

Anhang zu:

**Generelle Methodik
zur Abschnittsbildung
bei der Erhaltung
kommunaler Straßen**

von

Dietmar Berthold
Andreas Großmann

LEHMANN + PARTNER GmbH
Erfurt

**Berichte der
Bundesanstalt für Straßenwesen**

Straßenbau Heft S 184

FoPS | Verbesserung der
Verkehrsverhältnisse
der Gemeinden
FORSCHUNGSPROGRAMM STADTVERKEHR

bast

Anhang

- Anhang 1: Vergleich der Segmentierungsergebnisse mit durchgeführten bzw. geplanten Maßnahmen der Stadt Düsseldorf
- Anhang 2: Vergleich der Segmentierungsergebnisse mit vorhandenen Erhaltungsabschnitten der Stadt Hamburg
- Anhang 3: Beschreibung der Algorithmen

Anlage 1

Vergleich der Segmentierungsergebnisse
mit durchgeführten bzw. geplanten Maßnahmen
der Stadt Düsseldorf

Das in Anlage 1 verwendete Kartenmaterial wurde im Rahmen der Prüfung der Segmentierungsergebnisse von der Stadt Düsseldorf unter der Datenlizenz Deutschland – Land NRW/Vermessungs- und Katasteramt Düsseldorf bereitgestellt (www.govdata.de/dl-de/by-2-0).

In allen folgenden Ausschnitten stellen die schwarzen Balken durchgeführte oder geplante Erhaltungsmaßnahmen der Stadt Düsseldorf und die farbigen Linien die vom Verfahren gebildeten Segmente dar. Die Farbgebung entspricht der üblichen Farbe der Zustandsklassen (vgl. Kapitel 5.8). Es wurde die integrierte Gesamtsegmentierung nach Gebrauchswert und Substanzwert durchgeführt.

Beispiele für optimale Segmentierungsergebnisse

Optimal bedeutet, dass die gebildeten Segmente den Erhaltungsabschnitten der Stadt Düsseldorf komplett oder weitgehend entsprechen. Auf diesen Abschnitten wurden entweder in der Vergangenheit – d.h. seit der zu Grunde liegenden Zustandserfassung und -bewertung in den Jahren 2016/17– bereits Erhaltungsmaßnahmen durchgeführt oder es sind Erhaltungsmaßnahmen in den Folgejahren geplant.



Bild 0.1: Antoniusstraße, Maßnahme in 2017 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

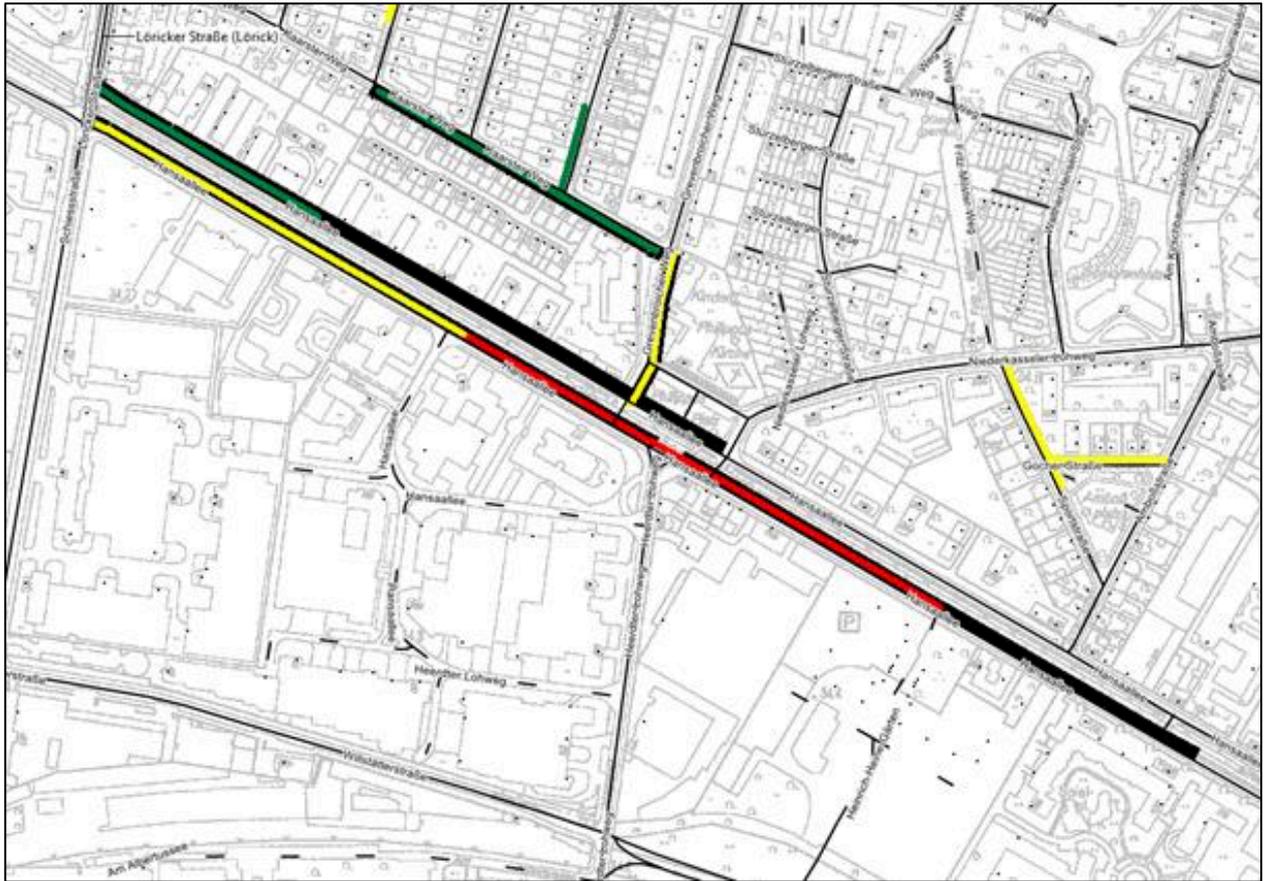


Bild 0.4: Hansaallee, Maßnahme in 2018-2019 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)



Bild 0.5: Drakestraße, Maßnahme in 2019 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

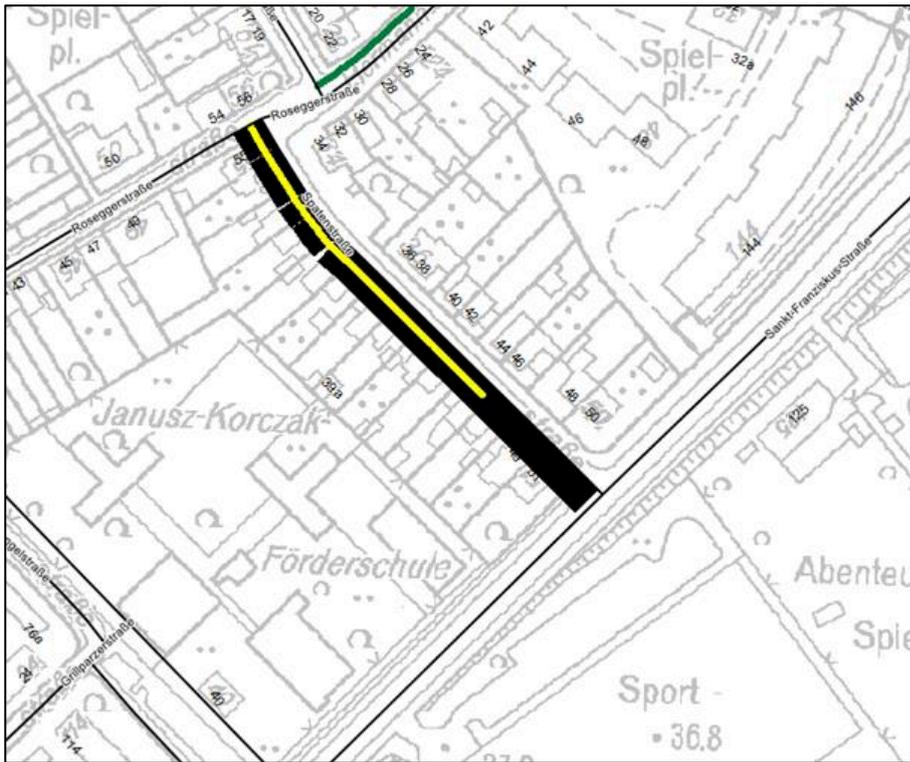


Bild 0.6: Spatenstraße, Maßnahme in 2018 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

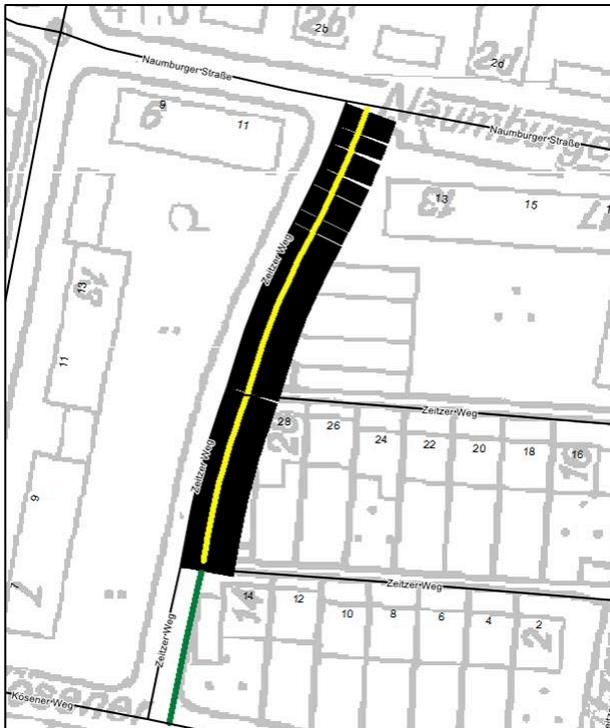


Bild 0.7: Zeitzer Weg, Maßnahme geplant in 2023 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)



Bild 0.8: An den Birken, Maßnahme geplant in 2025 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

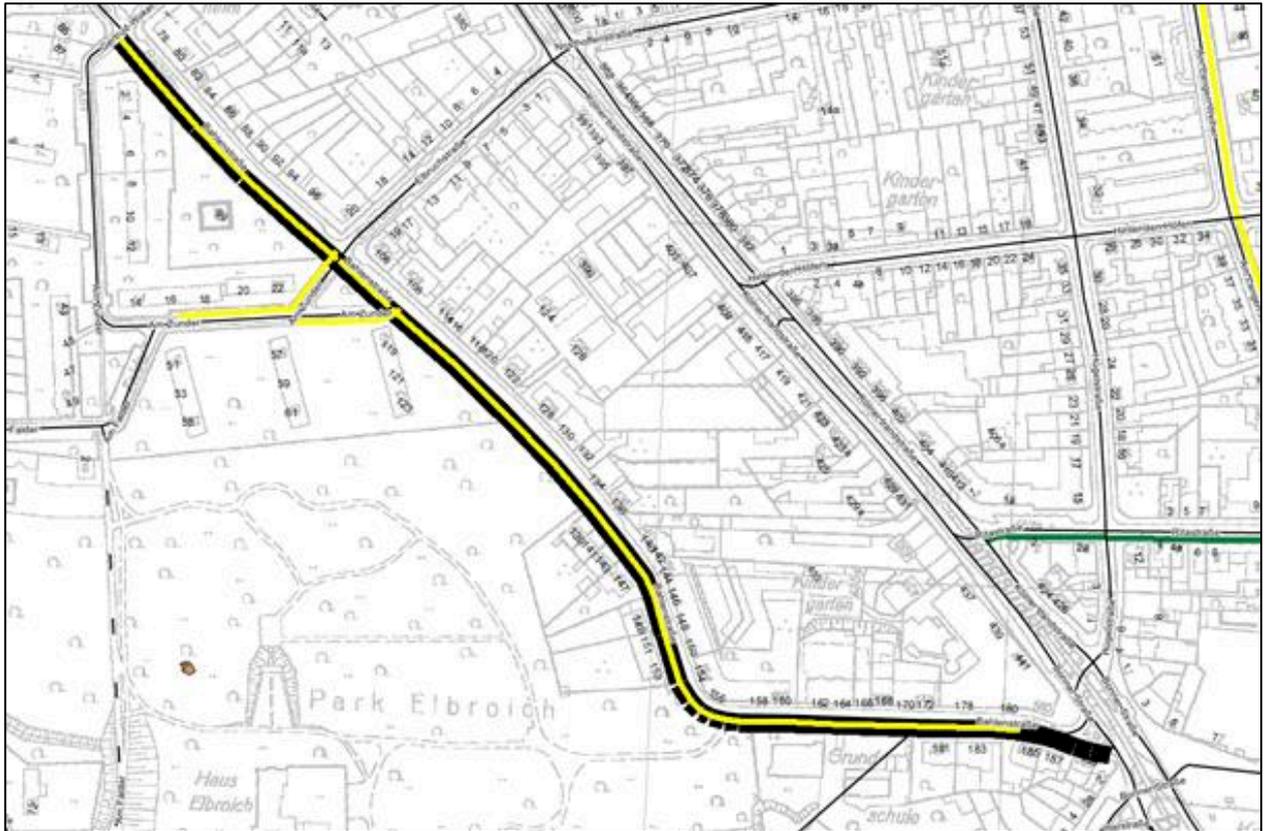


Bild 0.9: Bahlenstraße, Maßnahme geplant in 2023 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)



Bild 0.10: Harffstraße 2. BA, Maßnahme geplant in 2023 (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

Beispiele für abweichende Segmentierungsergebnisse

In diesen Fällen wurden durch das Verfahren Segmente gebildet, die deutlich von den Erhaltungsabschnitten und durchgeführten bzw. geplanten Maßnahmen der Stadt Düsseldorf abweichen. In einigen Fällen wurden die Changepoints nicht an den erwarteten Stellen oder gar nicht gesetzt. Ein Grund kann in der im Verfahren gesetzten Mindestsegmentlänge von 80 m bestehen; das Ergebnis gibt aber auch Hinweise auf weiteres Optimierungspotenzial des Segmentierungsverfahrens.

