



Emanuel Pichlbauer BSc

Signalqualitätsanalyse am Motorprüfstand

MASTERARBEIT

zur Erlangung des akademischen Grades
Diplom-Ingenieur

Masterstudium Technische Mathematik: Operations Research und Statistik

eingereicht an der

Technischen Universität Graz

Betreuer:

Univ.-Prof. Mag.rer.nat. Dr.rer.nat. Siegfried Hörmann
Institut für Statistik

Dr. DI Gerhard Pirker

DI Michael Wohlthan

DI Doris Schadler

Large Engine Competence Center (LEC)

Graz, Juli 2018

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am

.....

(Unterschrift)

DANKSAGUNG

Zu allererst möchte ich mich bei allen bedanken, die diese Masterarbeit ermöglicht haben. Zu Beginn möchte ich mich bei meinem Betreuer Univ.-Prof. Mag.rer.nat. Dr.rer.nat. Hörmann Siegfried bedanken, der immer wieder mit Ideen und Anregungen für einen stetigen Fortschritt gesorgt hat. Weiters gilt mein Dank Univ.-Prof.i.R. Dipl.-Ing. Dr.techn. Ernst Stadlober, der mich auf diese Arbeit aufmerksam gemacht und auch einen raschen Start ermöglicht hat. Diese Arbeit entstand am LEC (Large Engine Competence Center) unter der Aufsicht von Dr. DI Gerhard Pirker, DI Michael Wohlthän und DI Doris Schädler, die mich aufs herzlichste in der Firma aufgenommen und mir mit Rat und Expertise zur Seite gestanden haben

Mein größter Dank gilt meinen Eltern, die es mir durch ihre stets wohlwollende Unterstützung ermöglicht haben, dieses Studium zu absolvieren.

Abschließend möchte ich mich bei meinen Freunden und Studienkollegen bedanken, die mich sowohl in universitären Angelegenheiten als auch im privaten Bereich begleitet und für das richtige Verhältnis von Arbeit und Spaß gesorgt haben.


ABSTRACT

Due to increasing pressure on the economic factors time and costs concerning engine development, the requirements for measurement data quality gain more and more importance. Therefore a quality check, which also includes a fault detection, right at the engine test bench is indispensable. Since only measurement data that rest within a stationary state can be used for comparisons and evaluations, it is necessary to detect such a state in the first place. The method that is used in this master thesis uses a simple linear regression in a moving window. In this window, the slope of the line of best fit and its corresponding p-value are calculated and evaluated. Furthermore there will also be a target/actual performance, if there are any given values available. In the last instance the measurement data are evaluated. This evaluation contains a three step diagnosis that values the quality. This master thesis has been written in cooperation with the Large Engines Competence Center (LEC) under supervision of DI Dr. Gerhard Pirker, DI Doris Schadler and DI Michael Wohlthan.

KURZFASSUNG

Mit steigendem Kosten- und Zeitdruck in der Motorentwicklung wachsen die Anforderungen an die Messdatenqualität. Daher ist es unerlässlich eine Überprüfung dieser Qualität inklusive Fehlerdiagnose direkt am Motorprüfstand durchzuführen. Da die Messdaten nur in einem stationären Zustand zu Vergleichen und Bewertungen herangezogen werden können, muss vorher geprüft werden, ob ein stationäres Verhalten vorliegt. Im Rahmen dieser Arbeit wird ein Verfahren entwickelt, welches Stationarität beziehungsweise Beharrung eines Signals überprüft sowie etwaige Ausreißer detektieren kann. Für die Beharrungserkennung wird eine einfache lineare Regression über ein gleitendes Zeitfenster durchgeführt und die darin beschreibende Steigung der Ausgleichsgerade beziehungsweise der dazugehörige p-value bewertet. Falls Sollwerte zur Verfügung stehen, wird auch ein Soll-Ist-Vergleich durchgeführt. Zuletzt muss noch eine Überprüfung der Messpunkte vorgenommen werden. Die Evaluierung beinhaltet ein 3-Schritte Verfahren, welches den Messpunkt bewertet. Die Masterarbeit entstand mit dem Large Engines Competence Center (LEC) unter der Betreuung von DI Dr. Gerhard Pirker, DI Doris Schadler und DI Michael Wohlthan.

Inhaltsverzeichnis

1. Einleitung	1
1.1. Der Motorprüfstand	1
1.2. Aufgabenstellung	4
2. Daten	7
2.1. Messgrößen	7
3. Statistische Grundlagen	11
3.1. Explorative Datenanalyse	11
3.1.1. Definition und Erklärung der Autokorrelationsfunktion	18
3.2. Das einfache lineare Modell	31
3.3. Kleinste Quadrate Schätzer	31
3.4. Hypothesentests	32
3.5. Der p-Value	33
3.6. Der lm-Befehl des Statistik Programms 	33
4. Methoden der Beharrungserkennung	35
4.1. Die Methode von Wu, Chen, Zhou und Li	36
4.2. Eine neue Methode für den Motorprüfstand	46
4.3. Das Changepoint Problem	50
4.4. Korrelation & Tests	53
5. Die Messfreigabe	61
5.1. Der Soll-Ist-Vergleich	62
5.2. Die Beharrungserkennung	63
5.2.1. Probleme dieser Methode	63
5.2.1.1. Fehlinterpretation der p-Values	63
5.2.1.2. Kontinuierlicher schwacher Trend	65
5.2.1.3. Zyklische Daten	66
5.2.1.4. Variation der Streuung	67

5.2.2.	Lösung der Probleme	67
5.2.2.1.	Fehlinterpretation der p-Values	67
5.2.2.2.	Kontinuierlicher schwacher Trend	69
5.2.2.3.	Zyklische Daten	70
5.2.2.4.	Variation der Streuung	72
5.2.3.	Vorgegebene Streuung	75
5.2.4.	Ausreißer Detektion	77
5.3.	Die programmierte Beharrungserkennung	79
6.	Die Messpunktdiagnose	87
6.1.	Möglichkeiten zur Bewertung allgemeiner Daten	87
6.1.1.	Die Varianzanalyse - ANOVA	87
6.1.2.	Residuenanalyse - Jackknife Residuen	91
6.2.	Bewertungsschema eines Messpunkts	95
6.2.1.	Schritt 1: Soll-Ist-Vergleich	96
6.2.2.	Schritt 2: Überprüfung der Steigung	97
6.2.3.	Schritt 3: Ausreißer Detektion	98
7.	Resümee und Ausblick	103
A.	Pseudocode & Funktionen	105
	Literaturverzeichnis	108

Abbildungsverzeichnis

1.1. Der Motor im Prüfraum WIMMER (2014)	2
1.2. Überwachungsraum mit Blick auf Prüfraum WIMMER (2014)	2
1.3. Schematischer Aufbau eines Motorprüfstandes WIMMER (2014)	3
1.4. Motivationsbeispiel	5
3.1. T2strich ohne und mit Variation	12
3.2. AbgNox05mg ohne und mit Variation	13
3.3. Boxplots der stationären Bereiche von "T2strich" ohne Variation	15
3.4. Histogramme der stationären Bereiche von "T2strich" ohne Variation	16
3.5. QQ-Plot für "T2strich" ohne Variation	17
3.6. Stationärer Bereich von "T2strich"	18
3.7. Verlauf der Steigung	19
3.8. Korrelation für verschiedene "lags"	20
3.9. Histogramme der stationären Bereiche von "T2strich" mit Variation	21
3.10. Boxplots der stationären Bereiche von "Abgnox05mg" ohne Variation	22
3.11. QQ-Plot für "AbgNox05mg" ohne Variation	23
3.12. Stationärer Verlauf der Messgröße "AbgNox05mg"	24
3.13. Boxplots der stationären Bereiche von "Abgnox05mg" mit Variation	25
3.14. Korrelationsmatrix bezüglich Pearson	28
3.15. Korrelationsmatrix bezüglich Spearman	28
3.16. Scatterplots der hoch korrelierten Daten	30
4.1. Changepoints und ihre Liniensegmente aus WU et al. (2014)	37
4.2. Generischer Partikelfilter Algorithmus aus WU et al. (2014)	40
4.3. Algorithmus für schichtweises Sampling aus WU et al. (2014)	42
4.4. Partikelfilter Algorithmus aus WU et al. (2014)	43
4.5. Erkannte Beharrung der Datenzeitreihe	44
4.6. Erkannte Beharrung der Datenzeitreihe	45
4.7. Verwendete Methode auf echten Daten	46
4.8. Datenverlauf und die dazugehörige p-Value Zeitreihe	48

4.9. Erkannte Beharrung der Datenzeitreihe	49
4.10. Analysierter Change point	52
4.11. Kritische Werte der 95% Quantile	57
4.12. Anteil an fälschlicherweise verworfenen H_0	58
4.13. Anteil an verworfenen H_0	59
5.1. Übersicht der Vorgehensweise	61
5.2. Soll-Ist-Überprüfung des Datenverlaufs	62
5.3. Datenzeitreihe die zu einer Fehlinterpretation führt	64
5.4. Schwacher aber anhaltender Trend	65
5.5. Zyklischer Datenverlauf mit weißem Rauschen	66
5.6. Anstieg der Daten führt zu falsch berechneter Streuung	68
5.7. Anstieg der Datenwerte mit Vergleichswert	70
5.8. Zyklische Daten ohne durchgehender Beharrung	71
5.9. Zyklische Daten mit durchgehender Beharrung	72
5.10. Beharrungserkennung der einzelnen α 's	73
5.11. Verhalten der α 's bei konstanter geringer Streuung	74
5.12. Verhalten der α 's bei großer Streuung	75
5.13. Datensituation: Spezialfall erkannt	77
5.14. Beispiel für eine Ausreißer Detektion	79
5.15. Schematischer Überblick der 4 möglichen Fälle	80
5.16. Die 4 Fälle mit ihren Eigenschaften	83
5.17. Parameter T2strich ohne Sollwert	84
5.18. Parameter T2strich mit Sollwert	84
5.19. Parameter AbgNox05mg ohne Sollwert	85
5.20. Parameter AbgNox05mg mit Sollwert	85
6.1. Beispiel eines Messpunkts mit Ausreißer	89
6.2. Beispiel eines Messpunkts mit kontinuierlichem Anstieg	89
6.3. Beispiel für Versagen von ANOVA	90
6.4. Beispiel eines Messpunkts für Residuenanalyse	92
6.5. Studentisierte Jackknife Residuen des vorherigen Beispiels	93
6.6. Beispiel eines Messpunkts für Residuenanalyse	94
6.7. Studentisierte Jackknife Residuen des vorherigen Beispiels	95
6.8. Beispiel für Versagen eines reinen Soll-Ist-Vergleichs	96
6.9. Bewertung der Steigung	97
6.10. Steigung und Soll-Ist Vergleich ist nicht ausreichend	98

6.11. Gefundene Ausreißer	99
6.12. Beispiele für fehlerfreie Messpunkte	100
6.13. Beispiele für fehlerhafte Messpunkte	101

1. Einleitung

In dieser Einleitung wird die vorliegende Problemstellung erläutert. Da der Motorprüfstand eine zentrale Rolle spielt, wird im speziellen der Aufbau und sein Zweck etwas näher beschrieben. Nähere Details kann man in WIMMER (2014) nachlesen. In den einleitenden Kapiteln von SCHADLER (2016) findet man ebenfalls eine gute Zusammenfassung. Die Anforderungen an solche Prüfstände werden immer größer wie zum Beispiel werden Energieeinsparungen gefordert oder es müssen weitere Emissionsvorschriften eingehalten werden. Deswegen muss auch immer mehr Wert auf Optimierung gelegt werden. Da man geringstmögliche Zeit am Prüfstand verbringen will, müssen die Messungen sehr effizient durchgeführt werden. Aus diesem Wunsch heraus entstand die Aufgabenstellung der Diplomarbeit.

1.1. Der Motorprüfstand

Ein Motorprüfstand besteht aus einem Prüfraum (oder Prüfzelle) und einem angrenzenden Beobachtungsraum. Die Trennwand zwischen den beiden Räumen enthält eine Verbindungstür und üblicherweise ein großes Glasfenster.

