period 2 (2015-2019), Performance Plans for RP2“. Brussels: www.eurocontrol.int/articles/ses-performance-scheme-reference-period-2-2015-2019, (2.10.2016).
- EUROCONTROL (2014): „A Complexity Study of the Maastricht Upper Airspace Centre“. Brussels: EUROCONTROL Experimental Centre.
- EUROCONTROL (2015): „Advanced Flexible Use of Airspace“. Brussels: <https://www.eurocontrol.int/concept/advanced-flexible-use-airspace> (28.03.2020).
- EUROCONTROL (2016a): „Aeronautical Information Management“. Brussels: <https://www.eurocontrol.int/concept/aeronautical-information-management> (28.07.2016).
- EUROCONTROL (2016b): „ATM Cost-Effectiveness (ACE) 2014 Benchmarking Report with 2015-2019 outlook“. Brussels: Performance Review Commission.
- EUROCONTROL (2016c): „CAPAN Methodology - Sector Capacity Assessment“. Brussels: Network Manager.
- EUROCONTROL (2016d): „Monthly Adjusted Unit Rates“. Brussels: Central Route Charges Office, <https://www.eurocontrol.int/services/monthly-adjusted-unit-rates>.
- EUROCONTROL (2016e): „Performance Review Unit“. Brussels: <http://www.eurocontrol.int/articles/performance-review-unit> (24.08.2016).
- EUROCONTROL (2016f): „U.S. - Europe continental comparison of ANS cost-efficiency trends 2006-2014“. Brussels: Performance Review Unit.
- EUROCONTROL (2016g): „Monthly Network Operations Report“. Brüssel: Network Manager.
- EUROCONTROL (2017a): „FABs“. Brussels: <https://www.eurocontrol.int/functional-airspace-block-fabs-defragmenting-european-airspace> (20.08.2017).
- EUROCONTROL (2017b): „Demand Data Repository (DDR)“. <http://www.eurocontrol.int/ddr> (16.03.2017).
- EUROCONTROL (2018a): „ATM Cost-Effectiveness (ACE) Benchmarking Report for 2016“. Brussels: Performance Review Unit.
- EUROCONTROL (2018b): „Functional Airspace Block (FAB)“. Brussels: <https://www.eurocontrol.int/articles/functional-airspace-block-fab> (23.07.2018).
- EUROCONTROL (2018c): „NEST modelling tool“. Brussels: <http://www.eurocontrol.int/services/nest-modelling-tool>.
- EUROCONTROL (2018d): „Seven-Year Forecast - Flight Movements and Service Units 2018 - 2024“. Brussels: STATFOR.
- EUROCONTROL (2018e): „The Monthly Network Operations Report for September 2018“. Brussels: Network Manager.
- EUROCONTROL (2018f): „UAS ATM Integration - Operational Concept“. Brussels: Aviation Cooperation and Strategies Division.
- EUROCONTROL (2019a): „Air traffic management cost-effectiveness (ACE) benchmarking report for 2017“. Brussels: Performance Review Unit.



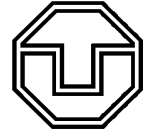
- EUROCONTROL (2019b): „European Aviation in 2040 - Challenges of Growth - Flight Forecast“. Brussels: STATFOR.
- EUROCONTROL (2019c): „FIR/UIR in the Lower Airspace“. Brussels: Network Manager. https://www.eurocontrol.int/sites/default/files/2019-08/23052019%20-%20firuir_lower_airspace_ectl.pdf (15.04.2020).
- EUROCONTROL (2019d): „FIR/UIR in the Upper Airspace“. Brussels: Network Manager. https://www.eurocontrol.int/sites/default/files/2019-08/23052019%20-%20firuir_upper_airspace_ectl.pdf (15.04.2020).
- EUROCONTROL (2019e): „Free Routes Airspace (FRA) Design Guidelines“. Brussels: Network Manager.
- EUROCONTROL (2019f): „New traffic record set: 37,228 flights in one day“. Brussels: Network Manager.
- EUROCONTROL (2019g): „Pan-European ANS Performance Data Portal“. Brussels, Performance Review Unit: <https://ansperformance.eu/data/> (24.08.2019).
- EUROCONTROL (2019h): „Performance Review Report - An Assessment of Air Traffic Management in Europe during the Calendar Year 2018“. Brussels: Performance Review Commission.
- EUROCONTROL (2019i): „Seven-Year Forecast: Autumn 2019 Update - Flight Movements and Service Units 2019-2025“. Brussels: STATFOR.
- EUROCONTROL (2019j): „The Central Route Charges Office and the Route Charges System“. Brussels, Central Route Charges Office: <https://www.eurocontrol.int/articles/what-are-route-charges> (09.04.2019).
- EUROCONTROL (2020a): „About us“. Brussels, <https://www.eurocontrol.int/about-us> (15.04.2020).
- EUROCONTROL (2020b): „Base of Aircraft Data“. Brussels: <https://simulations.eurocontrol.int/solutions/bada-aircraft-performance-model/> (17.02.2020).
- EUROCONTROL (2020c): „Composite flight-hour“. Performance Review Unit: <https://ansperformance.eu/definition/composite-flight-hour/> (29.04.2020).
- EUROCONTROL (2020d): „EUROCONTROL Area“. <https://ansperformance.eu/definition/eurocontrol-area/> (02.06.2020).
- EUROCONTROL (2020e): „Industry Moinitor, Issue 215“. Brussels: Network Manager.
- EUROCONTROL (2020f): „Minutes of en route ATFM delay per flight - RP2“. Brussels, Performance Review Unit: <https://www.eurocontrol.int/prudata/dashboard/metadata/minutes-of-en-route-atfm-delay-per-flight-ses-rp2/> (23.06.2020).
- EUROCONTROL (2020g): „Monthly Network Operations Report - Analysis January 2020“. Brussels.
- EUROCONTROL (2020h): „OneSky Online“. Brussels: ACE Working Group.
- EUROCONTROL (2020i): „Pan-European ANS Performance Data Portal - Traffic Complexity“. Brussels: Performance Review Unit.
- EUROCONTROL (2020j): „Principles for establishing the cost-base for en route charges and the calculation of the unit rates“. Brussels: Central Route Charges Office.