Andererseits gibt es Fälle, wo die Stadt eine Maßnahme über eine deutlich längere Strecke durchgeführt hat, z.B. Einsiedelstraße, wobei es dafür auch Gründe geben kann, die nicht durch den Zustand bedingt sind (z.B. Kanalbaumaßnahme o.ä.).



Bild 0.11: Gandersheimer Straße, links von Süd nach Nord verlaufend (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

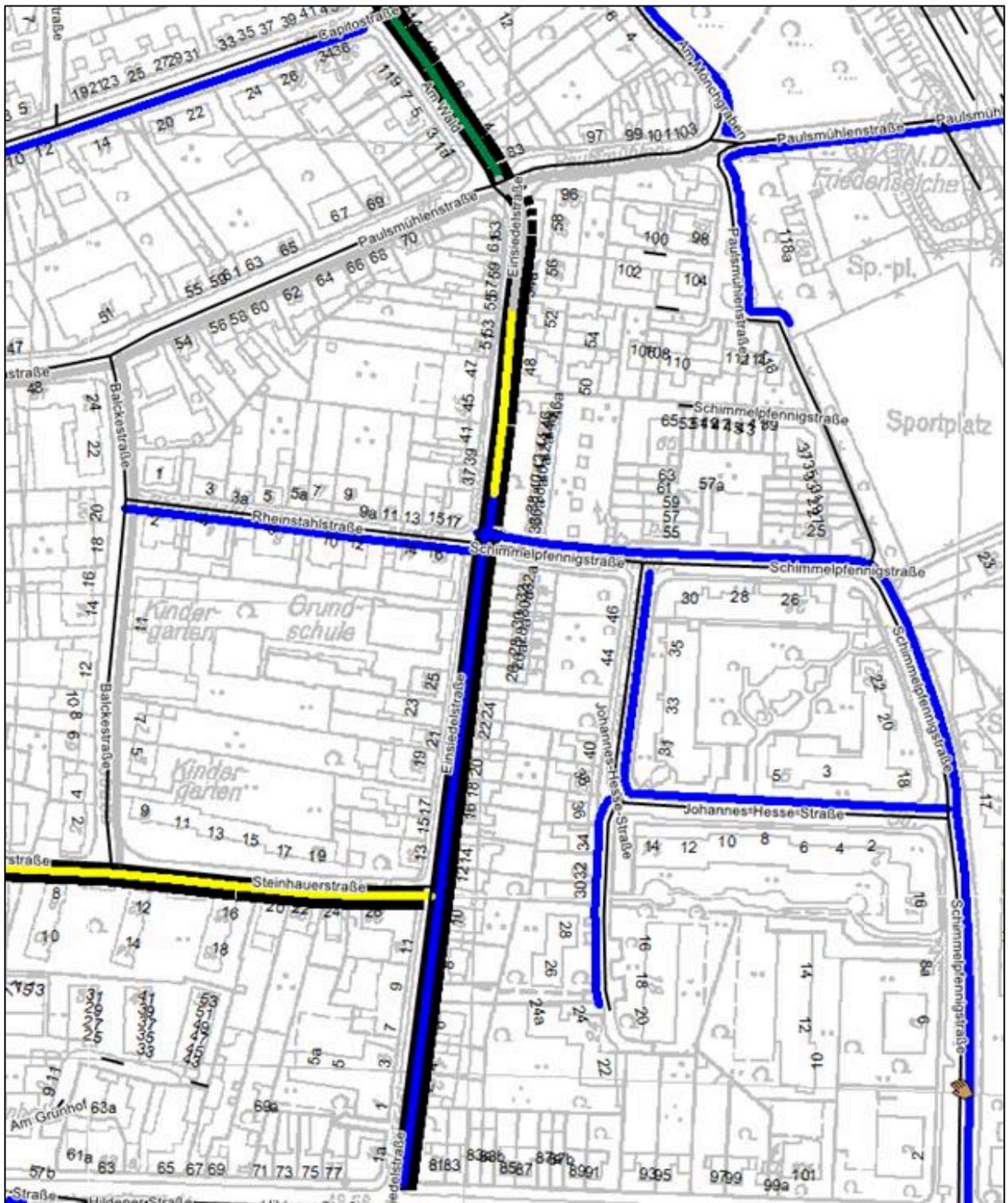


Bild 0.12: Einsiedelstraße, mittig von Süd nach Nord verlaufend (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

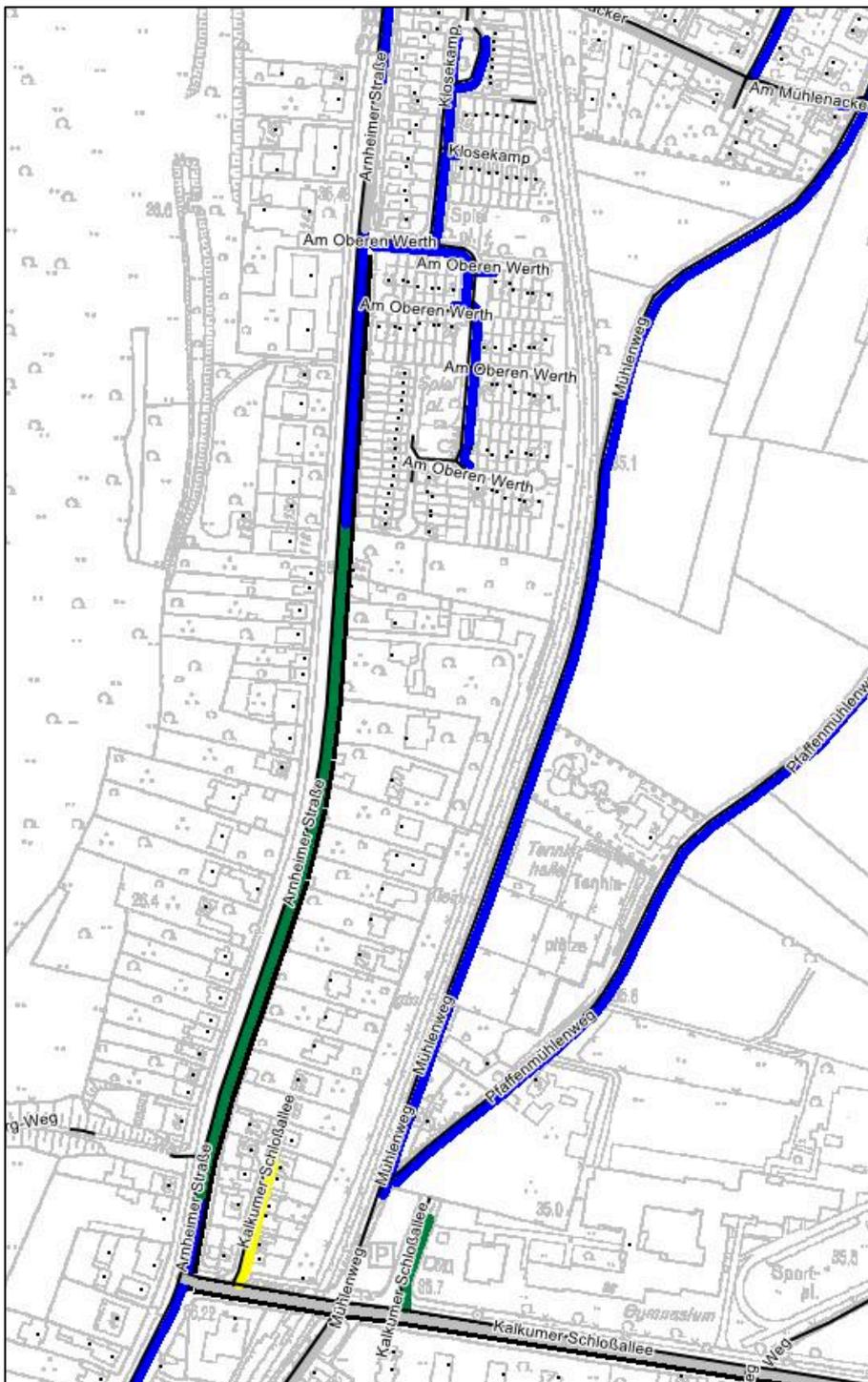


Bild 0.13: Anheimer Straße, links von Süd nach Nord verlaufend (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

Beispiele für neue Erkenntnisse aus dem Segmentierungsergebnis

Schließlich gab es noch Fälle, wo das Segmentierungsergebnis Erkenntnisse über erhaltungsbedürftige Strecken geliefert hat, die noch nicht als Erhaltungsabschnitte definiert waren.

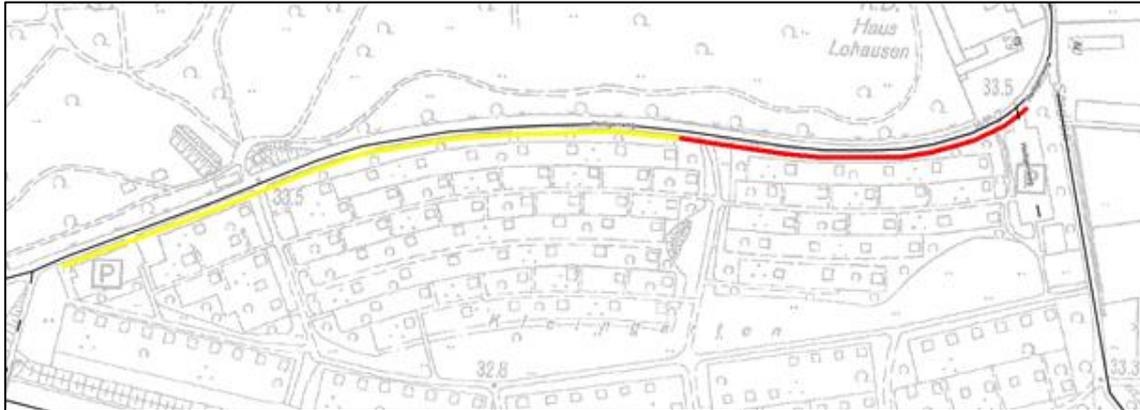


Bild 0.14: Heiligenweg (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)



Bild 0.15: Eulerstraße (© Quelle: Stadt Düsseldorf, dl-de/by-2-0)

Anlage 2

Vergleich der Segmentierungsergebnisse
mit vorhandenen Erhaltungsabschnitten
der Stadt Hamburg

Das in Anlage 2 verwendete Kartenmaterial wurde im Rahmen der Prüfung der Segmentierungsergebnisse von der Stadt Hamburg unter der Datenlizenz Deutschland - Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV), bereitgestellt (www.govdata.de/dl-de/by-2-0).

In den Kartenausschnitten stellen die hellroten Bänder die in der Stadt Hamburg vorhandenen Erhaltungsabschnitte dar. Diese waren auf Basis der Ergebnisse der ZEB 2014 im Rahmen eines anderen Projekts generiert worden.

Die farbigen Linien sind die vom hier untersuchten Verfahren gebildeten Segmente, jeweils beschriftet mit dem Mittelwert des GEB und des SUB. Die Farbgebung entspricht der üblichen Farbe der Zustandsklassen (vgl. Kapitel 5.8). Es wurde die integrierte Gesamtsegmentierung nach Gebrauchswert und Substanzwert durchgeführt.

Beispiel 1:



Bild 0.1: Maria-Louisen-Straße (© Quelle: Stadt Hamburg, dl-de/by-2-0)

Der vorhandene Erhaltungsabschnitt auf Fahrstreifen R1 ist 500 m lang und damit kürzer als das automatisch gebildete Segment, das sich hier über den gesamten Netzknotenabschnitt (638 m lang) erstreckt. Der Erhaltungsabschnitt hat einen mittleren SUB von 4,82. Das längere, automatisch gebildete Segment hat einen mittleren SUB von 4,36.

