



Präsidiatdepartement des Kantons Basel-Stadt

Statistisches Amt

Schutzblech oder Regenschirm?

Wetterbereinigte Zeitreihen des Veloaufkommens im Kanton Basel-Stadt

Statistiktage 2016, Neuenburg

Lukas Mohler, Statistisches Amt des Kantons Basel-Stadt



Warum wetterbereinigte Zeitreihen des Veloverkehrs?

Motivation:

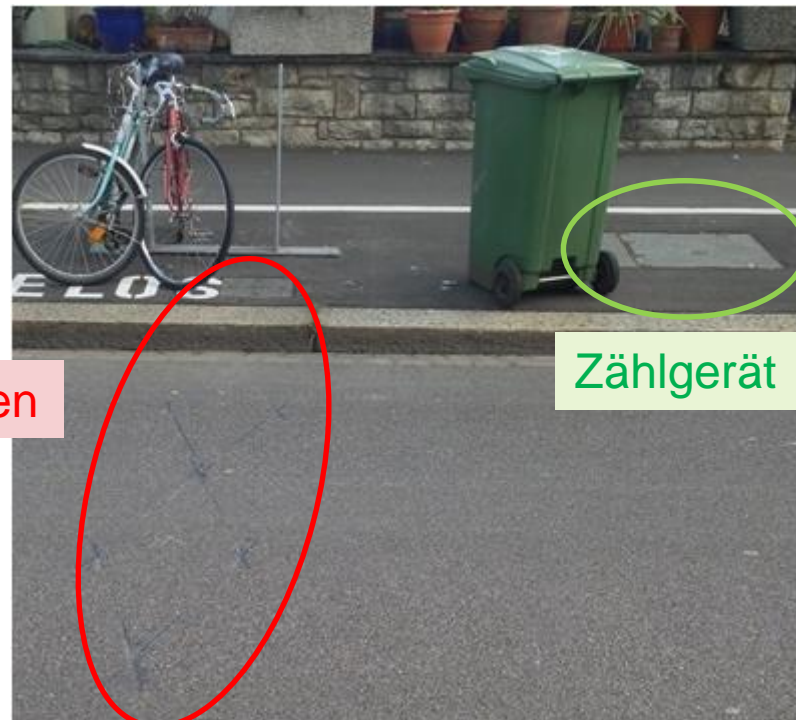
- Im Verkehrsleitbild des Kantons Basel-Stadt ist festgehalten:
 - dass steigender Veloverkehr einerseits die öffentlichen Verkehrsmittel zu Spitzenzeiten entlasten
 - und andererseits einen Teil des motorisierten Individualverkehrs ersetzen soll.
- Das Amt für Mobilität hat die Entwicklungen im Bereich Verkehr zu überprüfen und dem Regierungsrat Bericht zu erstatten.
- Das bedingt eine entsprechende vorausgehende Informationsgewinnung, z.B. die Messung der Verkehrszahlen.



Warum wetterbereinigte Zeitreihen des Veloverkehrs?

Motivation:

- Im Bereich des Veloverkehrs wird dafür unter anderem auf ein Netz von über 20 fest installierten Velozählstellen im ganzen Kantonsgebiet zurückgegriffen.





Warum wetterbereinigte Zeitreihen des Veloverkehrs?

Motivation:

- Das Amt für Mobilität erstellt daraus einen jährlichen Index der Velofrequenzen. Allerdings wurde folgendes festgestellt:
 - Der Indexstand schwankt von Jahr zu Jahr stark.
 - Eine Betrachtung von Wetterdaten weist darauf hin, dass diese Schwankungen wetterbedingt sein könnten.
- Es resultierte das Bedürfnis
 - mehr über den Wettereinfluss auf die Velofrequenzen zu erfahren
 - und einen Velo-Index konstruieren zu können, welcher bei einem «gleich bleibenden» Wetterverhalten zustande kommen würde.

➤ **Auftrag an das Statistische Amt Basel-Stadt**



Analysestrategie

Regressionsanalyse

- mit geeignetem Schätzer,
- welcher separat für jede Zählstelle und jede Fahrtrichtung
- die täglichen Velofrequenzen
- mittels Wettervariablen und anderen erklärenden Faktoren
- zu beschreiben versucht.