- EUROCONTROL (2020k): „Top Stats December 2019“. Brussels: Aviation Intelligence Unit.
- EUROCONTROL (2020l): „Traffic Complexity Score Dataset“. Performance Review Unit: http://ansperformance.eu/references/dataset/Traffic_Complexity_Score.html (10.02.2020).
- EUROCONTROL (2020m): „Traffic Complexity Score Dataset - Data description“. Brussels: Performance Review Unit.
- EUROCONTROL (2020n): „What is a slot“. Brussels: Network Manager, <https://www.eurocontrol.int/article/what-is-a-slot> (14.04.2020).
- EUROCONTROL, FAA (2016): „Comparison of Air Traffic Management-Related 2015 Operational Performance: U.S./Europe“.
- EUROCONTROL, FAA (2019a): „U.S. - Europe Comparison of ATM related operational performance 2017“. Brussels.
- EUROCONTROL, FAA (2019b): „U.S. - Europe continental comparison of ANS cost-efficiency trends 2006-2016“. Brussels: Performance Review Unit.
- EUROCONTROL, NATS (2000): „Investigating the Air Traffic Complexity - Potential impacts on workload and costs“. Brétigny-sur-Orge: EUROCONTROL Experimental Centre.
- European Commission (2004a): „Regulation (EC) No 549/2004 of the European Parliament and of the Council of 10 March 2004 laying down the framework for the creation of the Single European Sky“. In: *Official Journal of the European Union*, Vol. 96, pp. 1–8.
- European Commission (2004b): „Regulation (EC) No 550/2004 of the European Parliament and of the council of 10 March 2004 on the provision of air navigation services in the single European sky (the service provision Regulation)“. In: *Official Journal of the European Union*, Vol. 96, pp. 10–19.
- European Commission (2010a): „Commission Regulation (EU) No 1191/2010 of 16 December 2010 laying down a common charging scheme for air navigation services“. In: *Official Journal of the European Union*, Vol. 333, pp. 6–13.
- European Commission (2010b): „Commission Regulation (EU) No 255/2010 of 25 March 2010 laying down common rules on air traffic flow management“. In: *Official Journal of the European Union*, Vol. 80, pp. 10–16.
- European Commission (2010c): „Commission Regulation (EU) No 691/2010 of 29 July 2010 laying down a performance scheme for air navigation services and network functions and amending Regulation (EC) No 2096/2005 laying down common requirements for the provision of air navigation services“. In: *Official Journal of the European Union*, Vol. 201, pp. 1–22.
- European Commission (2011): „Commission Decision of 21 February 2011 setting the European Union-wide performance targets and alert thresholds for the provision of air navigation services for the years 2012 to 2014“. In: *Official Journal of the European Union*, Vol. 48, pp. 16–18.
- European Commission (2012a): „Commission Implementing Regulation (EU) No 923/2012 of 26 September 2012 laying down the common rules of the air and operational provisions regarding services and procedures in air navigation and amending Implementing Regulation (EU) No 1035/2011 and Regulations (EC) No 1265/2007, (EC) No 1794/2006, (EC) No 730/2006, (EC) No 1033/2006 and (EU) No 255/2010“. In: *Official Journal of the European Union*, Brussels Vol. 281, pp. 1–66.
- European Commission (2012b): „Commission Recommendation of 26 July 2012 on the implementation of performance plans and targets in consistency with the European Unionwide performance targets