In der Prüfzelle befindet sich der Verbrennungsmotor der über eine Verbindungswelle mit der Belastungsmaschine (oder Leistungsbremse) verbunden ist. Diese gesamte Konstruktion ist auf einer schwingungsfähigen Plattform aufgebaut. Das ist notwendig, um den Motor und das Gebäude in dem er sich befindet, schwingungstechnisch zu trennen. Zusätzlich muss der Prüfraum gasdicht sein und nicht unterkellert. Die Gründe für diese Richtlinien sind offensichtlich: die entstehenden Schwingungen des Motors müssen abgefangen werden da sonst Schäden am Gebäude entstehen können. Damit die Gesundheit der Mitarbeiter nicht gefährdet wird, ist die schon erwähnte Dichtheit unerlässlich.

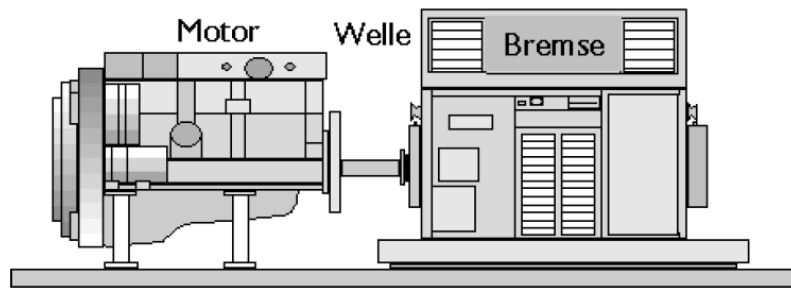


Abbildung 1.1.: Der Motor im Prüfraum WIMMER (2014)

Im Beobachtungsraum befinden sich Überwachungsanzeigen, Bedienfelder und auch das Automatisierungssystem. Die Steuer- und Bedienpulte dienen zur Aufnahme sämtlicher Elemente und Daten. Bis zu 5 Prüfstandscomputer können hierfür simultan zum Einsatz kommen.



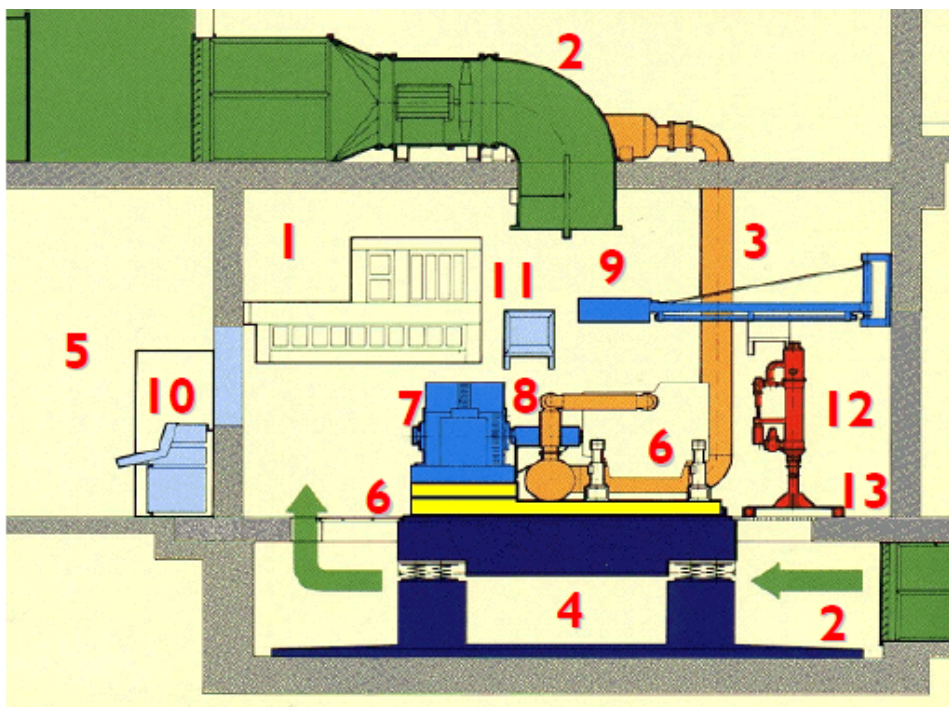
Abbildung 1.2.: Überwachungsraum mit Blick auf Prüfraum WIMMER (2014)

Der Prüfstand verfügt über ein eigenes Zuluft und Abluftsystem. Die verschiedensten Umgebungsbedingungen im Prüfraum erfolgen mittels Raumluftkonditionierung wie Wärmetauschern und Kälteaggregaten. Die Auspuffanlagen sind Teil der Haustechnik und sollen bei möglichst kleiner Lärmentwicklung die Abgasabfuhr sicherstellen und auch spezielle messtechnische Komponenten aufnehmen. Das Schwingfunda-

1.1. Der Motorprüfstand

ment auf dem sich der Versuchsträger (Motor) befindet, besteht üblicherweise aus Stahlguss und ist meistens auf einem luftgefederten Betonfundament mit Niveauregulierung verankert. Die Belastungsmaschine, welche direkt an den Motor über die Kurbelwelle angeflanscht ist, dient als Simulation realer Belastungszustände. Grundsätzlich unterscheidet man zwischen einer hydraulischen Leistungsbremse und einer elektrischen Leistungsbremse. Der Unterschied in den jeweiligen Bauarten liegt in der Übertragung des Drehmoments.

Ein schematischer Aufbau eines Motorprüfstandes sieht wie folgt aus:



Legende:

1	Prüfraum	11	Motorspezifische Messtechnik; z.B.:
2	Be- und Entlüftung		Kraftstoffwaage
3	Auspuffanlage		Drehkolbengaszähler
4	Schwingfundament		Blow-By- Meter
5	Beobachtungsraum		Ölverbrauchsmessgerät
6	Versuchsträger (Motor)	12	Konditioniereinrichtungen; z.B.:
7	Leistungsbremse		Ölkonditionierung
8	Wellenkupplung		Kühlmittelkonditionierung
9	Messgalgen		Kraftstoffkonditionierung
10	Bedienfeld mit: Prüfstandsautomatisierung Abgasanalyse Kalibriergasentnahme	13	Anschluss für Sekundärkühlmittel (Hauswasser)

Abbildung 1.3.: Schematischer Aufbau eines Motorprüfstandes WIMMER (2014)

1.2. Aufgabenstellung

Bei einem Motorprüfstand gibt es eine Vielzahl an überwachten Messgrößen beziehungsweise Parametern, die in einer Messreihe, bestehend aus allen Messgrößen und Messwerten erhoben werden.

Definition 1 *Ein Parameter am Motorprüfstand wird Messgröße genannt.*

Die Temperatur vor dem Motorgehäuse ist zum Beispiel eine Messgröße.

Definition 2 *Eine einzelne Realisierung oder ein einzelner beobachteter Wert einer Messgröße heißt Messwert.*

Definition 3 *Eine Messreihe besteht aus allen Messwerten aller Messgrößen.*

Definition 4 *Ein Messpunkt besteht (in der Regel) aus 150 aufeinanderfolgenden Messwerten aller zu überwachenden Messgrößen. Ein Messpunkt sollte immer in einem stationären Zustand sein. Stationarität bedeutet grob gesagt, dass ein Parameter abgesehen von einer statistischen Variation keine Trends oder Zyklen aufweist. Man spricht in diesem Fall auch von Beharrung*

Der Wunsch der Prüfstandsfahrer besteht nun darin, in einer Messreihe so viele Messpunkte wie möglich zu erheben und zur weiteren Analyse freizugeben. Man ist deswegen an Messpunkten interessiert, da die gesamte Aufzeichnung von Messreihen keinen Vergleich von Betriebszuständen ermöglicht. Die folgende Grafik veranschaulicht die Motivation der Arbeit. Ziel ist es den stationären Zustand so früh wie möglich zu erkennen um eine erfolgreiche Messung durchzuführen.

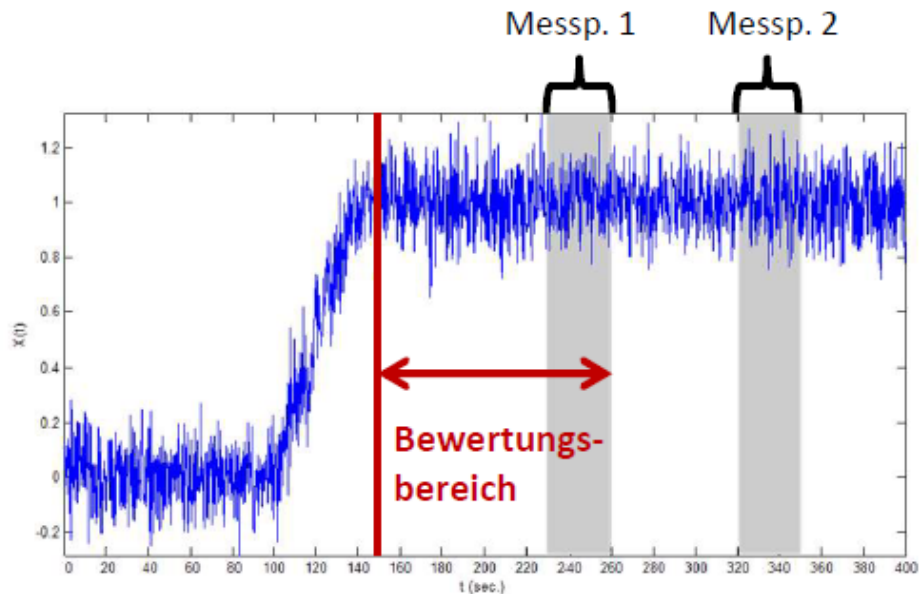


Abbildung 1.4.: Motivationsbeispiel

Dieses Beispiel in Abbildung 1.4 illustriert den Verlauf einer Messgröße und auch zwei ideale Messpunkte. In der Praxis weisen sehr wenige Messgrößen ein so offensichtliches Verhalten auf. Oftmals treten extremere Schwankungen auf oder der Parameter verlässt aus diversen Gründen den üblichen Wertebereich. Wenn diese Fälle während einer Messung auftreten, ist der erhaltene Messpunkt unbrauchbar. Ein Messpunkt kann daher auch wieder verworfen werden, sollten die Messwerte des Messpunkts ein nicht stationäres Verhalten aufweisen. Daher ist eine Diagnose des erhaltenen Messpunkts ebenfalls ein wichtiger Teil dieser Arbeit. Nun soll eine Methode entwickelt werden um eben jene Beharrung zu erkennen und anschließend eine Messung durchzuführen die nicht verworfen wird.

Um den wünschenswerten Beharrungszustand zu erkennen, soll im Rahmen dieser Arbeit ein Algorithmus entwickelt werden der auf statistischen Kenngrößen und deren Auswertung beruht. Zuerst wird ein Verfahren entwickelt, welches anschließend auf verschiedensten Ausgangssituationen sowie auch auf simulierten Daten angewendet wird, um die Funktionalität und auch die Güte zu demonstrieren. Zu guter Letzt werden die Messpunkte, die aufgrund der Beharrungserkennung vorgenommen werden, bewertet.

2. Daten

Um die Methode beziehungsweise das Verfahren sinnvoll erarbeiten und testen zu können, werden vom LEC bereits aufgezeichnete Messreihen zur Verfügung gestellt. Diese Messreihen beinhalten zu jeder Messgröße zwischen 60.000 und 450.000 Messwerte.

Zu jeder Messreihe gibt es auch ein Protokoll in dem sämtliche Details verzeichnet sind. Diese Details beinhalten unter anderem die verwendeten Bauteile des Motors wie den Zylinderkopf, sowie einige Einstellungen der Parameter beziehungsweise deren Betriebspunkte. Einen Betriebspunkt kann man sich als vorgegebene Richtwerte eines oder mehrerer Parameter beziehungsweise eine Einstellung eines oder mehrerer Parameter vorstellen. Wenn man einen Betriebspunkt anfährt, werden die Einstellung des Motors so lange angepasst, bis sämtliche Vorgaben des Betriebspunktes erfüllt sind. Man spricht von einer Parametervariation, wenn eine Messgröße verschiedene Messwerte annehmen soll. Des Weiteren sind auf dem Messprotokoll auch etwaige Anmerkungen zu finden, sollten wichtige Ereignisse vor oder während der Messreihe aufgetreten sein wie z.B. der Austausch von Motorkomponenten.

2.1. Messgrößen

Es wurden insgesamt 23 Parameter zur Verfügung gestellt, wovon im Rahmen dieser Arbeit allerdings nur die wichtigsten 9 untersucht und bearbeitet wurden. Diese werden nun kurz aufgelistet und erklärt. Dabei wird die Kurzbezeichnung angegeben, in welcher Einheit die Größe ihre Werte annimmt und auch kurz beschrieben wofür die Messgröße überhaupt steht.