Es resultieren

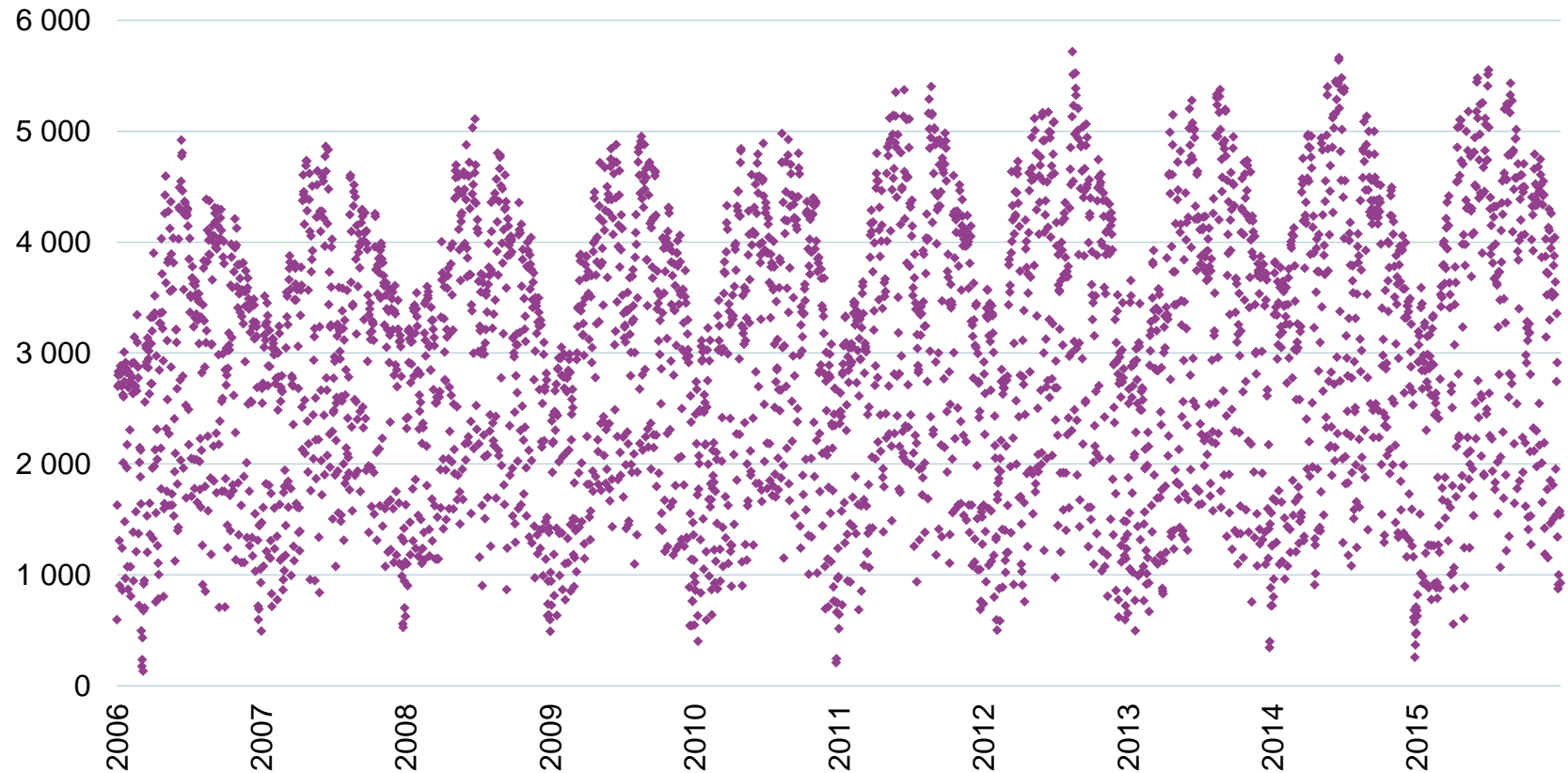
- geschätzte Koeffizienten, welche die Stärke des Wettereinflusses beschreiben
- sowie Voraussagen, wie die Velofrequenz mit «konstantem» Wetter ausgesehen hätte.



Datengrundlage

Abhängige Variable: Velofrequenzen an über 20 Zählstellen.

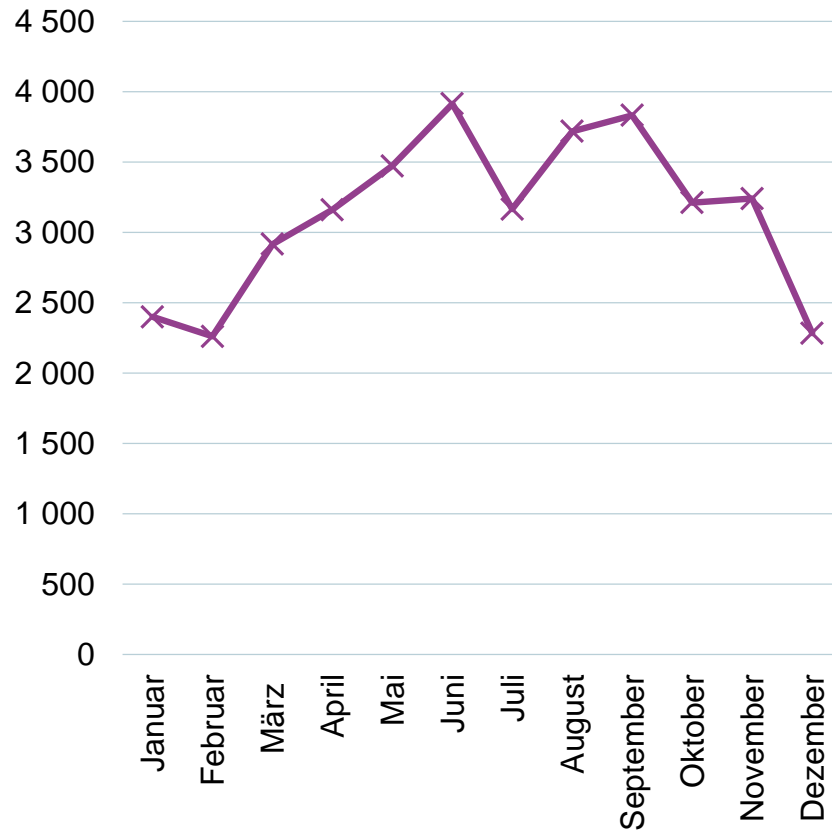
Grafik: Abbildung der Tagesfrequenzen (Wettsteinbrücke, Fahrtrichtung 1):