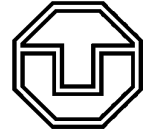
- adopted pursuant to Commission Regulation (EU) No 691/2010 and the preparation for the second performance reference period“. In: *Official Journal of the European Union*, Vol. 228, pp. 1–2.
- European Commission (2012c): „Single European Sky: 10 years on and still not delivering“. Brussels: https://ec.europa.eu/commission/presscorner/detail/en/IP_12_1089 (28.02.2020).
- European Commission (2013): „Commission Implementing Regulation (EU) No 390/2013 of 3 May 2013 laying down a performance scheme for air navigation services and network functions“. In: *Official Journal of the European Union*, Vol. 128, pp. 1–30.
- European Commission (2014a): „Commission Implementing Decision of 24 September 2014 on the extension of the designation of the Performance Review Body of the single European sky“. In: *Official Journal of the European Union*, Brüssel Vol. 281, pp. 5–8.
- European Commission (2014b): „Commission Implementing Decision of 11 March 2014 setting the Union-wide performance targets for the air traffic management network and alert thresholds for the second reference period 2015-19“. In: *Official Journal of the European Union*, Vol. 71, pp. 20–23.
- European Commission (2016): „Performance and Charging Scheme“. Mobility and Transport: http://ec.europa.eu/transport/modes/air/single_european_sky/performance_and_charging_en.htm (02.10.2016).
- European Commission (2018): „EU-wide target ranges for RP 3 - Annex 2. Air Navigation Service Providers: Advice on benchmarking of ANSPs and EU-wide cost targets“. Brussels: Performance Review Body.
- European Commission (2020a): „RP3 developments“. Mobility and Transport, Brussels: https://ec.europa.eu/transport/modes/air/single-european-sky/ses-performance-and-charging/rp3-developments_en (04.03.2020).
- European Commission (2020b): „Single European Sky II“. Mobility and Transport, Brussels: https://ec.europa.eu/transport/modes/air/single_european_sky/ses_2_da (04.03.2020).
- EUROSTAT (2019): „Transport Database“. <https://ec.europa.eu/eurostat/data/database> (24.08.2019).
- Ewers, H.-J., Tegner, H. (2002): „Volkswirtschaftliche Nutzen durch Wettbewerb im europäischen Air Traffic Management am Beispiel Deutschland“. Technische Universität Berlin, Fachgebiet Wirtschafts- und Infrastrukturpolitik.
- FABEC (2012): „FAB Performance Plan, Second Reference Period (2015-2019)“. http://www.fabec.eu/fabec_homepage/en/Performance/Performance%20Plan/FAB%20EC%20FAB%20RP2%20Performance%20Plan%20body_Revised_V2.0_clean_.pdf (02.10.2016).
- FABEC (2014): „ATM in Europe - It's all about performance“. Langen: Performance Management Group.
- FABEC (2015): „ACE Data Analysis“. Langen: FABEC Performance Management Group. Contributors: Thomas Standfuss, Christoph Czech, Thomas Hellbach.
- FABEC (2018a): „Facts & Figures“. <https://www.fabec.eu/about/facts-and-figures> (14.06.2018).
- FABEC (2018b): „Volatility in ATM: Cases, Challenges, Solutions“. Madrid: World ATM Congress 2018, Inter-FAB Panel, FABEC OPS Theatre.
- FABEC (2018c): „Volatility Task Force - Final Report“. Langen: Performance Management Group, Volatility Task Force.



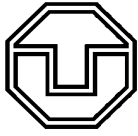
- FABEC (2019): „ACC Capacity Benchmarking Report - An analysis of difference in capacity and operational practices applied across selected FABEC ACCs“. Langen: FABEC Performance Management Group, PBTF. Contributors: Maja Marciniak-Cork, Thomas Hellbach, Juan Espinar-Nova, Thomas Standfuß, Stéphane Mariétan, Riccardo Massacci, and Christoph Czech.
- FABEC, BLUE MED FAB (2019): „Impact of Forecast Inaccuracy“. InterFAB Meeting 2019.
- Färe, R., Grosskopf, S., Norris, M., Z, Z. (1994): „Productivity growth, technical progress and efficiency changes in industrialised countries.“. In: *American Economic Review*, Vol. 84, pp. 66–83.
- Flight Service Bureau (2018): „Internationale OPS 2018“. OPS Group, Christchurch: [http://flightservicebureau.org/tag/france/\(14.06.2018\)](http://flightservicebureau.org/tag/france/(14.06.2018)).
- Forsyth, P., Gillen, D., Hüscherlath, K., Niemeier, H.-M., Wolf, H. (2016a): „Liberalization in Aviation: Competition, Cooperation and Public Policy“. London: Routledge.
- Forsyth, P., Gillen, D., Muller, J., Niemeier, H.-M. (2016b): „Airport Competition: The European Experience“. London, New York: Routledge.
- Forsyth, P., Gillen, D. W., Knorr, A., Mayer, O. G., Starkie, D. (2017): „The Economic Regulation of Airports: Recent Developments in Australasia, North America and Europe“. London: Routledge.
- Forsyth, P., Niemeier, H.-M. (2016): „Wider Economic Benefits or Catalytic Effects of Air Transport? How to assess airport investments?“. Berlin: Konferenz Verkehrsökonomik und -politik.
- Fried, H. O., Lovell, C. A. K., Schmidt, S. S. (2008): „The Measurement of Productive Efficiency and Productivity Growth“. Oxford: Oxford University Press.
- Fritzsche, S. (2010): „Das europäische Luftverkehrsrecht und die Liberalisierung des transatlantischen Luftverkehrsmarktes“. Berlin: Berliner Wissenschafts-Verlag.
- Gehrke, M. (2019): „Angewandte empirische Methoden in Finance & Accounting“. München: De Gruyter Oldenbourg.
- Gelhausen, M. C., Berster, P., Wilken, D. (2018): „A new direct demand model of long-term forecasting air passengers and air transport movements at German airports“. In: *Journal of Air Transport Management*, Vol. 71, pp. 140–152.
- Gelhausen, M. C., Berster, P., Wilken, D. (2019): „Airport Capacity Constraints and Strategies for Mitigation: A Global Perspective“. London: Academic Press.
- Gerdes, I., Klawonn, F., Kruse, R. (2004): „Evolutionäre Algorithmen“. Wiesbaden: Vieweg + Teubner Verlag.
- Gerdes, I., Temme, A., Schultz, M. (2016): „Dynamic Airspace Sectorization using Controller Task Load“. Delft: SESAR Innovation Days (SID).
- Gerdes, I., Temme, A., Schultz, M. (2018): „Dynamic airspace sectorisation for flight-centric operations“. In: *Transportation Research Part C: Emerging Technologies*, Vol. 95, pp. 460–480.
- Gladen, W. (2014): „Performance Measurement – Controlling mit Kennzahlen“. 6. Auflage, Wiesbaden: Gabler.
- Hafner, C. (2013): „Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility“. Berlin, Heidelberg: Springer Science & Business Media.