1.
 - Kurzbezeichnung: mGas
 - Angabe in: kg/h
 - Kurzbeschreibung: Kraftstoffmassenstrom. Gibt an wie viel Kraftstoff

2. Daten

sich pro Stunde durch einen Querschnitt bewegt

2.
 - Kurzbezeichnung: mLuft
 - Angabe in: kg/h
 - Kurzbeschreibung: Luftmassenstrom. Gibt an wie viel Luft sich pro Stunde durch einen Querschnitt bewegt

3.
 - Kurzbezeichnung: T2strich
 - Angabe in: °C
 - Kurzbeschreibung: Entspricht der Lufttemperatur vor dem Motor

4.
 - Kurzbezeichnung: p2str
 - Angabe in: bar
 - Kurzbeschreibung: Entspricht dem Luftdruck vor dem Motor

5.
 - Kurzbezeichnung: Mdflansch
 - Angabe in: Nm
 - Kurzbeschreibung: Motordrehmoment. Die Drehwirkung einer Kraft auf einen Körper

6.
 - Kurzbezeichnung: AbgHC05mg
 - Angabe in: mg/mn³
 - Kurzbeschreibung: Unverbrannter Kraftstoff im Abgas auf 5% O₂.

7.
 - Kurzbezeichnung: AbgCO05mg
 - Angabe in: mg/mn³
 - Kurzbeschreibung: Kohlenmonoxid-Konzentration im Abgas auf 5% O₂ korrigiert.

8.
 - Kurzbezeichnung: AbgNox
 - Angabe in: ppm
 - Kurzbeschreibung: Stickoxide im Abgas

9.
 - Kurzbezeichnung: AbgNox05mg
 - Angabe in: ppm
 - Kurzbeschreibung: Stickoxide im Abgas auf 5% O₂ (5% Sauerstoff äquivalent)

Weitere Parameter, die zwar für diese Problemstellung nicht untersucht werden aber trotzdem eine große Bedeutung haben, sind unter anderem die Drehzahl oder λ . Hierbei bezeichnet λ das Luft-Kraftstoff-Verhältnis. Es misst die tatsächlich vorhandene Luftmasse im Verhältnis zur stöchiometrischen Luftmasse. Die stöchiometrische Luftmasse ist jene Masse an Luft die für eine vollständige Verbrennung des Kraftstoffes benötigt wird. Ein Wert $\lambda > 1$ bedeutet daher Luftüberschuss, d.h. es war mehr Luft als benötigt vorhanden. Diese und weitere Größen können genauer in BASSHUYSEN et al. (2015) oder BÖGE et al. (2015) nachgelesen werden. Auch in SCHADLER (2016) wird auf den Seiten 8-10 eine ausgezeichnete Kurzfassung angegeben. Nun wird eine explorative Datenanalyse durchgeführt um das Verhalten der Messgrößen herausfinden und bewerten zu können.

3. Statistische Grundlagen

In diesem Kapitel werden grundlegende Analysen der Daten durchgeführt und im Anschluss die wichtigsten statistischen Überlegungen kurz erklärt. Zu Beginn wird anhand einer explorative Datenanalyse versucht, Erkenntnisse über die Daten zu sammeln die für spätere Überlegungen wichtig sind. Danach erhält man einen Überblick über die benötigten statistischen Werkzeuge, die für diese Arbeit von essentieller Bedeutung sind.

3.1. Explorative Datenanalyse

Bei einer explorativen Datenanalyse geht es um die Bewertung und empirische Analyse von Daten. Für gewöhnlich würde man einen Datensatz mithilfe von Boxplots und Histogrammen darstellen und versuchen das Streuverhalten, die Schiefe, den Wertebereich und sonstige Informationen herauszulesen und zu interpretieren. In diesem Fall ist ein solches Vorgehen nicht sinnvoll da es bereits eine wesentliche Vorinformation gibt; Die Daten entsprechen einer Zeitreihe, sind normalerweise hoch korreliert über die Zeit und werden sogar bewusst auf verschiedenen Wertebereichen variiert. Daher ist ein standardmäßiges Vorgehen nicht möglich und man muss einen etwas anderen Zugang wählen.

Im ersten Schritt ist es sinnvoll, visuell die verschiedenen Zeitreihen zu betrachten. Von den zuvor genannten Größen werden zum Zwecke der Illustration nun zwei ausgewählt und als Zeitreihe dargestellt. Diese beiden Messgrößen werden aufgrund ihres unterschiedlichen Verhaltens ausgewählt und bewertet. Andere Parameter können ähnlich analysiert werden. Für diese beiden ausgewählten Größen ("T2strich" und "AbgNox05mg") gibt es Messreihen mit bewusster Variation des Wertebereichs. Das bedeutet, es wurden absichtlich verschiedenen Wertebereiche $[a_i, b_i]$ angenommen. Daher beurteilt man eine Zeitreihe einmal mit und einmal ohne Variation.

3. Statistische Grundlagen

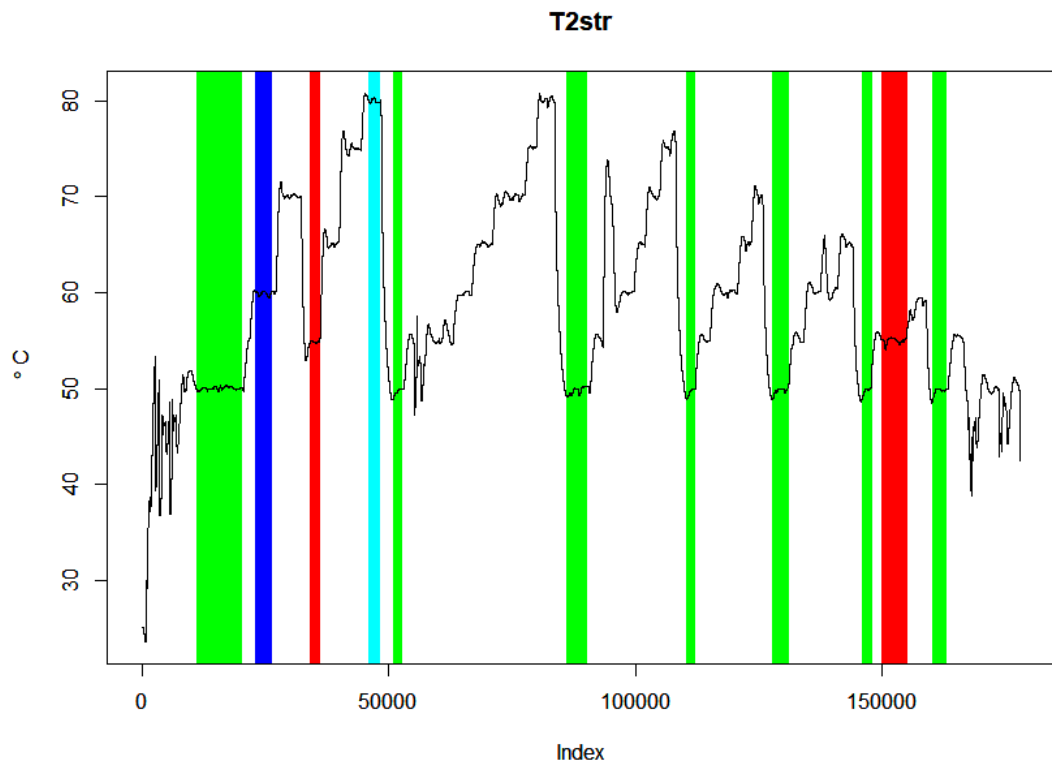
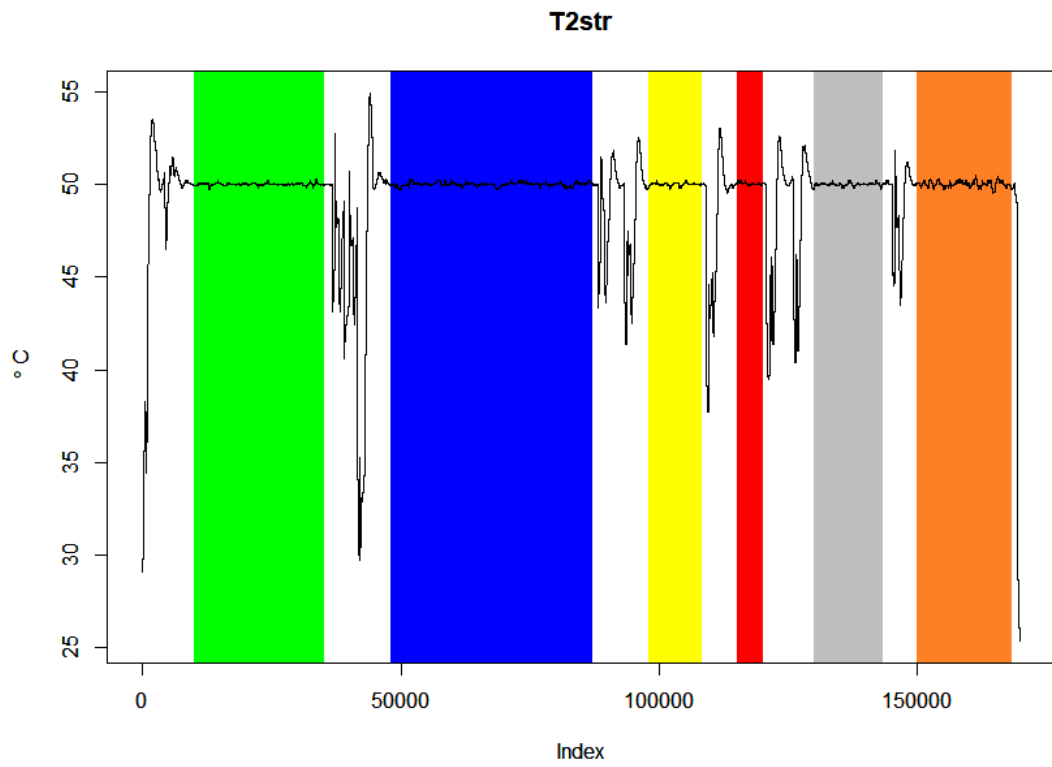