Dies ist eine direkte Folge der gewählten Mindestsegmentlänge $L_{\text{Min}} = 200$ m. Das zeigt der QGIS-Screenshot des Segmentierungsverfahrens in Bild 0.2. Die Auswerteabschnitte am Anfang des Abschnitts mit vorwiegend gutem SUB weisen eine Länge von 180 m auf. Diese können wegen der Mindestsegmentlänge kein separates Segment bilden und werden daher in das benachbarte Segment integriert. Dadurch vermindert sich dessen Mittelwert auf 4,36, was auch Auswirkung auf die Zuordnung zur Zustandsklasse hat.

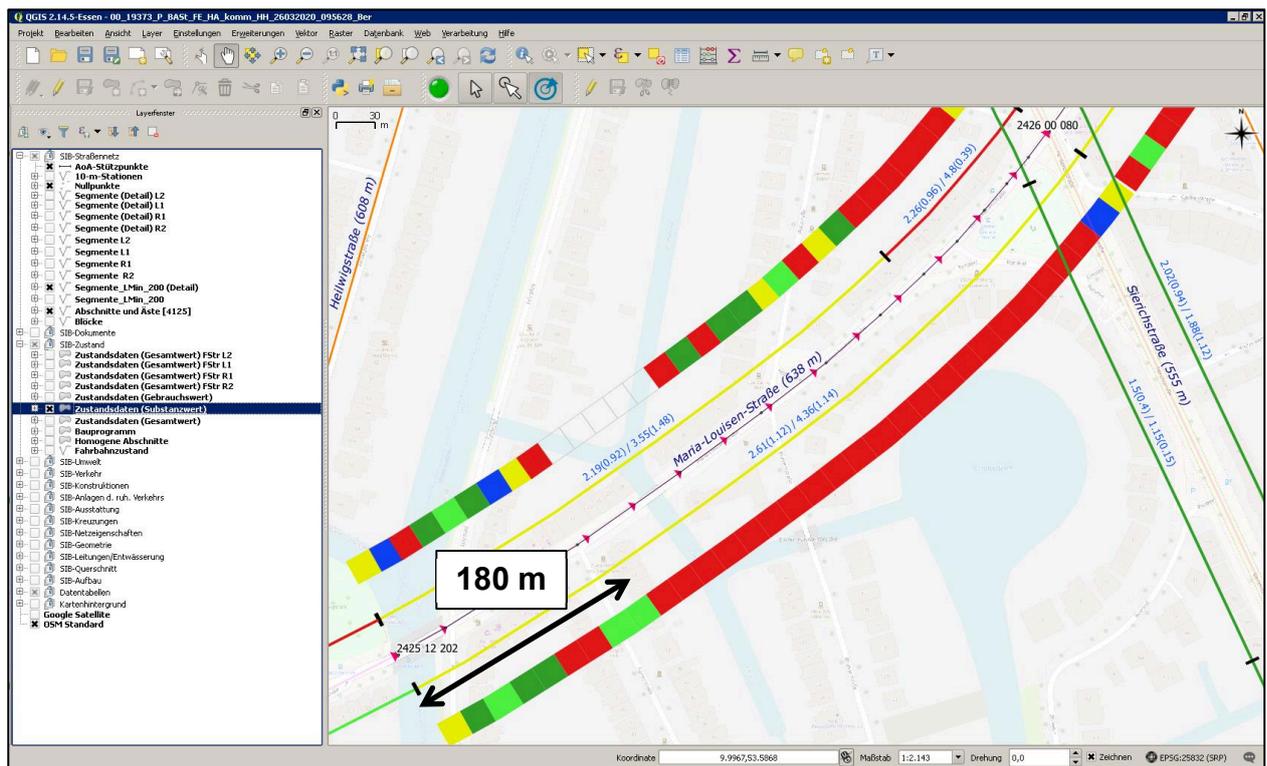


Bild 0.2: Maria-Louisen-Straße, Segmentierungsverfahren (Quelle: © OpenStreetMap und Mitwirkende, CC-BY-SA)

Beispiel 2:

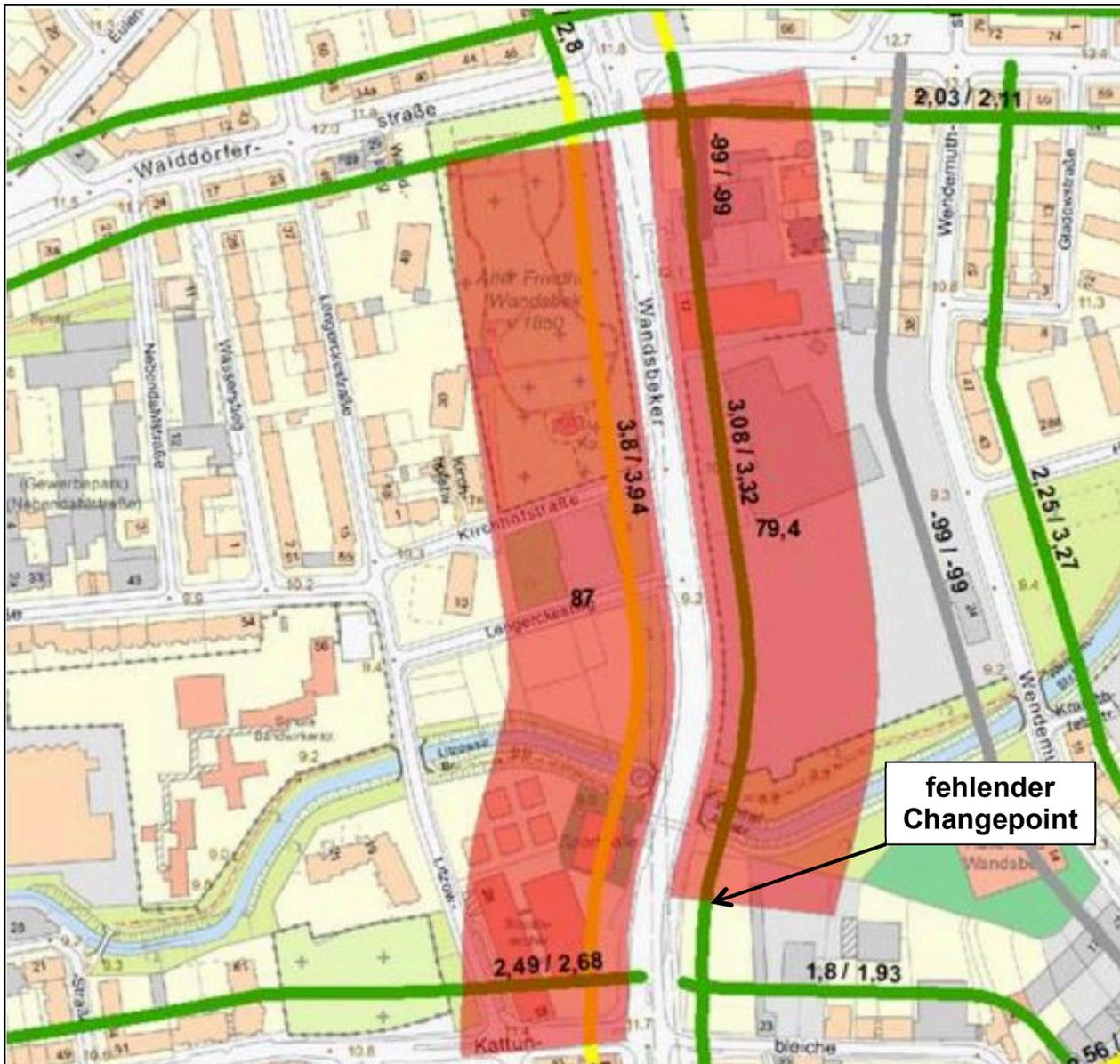


Bild 0.3: Wandsbeker Allee (Quelle: © Stadt Hamburg, dl-de/by-2-0)

Bei Fahrstreifen R1 (rechts im Bild) bildet der vorhandene Erhaltungsabschnitt die Situation in der Tat besser ab als das automatisch berechnete Segment. Aufgrund des Verlaufs des SUB-Wertes (siehe Bild 0.4) wäre hier das Setzen eines zusätzlichen Change points geboten gewesen. Dieser ist jedoch beim Signifikanztest (t-Test) eliminiert worden. Möglicherweise lässt sich hieraus Optimierungspotenzial für das Verfahren ableiten. Ergänzend wird angemerkt, dass ein netzweit arbeitendes Verfahren nicht für jede Zustandsdatenkonstellation das optimale Ergebnis liefern kann.

Bei Fahrstreifen L1 (links im Bild) ist der Unterschied zwischen Erhaltungsabschnitt und Segment wiederum durch die Mindestsegmentlänge $L_{\text{Min}} = 200$ m bedingt, denn auf 180 m Länge am Abschnittsbeginn liegt ein „guter“ SUB-Wert vor. Daher wurde zunächst ein vorläufiger Change point gesetzt, der wegen des Mindestsegmentlängenkriteriums allerdings

wieder entfallen ist. Die ersten 180 m werden dadurch in das nachfolgende Segment integriert und vermindern den Mittelwert des SUB. Es handelt sich um die gleiche Ursache wie oben.

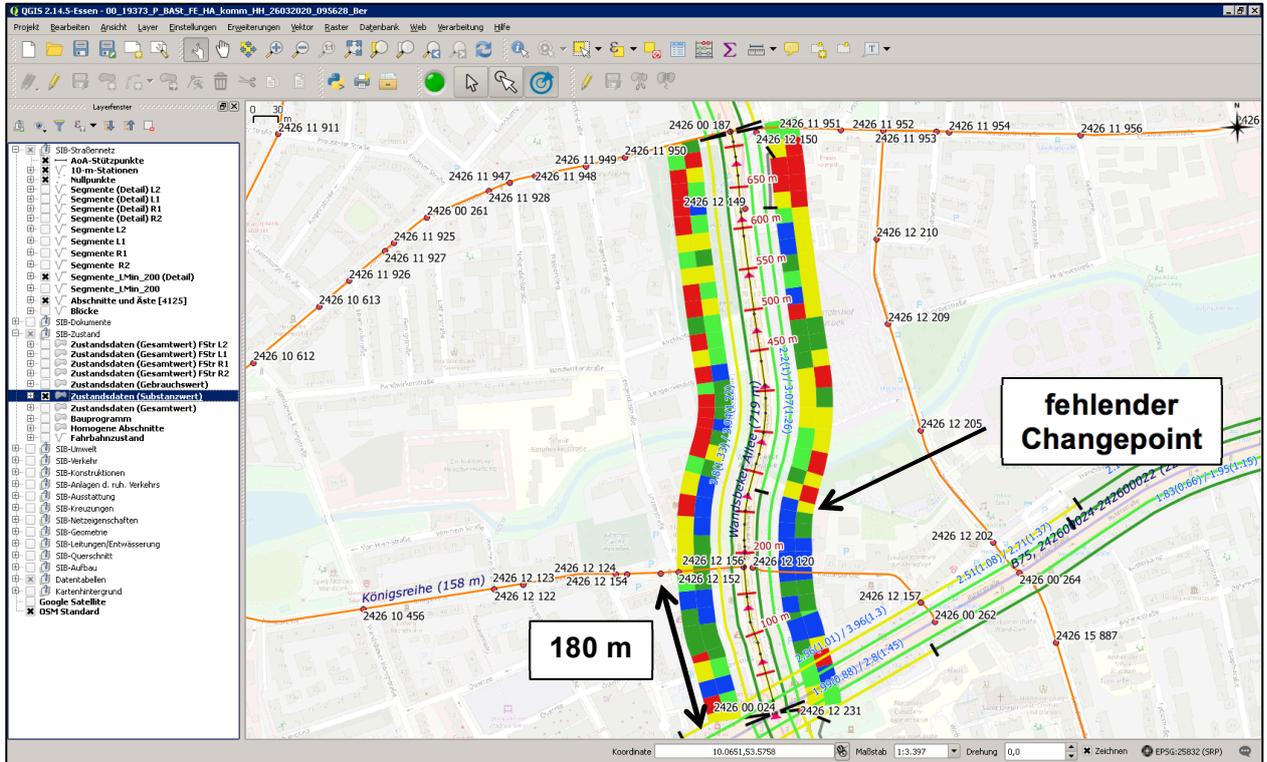


Bild 0.4: Wandsbeker Allee, Segmentierungsverfahren (Quelle: © OpenStreetMap und Mitwirkende, CC-BY-SA)

Beispiel 3:

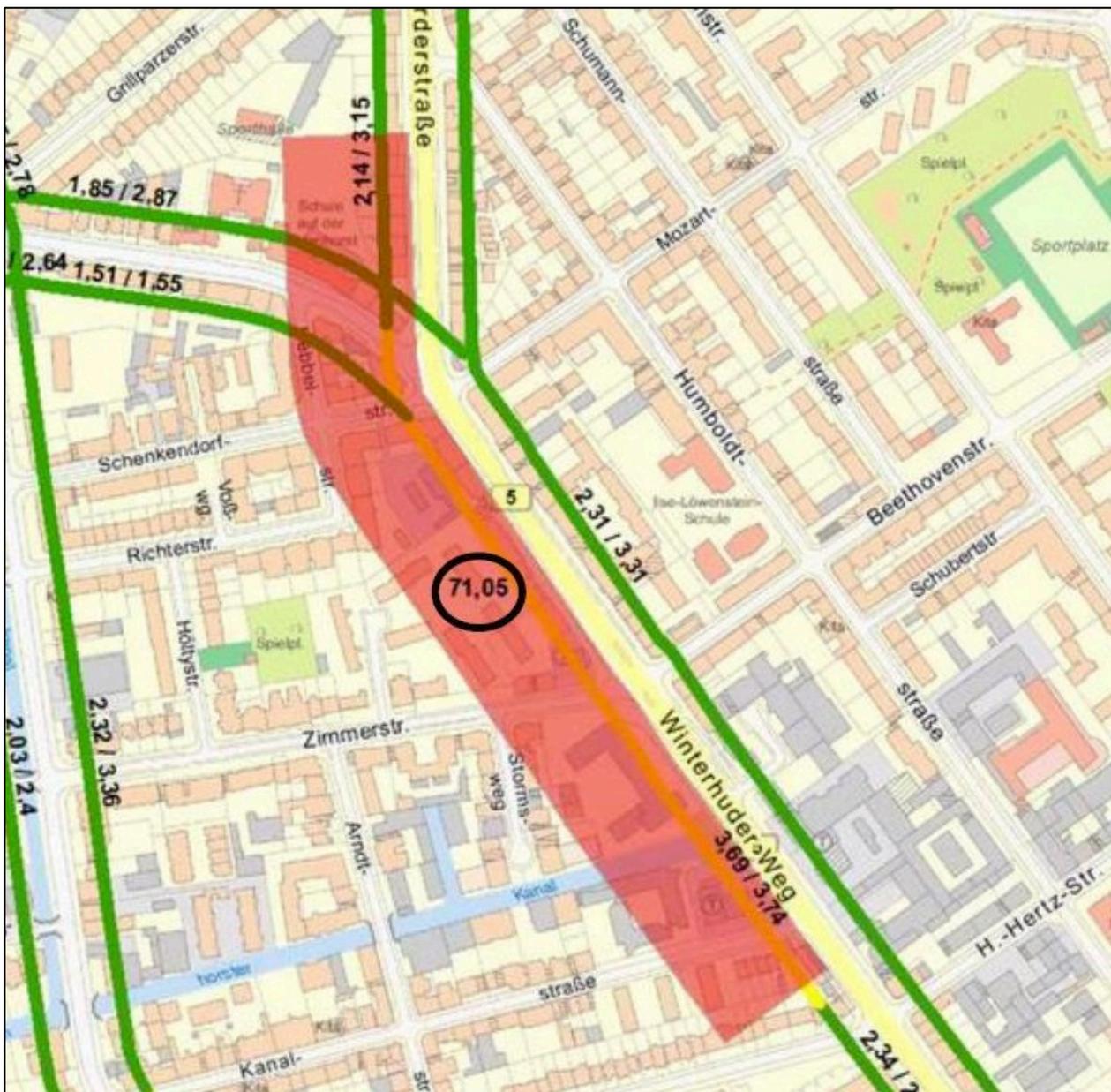


Bild 0.5: Lerchenfeld (Quelle: © Stadt Hamburg, dl-de/by-2-0)

Bei diesem Teilabschnitt der Straße B5 (Lerchenfeld, Fahrstreifen L1) sind Erhaltungsabschnitt und Segment ebenfalls unterschiedlich lang (680 m bzw. 560 m); das Segment endet im Norden deutlich früher als der Erhaltungsabschnitt. Die Ursache liegt wiederum an der geforderten Mindestsegmentlänge von 200 m. Durch das automatische Verfahren war in der Tat ein weiterer Changepoint, ungefähr am nördlichen Ende des Erhaltungsabschnitts gelegen, identifiziert worden (siehe Bild 0.6). Da jedoch auf diese Weise ein kurzes Segment der Länge 140 m entstanden war, ist dieser Changepoint in einem späteren Verfahrensschritt wieder eliminiert worden. Das zu kurze Segment wurde – auf Basis eines Mittelwertvergleichs – dem nördlich benachbarten Segment hinzugefügt.

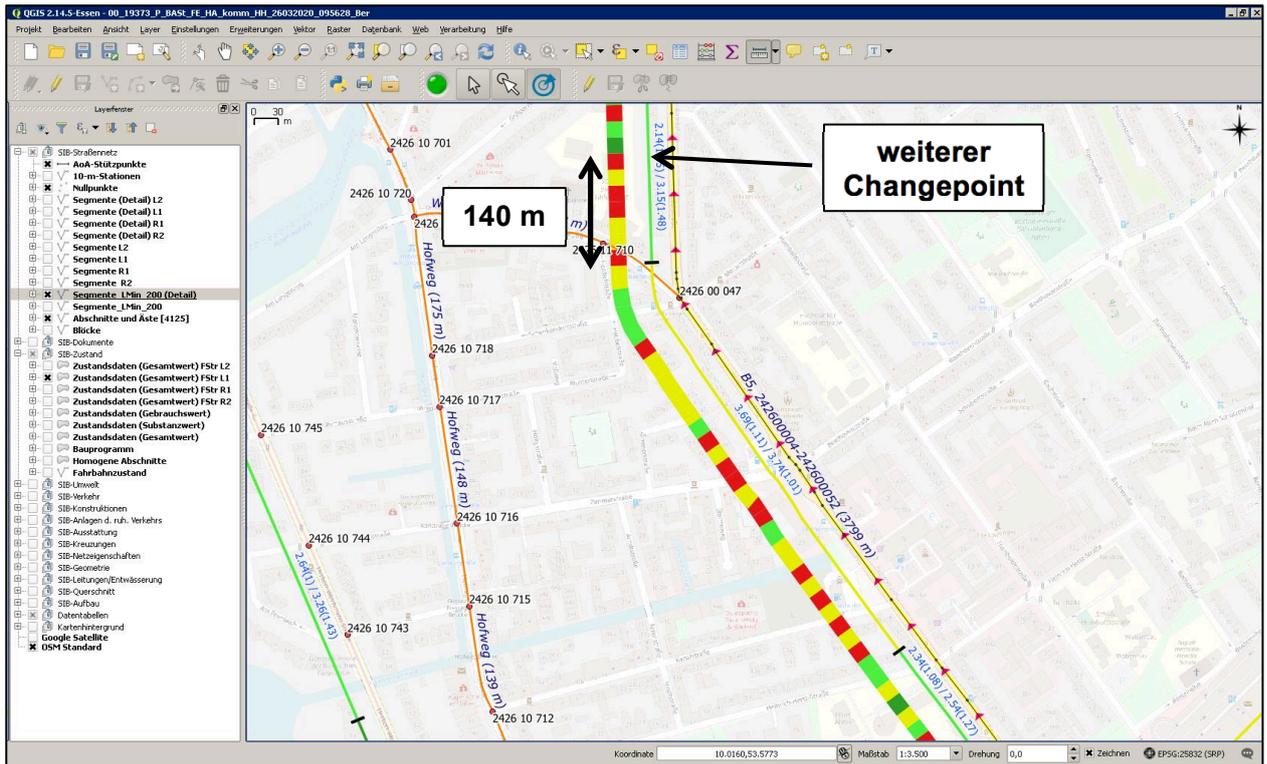


Bild 0.6: Lerchenfeld, Segmentierungsverfahren (Quelle: © OpenStreetMap und Mitwirkende, CC-BY-SA)

Anlage 3

Beschreibung
der
Algorithmen

0. Allgemeines

Die Beschreibung der Algorithmen untergliedert sich in die beiden Teile

1. Bildung der Abschnittsfolgen (entsprechend Kapitel 4.2.2) und
2. Beschreibung des Segmentierungsverfahrens (entsprechend Kapitel 5).

Anstelle von "Abschnittsfolge" wird im Folgenden häufig auch der Begriff "Block" verwendet, insbesondere in den Flowcharts.

1. Bildung der Abschnittsfolgen (Blöcke)

Die Blockbildung ist ein Teil des Preprozessings der Eingangsdaten. Dies erfolgt ausschließlich auf Basis der TPO-Daten: Netz inkl. dessen Liniengeometrie und den administrativen Daten, wie z.B. Dienststellen. Es wird davon ausgegangen, dass eventuell vorhandene Fehler und Inkonsistenzen in diesen Daten sowie auch Widersprüche zur Ergebnistabelle der Zustandsdaten bereits bereinigt wurden. Die Zustandsdaten an sich fließen jedoch nicht in die Blockbildung ein, sondern erst in das Segmentierungsverfahren.

Die Blockbildung ist ein fester Bestandteil des Gesamtverfahrens, da hierbei auch definiert wird, ob eine abschnittsübergreifende oder eine abschnittsbezogene Segmentierung erfolgt. Dabei gilt bei der

- abschnittsbezogenen Segmentierung: ein Block = ein Abschnitt oder Ast (AoA);
- abschnittsübergreifenden Segmentierung: ein Block = eine zusammenhängende lineare Abfolge von AoA derselben Straße.

Für die abschnittsübergreifende Segmentierung wird vorausgesetzt, dass die örtlich vorhandene Abfolge der AoA innerhalb einer Straße in einem numerischen Sortierfeld der Netzdaten korrekt gespeichert ist (z.B. Abschnittsfolgennummer nach ASB 2017, Kernsystem, Abschnitt 3.7). Für die konkrete Gestaltung der Abschnittsfolgennummer gibt es in der Praxis der Straßeninformationssysteme eine gewisse Vielfalt. Varianten können z.B. sein:

- Durchnummerierung der regulären¹ AoA einer Straße in 10er oder 100er oder 1.000er-Schritten;
- bei einer örtlichen Unterbrechung der Abschnittsfolge derselben Straße wird 1.000.000 addiert (Nummerierungsverfahren in TT-SIB[®]);
- nichtreguläre Abschnitte und Äste (im Fall eines überörtlichen Netzes) werden in 1er-Schritten nummeriert, jeweils folgend auf die Abschnittsfolgennummer des vorhergehenden AoA;
- nichtreguläre Abschnitte und Äste haben eine besondere Kennzeichnung in einem separaten Feld Abschnittsart.

In aller Regel sind die vergebenen Abschnittsfolgennummern aber geeignet, den örtlichen Zusammenhang der AoA abzubilden. Die hier beschriebene Blockbildung geht von einer Befüllung der Abschnittsfolgennummer und der Abschnittsart nach der Regel "mblock" = 1 aus

¹ Als nichtreguläre Abschnitte gelten Seitenarme oder im überörtlichen Bereich die getrennt verlaufenden Fahrbahnen. Alle anderen Abschnitte werden regulär genannt.

Tabelle 0.1 aus (falls nicht so vorliegend, muss diese Variante im Rahmen der Eingangsprüfung der Daten hergestellt werden).

Liegen speziell vordefinierte Abschnittsfolgen bereits vor (z.B. als Einteilung jeder Straße in die gewünschten Teilstrecken bereits in der SIB) oder ist die abschnittsbezogene Segmentierung vorgesehen (Abschnittsfolgennummer spielt dann keine Rolle), reduziert sich die Blockbildung auf einen minimalen Schritt.

Der Algorithmus zur Blockbildung verwendet als Eingangsgröße die Variable "mblock" (= Merkmal Blockbildung).

"mblock"	Feld Abschnittsfolgennummer "sort"	Feld Abschnittsart "art"
1	alle AoA in 100er-Schritten (pro Straße)	Kennzeichnung von Seitenarmen und anderen nichtregulären Abschnitten und Ästen durch ein Flag
2	speziell vordefinierte Abschnittsfolgen	
3	abschnittsbezogene Segmentierung	

Tabelle 0.1 Blockbildungsmerkmale

Grundsätzlich wird für die Blockbildung eine Tabelle "net_aoa" mit minimal folgenden Spalten als Input benötigt:

Feldname	Feldtyp	Beschreibung
strasse	varchar	Straßenbezeichnung (Klasse + Nummer + ggf. Zusatzbuchstabe) nach ASB
strname	varchar	Straßenname
sort	integer	Abschnittsfolgennummer, siehe Tabelle oben
art	varchar	OneOf('Ast','Seit','gvF1','gvF2','reg'), AoA-Art nach ASB (siehe auch ASB 2017, Kernsystem, Abschnitt 3.7), Information stammt aus der SIB: Seit = Seitenarm, gvF1/gvF2 = getrennt verlaufende Fahrbahn, Zweig 1 bzw. Zweig 2, reg = regulärer Abschnitt
aoa	varchar	AoA= vnp + nnp (= eindeutiger fachlicher Schlüssel der Tabelle)
vnp	varchar	Von-Nullpunkt (als redundante Information)
nnp	varchar	Nach-Nullpunkt (als redundante Information)
aalen	integer	AoA-Länge in m
ds	varchar	Kennziffer der für den AoA an der Station 0 zuständigen Dienststelle
geom	geometry	Liniengeometrie des AoA (falls Darstellung im GIS gewünscht)

Tabelle 0.2 Tabelle "net_aoa"

Regeln für die Blockbildung:

- Nichtreguläre Abschnitte und Äste werden stets als jeweils einzelner Block betrachtet ('Ast', 'gvF1', 'gvF2' gibt es regelmäßig nur bei überörtlichen Straßennetzen).
- Die Zuordnung der AoAs und damit nachfolgend auch der Blöcke zu Amtsbezirken/ Dienststellen (DS) erfolgt der Einfachheit halber AoA-bezogen, nicht stationsscharf.
- Im Fall "mblock" = 1 ist ein wichtiges Kriterium, dass der Nach-Nullpunkt "nnp" dem Von-Nullpunkt "vnp" des nachfolgenden Abschnitts entsprechen muss, damit der Zusammenhang der Abschnittsfolge gewährleistet ist.