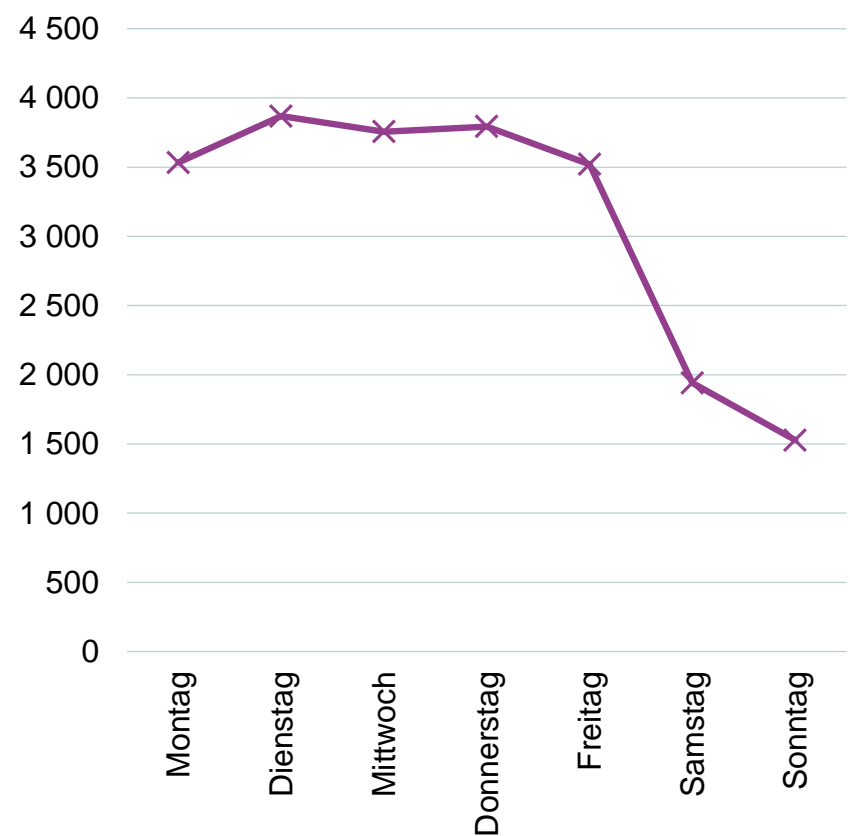


Datengrundlage

Jahresganglinien:



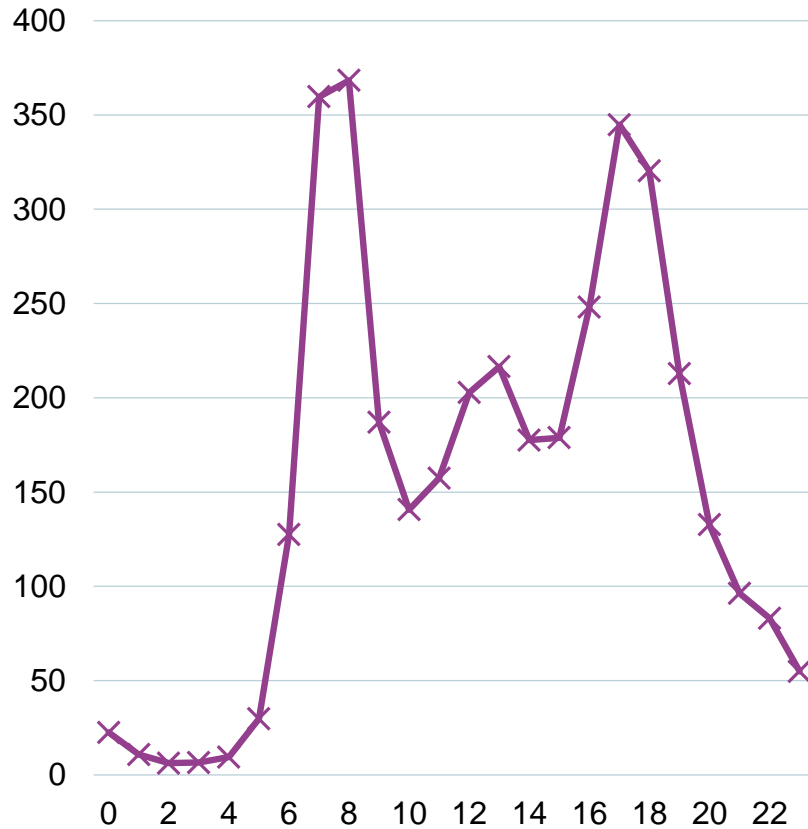
Wochenganglinien:



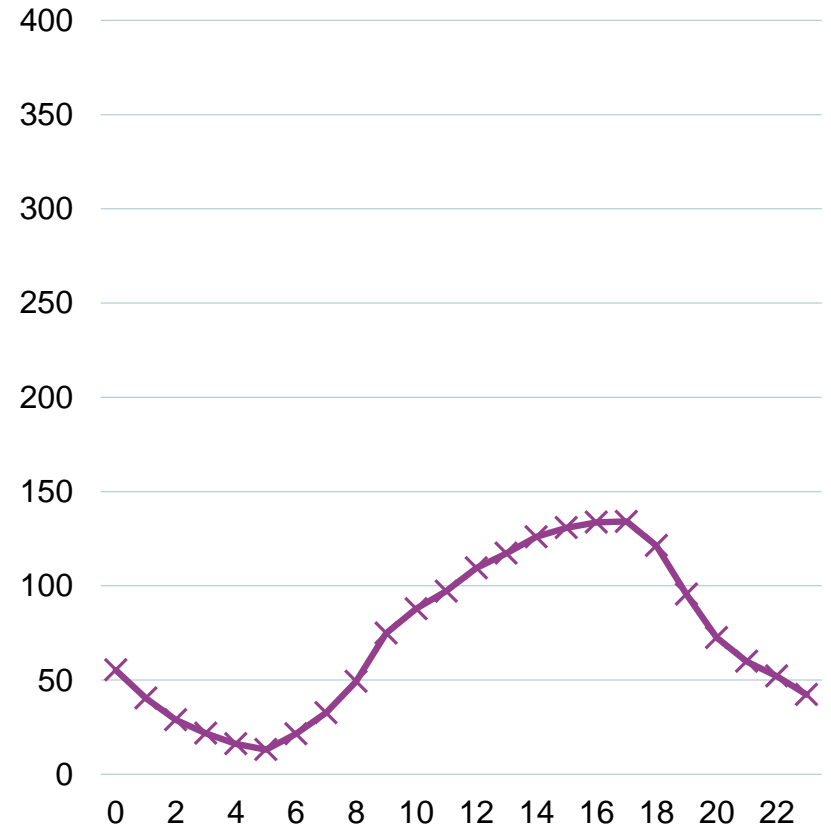


Datengrundlage

Tagesganglinien Mo-Fr:



Tagesganglinien Sa-So:





Datengrundlage

Erklärende Variablen:

- Verfügbare Wetterdaten für Basel-Stadt:

Typ	Frequenz
Sonnenscheindauer	stündlich
Regenmenge	stündlich
Mittlere Temperatur	stündlich
Maximale Temperatur	stündlich
Minimale Temperatur	stündlich
Maximale Windstärke	stündlich
Schneehöhe morgens	täglich
Schneehöhe abends	täglich



Datengrundlage

Erklärende Variablen:

- Eigene «Veranstaltungsdatenbank» Basel-Stadt (Auswahl):

Feiertage, Messen	Sport
Feiertage	Fussballspiele, Eishockeyspiele
Ferien	Swiss Indoors, TopVolley, Badminton, CSI
Semesterferien	Stadtlauf, Marathon, Birslauf, Bruggelauf
Grosse Messen und Kongresse	Quer durch Basel, Waldlauf Lange Erlen
	Rheinschwimmen, SlowUp
Kultur	Sonstiges
Tattoo, Baloise Session, Bscene	Herbstmesse, Weihnachtsmarkt
Grosse Konzerte (St.Jakob)	Verkaufsoffene Sontage
Summerblues, Bebbi si Jazz, Jazzfestival	Feuerwerke (Silvester, August)
Imagine, JKF, Im Fluss, Basel Open Air	Sportnacht, Museumsnacht
Fasnacht, Schulfasnacht, Bummelsonntage	Vogel Gryff, Klosterbergfest, Hafenfest



Methodisches Vorgehen

Variablen für das Modell zur Erklärung der Velofrequenzen:

- Abhängige Variable:
 - Tägliche Velofrequenz
- Erklärende Variablen:
 - Tagesmerkmale: Wochenende, Monat, Feiertag, Ferien, Veranstaltungen
 - Wetterdaten: Sonnenscheindauer, Regenmenge, mittlere Temperatur, maximale Windgeschwindigkeit, Schneehöhe morgens und abends
 - Transformationen von / Interaktionsvariablen zwischen Wetterdaten sowie zwischen Wetterdaten und Tagesmerkmalen:
 - Beispiele: Interaktion mittlere Temperatur mit Sonnenscheindauer, Interaktion Regenmenge mit Wochenend-Dummy, Temperatur im Quadrat, Wetter am frühen Morgen, gelaggte Werte von Wettervariablen usw.



Methodisches Vorgehen

Verschiedene Herausforderungen:

1. Spezifikation der Schätzgleichung:

- Separate Schätzung für jede Zählstelle: Immer die gleiche Spezifikation?
 - Verschiedenste Möglichkeiten bei der Auswahl der erklärenden Variablen (Interaktionen, Terme höherer Ordnung)
 - Verschiedene Transformationsmöglichkeiten der Variablen (log, sqrt,...)
- Für jede Zählstelle werden Dutzende verschiedene Spezifikationen geschätzt mit unterschiedlichen erklärenden Variablen und unterschiedlicher Transformation der abhängigen Variable.



Methodisches Vorgehen

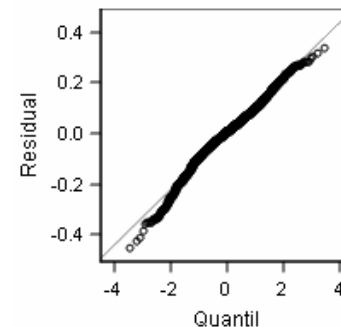
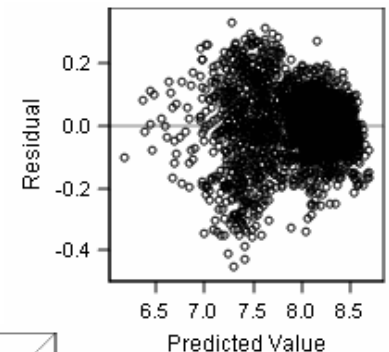
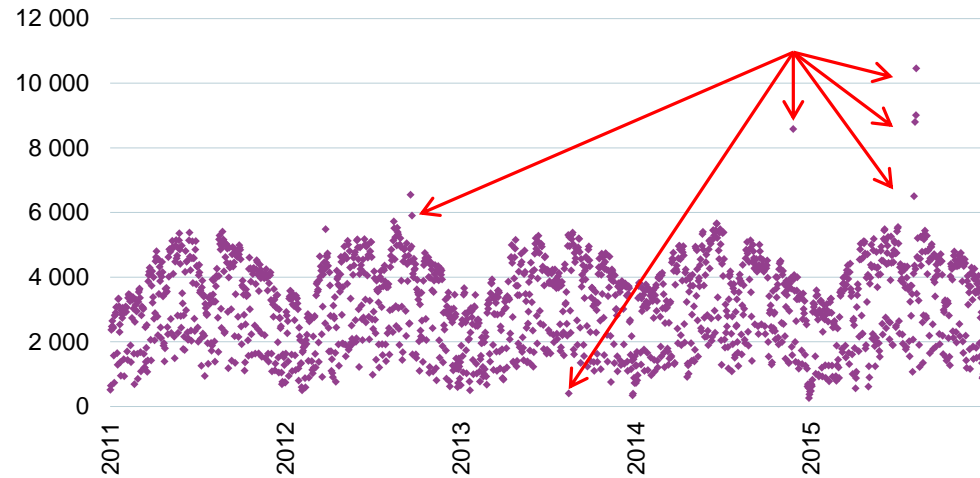
Verschiedene Herausforderungen:

2. Der Datensatz enthält viele potentielle Ausreisser. Lösungsansätze:

- Zusätzliche erklärende Variablen! (=> insb. Veranstaltungen)
- Wahl der Schätzmethode:
 - OLS plus Ausreisser-Identifikation und -Eliminierung
 - Robuste Regression (MM-Estimation)
 - Poisson-Regression, robuste Poisson-Regression

➤ Diagnostik:

- Residual plots, Q-Q-Plots, ...
- Adj. R^2 , AIC, ...
- Vergleich Koeffizienten und Voraussagen





Methodisches Vorgehen

Finale Schätzspezifikation: «so simpel wie möglich, so komplex wie nötig»

- OLS mit Ausreisser-Eliminierung:
 - Diagnostische Tests i.O.
 - Resultierende Koeffizienten und Predictions sehr nahe an robuster Regression- und Poisson-Regression, keine Konvergenzprobleme
 - Einfachste Interpretation der Koeffizienten
- Relativ «einfache» Schätzspezifikation:
 - Log-linear-Modell erfüllt Verteilungsannahmen; Resultate sind einfach interpretierbar.
 - Interaktionen bringen kaum Verbesserung, Beurteilung des Wetteinflusses wäre aber komplizierter
 - Terme höherer Ordnung bringen keine Verbesserung, auch diese würden Interpretation erschweren.



Resultate

Einfluss des Wetters auf die Velobewegungen

- Über 20 Zählstellen, meist zwei Richtungen: Resultate variieren, Muster ist aber überall ähnlich: Sonne +, Temperatur +, Regen -, Schnee -
- Mittelwerte und Mediane der wichtigsten Wetterkoeffizienten über alle Zählstellen und Fahrrichtungen:

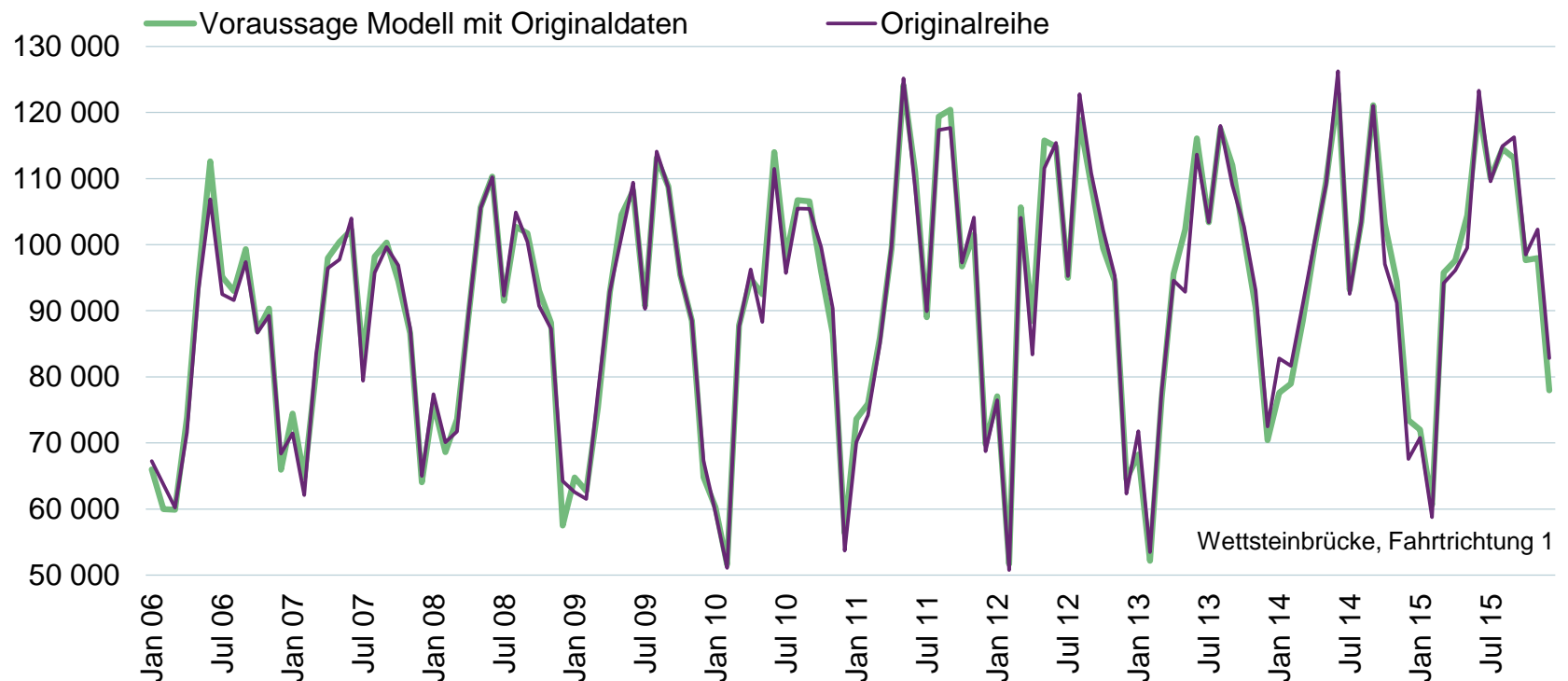
Koeffizient	Mittelwert	Median
Regen in mm	-1.6%	-1.6%
Sonnenscheindauer in Stunden	1.6%	1.5%
Mittlere Temperatur in °C	0.2%	1.0%
Schneemenge morgens in cm	-3.0%	-1.7%
Schneemenge abends in cm	-2.3%	-1.8%
Maximale Windstärke in km/h	0.0%	0.0%