Abbildung 3.1.: T2strich ohne und mit Variation

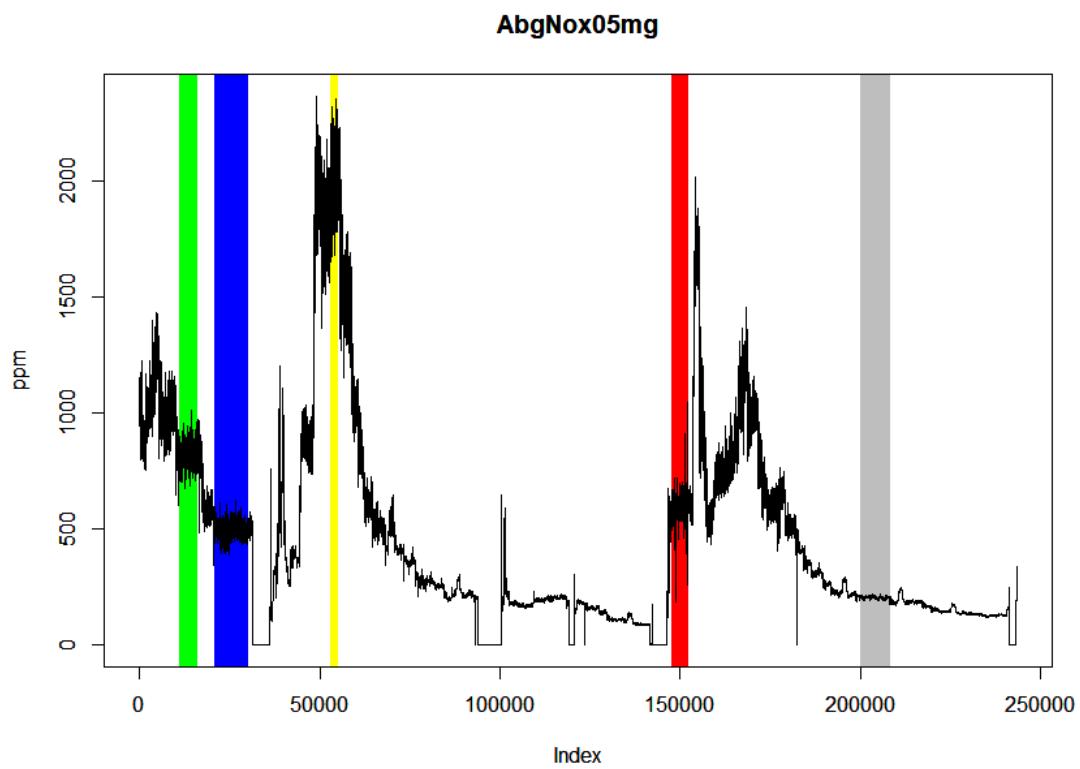
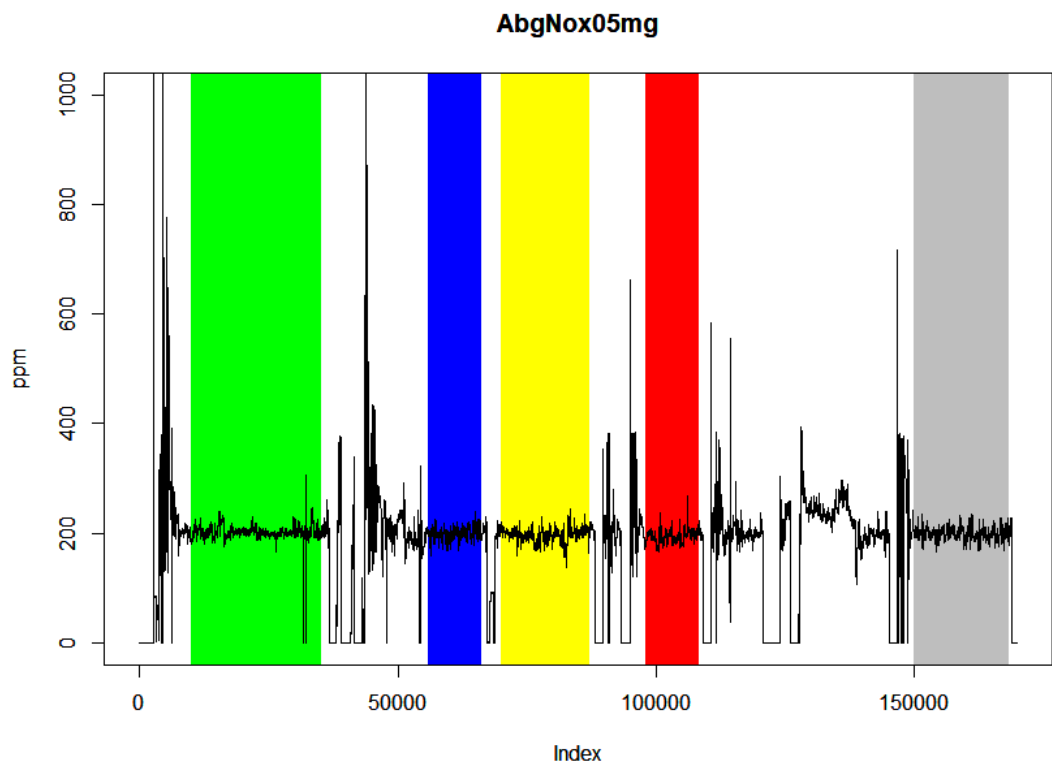


Abbildung 3.2.: AbgNox05mg ohne und mit Variation

3. Statistische Grundlagen

In den folgenden Grafiken wurden die zeitlichen Verläufe dargestellt und die visuell ersichtlichen stationären Bereiche gekennzeichnet. Es wurden aber aus Übersichtsgründen nicht immer alle Bereiche gekennzeichnet, speziell wenn diese variieren. Wir versuchen nun die gesamte Zeitreihe in mehrere kleine Zeitreihen zu unterteilen, in denen Stationarität plausibel erscheint. Erst in eben jenen Bereichen kann eine sinnvolle univariate Datenanalyse durchgeführt werden.

Für die farblich gekennzeichneten Bereiche (und auch in einigen nicht eingezeichneten) erscheint die Annahme von Stationarität gerechtfertigt zu sein. Für eben diese Bereiche kann nun eine univariate Datenanalyse durchgeführt werden. Dabei wird bei den Zeitreihen ohne Variation Wert daraufgelegt, ob sich die einzelnen Bereiche stark voneinander unterscheiden. Bei den Zeitreihen mit Variation werden die Abschnitte mit demselben Wertebereich zusammengefasst und mit den verschiedenen Wertebereichen verglichen.

Ziel dieser Analyse ist es, das Verhalten der Parameter in einem stationären Zustand bewerten zu können. Insbesondere wird dabei Wert auf die Verteilung und auf eine mögliche Unabhängigkeit der Messwerte gelegt. Diese Informationen sind relevant um die jeweilige Methode zur Beharrungserkennung geeignet zu kalibrieren und statistisch zu rechtfertigen. Angenommen in einem stationären Zustand verhalten sich die Messgrößen wie unabhängige normalverteilte Zufallsvariablen, dann werden wir sehen, dass lineare Modelle aus der klassischen Regression dazu verwendet werden können, diese stationären Daten zu beschreiben. Ein weiterer wichtiger Punkt ist das Streuverhalten. Schwanken die Messgrößen in unterschiedlichen Wertebereichen unterschiedlich stark, dann muss darauf Rücksicht genommen werden.

Im Folgenden werden nun die beiden genannten Messgrößen untersucht um Aufschluss über die soeben genannten Verhaltensmuster zu bekommen. Dazu betrachtet man zuerst die Boxplots und die Histogramme um eine mögliche Verteilung erahnen zu können. Idealerweise erhofft man sich eine Normalverteilung welche man im Anschluss mithilfe des Shapiro Wilk Tests und mittels QQ-Plots überprüfen kann.

Zunächst betrachten wir die Boxplots bzgl. der Messgröße "T2strich" als keine Variation vorlag.

3.1. Explorative Datenanalyse

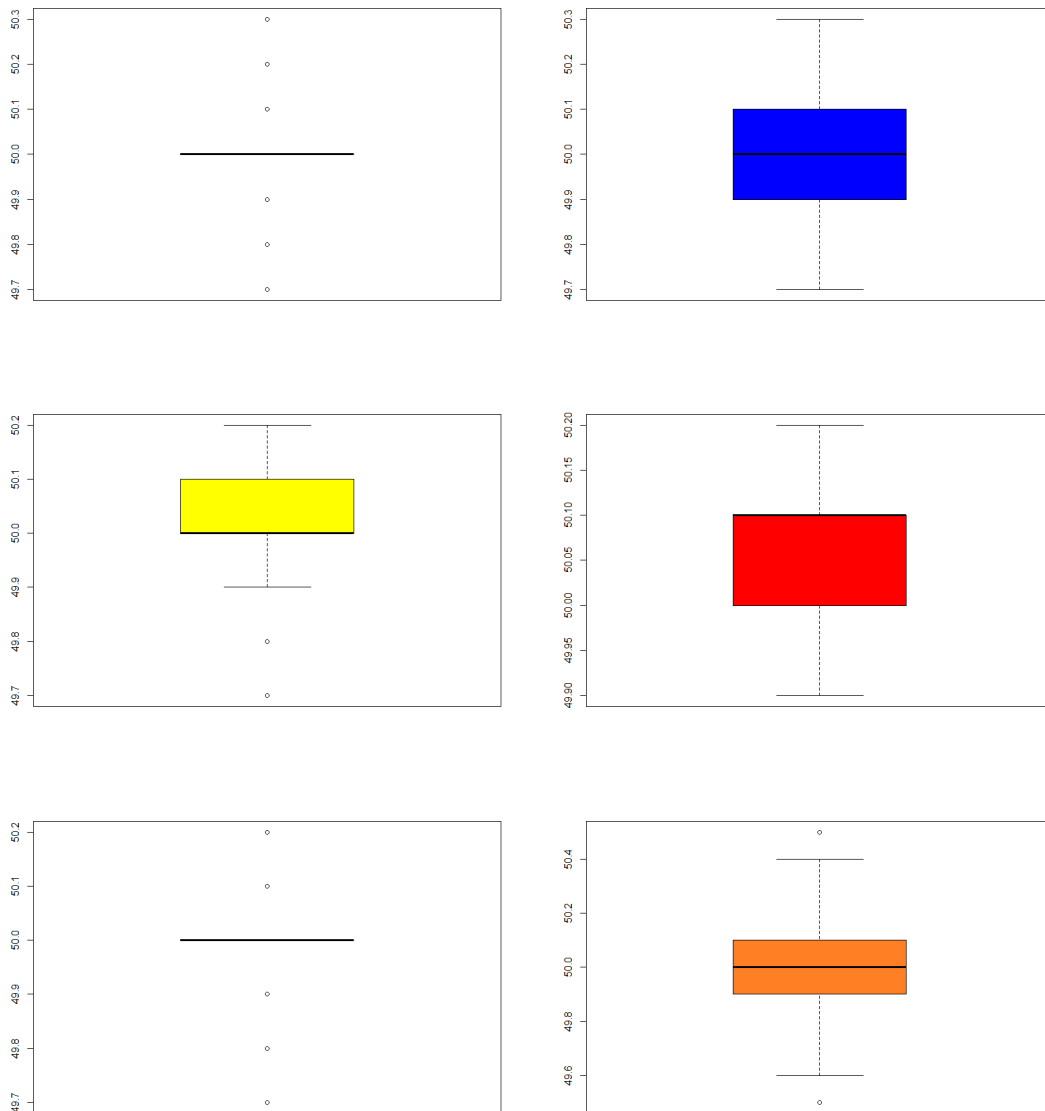


Abbildung 3.3.: Boxplots der stationären Bereiche von "T2strich" ohne Variation

Auf den ersten Blick scheinen die einzelnen Abschnitte ein unterschiedliches Verhalten aufzuweisen. Wenn man allerdings die Boxplots etwas genauer betrachtet, so fällt auf, dass der Median in allen Plots sehr ähnlich ist. Auch die 25% beziehungsweise 75% Quantile unterscheiden sich kaum. Der Grund für das eigenartige Aussehen einiger Boxplots liegt an der Beschaffenheit der Daten. Diese wurden bei dieser Messgröße gerundet. Man hat also diskrete Werte und daher stimmt auch oft das

3. Statistische Grundlagen

25% beziehungsweise 75% Quantil mit dem Median überein. Im ersten und fünften Boxplot stimmen sogar q_{25} und q_{75} überein, was zu einer iqr-Standardabweichung von 0 führt. Der Vollständigkeit halber werden unten auch noch die Histogramme angegeben mit den Kenngrößen \bar{x} und s^2