Das Ergebnis der Blockbildung ist die Tabelle "net_blk".

Feldname	Feldtyp	Beschreibung
blk	integer	Block-ID (eindeutiger Schlüssel der Tabelle)
strasse	varchar	Straßenbezeichnung (Klasse + Nummer + ggf. Zusatzbuchstabe) nach ASB (ein Block ist niemals straßenübergreifend)
vonsort	integer	Abschnittsfolgenummer, an welcher der Block beginnt
bissort	integer	Abschnittsfolgenummer, an welcher der Block endet
art	varchar	Abschnittsart des Blocks (ein Block gehört immer zu derselben Art)
anzaoa	varchar	Anzahl AoAs im Blocks
blklen	integer	Länge des Blocks in m
geom	geometry	Liniengeometrie des Blocks (falls Darstellung im GIS gewünscht)

Tabelle 0.3 Tabelle "net_blk"

Im Fall der abschnittsübergreifenden Segmentierung ("mblock" = 1) wird diese Tabelle mittels sequenzieller Gruppierung² der Tabelle "net_aoa" erstellt. Dazu ist diese Tabelle nach den Feldern "strasse", "art", "sort" und "ds" zu sortieren. Die Gruppierkriterien sind: Gleiche "strasse", gleiche "art", gleiche "ds" und "nnp" ist gleich "vnp" des nachfolgenden Records. Die dabei entstehenden Gruppen sind die Blöcke im Sinne des Verfahrens. Jede Gruppe bekommt eine numerische Block-ID zugeordnet.

In den Fällen "mblock" = 2 oder 3 wird die Tabelle "net_blk" direkt aus Tabelle "net_aoa" abgeleitet. Das Gruppierkriterium ist hier: Gleiche Abschnittsfolge bzw. gleicher "aoa". Jede vordefinierte Abschnittsfolge bzw. jeder AoA bildet einen eigenen Block und erhält eine Block-ID.

² Die sequenzielle Gruppierung ist eine besondere Form der Gruppierung von Records in einer sortierten Tabelle. Dabei wird die Tabelle in der gewünschten Sortierreihenfolge sequenziell auf Änderungen des Gruppierkriteriums durchsucht. Bei jeder Änderung des Gruppierkriteriums beginnt eine neue Gruppe. Die Gruppierkriterien können auch Verweise auf vorhergehende oder nachfolgende Records beinhalten (bezogen auf die Sortierreihenfolge). Die einfache SQL-Gruppieranweisung GROUP BY bildet im Gegensatz dazu Gruppen ohne Beachtung einer Sortierreihenfolge.

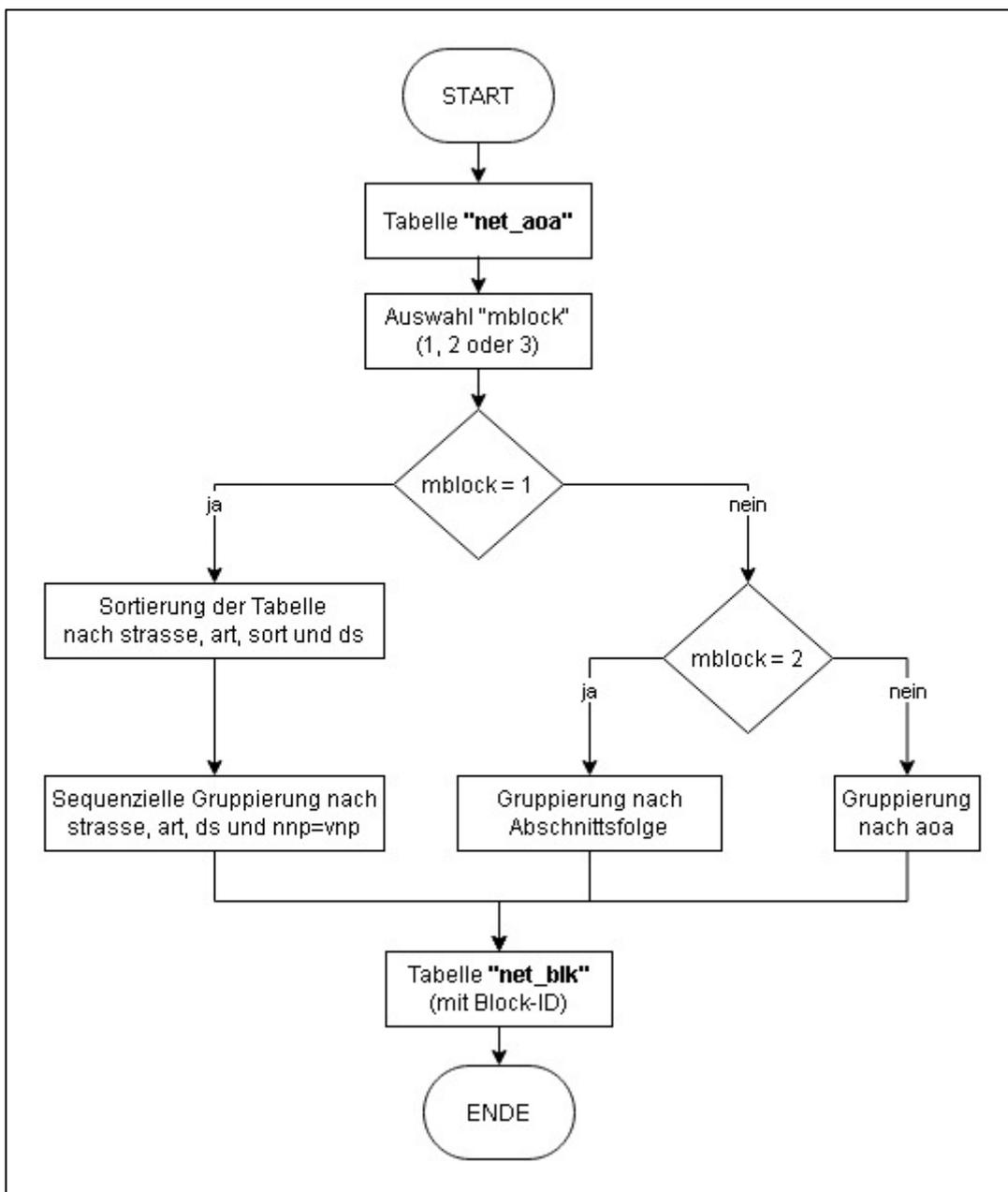


Bild 0.1: Flowchart Blockbildung

2. Beschreibung des Segmentierungsverfahrens

Der grundsätzliche Ablauf des Segmentierungsverfahrens wurde in Bild 5.3 grob skizziert. Die folgenden Ausführungen dienen der Untersetzung der einzelnen Verfahrensschritte.

Als Input werden die Tabellen "net_aoa", "net_blk" und "obj_erg" benötigt. Die beiden ersten wurden bereits unter Bildung der Abschnittsfolgen (Blöcke) beschrieben.

Die Tabelle "obj_erg" ist die Tabelle mit den Ergebnissen der Zustandsbewertung. Falls die Zustandserfassung nach ZTV ZEB-StB erfolgte, kann dafür die Ergebnistabelle (ERG.dbf) verwendet werden. Liegen die Ergebnisse der Zustandsbewertung in einer abweichenden tabellarischen Form vor, so ist für "obj_erg" im Preprozessing eine ERG-Struktur nach ZTV ZEB-StB herzustellen. Es ist für das Verfahren unerheblich, welche Länge die Auswerteabschnitte haben (10 m, 20 m, 100 m). Grundsätzlich müssen in "obj_erg" mindestens folgende Spalten vorhanden sein:

Feldname	Feldtyp	Beschreibung
aoa	varchar	AoA
fstr	varchar(2)	Fahrstreifen in der Form ..., L2, L1, R1, R2, ...
vst	integer	Von-Station des Auswertabschnitts
bst	integer	Bis-Station des Auswertabschnitts
zp1	numeric	Zustandswert des ersten Zustandsparameters (z.B. GW)
zp2	numeric	Zustandswert des zweiten Zustandsparameters (z.B. SUB)
zp3	numeric	Zustandswert des dritten Zustandsparameters (z.B. GEB)
zp4	numeric	Zustandswert des vierten Zustandsparameters (z.B. TWRIO)
...		Spalten mit weiteren Zustandsparametern

Tabelle 0.4 Tabelle "obj_erg"

2.1 Auswahl Netzbereich und Zustandsparameter

Die Auswahl des zu segmentierenden Netzbereichs durch den Anwender erfolgt über die gebildeten Blöcke. Die Tabelle "net_blk" ist mit der Tabelle "net_aoa" durch einen JOIN verknüpft, so dass alle zum Block gehörigen AoA damit ebenfalls ausgewählt sind. Tabelle "net_aoa" ist wiederum durch einen JOIN mit der Tabelle "obj_erg" der Zustandsdaten verknüpft.

Als weitere Nutzerauswahl müssen die Zustandsparameter, nach denen segmentiert werden soll, und der gewünschte Fahrstreifen angegeben werden. Falls Zustandsdaten für mehrere Fahrstreifen vorliegen, muss das Segmentierungsverfahren nacheinander für jeden gewünschten Fahrstreifen durchgeführt werden. In der aktuellen Implementierung ist vorgesehen, dass die Ergebnisse der Segmentierung am Ende in einer Gesamttabelle für alle Fahrstreifen ausgegeben werden.

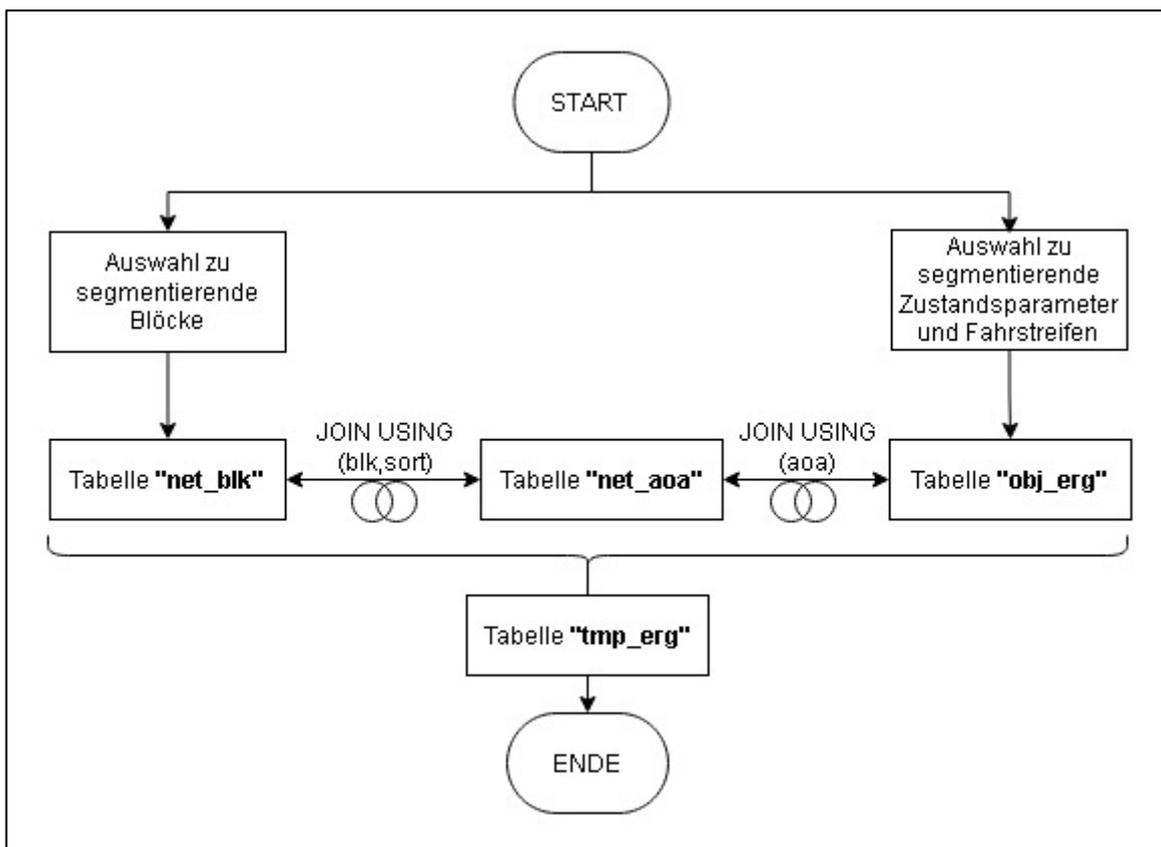


Bild 0.2: Flowchart Auswahl Netz und Zustandsparameter

Nachdem der Netzbereich, die Zustandsparameter und der Fahrstreifen gewählt wurden, wird eine temporäre Tabelle "tmp_erg" angelegt und mit der entsprechenden Teilmenge der "obj_erg" befüllt. Diese dient im Weiteren als Arbeitstabelle für das Verfahren.