Resultate

Voraussagen des Modells

- Voraussage mit den tatsächlich beobachteten Originaldaten: Wie gut kann unser Modell die beobachteten Werte abbilden? => $R^2 > 90\%$

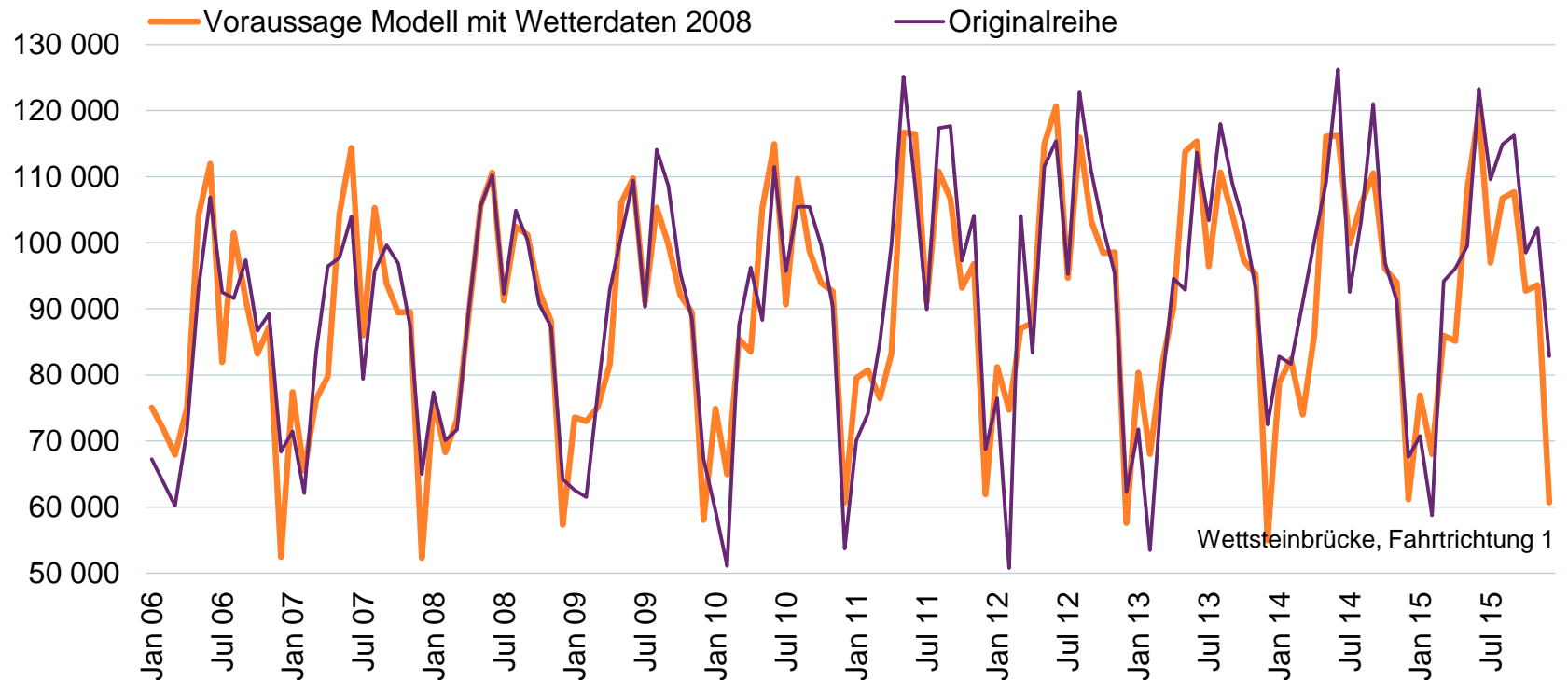




Resultate

Voraussagen des Modells

- Voraussagen des Modells mit «künstlichen» Wetterdaten? => Wetterdaten 2008 für alle Jahre, Wetterverhalten bleibt «konstant».