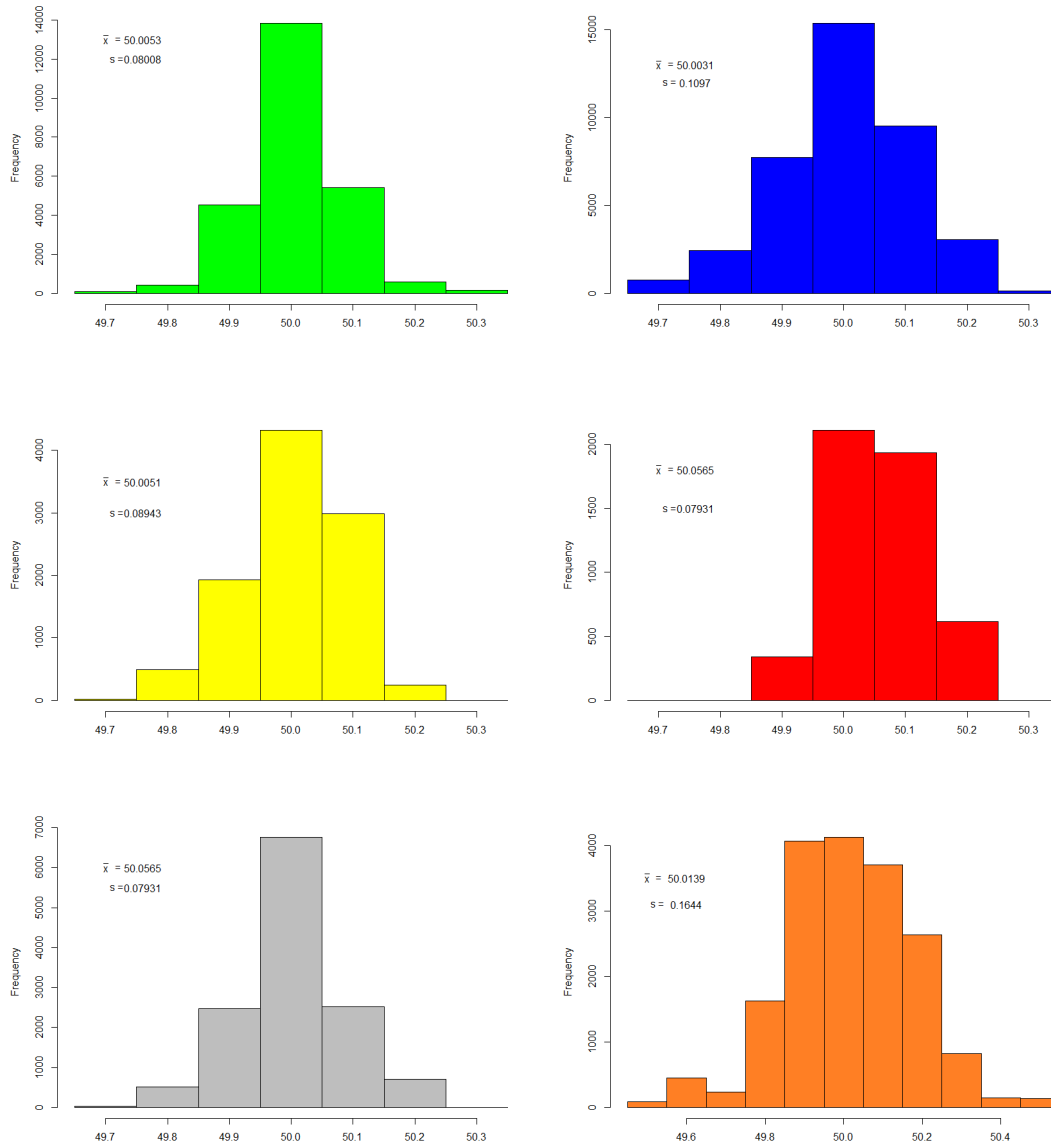


Abbildung 3.4.: Histogramme der stationären Bereiche von "T2strich" ohne Variation

Es kann deutlich erkannt werden, dass das Verhalten des Parameters sich auf diesem

Niveau nicht ändert. Speziell die Streuung ist auffällig ähnlich. Nun betrachtet man die QQ-Plots und die Aussage des Kolmogorov Smirnov Tests für eine Normalverteilung. Der Shapiro-Wilk Test ist aufgrund der riesigen Datenmenge leider nicht verwendbar. Weiters wird nur ein QQ-Plot angeführt und nur einmal das Ergebnis des Tests angegeben, da sich die Ergebnisse nicht voneinander unterscheiden.

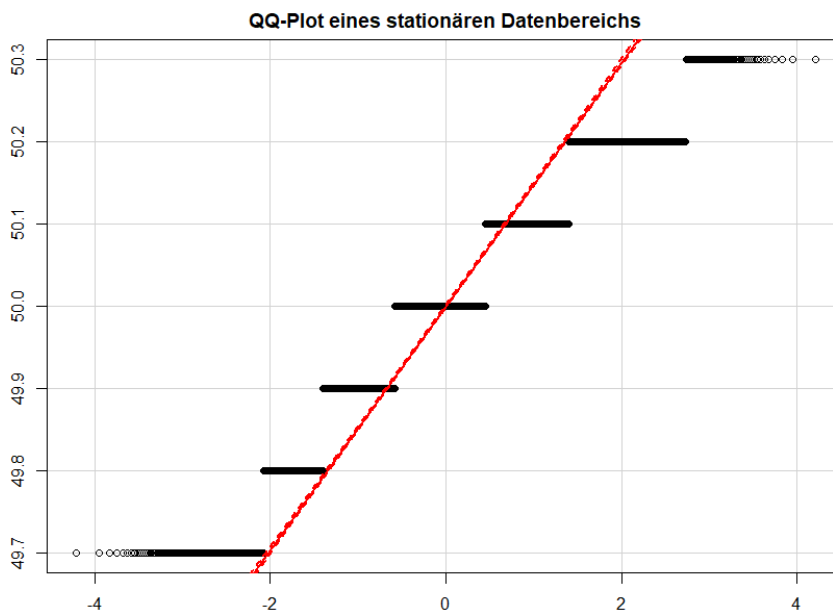


Abbildung 3.5.: QQ-Plot für "T2strich" ohne Variation

One-sample Kolmogorov-Smirnov test

data: Daten

D = 0.2329, p-value < $2.2e-16$

Obwohl die Annahme der Normalverteilung offensichtlich verletzt ist - unsere Beobachtungen sind diskret - ersehen wir aus dem Q-Q-Plot, dass die Daten nicht außergewöhnlich schief oder heavy-tailed sind. Wir erwarten daher, dass die Ergebnisse einer linearen Regression so beurteilt werden können als wenn die Normalverteilungsannahme erfüllt wäre. Die Annahme der Unabhängigkeit hingegen muss verworfen werden: Der Wertebereich des Parameters ist diskret und die einzelnen Werte bleiben sehr lange konstant. Des Weiteren wird ein diskreter Wert immer nur um eine Stufe erhöht oder verringert. Das bedeutet, ist der aktuelle Wert 50, so

3. Statistische Grundlagen

kann der nächste Wert nur 49.9, 50.0 oder 50.1 sein. Zur Illustration betrachten wir einen zeitlichen Verlauf als der Parameter stationär war.



Abbildung 3.6.: Stationärer Bereich von "T2strich"

Diese Grafik verdeutlicht das soeben erklärte Verhalten. Um von einer Unabhängigkeit ausgehen zu können, dürfte kein Muster erkennbar sein, beziehungsweise die einzelnen Werte müssten zufällig angenommen werden. Speziell die letzten 2000 Messwerte sprechen gegen eine solche Annahme. Um den Begriff der Abhängigkeit besser zu verstehen wird nun kurz die Autokorrelation erklärt

3.1.1. Definition und Erklärung der Autokorrelationsfunktion

Sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ gegeben und $T \subset \mathbb{R}$. Genauere Erklärungen eines Wahrscheinlichkeitsraums findet man in HARTMANN (2012). Eine Familie $(X_t)_{t \in T}$ von Zufallsvariablen nennt man einen stochastischen Prozess. Die Autokorrelation beschreibt die Abhängigkeit des Prozesses. Sei $(X_t)_{t \in \mathbb{N}}$ ein stationärer stochastischer Prozess. Für zwei beliebige Zeitpunkte t_1, t_2 ist die Autokovarianzfunktion γ für $t_1, t_2 \in \mathbb{Z}$ gegeben als

$$\gamma(t_1, t_2) = \text{Cov}(X_{t_1}, X_{t_2})$$

3.1. Explorative Datenanalyse

O.B.d.A. sei $t_1 < t_2$ und damit $k = t_2 - t_1 > 0$. $\gamma(k) = \text{Cov}(X_t, X_{t+k})$ beschreibt nun die Korrelation zwischen Elementen des stochastischen Prozesses mit Distanz oder auch "lag" k . Die Autokorrelationsfunktion $\rho(k)$ ist nun definiert als $\frac{\gamma(k)}{\gamma(0)}$. Nun wird kurz ein Beispiel angegeben um soeben erklärte Begriffe zu verdeutlichen. Es wurde von einem Teil einer Messgröße die Steigung von aufeinanderfolgenden Messwerten berechnet. Da die Messwerte hoch korreliert sind, gilt dies auch für die Steigung der Messwerte. Der Verlauf sieht folgendermaßen aus:

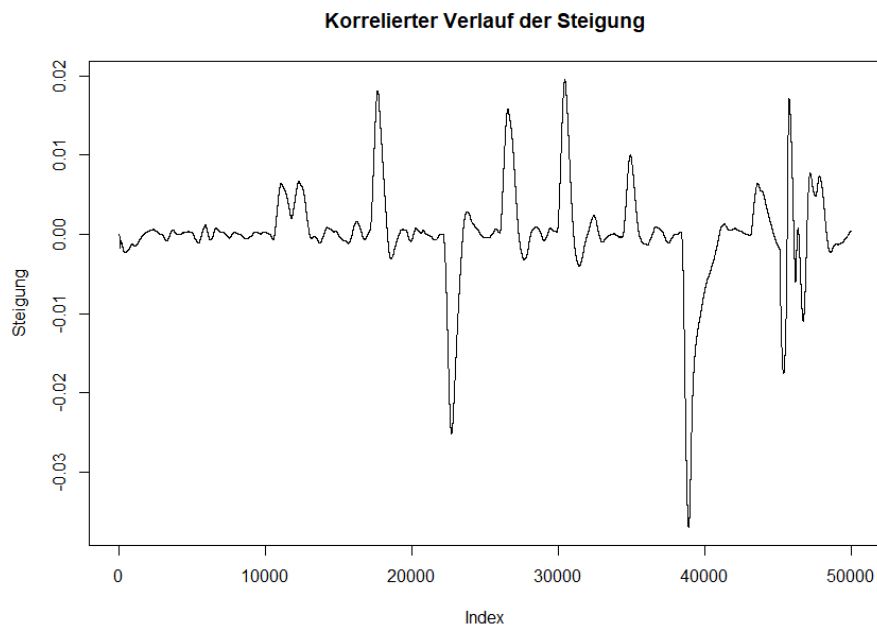


Abbildung 3.7.: Verlauf der Steigung

Es ist deutlich zu erkennen, dass die Daten hoch korreliert sind und einem Art stationären Verlauf folgen. Nun wird die Autokorrelation mit verschiedenen "lags" also mit verschiedenen k 's aufgerufen. Die Werte der Autokorrelation befinden sich im Intervall $[-1, 1]$ wobei große Absolutwerte für hohe Korrelation sprechen.

3. Statistische Grundlagen

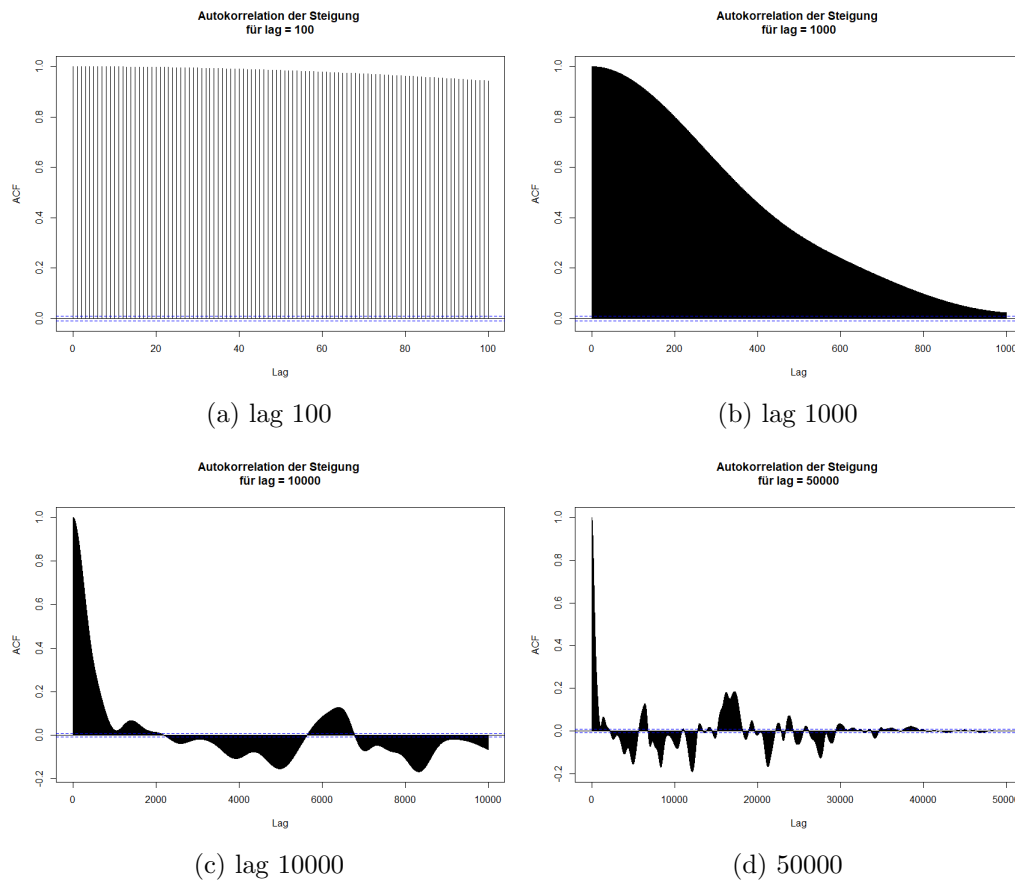


Abbildung 3.8.: Korrelation für verschiedene "lags"

Auf diesen 4 Bildern ist deutlich zu sehen wie stark die Korrelation der Steigung wirklich ist. Eine Distanz von 1 - 101 Werten (lag 100) weist immer noch einen sehr hohen Wert auf. Das bedeutet, dass selbst wenn man 100 Werte entfernt ist, so sind der 1. und der 100. Wert immer noch sehr stark voneinander abhängig. Erst wenn man die Entfernung auf 1000 erhöht nähert sich die Autokorrelation dem wünschenswerten Bereich nahe der 0. Aber erst, wenn man die Distanz wirklich auf den maximalen Wert setzt bekommt man einigermaßen zufriedenstellende ρ Werte. Grob gesagt gibt die Autokorrelation also an, wie viel Distanz zwischen zwei Elementen des stationären Prozesses liegen muss, um die Korrelation der beiden Elemente minimal zu haben.