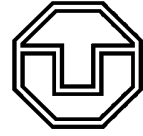
- Hagenfeldt, V. (2011): „EC Competition Law - the Essential Facilities Doctrine“. Ravensburg: GRIN Verlag.
- Hammerschmidt, M. (2006): „Effizienzanalyse im Marketing“. 1. Auflage, Wiesbaden: Deutscher Universitätsverlag.
- Haneke, U., Trahasch, S., Zimmer, M., Felden, C. (2019): „Data Science: Grundlagen, Architekturen und Anwendungen“. Heidelberg: dpunkt.verlag.
- Hasche, M., Arndt, A., Braun, T., Eichinger, A., Pansch, H., Wagner, C. (2008): „Katalytische volks- und regionalwirtschaftliche Effekte des Luftverkehrs in Deutschland“. Darmstadt: European Center for Aviation Development.
- Helios (2013): „ATM costefficiency targets: A guide for European ANSPs and NSAs“. Helios Adviser - Analysis & commentary for decision makers in aviation: <https://askhelios.s3.amazonaws.com/uploads/document/file/44/heliosadviser-issue6.pdf> (02.10.2016).
- Hergert, M., Thießen, F. (2014): „Fehlprognosen im Luftverkehr - Untersuchung zur Qualität von Luftverkehrsprognosen am Beispiel der Intraplan Consult GmbH“. Technische Universität Chemnitz, Fakultät für Wirtschaftswissenschaften.
- Hirschhausen, C. von, Nieswand, M., Wilhelm, A., Hess, B. (2007): „Wissenschaftliche Benchmarking-Methoden im ÖPNV - Methodische Ansätze und internationale Erfahrungen“. In: *Efficiency Analysis Working Papers*, Technische Universität Dresden.
- Hofert, M., Kojadinovic, I., Mächle, M., Yan, J. (2019): „Elements of Copula Modeling with R“. 1. Auflage, Cham: Springer Nature Switzerland.
- Hoff, A. (2007): „Second stage DEA: Comparison of approaches for modelling the DEA score“. In: *European Journal of Operational Research*, Vol. 181, pp. 425–435.
- Hoffmann, C. (2006): „Die Data Envelopment Analysis (DEA) und ihre Anwendungsmöglichkeiten zur vergleichenden Effizienzanalyse im Forstwesen“. Universität für Bodenkultur Wien, Department für Wirtschafts- und Sozialwissenschaften, Institut für Agrar- und Forstökonomie.
- Hsiao, C. (2014): „Analysis of Panel Data“. Cambridge University Press.
- Hutzschenreuter, T. (2015): „Allgemeine Betriebswirtschaftslehre“. 6. Auflage, Wiesbaden: Gabler Verlag.
- ICAO (2006): „DOC 9734 - PART C Oversight Manual - The Establishment and Management of a State's Aviation Security Oversight System, Appendix B“. 2. Auflage, Montreal. <https://standards.globalspec.com/std/1082535/9734%20PART%20C> (27.05.2020).
- ICAO (2016): „DOC 4444 - Air Traffic Management - Procedures for Air Navigation Services“, 16. Auflage. Montreal.
- ICAO (2018): „Cost Effectiveness (ATCO Productivity)“. Moskau: ICAO EUR Regional Performance Framework Workshop.
- Irizarry, R. A. (2019): „Introduction to Data Science: Data Analysis and Prediction Algorithms with R“. Boca Raton: CRC Press.
- Islami, A., Sun, M., Chaimatanan, S., Delahaye, D. (2017): „Optimization of military missions impact on civilian 4D trajectories“. Tokyo: ENRI International Workshop on ATM/CNS (EIWAC).



- Kaiser, H. F. (1958): „The varimax criterion for analytic rotation in factor analysis“. In: *Psychometrika*, Vol. 23, pp. 187–200.
- Kaiser, H. F. (1974): „An index of factorial simplicity“. In: *Psychometrika*, Vol. 39, pp. 31–36.
- Kaltenhäuser, S., Morlang, F., Luchkova, T., Hampe, J., Sippel, M. (2017): „Facilitating Sustainable Commercial Space Transportation Through an Efficient Integration into Air Traffic Management“. In: *New Space*, Vol. 5 (4), pp. 244–256.
- Keitel, S. (2015): „Auswirkungen disaggregierter Regulierung auf die Kapazität von Verkehrssektoren: Eine wettbewerbstheoretische und empirische Untersuchung am Beispiel des Luftverkehrssektors“. Berlin: Universitätsverlag der TU Berlin.
- Kern, S. (2019): „Analyse kapazitätssteigernder Maßnahmen für generische Start- und Landebahnsysteme“. TU Braunschweig.
- Kerner, B. S. (1998): „Experimental Features of Self-Organization in Traffic Flow“. In: *Physical Review Letters*, Vol. 81 (17), pp. 3797–3800.
- Kjærulff, U. B., Madsen, A. L. (2012): „Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis“. 1. Auflage, New York: Springer.
- Knieps, G. (2008): „Wettbewerbsökonomie: Regulierungstheorie, Industrieökonomie, Wettbewerbspolitik“. 3. Auflage, Berlin, Heidelberg: Springer-Verlag.
- Korn, B., Edinger, C., Tittel, S., Pütze, T., Mohrhard, B. (2010): „Sectorless ATM - Analysis and Simulation Results“. Nice: International Council of Aeronautical Sciences Conference (ICAS).
- Kreuz, M. (2015): „Modellierung von Flugsicherungsprozessen auf Basis von System Dynamics“. Braunschweig: TU Braunschweig.
- Krishnapuram, R., Keller, J. M. (1993): „A possibilistic approach to clustering“. In: *IEEE Transactions on Fuzzy Systems*, Vol. 1 (2), pp. 98–110.
- Kumbhakar, S. C., Lovell, C. A. K. (2003): „Stochastic Frontier Analysis“. Cambridge: Cambridge University Press.
- Kumbhakar, S. C., Wang, H.-J., Horncastle, A. P. (2015): „A Practitioner's Guide to Stochastic Frontier Analysis Using Stata“. Cambridge: Cambridge University Press.
- Last, A.-K., Wetzel, H. (2009): „Effizienzmessverfahren – Eine Einführung“. In: *Working Paper Series in Economics, University of Lüneburg*, University of Lüneburg Vol. 145.
- Laursen, T. (2018): „Volatility in air traffic and its impact on ATM performance“. Warsaw: FABEC Workshop on Volatility in air traffic and its impact on ATM Performance.
- LeSage, J., Pace, R. K. (2009): „Introduction to Spatial Econometrics“. Boca Raton: CRC Press.
- Liebert, V. P. (2011): „Airport Benchmarking: An Efficiency Analysis of European Airports from an Economic and Managerial Perspective“. Jacobs University Bremen.
- Lo, C. P., Yeung, A. K. W. (2006): „Concepts and Techniques of Geographic Information Systems“. 2. Auflage, University of Michigan: Pearson Prentice Hall.
- Lovell, C. A. K. (1992): „Production Frontiers and Production Efficiencies“. In: *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford: Oxford University Press.