Feldname	Feldtyp	Beschreibung
id	integer	Auswertabschnitts-ID
blk	integer	Block-ID
sort	integer	Abschnittsfolgennummer des AoA
aoa	varchar	AoA
fstr	varchar(2)	Fahrstreifen in der Form ..., L2, L1, R1, R2, ...
vst	integer	Von-Station des Auswertabschnitts
bst	integer	Bis-Station des Auswertabschnitts
zp	numeric	Zustandswert des ersten gewählten Zustandsparameters
zq	numeric	Zustandswert des zweiten gewählten Zustandsparameters
...		Spalten mit weiteren Zustandsparametern (für Programmerweiterung)
cp	integer	Flag (0/1) für die Changepoint-Eigenschaft

Tabelle 0.5 Tabelle "tmp_erg"

Die Spalte "cp" spielt in weiteren Verlauf des Verfahrens die Rolle des Indikators, der anzeigt, ob an der Von-Station des Auswerteabschnitts ein Changepoint vorliegt.

An dieser Stelle sei angemerkt, dass die weiteren Schritte für den Fall der konstanten Auswerteabschnittslänge gelten. Liegen Zustandsdaten auf variablen Auswerteabschnittslängen, z.B. aus einer visuell-sensitiven Zustandsdatenerfassung, vor, so muss die Behandlung von Missings, die Glättung, die Identifikation der Changepoints und die Berücksichtigung einer Mindestsegmentlänge auf eine abweichende Weise erfolgen. Diese Aspekte wurden im vorliegenden Projekt nicht im Detail umgesetzt (vgl. dazu Kapitel 4.3.5).

2.2 Behandlung von Missings

Die Folge jedes gewählten Zustandsparameters der Tabelle "tmp_erg" wird auf fehlende Werte (Missings) untersucht. Missings in der Folge der Zustandsparameter sind üblicherweise durch Werte < 0 gekennzeichnet. Damit liegt ein einfach auszuwertendes Kriterium für einen gültigen bzw. ungültigen Wert vor. Jedem Zustandsparameter "z*" kann das Flag "z*_valid" mit den Werten TRUE oder FALSE zugeordnet werden.

Auf Basis des nutzerspezifischen Eingabewertes M_{\max} wird entschieden, ob fehlende Werte interpoliert werden oder ob eine Bildung von Unterblöcken erfolgen soll. Unterblöcke teilen die oben generierten Blöcke nochmals auf. Technisch erfolgt dies durch Vergabe von neuen Block-IDs, wobei die bisherige Block-ID des aufzuteilenden Blocks eliminiert wird.

Zunächst muss jedoch die Anzahl zusammenhängender ungültiger Werte ermittelt werden, damit das Kriterium M_{\max} angewendet werden kann. Dies geschieht wiederum mittels sequenzieller Gruppierung. Die Sortierordnung ist die örtliche Abfolge der Auswerteabschnitte im Zuge eines Blocks. Das Gruppierkriterium lautet: Gleiche Block-ID und gleicher Wert von "z*_valid".

Liegen höchstens M_{\max} zusammenhängende ungültige Werte vor, so werden die fehlenden Werte z_p bzw. z_q durch lineare Interpolation unter Verwendung der benachbarten gültigen Werte oder durch konstante Fortsetzung des ersten/letzten gültigen Wertes am Rand des Blocks ermittelt. Anschließend erfolgt eine Harmonisierung der noch immer ungültigen Werte über beide Zustandsparameterfolgen: Dort, wo eine der beiden Zustandsparameterfolge einen ungültigen Wert hat, wird der Wert in der anderen Zustandsparameterfolge ebenfalls als ungültig geflagt. Für jede danach noch verbliebene Teilfolge mit mehr als M_{\max} zusammenhängenden ungültigen Werten wird der Block in Unterblöcke eingeteilt.

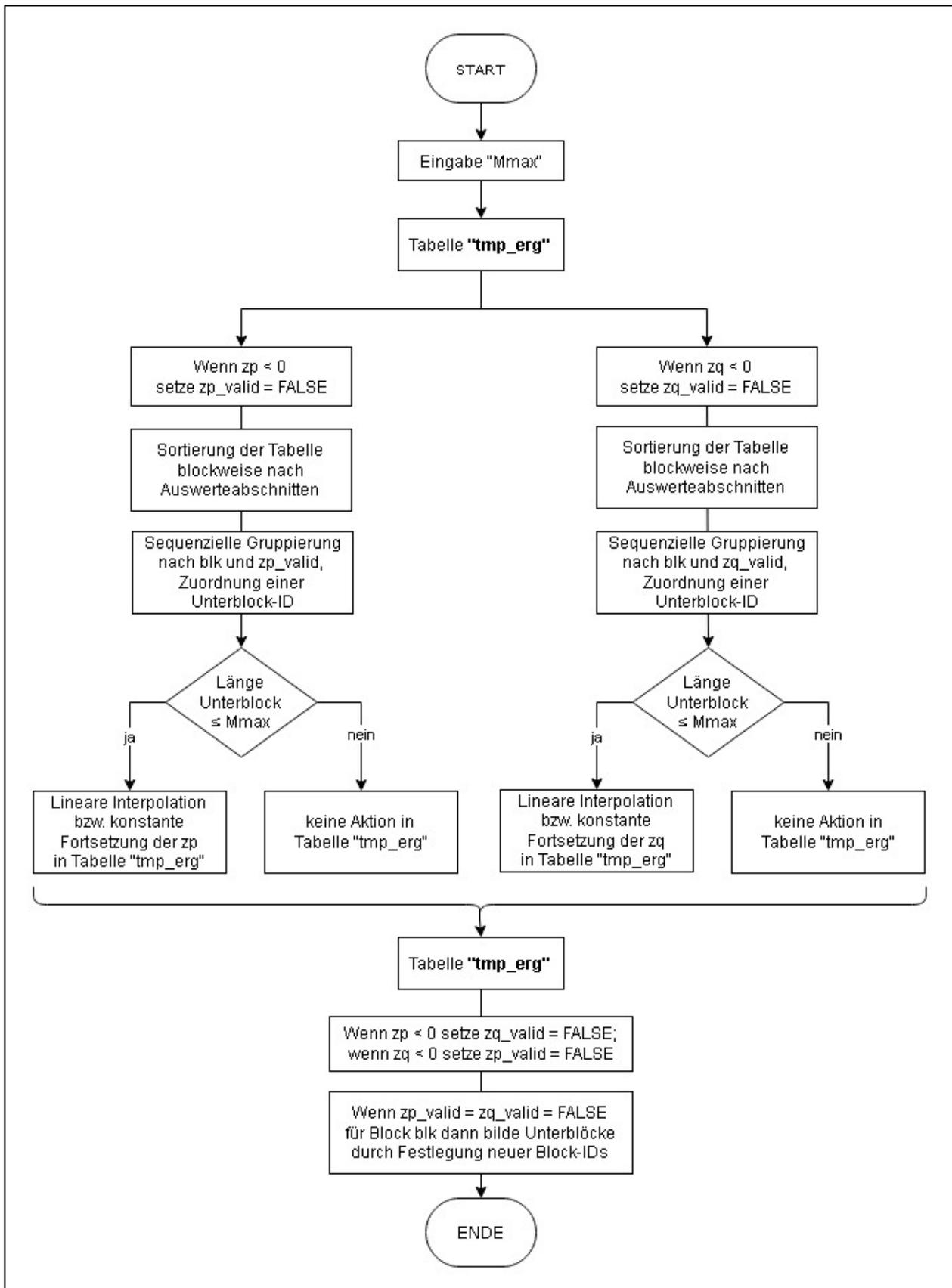


Bild 0.3: Flowchart Behandlung von Missings

2.3 Test auf Notwendigkeit der Segmentierung

Mit der Festlegung aller Blöcke und Unterblöcke (im Weiteren nur als Block bezeichnet) steht deren endgültige Länge fest. Auf Basis des nutzerspezifischen Eingabewerts der Mindestsegmentlänge L_{Min} werden in diesem Teilschritt alle Blöcke von der Segmentierung ausgeschlossen, deren Länge L_{Min} nicht übersteigt. Von der Segmentierung werden ferner noch alle Blöcke/Unterblöcke ausgeschlossen, für die keine Zustandsdaten vorliegen. Dieser Fall tritt auf, wenn die zugrundeliegenden Netzdaten umfassender sind als der Netzbereich, auf dem die Zustandserfassung durchgeführt wurde, oder wenn Unterblöcke wegen fehlender Zustandsdaten gebildet werden mussten (Lücke größer als M_{Max}).

Die nicht zu segmentierenden Blöcke werden in eine separate Tabelle "tmp_erg_nosgm" ausgelagert, die dieselbe Struktur wie die "tmp_erg" besitzt. Diese Blöcke werden am Ende des Gesamtverfahrens der Ergebnisdatei hinzugefügt.

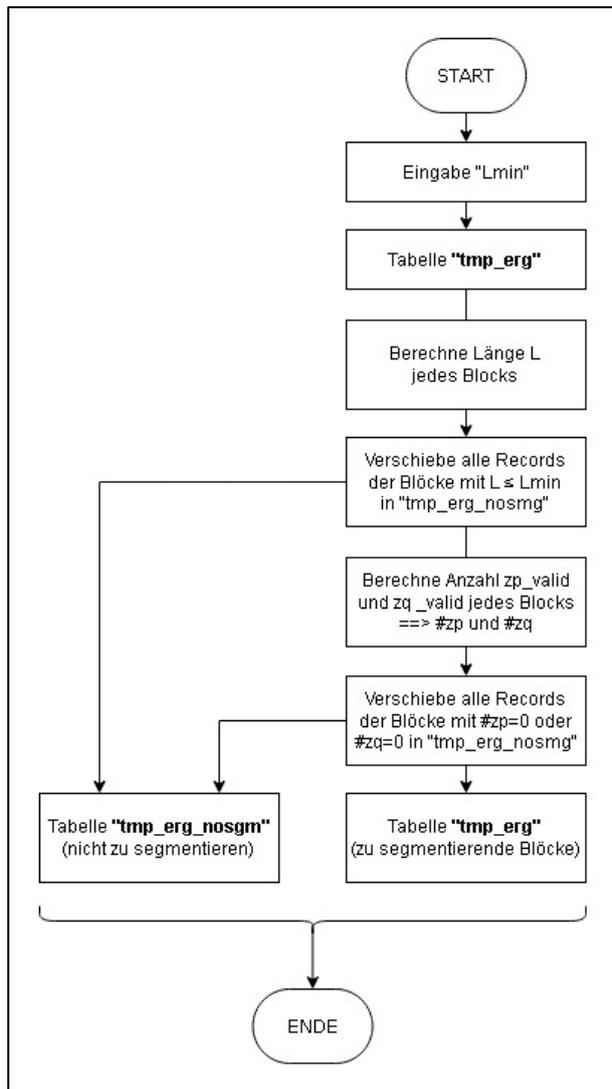


Bild 0.4: Flowchart Test auf Notwendigkeit der Segmentierung

Für die in der "tmp_erg" verbliebenen Blöcke startet nun das eigentliche Segmentierungsverfahren. Die Folgen der zu einem bestimmten Block gehörigen Zustandsparameter z_p und z_q werden im Weiteren als Messwertfolgen $\{z_{p_i}\}$ und $\{z_{q_i}\}$ der Länge n bezeichnet. Diese werden stets je Block verarbeitet (die Länge n hängt vom jeweils betrachteten Block ab).

2.4 Glättung

Der Glättungsschritt besteht aus mehreren Stufen. Für die Anwendung der morphologischen Glättung (formbasierte Glättung) auf die Messwertfolgen $\{z_{p_i}\}$ und $\{z_{q_i}\}$ wurde eine PostgreSQL-Funktion entwickelt. Auf die morphologisch geglätteten Folgen wird anschließend noch ein gleitendes Mittel angewendet.