Nun können wir die explorative Datenanalyse fortsetzen. Vergleicht man nun die einzelnen Histogramme dieser Messgröße auf den einzelnen Stufen der Variation, so erhält man folgende Histogramme.

3.1. Explorative Datenanalyse

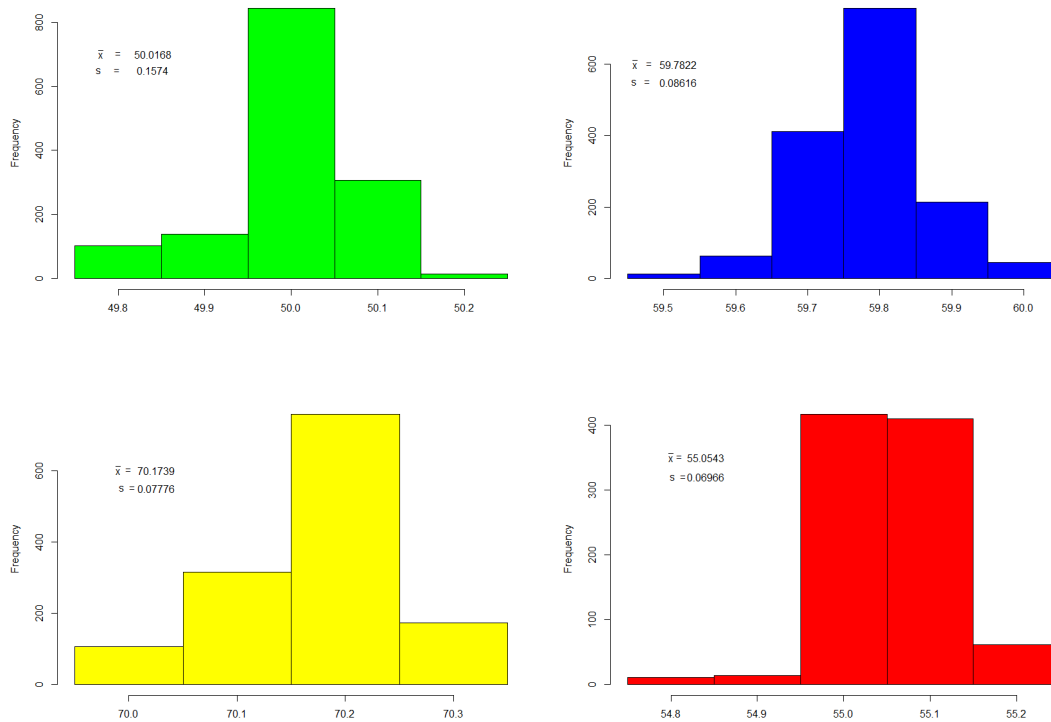


Abbildung 3.9.: Histogramme der stationären Bereiche von "T2strich" mit Variation

Das Verhalten des Parameters ändert sich in den einzelnen Wertebereichen kaum. Sowohl im Verhalten als auch in der Variation hat man in allen 4 Histogrammen ein ähnliches Bild. Daher kann davon ausgegangen werden, dass diese Messgröße eine konstante Streuung hat, unabhängig vom Wertebereich. Versucht man hier ebenfalls auf eine Normalverteilung zu testen, so erhält man sehr ähnliche Resultate wie zuvor. Man kann also annehmen, dass man zumindest für diesen Parameter von einer Normalverteilung ausgehen kann (beziehungsweise von einer Binomialverteilung) wenn ein stationärer Zustand vorliegt.

Dasselbe wird auch mit der zweiten Messgröße durchgeführt. Wiederum betrachten wir zuerst die einzelnen stationären Abschnitte, wenn keine Variation durchgeführt wurde.

3. Statistische Grundlagen

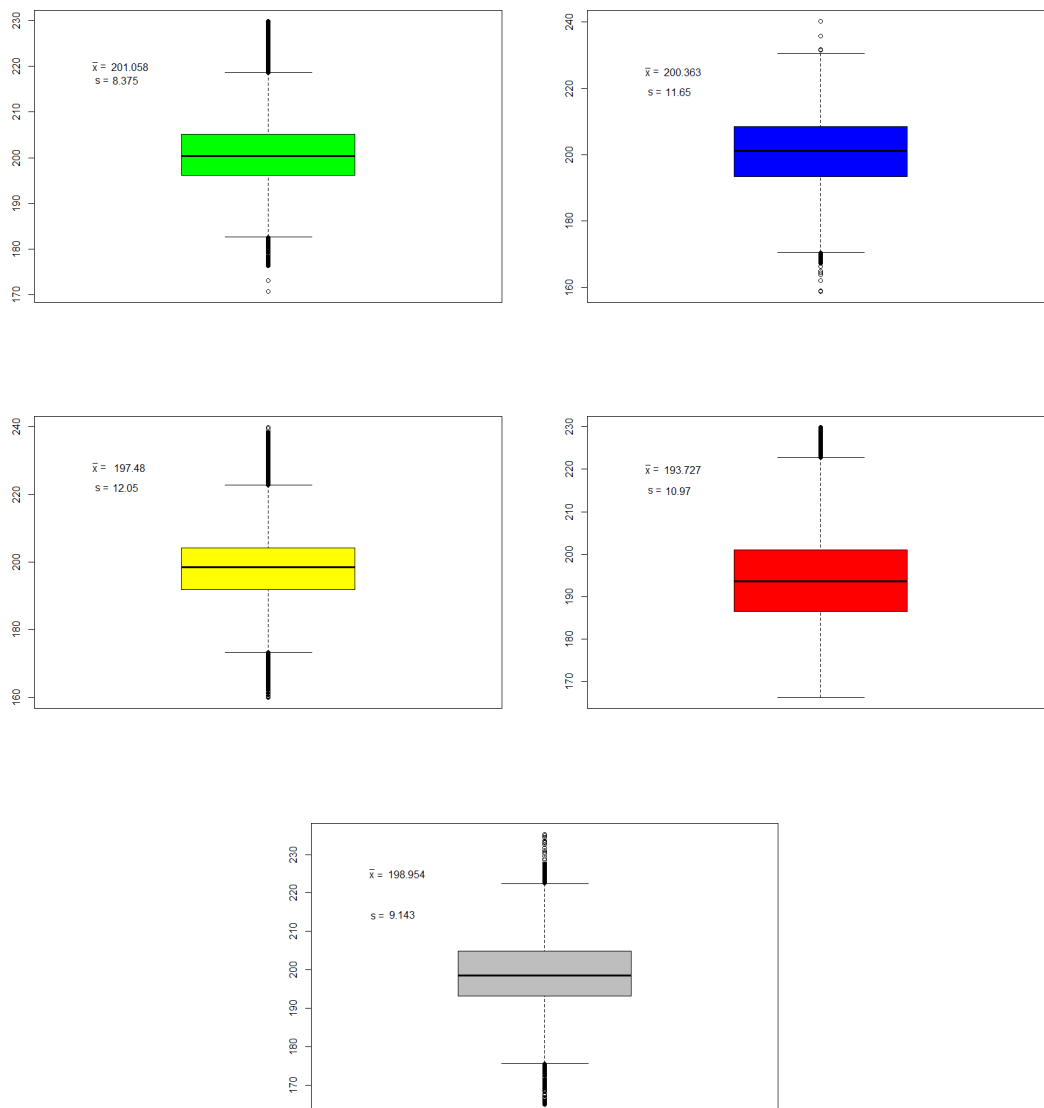


Abbildung 3.10.: Boxplots der stationären Bereiche von "Abgnox05mg" ohne Variation

Diese Messgröße weist in jedem stationären Abschnitt etliche Ausreißer auf. Für "AbgNox05mg" ist dies allerdings kein ungewöhnliches Verhalten, daher sollten die vielen Werte außerhalb der iqr-Standardabweichung nicht überbewertet werden. Die einzelnen Abschnitte scheinen symmetrisch verteilt zu sein mit einer ähnlichen Streuung. Wiederum wird untersucht, ob diese Daten aus einer Normalverteilung stammen könnten.

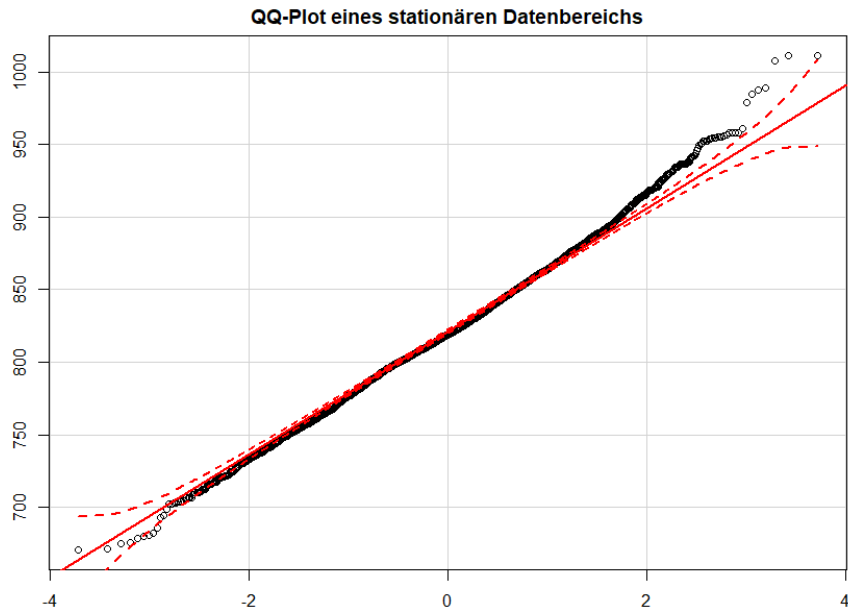


Abbildung 3.11.: QQ-Plot für "AbgNox05mg" ohne Variation

One-sample Kolmogorov-Smirnov test

data: Daten

D = 0.024738, p-value = 0.004393

Der QQ-Plot zeigt auf, dass doch einige Messwerte von der Normalverteilung abweichen. Da die Datenmenge ausgesprochen groß ist, wird der Vertrauensbereich dementsprechend kleiner. Allerdings erkennt man auch, dass die Abweichungen nicht eklatant sind insbesondere, wenn man weiß, dass diese Messgröße viele Ausreißer besitzt. Der KS Test verwirft zwar die Annahme einer Normalverteilung, doch als Anwender entscheiden wir uns trotzdem dafür, dass die Annahme der Normalverteilung in einem stationären Zustand aufrecht erhalten bleibt. Die Unabhängigkeit muss allerdings auch hier verworfen werden. Wiederum gibt der zeitliche Verlauf den Ausschlag.