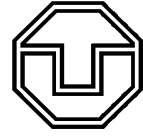
- Luchkova, T., Morlang, F., Kaltenhäuser, S. (2016): „Air Traffic Impact Analysis Design for a Suborbital Point-to-Point Passenger Transport Concept“. Daytona Beach: Space Traffic Management Conference (STM).
- Luchkova, T., Vujasinovic, R., Lau, A., Schultz, M. (2015): „Analysis of Impacts an Eruption of Volcano Stromboli could have on European Air Traffic“. Lisbon: ATM Seminar.
- Lütjens, K., Radde, M., Liedtke, G., Maertens, S., Standfuß, T., Scheier, B., Viergutz, K. (2018): „Innovationen im Zuge der Digitalisierung des Personenverkehrs“. In: *Wirtschaftsdienst*, Vol. 98 (7), pp. 512–518.
- Majumdar, A., Ochieng, W. Y. (2007): „Air Traffic Control Complexity and Safety: Framework for Sector Design Based on Controller Interviews of Complexity Factors“. In: *Transportation Research Record Journal of the Transportation Research Board*, Vol. 1, pp. 70–80.
- Meeusen, W., Broeck, J. van den (1977): „Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error“. In: *International Economic Review*, Vol. 18 (2), pp. 435–444.
- Mensen, H. (2014): „Moderne Flugsicherung: Organisation, Verfahren, Technik“. 4. Auflage, Berlin, Heidelberg: Springer.
- Michalewicz, Z. (1996): „Genetic Algorithms + Data Structures = Evolution Programs“. 3. Auflage, Berlin, Heidelberg: Springer.
- Müller, A. C., Guido, S. (2017): „Einführung in Machine Learning mit Python: Praxiswissen Data Science“. Sebastopol: O'Reilly.
- Neiva, R. (2014): „Economic Efficiency of Air Navigation Service Providers: An Assessment in Europe“. Fairfax: George Mason University.
- NERA (2006): „Cost Benchmarking of Air Navigation Service Providers: A Stochastic Frontier Analysis“. London.
- Niemeier, H.-M. (2001): „On the use and abuse of Impact Analysis for airports: A critical view from the perspective of regional policy“. In: *Regional input/output analysis*, Baden Baden: Nomos Verlag.
- Nuic, A., Poles, D., Mouillet, V. (2010): „BADA: An advanced aircraft performance model for present and future ATM systems“. In: *International Journal of Adaptive Control and Signal Processing*, Vol. 24, pp. 850–866.
- Nyhan, R. C., Martin, L. L. (1999): „Comparative Performance Measurement – A Primer on Data Envelopment Analysis“. In: *Public Productivity & Management Review*, Vol. 22, pp. 348–364.
- Ong, L. L. (2003): „Purchasing Power Parity: A Survey of the Issues“. In: London: Palgrave Macmillan pp. 1–28.
- Oster, C. V., Strong, J. S. (2017): „Managing the Skies: Public Policy, Organization and Financing of Air Traffic Management“. New York: Routledge.
- Pascoe, S., Kirkley, J. E., Gréboval, D., Morrison-Paul, C. J. (2003): „Measuring and assessing capacity in fisheries 2. Issues and methods“. In: *FAO Fisheries Technical Paper*, Vol. 433 (2).
- Pfähler, W., Niemeier, H.-M., Mayer, O. G. (1999): „Airports and Air Traffic: Regulation, Privatisation, and Competition“. New York: Peter Lang GmbH, Internationaler Verlag der Wissenschaften.
- Pflaumer, P., Heine, B., Hartung, J. (2017): „Statistik für Wirtschafts- und Sozialwissenschaften: Deskriptive Statistik“. 4. Auflage, Berlin: Walter de Gruyter GmbH & Co KG.