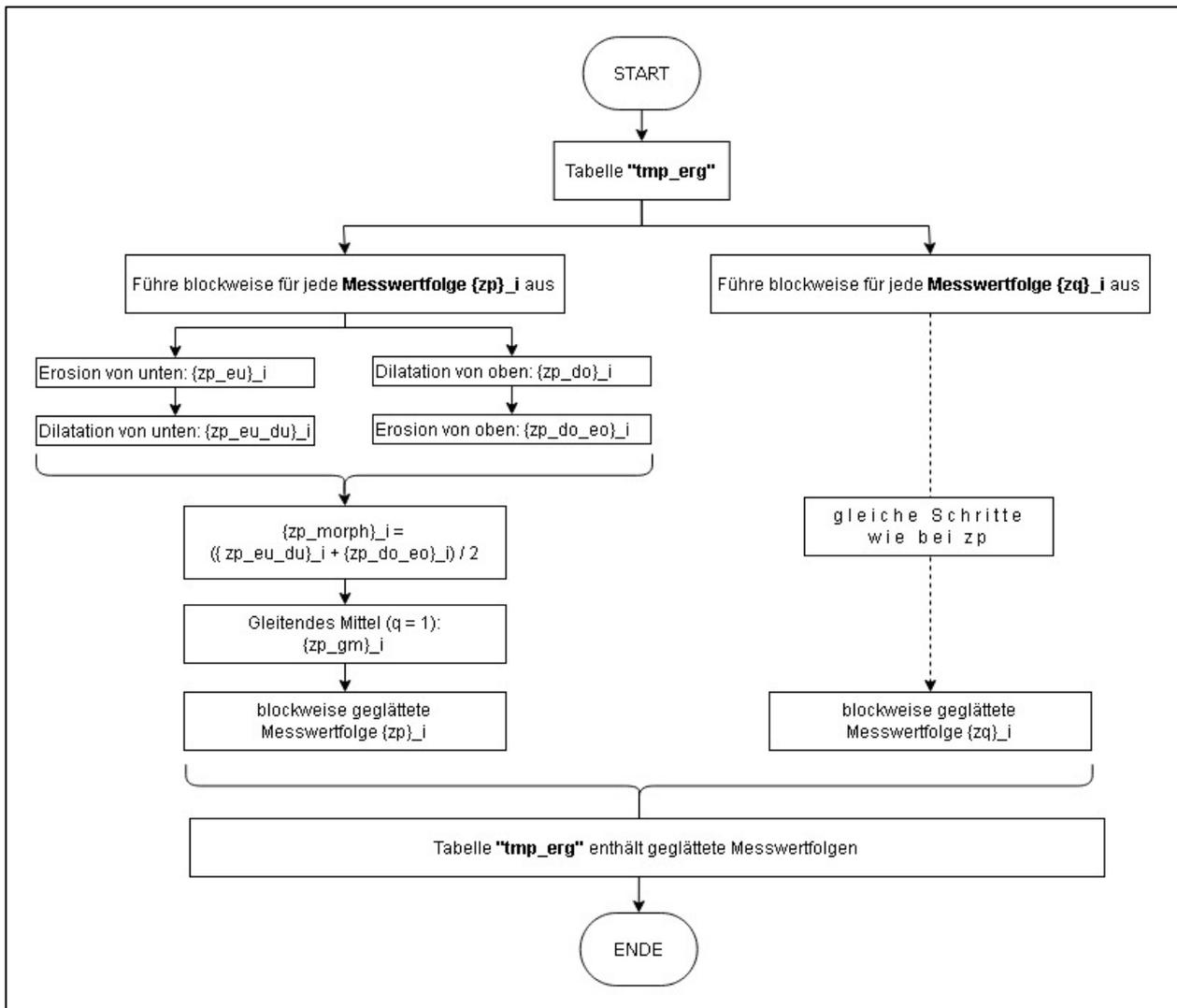


Bild 0.5: Flowchart Glättung

Die Formeln für die morphologischen Operationen lauten:

$$\text{Erosion von unten: } x_{eu_i} = \text{Min}_{-p \leq \theta \leq p} (x_{i+\theta} - s(\theta)),$$

$$\text{Dilatation von unten: } x_{du_i} = \text{Max}_{-p \leq \theta \leq p} (x_{i+\theta} + s(\theta)),$$

$$\text{Dilatation von oben: } x_{do_i} = \text{Min}_{-p \leq \theta \leq p} (-x_{i+\theta} - s(\theta)),$$

$$\text{Erosion von oben: } x_{eo_i} = \text{Max}_{-p \leq \theta \leq p} (-x_{i+\theta} + s(\theta)),$$

wobei die Minima und Maxima über die ganzzahligen Werte von θ gebildet werden. Als strukturierendes Element $s(x)$ wurde der Halbkreis mit Radius $p = 3$ verwendet: $s(x) = \sqrt{p^2 - x^2}$ falls $-p \leq x \leq p$ und $s(x) = 0$ sonst.

Das gleitende Mittel mit $q = 1$ berechnet sich mittels $x_{gm_i} = \frac{1}{3} \sum_{r=-1}^1 x_{i+r}$, $i = 2, \dots, n-1$. Für $i = 1$ und $i = n$ wurde nur der Mittelwert über die beiden "Randsummanden" gebildet.

2.5 Changepoint-Identifikation

Die Feststellung der potenziellen Changepoints erfolgt getrennt für jeden zu segmentierende Messwertfolge $\{z_{p_i}\}$ und $\{z_{q_i}\}$. Dafür kommen die Verfahren CUSUM und ADA zur Anwendung. Festgestellte potenzielle Changepoints werden in Tabelle "tmp_erg" in Spalte "cp" mit dem Flag = 1 gekennzeichnet (sonst Flag = 0). Der jeweils erste Record eines Block erhält stets Flag = 1.

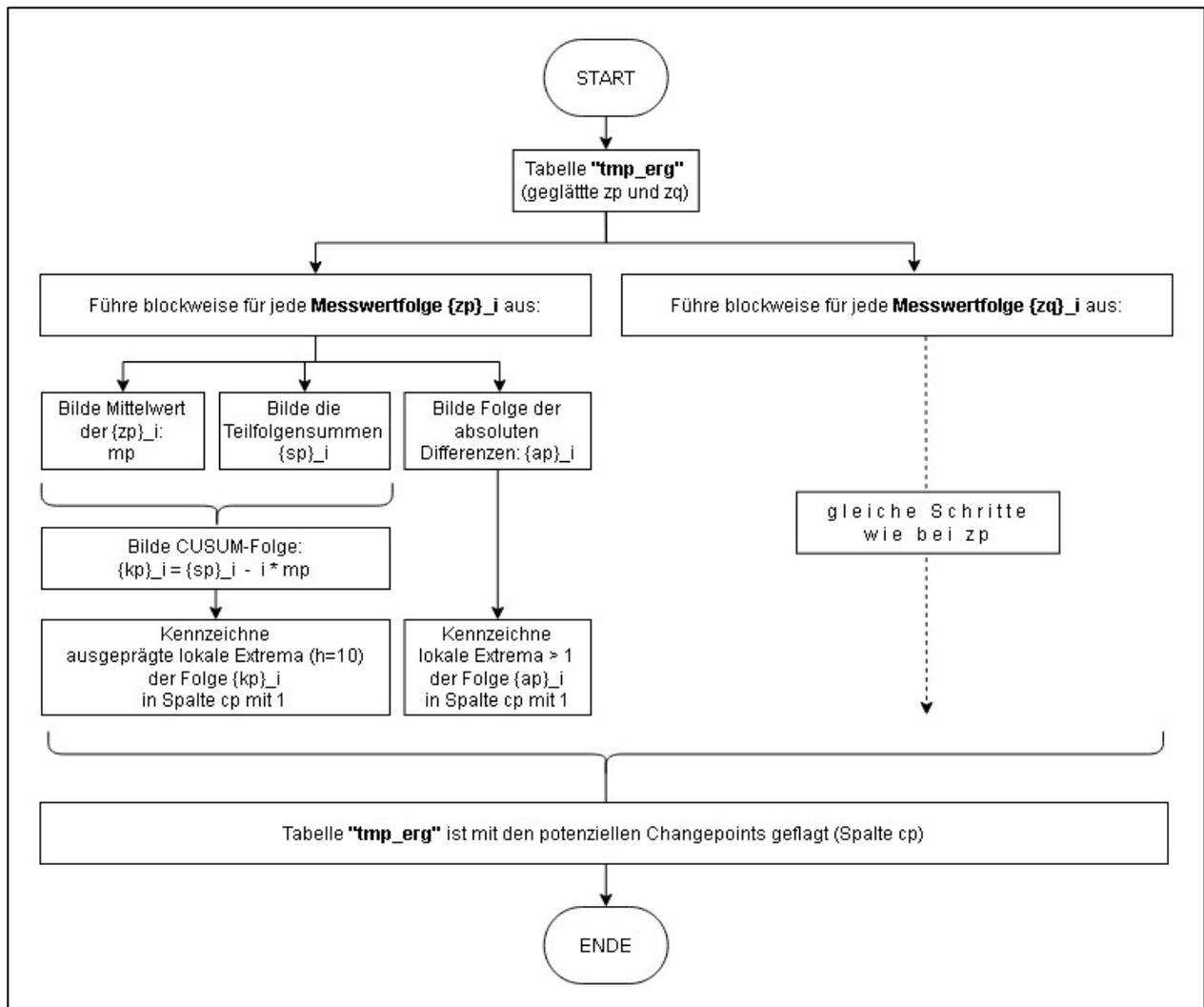


Bild 0.6: Flowchart Changepoint-Identifikation

Die Formeln lauten (hier für $\{z_p\}$ angegeben, analoges gilt für $\{z_q\}$):

$$\text{Mittelwert} \quad mp = \frac{1}{n} \sum_{j=1}^n (z_p)_j,$$

$$\text{Teilfolgensumme:} \quad (sp)_i = \sum_{j=1}^i (z_p)_j, \quad i = 1, \dots, n,$$

$$\text{CUSUM-Folge:} \quad (kp)_i = (sp)_i - i * mp, \quad i = 1, \dots, n,$$

$$\text{ADA-Folge mit } d = 2: \quad (ap)_i = |(z_p)_{i+2} - (z_p)_{i-2}|, \quad i = 1, \dots, n,$$

wobei $(z_p)_{-1} = (z_p)_0 = (z_p)_1$ und $(z_p)_{n+2} = (z_p)_{n+1} = (z_p)_n$ angenommen wird.

Anmerkung: Im Fall der variable Auswertabschnittslänge wäre eine alternative Vorgehensweise wie folgt. Für die Messwertfolge kommt hier neben dem Zustandswert z_i als weitere Größe noch die Länge d_i des Auswertabschnitts ins Spiel (jeweils zu bilden als Differenz der Stationswerte $bst - vst$). Wie in Kapitel 4.3.5 dargestellt, kann das Produkt $z_i * d_i$ als Maß für die Schädigung auf diesem Auswertabschnitt betrachtet werden. Als zu segmentierende Messwertfolge wird daher die Folge $\{x_i\}$ mit den Gliedern $x_i = z_i * d_i, i = 1, \dots, n$, betrachtet. Für die Identifikation der Changepoints wird nun analog zu CUSUM vorgegangen.

$$\text{Mittelwert} \quad m = \frac{1}{n} \sum_{j=1}^n z_j * d_j ,$$

$$\text{Teilfolgensumme:} \quad s_i = \sum_{j=1}^i z_j * d_j, \quad i = 1, \dots, n ,$$

$$\text{CDA-Folge:} \quad k_i = s_i - i * m, \quad i = 1, \dots, n .$$

Dieses Verfahren ist als CDA (Cumulative Differences Approach) bekannt (vgl. [15]).

Ausgeprägte lokale Extrema (hier $h = 10$) der obigen CUSUM-Folge $(kp)_i$ werden bestimmt, indem für jedes Folgeglied ermittelt wird, ob alle 10 links und rechts benachbarten Folgeglieder kleiner resp. größer sind als der Wert selbst.

Auf Basis der gefundenen vorläufigen Changepoints (Flag = 1 in Spalte cp der Tabelle "tmp_erg") können nun durch Bildung der laufenden Summe der Flags und anschließende Gruppierung die vorläufigen Segmente gebildet werden. Die folgende beispielhafte Darstellung zeigt die Vorgehensweise. Die Spalten "blk", "id" und "cp" stellen einen Teilbereich der Tabelle "tmp_erg" dar, die Spalte "laufende Summe der cp" existiert nur temporär.

blk	id	cp	laufende Summe der cp
1	1	1	1
1	2	0	1
1	3	0	1
1	4	0	1
1	5	1	2
1	6	0	2
1	7	1	3
1	8	0	3
1	9	0	3
1	10	1	4
1	11	0	4
1	12	0	4
...

Tabelle 0.6 Bildung der Segmente durch Gruppierung

Das Gruppierungsergebnis wird in der Tabelle "tmp_sgm" gespeichert, deren Records die (momentan noch vorläufigen) Segmente sind. Die Sortierung in der Tabelle ist so festzulegen, dass die Segmente in ihrer örtlichen Abfolge vorliegen, da im nächsten Schritt jeweils benachbarte Segmente miteinander verglichen werden.