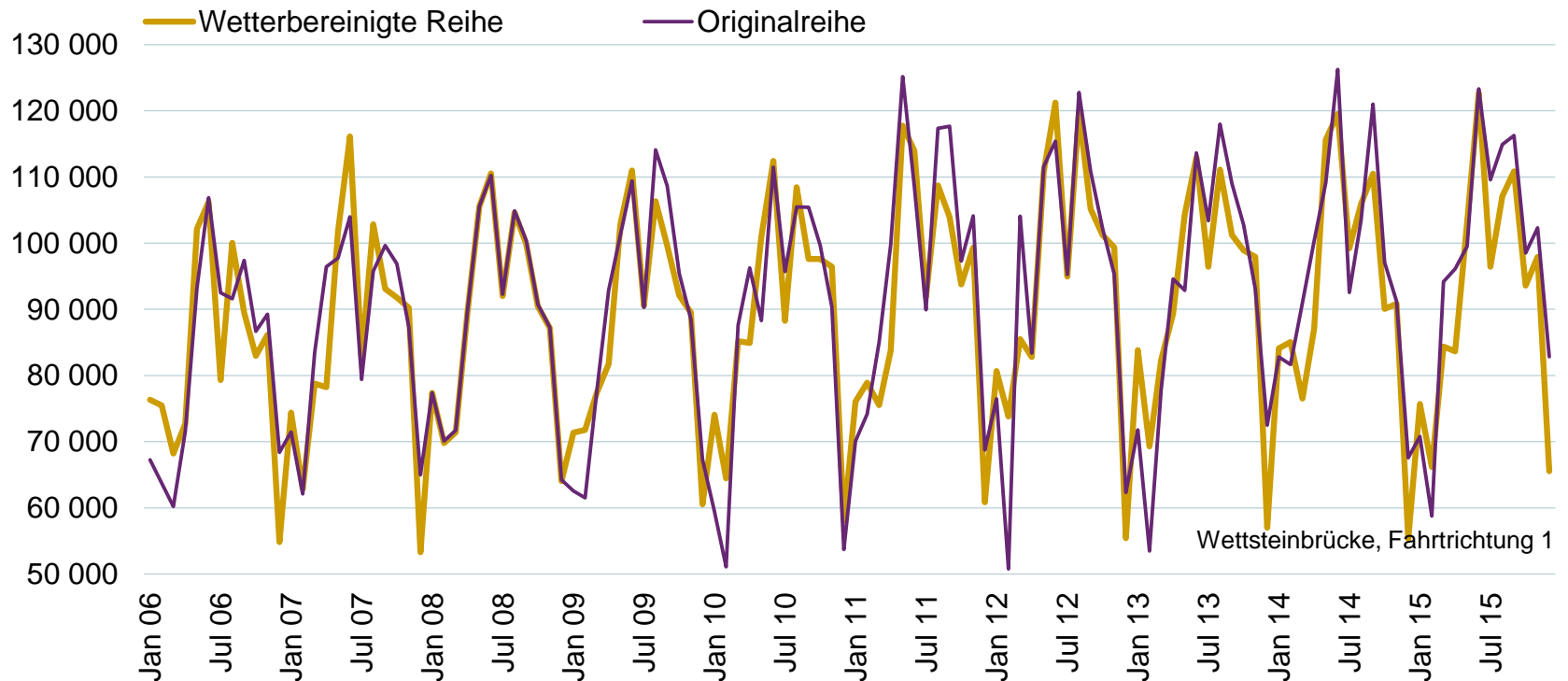




Resultate

Voraussagen des Modells

- Wetterbereinigte Daten =
Originalwerte – (Voraussage Originalwetter – Voraussage Wetter 2008)