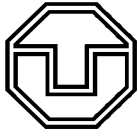
- Pompl, W. (2013): „Luftverkehr: Eine ökonomische und politische Einführung“. 3. Auflage, Berlin, Heidelberg: Springer.
- Rhoades, S. A. (1993): „The Herfindahl-Hirschman index“. In: *Federal Reserve Bulletin*, pp. 188–189.
- Rosenow, J., Förster, S., Lindner, M., Fricke, H. (2017a): „Impact of Multi-criteria Optimized Trajectories on European Air Traffic Density, Efficiency and the Environment“. Seattle: ATM Seminar.
- Rosenow, J., Fricke, H., Luchkova, T., Schultz, M. (2018a): „Minimizing contrail formation by rerouting around dynamic ice-supersaturated regions“. In: *Aeronautics and Aerospace Open Access Journal*, Vol. 2 (3), pp. 105–111.
- Rosenow, J., Fricke, H., Schultz, M. (2017b): „Air Traffic Simulation with 4D multi-criteria optimized trajectories“. In: *Proceedings of the Winter Simulation Conference*, pp. 2589–2600.
- Rosenow, J., Lindner, M., Fricke, H. (2015): „Assessment of air traffic networks considering multi-criteria targets in network and trajectory optimization“. Rostock: Deutscher Luft- und Raumfahrtkongress.
- Rosenow, J., Schultz, M. (2018): „Coupling of turnaround and trajectory optimization in an air traffic simulation based on delay costs“. Gothenburg: Winter Simulation Conference (WSC).
- Rosenow, J., Strunck, D., Fricke, H. (2018b): „Free Route Airspaces in Functional Air Space Blocks“. Salzburg: SESAR Innovation Days (SID).
- Salaün, E., Gariel, M., Vela, A., Feron, E., Clarke, J.-P. (2010): „Airspace Complexity Estimations Based on Data-Driven Flow Modeling“. Anaheim: AIAA Guidance, Navigation, and Control Conference.
- Schelbert, K., Vogel, M., Thiel, C., Fricke, H. (2012): „Adapting Enroute ATM Complexity Metrics for Terminal Airspace Safety Assessments - Factor and correlation analyses based on radar data of a European hub airport“. Berkeley: International Conference for Research in Air Transportation (ICRAT).
- Schendera, C. F. (2010): „Clusteranalyse mit SPSS: Mit Faktorenanalyse“. München: Oldenbourg Verlag.
- Schladebach, M. (2007): „Luftrecht“. Tübingen: Mohr Siebeck Verlag.
- Schultz, M., Gerdes, I., Standfuss, T., Temme, A. (2017): „Future airspace design by dynamic sectorization“. Tokyo: ENRI International Workshop on ATM/CNS (EIWAC).
- Schultz, M., Reitmann, S. (2019): „Machine learning approach to predict aircraft boarding“. In: *Transportation Research Part C Emerging Technologies*, Vol. 98, pp. 391–408.
- Schultz, M., Standfuss, T., Gerdes, I., Temme, A. (2018a): „Efficiency gains through Flight Centered Air Traffic Management Approach - How to meet future needs with a dynamic airspace sectoring“. Warsaw: FABEC Workshop on Volatility in air traffic and its impact on ATM Performance.
- Schultz, M., Temme, A., Kügler, D. (2018b): „Unbemannter Frachttransport im Luftverkehrssystem“. In: *Internationales Verkehrswesen*, Vol. 1, pp. 40–42.
- Schulze, P. M. (1995): „Zur Messung räumlicher Autokorrelation“. In: *Review of regional research*, Vol. 14, pp. 57–78.
- Schwedes, O. (2018): „Verkehrspolitik: Eine interdisziplinäre Einführung“. 2. Auflage, Wiesbaden: Springer-Verlag.



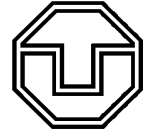
- Schwenk, W., Schwenk, R. (1998): „Aspects of International Cooperation in Air Traffic Management“. Alphen aan den Rijn: Kluwer Law International B.V.
- Sengupta, J. K. (1982): „Efficiency measurement in stochastic input-output systems“. In: *International Journal of Systems Science*, Vol. 13, pp. 273–285.
- Sergeeva, M., Delahaye, D., Zerrouki, L., Schede, N. (2015): „Dynamic Airspace Configuration Generated by Evolutionary Algorithms“. Prague: Digital Avionics Systems Conference (DASC).
- Sherali, H. D., Hill, J.M. (2011): „Configuration of airspace sectors for balancing air traffic controller workload“. In: *Annals of Operations Research*, Vol. 203, pp. 3–31.
- Simar, L., Wilson, P. W. (2007): „Estimation and inference in two-stage, semi-parametric models of production processes“. In: *Journal of Econometrics*, Vol. 136, pp. 31–64.
- Snedecor, G. W., Cochran, W. G. (1989): „Statistical Methods“. Ames: Iowa State University Press.
- Spaho, A. (2015): „Determinants of Banks' Efficiency in Albania - Data Envelopment Analysis and Tobit Regression Analysis“. Istanbul: International Conference on Social Sciences (ICOSS).
- Springer (2014): „Kompakt-Lexikon Wirtschaftsmathematik und Statistik“. Wiesbaden: Springer Fachmedien.
- Sridhar, B., Sheth, K. S., Grabbe, S. (1998): „Airspace Complexity and its Application in AirTraffic Management“. Orlando: ATM Seminar.
- Standfuss, T. (2020): „Complexity Metrics in ANSP Benchmarking – Applicability and potential improvements of the PRU Metric“. ASB Information Paper. Langen: FABEC Performance Management Group.
- Standfuss, T., Czech, C., Hellbach, T. (2016a): „PRU Data Analysis“. FABEC Performance Management Group Benchmarking Taskforce.
- Standfuss, T., Deltuvaite, V., Whittome, M., Fichert, F. (2020): „Performance Target Setting for Air Traffic Management - Opportunities and Limitations of an Academic Assessment in a Complex Industry“. Rome: FABEC Research Workshop Interdependencies within ATM Performance in the context of a dynamic environment (Akzeptiert, noch nicht veröffentlicht).
- Standfuss, T., Fichert, F., Schultz, M. (2017a): „Air Traffic Management Performance Benchmarking - A disaggregated approach“. Tokyo: ENRI International Workshop on ATM/CNS (EIWAC).
- Standfuss, T., Fichert, F., Schultz, M. (2017b): „Efficiency gains through Functional Airspace Blocks? An analysis of Economies of Scale in European Air Traffic Management“. Barcelona: International Transportation Economics Association Conference (ITEA).
- Standfuss, T., Fichert, F., Schultz, M. (2018a): „From static sectorization to dynamic airspace structures - Operational and economic benefits of flight centered air traffic management“. Hongkong: International Transportation Economics Association Conference (ITEA).
- Standfuss, T., Fichert, F., Schultz, M. (2018b): „Input and Output measurement in Air Navigation Service Provider Performance Benchmarking - Implementing composite indicators for efficiency analysis using European data“. Seoul: Air Transport Research Society Conference (ATRS).
- Standfuss, T., Fichert, F., Schultz, M., Stratis, P. (2019a): „Efficiency losses through Fragmentation? Scale effects in European ANS Provision“. In: *Competition and Regulation in Network Industries*, Vol. 20 (4), pp. 275–289.