Feldname	Feldtyp	Beschreibung
sgm	integer	Segment-ID (1,...,m)
blk	integer	Block-ID des Blocks, zu dem das Segment gehört
minid	integer	Auswertabschnitts-ID des Segmentanfangs
maxid	integer	Auswertabschnitts-ID des Segmentendes
sgmlen	integer	Segmentlänge in m
n	integer	Anzahl Auswertabschnitte im Segment
mzp	numeric	Mittelwert von zp im Segment
szp	numeric	Standardabweichung von zp im Segment
mzq	numeric	Mittelwert von zq im Segment
szq	numeric	Standardabweichung von zq im Segment

Tabelle 0.7 Tabelle "tmp_sgm"

2.6 Changepoint-Verifikation

Die Changepoint-Verifikation nutzt den t-Test (Hypothesentest), um die Signifikanz der Mittelwertdifferenzen jeweils zweier benachbarter Segmente i und $i+1$, $i = 1, \dots, m-1$, festzustellen. Ist die Mittelwertdifferenz bei mindestens einem der beteiligten Zustandsparameter signifikant, gilt der Changepoint durch den t-Test als bestätigt. Falls die Mittelwertdifferenzen aller beteiligten Zustandsparameter nicht signifikant sind, wird der Changepoint entfernt.

Die Formeln lauten (hier für zp angegeben, analoges gilt für zq):

$$\text{Betrag der Mittelwertdifferenz: } M_{i,i+1} = |(mzp)_i - (mzp)_{i+1}|$$

$$\text{Gesamtstandardabweichung: } A_{i,i+1} = \sqrt{(szp)_i^2/n_i + (szp)_{i+1}^2/n_{i+1}}$$

$$\text{Testgröße: } T_i = M_{i,i+1}/A_{i,i+1}$$

Für die Testentscheidung ist das zum jeweiligen Segmentpaar ($i, i+1$) gehörige Testquantil Q_i zu bestimmen. Dieser Wert hängt nur von n_i und n_{i+1} ab. Er lässt sich der Quantiltabelle "stud_distr" der t-Verteilung entnehmen, die in einschlägigen Statistiklehrbüchern bzw. Internetquellen verfügbar ist (auch als Student-Verteilung bekannt).

Feldname	Feldtyp	Beschreibung
f	integer	Anzahl Freiheitsgrade
side	integer	Art des t-Testes (einseitig, zweiseitig), Wert 1 oder 2
alpha	numeric	Signifikanzniveau 0.01, 0.02, 0.05, 0.1 (häufig verwendete Werte)
quantil	numeric	Wert des Quantils als Funktion von n, side und alpha

Tabelle 0.8 Tabelle "stud_distr"

Der Freiheitsgrad bestimmt sich bei diesem Test zu $f_i = n_i + n_{i+1} - 1$. Für die Auswahl des Wertes des Testquantils Q_i werden noch die fixen Parameter "side" und "alpha" benötigt. Im Verfahren wurden side = 2 (da zweiseitiger t-Test) und alpha = 0.05 verwendet.

f	side	alpha	quantil
...
4	2	0.05	2.776
5	2	0.05	2.571
6	2	0.05	2.447
7	2	0.05	2.365
8	2	0.05	2.306
9	2	0.05	2.262
10	2	0.05	2.228
...

Tabelle 0.9 Ausschnitt aus dem Inhalt der Tabelle "stud_distr"

Die Testentscheidung selbst besteht darin, dass für jeden beteiligten Zustandsparameter z_p und z_q geprüft wird, ob $T_i \leq Q_i$ gilt. Ist dies bei allen Zustandsparametern der Fall, dann ist der die beiden Segmente trennende Changepoint nicht signifikant und wird rückwirkend in Tabelle "tmp_erg" durch Setzen des CP-Flag = 0 entfernt.

Da durch Verändern der CP-Flags eine neue (vorläufige) Segmentierung entsteht, wird der t-Test iterativ wiederholt. In der Regel sind 4 bis 5 Durchläufe erforderlich.

Nach dem zweiten und ggf. vierten Durchlauf des t-Testes erfolgt eine Bewertung der (vorläufigen) Segmente bzgl. der Mindestsegmentlänge. Bei jedem Segment, dessen aktuelle Länge unterhalb L_{Min} liegt, wird einer der beiden begrenzenden Changepoint entfernt. Für die Entscheidung, welcher der beiden entfernt wird, werden wiederum Mittelwertdifferenzen bewertet. Auf der Seite, wo die Mittelwertdifferenz kleiner ist, wird das CP-Flag = 0 gesetzt.

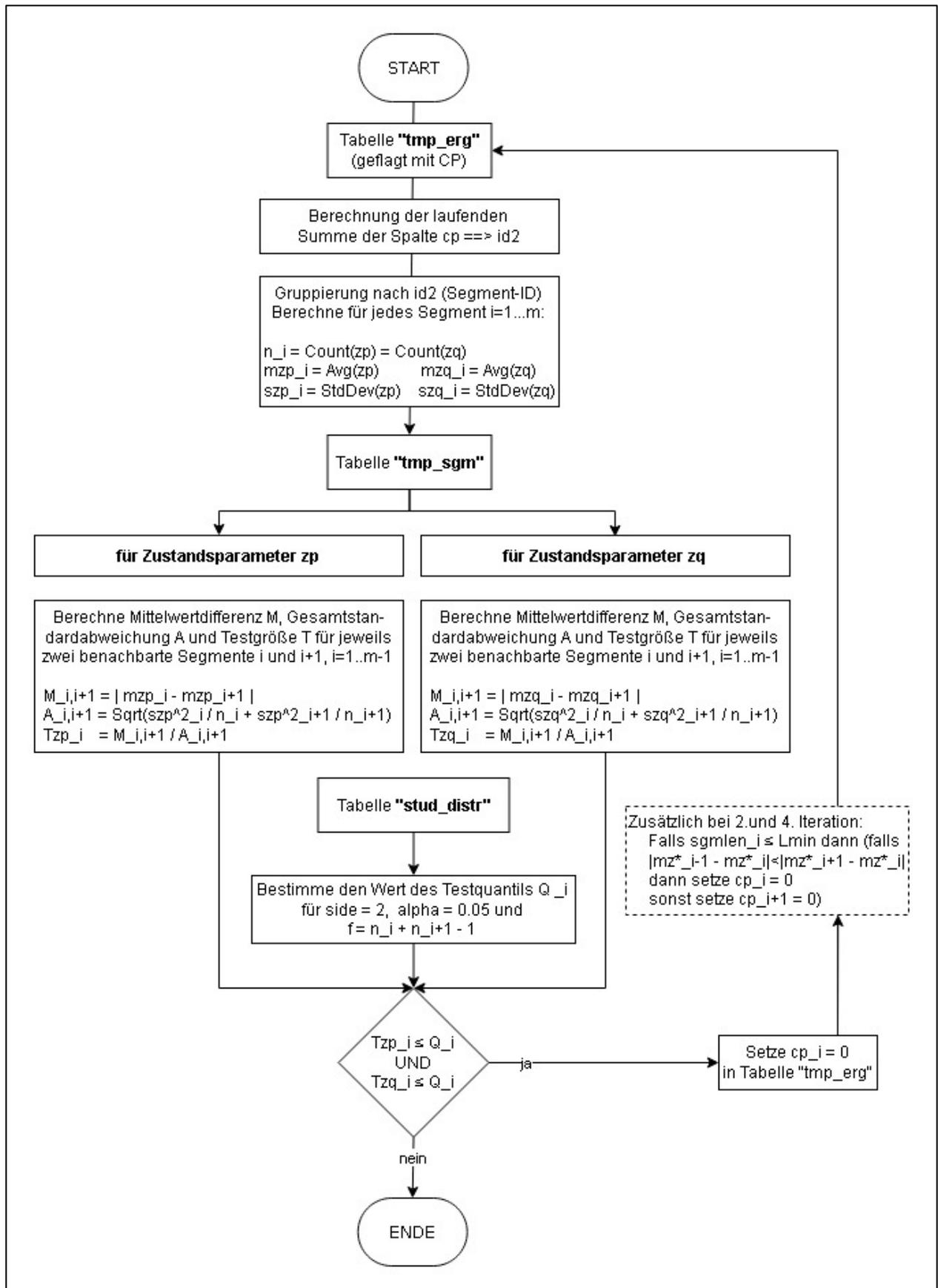


Bild 0.7: Flowchart Changepoint-Verifikation

2.7 Zusammenführung aller Ergebnisse

Sind nach der Changepoint-Verifikation nur noch signifikante Changepoints vorhanden, enthält die Tabelle "tmp_sgm" das Segmentierungsergebnis aller zu segmentierenden Blöcke.

Diese Tabelle muss nun mit der Tabelle "tmp_erg_nosgm" der nicht zu segmentierenden Blöcke (Blöcke, welche die Mindestsegmentlänge nicht übersteigen, und Blöcke ohne Zustandsdaten) zusammengeführt werden. In "tmp_erg_nosgm" bildet ein kompletter Block jeweils ein Segment. Für Blöcke, die Zustandsdaten besitzen, müssen die charakteristischen Werte (Mittelwerte und Standardabweichungen über alle Auswerteabschnitte) berechnet werden. Dies geschieht durch Gruppierung nach der Block-ID und Einfügen in Tabelle "tmp_sgm".

Danach liegt das finale Segmentierungsergebnis in Tabelle "tmp_sgm" vor.

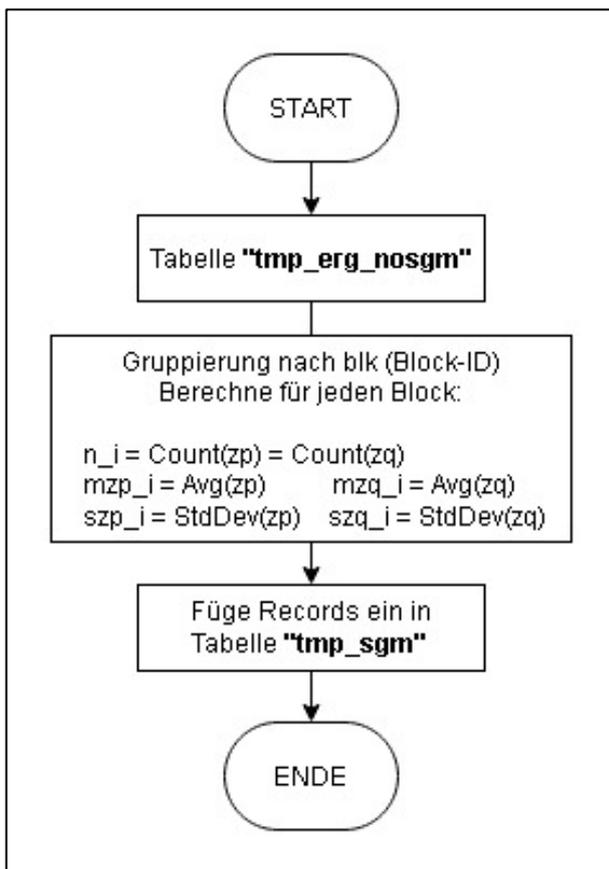


Bild 0.8: Flowchart Zusammenführung aller Ergebnisse

Zur Ausgabe an den Nutzer werden nun noch die ASB-Angaben hinzugefügt. Da es sich bei einem Segment um eine Strecke im Sinne der ASB handelt, müssen Anfang und Ende als zwei Streckenpunkte definiert werden, d.h. jeweils mit AoA und Station. Bei einer abschnittsbezogenen Segmentierung sind Anfangs- und End-AoA jeweils identisch. Die Ausgabetable "obj_sgm" hat folgende Struktur und stellt lediglich einer Erweiterung der "tmp_sgm" um die hellgrau hinterlegten Spalten dar. Die Inhalte dieser Spalten werden durch JOINS mit den Tabellen "net_aoa", "net_blk" und "obj_erg" ermittelt.

Feldname	Feldtyp	Beschreibung
sgm	integer	Segment-ID (1,...,m)
blk	integer	Block-ID des Blocks, zu dem das Segment gehört
minid	integer	Auswertabschnitts-ID des Segmentanfangs
maxid	integer	Auswertabschnitts-ID des Segmentendes
sgmlen	integer	Segmentlänge in m
nosgm	integer	Flag, ob das Segment zu den nicht segmentierten Blöcken gehört
strasse	varchar	Straßenbezeichnung (Klasse + Nummer + ggf. Zusatzbuchstabe)
strname	varchar	Straßenname
fstr	varchar(2)	Fahrstreifen in der Form ..., L2, L1, R1, R2, ...
sort	integer	Abschnittsfolgennummer zu aoa0
aoa0	varchar	AoA, in dem das Segment beginnt
stat0	integer	Station im aoa0, wo das Segment beginnt
aoa1	varchar	AoA, in dem das Segment endet
stat1	integer	Station im aoa1, wo das Segment endet
n	integer	Anzahl Auswertabschnitte im Segment
mzp	numeric	Mittelwert von zp im Segment
szp	numeric	Standardabweichung von zp im Segment
mzq	numeric	Mittelwert von zq im Segment
szq	numeric	Standardabweichung von zq im Segment

Tabelle 0.10 Tabelle "obj_sgm"