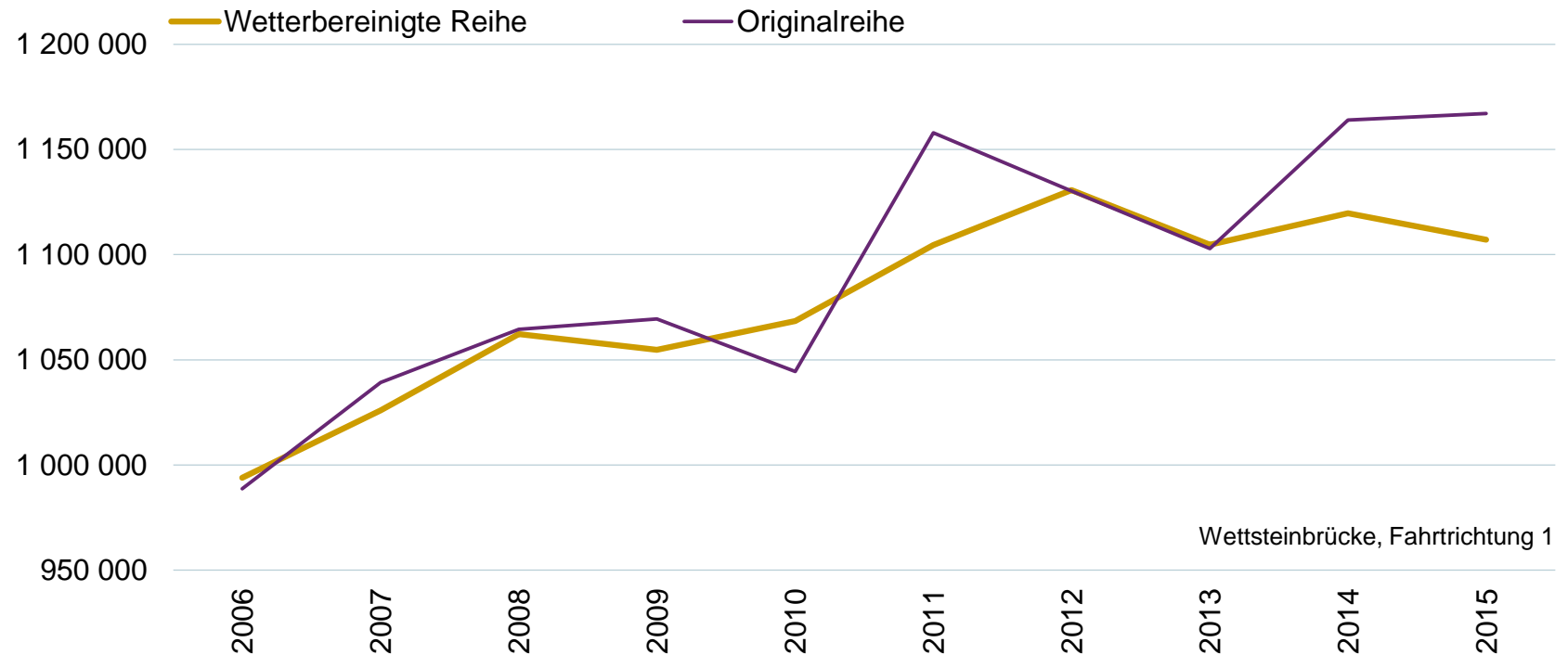




Resultate

Voraussagen des Modells

- Jahresaggregate der wetterbereinigten Zahlen:

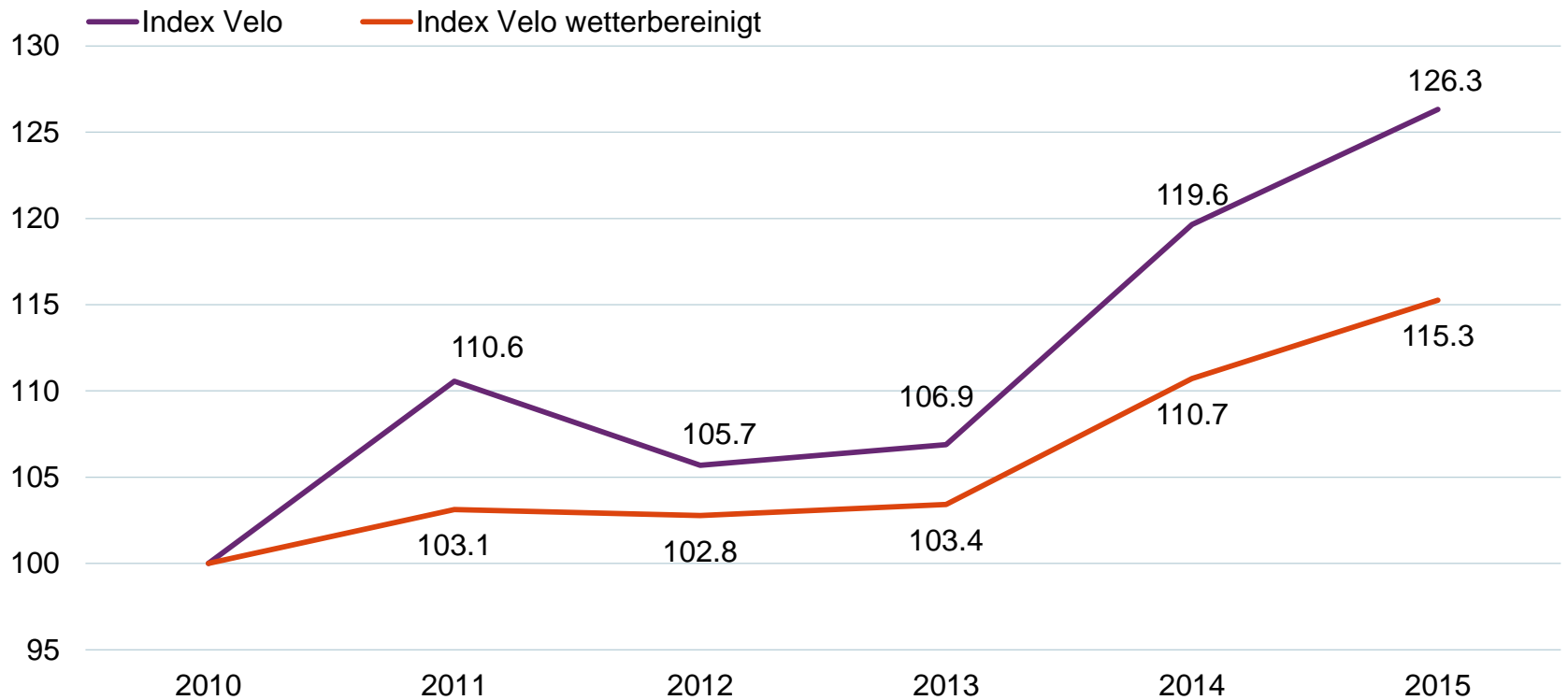




Resultate

Wetterbereinigter Velo-Index

- Verschiedene Zählstellen zu einem Index zusammengefasst:





Zusammenfassung und Schluss

Bedürfnis des Amts für Mobilität nach:

- Informationen über den Wettereinfluss auf die Velofrequenzen und
- Velo-Index, der nicht vom Wetter beeinflusst wird.

Lösungsansatz: Regressionsanalyse, welche die Velofrequenzen mit

- Wetterdaten sowie
- Tagesmerkmalen und Veranstaltungen zu erklären versuchen.

Resultat:

- Mittlerer Einfluss der Wettervariablen für jede Zählstelle
- Wetterbereinigte Zeitreihen für jede Zählstelle