- Standfuss, T., Fichert, F., Whittome, M. (2018c): „Adapting Capacity of Air Navigation Service Provision in Europe - Between scylla and charybdis“. Belo Horizonte: International Council of Aeronautical Sciences Conference (ICAS).
- Standfuss, T., Gerdes, I., Temme, A., Schultz, M. (2016b): „Dynamische Luftraumoptimierung“. Braunschweig: Deutscher Luft- und Raumfahrtkongress (DLRK).
- Standfuss, T., Gerdes, I., Temme, A., Schultz, M. (2018d): „Dynamic Airspace Optimisation“. In: *CEAS Aeronautical Journal*, Vol. 9 (3), pp. 517–531.
- Standfuss, T., Rosenow, J. (2020): „Applicability of Current Complexity Metrics in ATM Performance Benchmarking and Potential Benefits of Considering Weather Conditions“. San Antonio: Digital Avionics Systems Conference (DASC).
- Standfuss, T., Schultz, M. (2018): „Performance Assessment of European Air Navigation Service Providers“. London: Digital Avionics Systems Conference (DASC).
- Standfuss, T., Whittome, M. (2018): „Benchmarking EU / US – Operational Heterogeneities“. ASB Information Paper. Langen: FABEC Performance Management Group.
- Standfuss, T., Whittome, M. (2019a): „Benchmarking EU / US – Operational Heterogeneities, Update 2019“. ASB Information Paper. Langen: FABEC Performance Management Group.
- Standfuss, T., Whittome, M. (2019b): „Forecast Quality of STATFOR Predictions“. ASB Information Paper. Langen: FABEC Performance Management Group.
- Standfuss, T., Whittome, M., Ruiz-Gauna, I., Knabe, F. (2019b): „Volatility in Air Traffic Management – How changes in traffic patterns affect efficiency in service provision“. Tokyo: ENRI International Workshop on ATM/CNS (EIWAC).
- Statista (2020): „Europäische Union & Euro-Zone: Wachstum des realen Bruttoinlandsprodukts (BIP) von 2009 bis 2019“. <https://de.statista.com/statistik/daten/studie/156282/umfrage/entwicklung-des-bruttoinlandsprodukts-bip-in-der-eu-und-der-eurozone/> (27.05.2020).
- Stepan, A., Fischer, E. O. (2014): „Betriebswirtschaftliche Optimierung: Einführung in die quantitative Betriebswirtschaftslehre“. 8. Auflage, Berlin: Walter de Gruyter GmbH & Co KG.
- Sunil, E., Hoekstra, J., Ellerbroek, J., Bussink, F., Nieuwenhuisen, D., Vidosavljevic, A., Kern, S. (2015): „Metropolis: Relating Airspace Structure and Capacity for Extreme Traffic Densities“. Lisbon: ATM Seminar.
- Temme, A. (2002): „Objective Function based Fuzzy Clustering in Air Traffic Management“. Magdeburg: Otto-von-Guericke Universität.
- Temme, A. (2016): „Unmanned Freight Operations“. Braunschweig: Deutscher Luft- und Raumfahrtkongress (DLRK).
- Tone, K. (2001): „A slacks-based measure of efficiency in data envelopment analysis“. In: *European Journal of Operational Research*, Vol. 130, pp. 498–502.
- Urban, D., Mayerl, J. (2018): „Angewandte Regressionsanalyse: Theorie, Technik und Praxis“. 5. Auflage, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Varian, H. R. (2016): „Grundzüge der Mikroökonomik“. 9. Auflage, Berlin: Walter de Gruyter GmbH & Co KG.



- Vasigh, B., Tacker, T., Fleming, K. (2008): „Introduction to Air Transport Economics: From Theory to Applications“. Aldershot, Burlington: Ashgate Publishing.
- Verbeek, R. J. D., Visser, H. G. (2016): „Why aircraft will fly more fuel-efficiently on FRIDAY - The FRIDAY route charges method“. Philadelphia: International Conference for Research in Air Transportation (ICRAT).
- Vlinen, T., Nissi, J., Ala-Risku, T., Smaros, J. (2017): „Measuring Forecast Accuracy“. Helsinki: RELEX. <https://www.relexsolutions.com/resources/measuring-forecast-accuracy/> (23.09.2019).
- Welc, J., Esquerdo, P. J. R. (2017): „Applied Regression Analysis for Business: Tools, Traps and Applications“. Cham: Springer International Publishing.
- Whittome, M., Standfuss, T. (2018): „Infrastructure Financing of European Airspaces - Computation of US Passenger Taxes on Europe“. Studie für die Inter-FAB Performance Initiative: Langen.
- Wooldridge, J. M. (2002): „Econometric Analysis of Cross Section and Panel Data“. Cambridge: MIT Press.
- Worldbank (2019): „World Bank Open Data - Free and open access to global development data“. <https://data.worldbank.org/> (20.04.2019).
- Xu, R., Wunsch, D. (2009): „Clustering“. Hoboken: John Wiley & Sons (IEEE Press Series on Computational Intelligence).
- Yamagata, Y., Seya, H. (2019): „Spatial Analysis Using Big Data: Methods and Urban Applications“. Cambridge: Academic Press.
- Zhang, T., Garvey, E. (2008): „A comparative analysis of multi-output frontier models“. In: *Journal of Zhejiang University SCIENCE A*, Vol. 9, pp. 1426–1436.
- Zhu, J. (2014): „Quantitative Models for Performance Evaluation and Benchmarking“. New York: Springer.



Danksagung

Mein Interesse am Luftverkehr wurde bereits in Kindertagen geweckt. Seitdem sind Luftfahrtspezifische Themen ein fester Bestandteil meines Lebens, so auch im beruflichen Werdegang. Mit der vorliegenden Dissertation wird ein wesentlicher Abschnitt abgeschlossen und ein lang verfolgtes Ziel erreicht. Daher möchte ich mich an dieser Stelle bei allen Personen bedanken, die mich auf diesem Weg begleitet haben.

Mein erster Dank gilt meinem Doktorvater, Prof. Georg Hirte, für die Betreuung der Doktorarbeit. Seine Expertise war jederzeit eine große Unterstützung, um die angewendeten Methoden sowie erzielten Ergebnisse zu diskutieren. Zudem bin ich für die problemlose Übernahme der Betreuung in einem fortgeschrittenen Status meiner Dissertation sehr dankbar. Die gemeinsamen Konsultationen ermöglichten eine stetige Verbesserung der Dissertationsschrift hinsichtlich Struktur und Inhalt. Ein großer Dank gilt auch meinem Zweitgutachter, Prof. Frank Fichert. Der intensive wissenschaftliche Austausch über Forschungsfragen sowie die gemeinsame Einwicklung von Ideen ermöglichten sowohl alternative Denkansätze als auch eine Vielzahl an daraus resultierenden Publikationen. Hier bin ich sehr dankbar dafür, dass trotz seiner gelegentlichen Kopfschmerzen über meine englischen Formulierungen jederzeit die fachliche Unterstützung garantiert war. Ich bedanke mich zudem für die Unterstützung von Prof. Bernhard Wieland sowie Prof. Dirk Kügler, welche die Dissertation ermöglichten und in einem früheren Stadium begleitet haben.

Ein großes Dankeschön geht an meinen fachlichen Betreuer, PD Dr. Michael Schultz. Er konnte mich jederzeit, jedoch insbesondere in schwierigen Phasen, dazu motivieren, weiter am Ball zu bleiben und hat damit einen wesentlichen Anteil am Abschluss des Dissertationsvorhabens. Sowohl seine fachliche Expertise als auch den zwischenmenschlichen Umgang empfand ich stets als große Unterstützung.

Ich bedanke mich weiterhin bei allen Personen, die mir den Weg zur Dissertation geebnet haben. Zunächst gilt mein Dank Dr. Egbert Gärtner, der mich zu diesem Vorhaben während meiner Werkstudententätigkeit bei EADS motivierte. Des Weiteren bedanke ich mich bei Prof. Christos Evangelinos, der einen wesentlichen Einfluss auf die Themenfindung der vorliegenden Doktorarbeit hatte. Die Umsetzung des Dissertationsvorhabens wurde außerdem durch die Deutsche Flugsicherung unterstützt. Daher danke ich insbesondere Dr. Matthias Whittome und Christoph Czech für ihren organisatorischen und fachlichen Support sowie die Vermittlung an das Deutsche Zentrum für Luft und Raumfahrt.

Dankbar bin ich für die wissenschaftliche Unterstützung meiner Kollegen vom DLR, insbesondere Dr. Annette Temme und Franz Knabe sowie meinen Doktoranden-Kollegen Dr. Michael Kreuz, Dr. Stefan Kern, Stefan Reitmann, Alexander Gillissen und Max Friedrich. Auch danke ich meinen Kollegen an der TU Dresden. Dabei möchte ich insbesondere meinem Institutsleiter Prof. Hartmut Fricke sowohl für die reibungslose organisatorische und thematische Integration ins Team als auch für seine fachliche Expertise danken.



Ich möchte an dieser Stelle auch die operative Expertise der konsultierten Flugsicherungsunternehmen sowie den hervorragenden, inhaltlichen Austausch mit der DFS sowie FABEC herausstellen. Dabei gilt mein besonderer Dank Thomas Hellbach, Juan Espinar Nova und Jörg Buxbaum sowie allen Mitgliedern der Performance Management Group.

Der Abschluss der Dissertation wäre ohne den Rückhalt meiner Familie und Freunde nicht möglich gewesen. Daher bedanke ich mich vor allem bei meiner Lebensgefährtin Nina Fillmann sowie meinen Eltern für ihr Verständnis und die bedingungslose Unterstützung während meiner Zeit als Doktorand. Zuletzt danke ich meinen engsten Freunden Luise Lässig, Katrin Schönbach, Rene Dauven und Marcel Garon, die mir auch in den schweren Tagen immer zur Seite standen.