3. Statistische Grundlagen

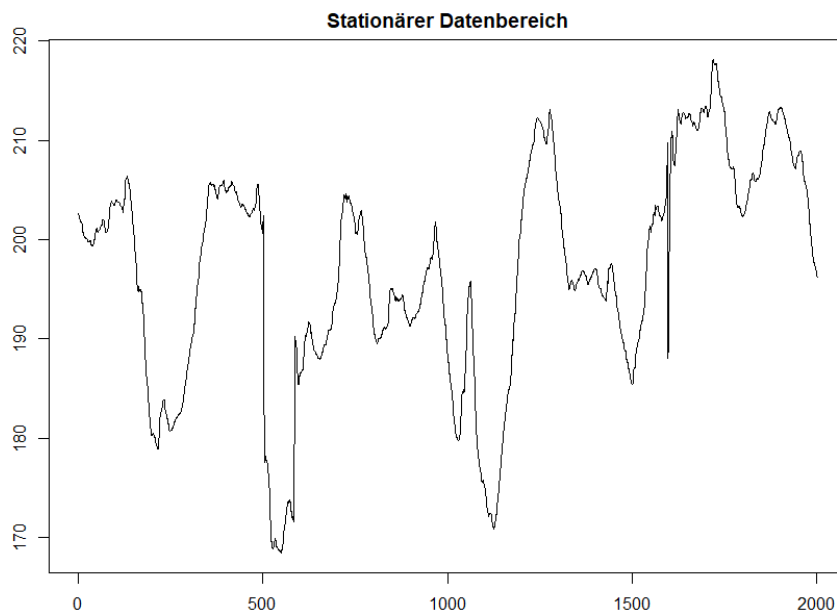


Abbildung 3.12.: Stationärer Verlauf der Messgröße "AbgNox05mg"

Diese Grafik spricht eindeutig gegen eine Unabhängigkeit der Daten. Man könnte hier von einem Art zyklischen Verhalten ausgehen, denn es scheint als ob die Messwerte immer wieder einen ansteigenden und anschließend einen abfallenden Trend folgen. Wie bei der Messgröße zuvor kann man daher zwar von einer Normalverteilung ausgehen, allerdings muss die Unabhängigkeit verworfen werden. Dieses Ergebnis bleibt bestehen, wenn man nun diese Messgröße während einer Variation untersucht, daher werden die Grafiken und Resultate kein weiteres Mal angeführt. Allerdings müssen die Boxplots dieser Größe betrachtet und speziell die Streuung bewertet werden.

3.1. Explorative Datenanalyse

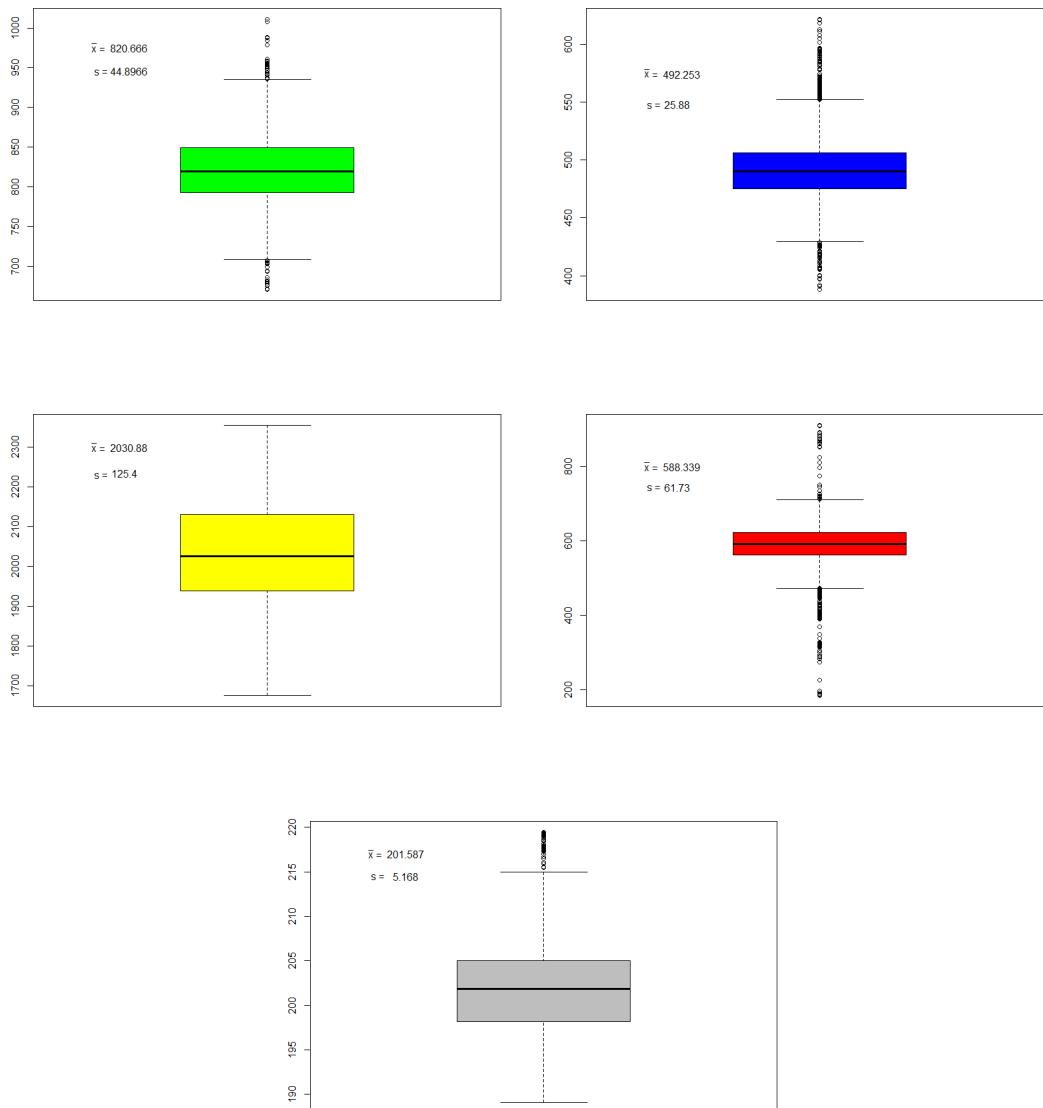


Abbildung 3.13.: Boxplots der stationären Bereiche von "Abgnox05mg" mit Variation

Deutlich zu erkennen ist die Änderung der Streuung der einzelnen Abschnitte. Je größer die absoluten Werte des Wertebereichs, desto größer auch die Streuung. Diese Eigenschaft ist auch bei poissonverteilten Zufallsvariablen zu sehen. Man muss also von beiden Situationen ausgehen: Es gibt Parameter die eine konstante Streuung aufweisen und es gibt Parameter deren Streuung abhängig vom Wertebereich der Messung ist. Diese Erkenntnis ist für den zu entwickelnden Algorithmus von ent-

3. Statistische Grundlagen

scheidender Bedeutung.

Nach dieser univariaten Datenanalyse wird eine bivariate Datenanalyse durchgeführt. Ziel einer bivariaten Datenanalyse ist es paarweise hoch korrelierte Daten zu erkennen. Es gibt in der Statistik zwei besonders wichtige Korrelationskoeffizienten. Den Korrelationskoeffizienten von Spearman und den Korrelationskoeffizienten von Pearson. Man nimmt dazu eine bivariate Stichprobe der Form $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ an. Der Koeffizient von Pearson errechnet sich wie folgt:


$$r_P = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Der Korrelationskoeffizient von Spearman wird mittels den Rängen der jeweiligen Stichproben errechnet. Der Rang eines Stichprobenelements entspricht seiner Position im Falle einer geordneten Stichprobe. In anderen Worten, falls x_i das k -größte Element ist, so entspricht der Rang von x_i gleich k . Der Rang wird also mit r_{x_i} beziehungsweise mit r_{y_i} bezeichnet und \bar{r}_x beziehungsweise \bar{r}_y entsprechen dem Mittelwert der Ränge von x beziehungsweise y .

Der Mittelwert der Ränge ist nur vom Datenumfang abhängig, denn man verteilt nur Positionen und das sind (i.A.) die natürlichen Zahlen $1, \dots, n$. Der Rangmittelwert kann daher ebenfalls geschrieben werden als $\bar{r}_x = \frac{1}{n} \sum_{i=1}^n i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$. Dieses Ergebnis bleibt aufrecht, wenn sich zwei Elemente den selben Rang teilen. Sollten zwei Elemente ident sein würden sie beispielsweise Rang 9 und 10 zugewiesen bekommen. Da sie aber ununterscheidbar sind wird jedem Element der Mittelwert der beiden verfügbaren Ränge zugeordnet. Daher wird die Gesamtsumme der Ränge nicht verändert und der Mittelwert aller Ränge bleibt derselbe. Nun kann man den Korrelationskoeffizienten von Spearman angeben.

$$r_S = \frac{\sum_{i=1}^n (r_{x_i} - \bar{r}_x)(r_{y_i} - \bar{r}_y)}{\sqrt{\sum_{i=1}^n (r_{x_i} - \bar{r}_x)^2 \sum_{i=1}^n (r_{y_i} - \bar{r}_y)^2}}$$

Im Folgenden wurden die soeben vorgestellten Koeffizienten für die wichtigsten Messgrößen ermittelt. Dazu wird eine Messreihe verwendet, in denen diese Parameter zum größten Teil stationär sind und nicht variiert werden. Andernfalls wäre der empirische Mittelwert natürlich ein verzerrter Schätzer und die Streuung stark verfälscht. Es gibt noch weitere Koeffizienten die man für diese Situationen verwenden könnte, wie zum Beispiel den Quadrantenkorrelationskoeffizient.

Für jedes Paar wird der Koeffizient berechnet und in die Korrelationsmatrix eingetragen. Die folgenden beiden Grafiken wurden mithilfe des Programms  und dem Befehl "corrplot" und "cor" generiert. Die Funktion "corrplot" enthält noch den Zusatz "hclust", welcher bewirkt, dass hochkorrelierte Paare automatisch nebeneinandergelegt werden. Dadurch entstehen Blöcke entlang der Diagonale. Die Elemente aus denen ein Block entsteht sind hoch korreliert. Die Matrizen für Pearson und Spearman sehen wie folgt aus:

3. Statistische Grundlagen

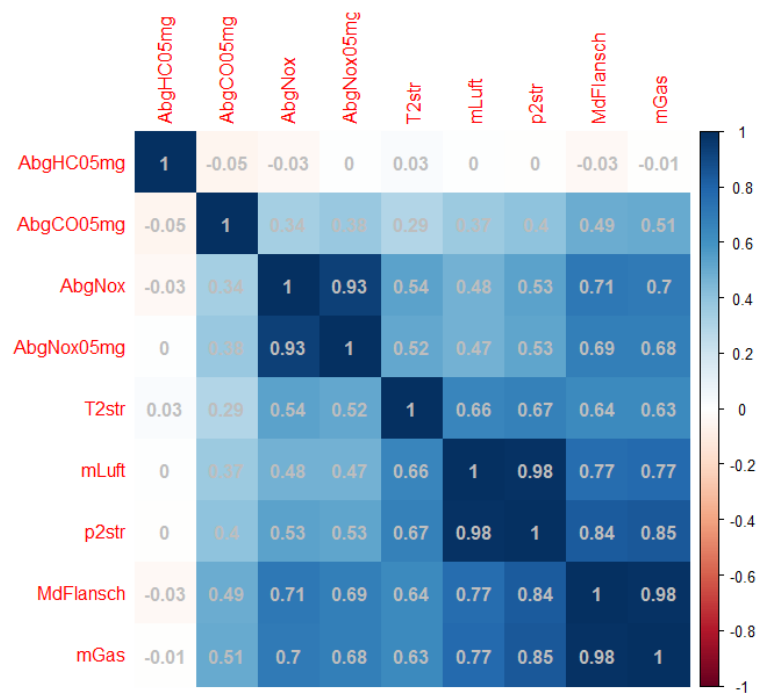


Abbildung 3.14.: Korrelationsmatrix bezüglich Pearson

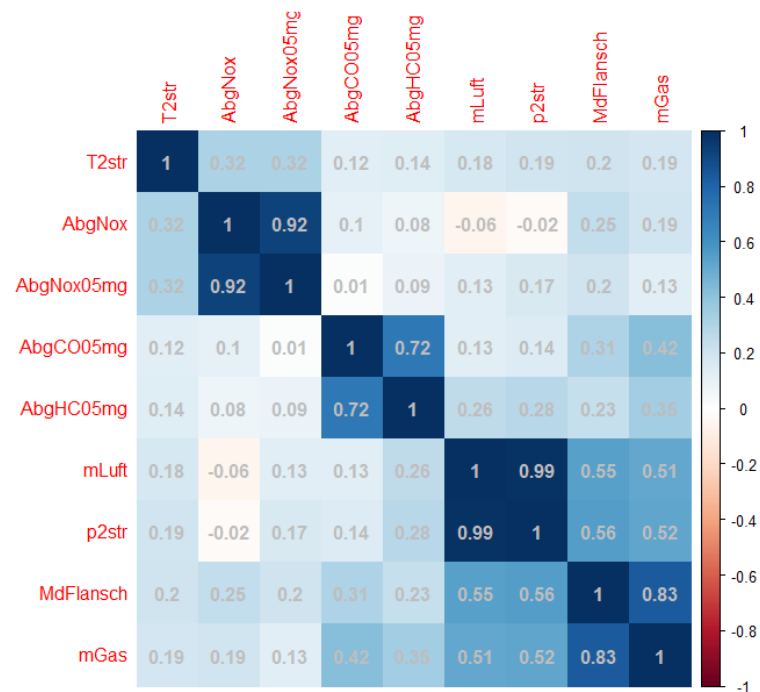


Abbildung 3.15.: Korrelationsmatrix bezüglich Spearman

3.1. Explorative Datenanalyse

In beiden Diagrammen liegt meist positive Korrelation vor. In den wenigen Fällen der negativen Korrelation ist diese so gering, dass man eher von linearen Unabhängigkeit spricht. Weitere Gemeinsamkeiten sind die hoch korrelierten Messgrößen "AbgNox" mit "AbgNox05mg" und "mLuft" mit "p2strich". Ein deutlicher Unterschied dieser beiden Koeffizienten liegt allerdings in der Bewertung der Messgröße "AbgHC05mg". Während Pearson diesen Parameter mit keinem anderen korreliert, errechnet Spearman eine hohe Korrelation zwischen "AbgHC05mg" und "AbgCO05mg". Des weiteren kann im Korrelationskoeffizienten von Spearman eine tatsächliche Blockdiagonalmatrix erkannt werden, wohingegen in der Korrelationsmatrix bezüglich Pearson generell hohe Werte zu finden sind. Der zeitliche Verlauf wird zwar besser durch r_P beschrieben, allerdings ist dieser Koeffizient nicht sehr robust. Der Korrelationskoeffizient r_S hingegen ist nicht stark empfindlich für Ausreißer, da der größte Wert der Beobachtungen immer denselben Rang erhält, irrelevant wie weit er von der Datenmenge entfernt. Daher könnte dieser Koeffizient besser geeignet sein.

Der Grund wieso wir uns für den Korrelationskoeffizienten interessieren ist schnell erklärt. Im Verlauf der Arbeit wird eine Methode zum Lösen der Aufgabenstellung entwickelt. Dabei wird jeder Parameter einzeln auf Eigenschaften überprüft. Da man in der Praxis allerdings mehrere Parameter gleichzeitig überwachen muss, kann man bei hoch korrelierten Messgrößen einen Repräsentanten auswählen der überprüft wird und spart sich dadurch enorm viel Rechenzeit und auch Speicherplatz.

Nun betrachten wir noch kurz die drei kleinen Blöcke der Pearson Matrix mittels eins Scatterplots, um eine lineare Abhängigkeit zu erkennen. Die einzelnen Scatterplots setzen sich aus 4 Bildern zusammen. Die beiden Grafiken auf der Diagonale entsprechen der geschätzten Dichte und dem Histogramm der beiden Messgrößen. Das linke untere Bild ist der Scatterplot und das rechte obere Bild besteht nur aus einer Zahl, bei der es sich um den Korrelationskoeffizienten von Pearson handelt. Es wird ersichtlich, dass nur zwei der drei Blöcke dem Korrelationskoeffizienten entsprechen. Der Grund für den großen Wert des Koeffizienten, ist die Anfälligkeit von Ausreißern.

3. Statistische Grundlagen

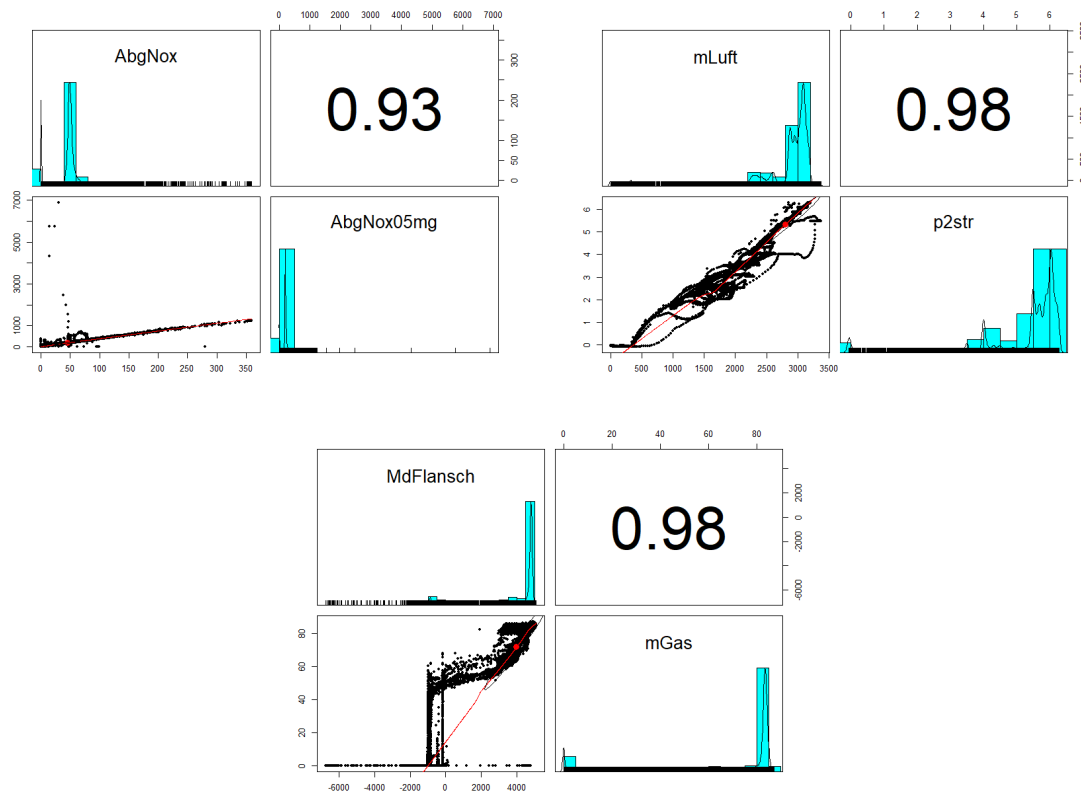


Abbildung 3.16.: Scatterplots der hoch korrelierten Daten

Nach dieser explorativen Datenanalyse hat man also folgende Erkenntnisse

- Die Varianz ist teils nicht homogen.
- In einem stationären Zustand sind die Daten approximativ normalverteilt, allerdings nicht unabhängig.
- Es gibt einige Messgrößen die stark korrelieren und daher müssen nicht immer alle Parameter untersucht werden.

Diese Informationen werden im Verlauf der Arbeit immer wieder benötigt.

Nun folgen die wichtigsten Begriffe und Methoden, die für die Methode der Beharungserkennung benötigt werden.

3.2. Das einfache lineare Modell

Für ein einfaches lineares Regressionsmodell benötigt man folgende Datenpunkte $(Y_1, x_1), \dots, (Y_n, x_n)$ wobei Y_i die sogenannte Responsevariablen sind und x_i feste und bekannte Konstanten, auch Prädiktor genannt. Man versucht nun die Responsevariable durch eine lineare Funktion in Abhängigkeit des Prädiktors und einem Zufallsfehler zu beschreiben.

$$Y_i = \beta_0 + \beta_1 \cdot x_i + \epsilon_i \quad i = 1, \dots, n.$$

Die ϵ_i beschreiben den nicht beobachtbaren statistischen Fehler. Es handelt sich um Zufallsvariablen mit der Eigenschaft $\mathbb{E}(\epsilon_i) = 0$ und $\text{Var}(\epsilon_i) = \sigma^2$. Man nimmt häufig an, dass $\epsilon_i \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$. Die Konstante der linearen Funktion β_0 wird Intercept genannt und die Steigung β_1 wird als Slope-Parameter bezeichnet. Diese beiden Parameter sind unbekannt und müssen geschätzt werden. Für den Erwartungswert der Response gilt

$$\mathbb{E}(Y_i) = \mathbb{E}(\beta_0 + \beta_1 \cdot x_i + \epsilon_i) = \beta_0 + \beta_1 \cdot x_i = \mu(x_i).$$

Weil man eine lineare Funktion mit nur einem Prädiktor modelliert, nennt man dieses Modell einfaches lineares Regressionsmodell.

3.3. Kleinste Quadrate Schätzer

Das kleinste Quadrate Kriterium wählt jene Gerade $\beta_0 + \beta_1 x$, welche die Summe der vertikalen Distanzen der Punkte (x_i, Y_i) und $(x_i, \beta_0 + \beta_1 \cdot x_i)$ minimiert. Das Ergebnis sind eindeutige kleinste Quadrate Schätzer $\hat{\beta}_0$ und $\hat{\beta}_1$, die

$$SSE(\beta_0, \beta_1) = \sum_{i=1}^n (Y_i - \mu(x_i))^2 = \sum_{i=1}^n (Y_i - (\beta_0 + \beta_1 \cdot x_i))^2$$

minimieren. Hier steht SSE für **S**um of **S**quared **E**rrors. Die entsprechenden Schätzer erhält man einfach durch differenzieren als

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i \cdot Y_i - n \cdot \bar{x} \cdot \bar{Y}}{\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{s_{xY}^2}{s_x^2},$$

3. Statistische Grundlagen

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{x}.$$

Man kann zeigen, dass diese Schätzer wirklich minimal sind. Der Vollständigkeit halber wird nun dieses Ergebnis in Matrixschreibweise dargestellt da es zu einem späteren Zeitpunkt benötigt wird. $\hat{\beta} = (X^\top X)^{-1} X^\top Y$. In unserem speziellen Fall wird man folgenden Vektor $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)^\top$ und folgende Matrix $X = (\mathbb{1}, x)^\top$ betrachten.

3.4. Hypothesentests

Sei $(F_\theta | \theta \in \Theta)$ eine Familie von Verteilungen. Eine statistische Hypothese ist eine Aussage über den Parameter θ der Form $H: \theta \in \Theta_0 \subseteq \Theta$ wobei $\Theta \subseteq \mathbb{R}^k$ der sogenannte Parameterraum ist. Es gibt immer zwei komplementäre Hypothesen.

1. H_0 : Die Nullhypothese ist jene Aussage über den Populationsparameter die überprüft wird.
2. H_1 : Die Alternativhypothese ist das Komplement der Nullhypothese. Sollte man die Nullhypothese verwerfen, so gilt automatisch die Alternativhypothese als die Wahrheit.

Sei Θ der gesamte Raum des Parameters θ , der die Daten beschreibt. Die allgemeine Form der Hypothesen sieht wie folgt aus:

$$H_0 : \theta \in \Theta_0 \quad \text{gegen} \quad H_1 : \theta \in \Theta_0^c.$$

Nach Beobachten und Bewerten der Stichprobe entscheidet man ob die Nullhypothese als wahr angenommen werden kann oder ob sie zugunsten der Alternativhypothese verworfen werden sollte. Ein Hypothesentest ist also eine Regel, die folgendes spezifiziert:

1. Für welche Werte basierend auf der Stichprobe man H_0 als wahr akzeptiert und daher behält.
2. Für welche Werte basierend auf der Stichprobe man H_0 verwirft und daher H_1 als Wahrheit akzeptiert.

Der Bereich für den die Nullhypothese verworfen wird nennt man kritischer Bereich oder Verwerfungsbereich. Der Bereich für den die Nullhypothese nicht verworfen

werden kann, also das Komplement des Verwerfungsbereichs, nennt man Annahmereich oder Akzeptanzbereich.

Eine Möglichkeit die Richtigkeit der Nullhypothese zu bewerten ist der sogenannte Likelihood Ratio Test (LRT). Näheres hierzu und Begriffe wie Powerfunktion und Level α Tests findet man in CASELLA et al. (2006).

3.5. Der p-Value


Ein wichtiger Begriff den wir später verwenden werden ist der sogenannte p-Value. Der p-Value ist eine Teststatistik $p(X)$ mit Wertebereich $[0,1]$. Kleine p-Values legen die Richtigkeit der Alternativhypothese nahe. Ein p-Value ist gültig falls für jedes $\theta \in \Theta_0$ und jedes $0 \leq \alpha \leq 1$

$$\mathbb{P}_\theta(p(X) \leq \alpha) \leq \alpha$$

gilt. Für gültige p-Values kann ein Level α Test konstruiert werden der H_0 genau dann verwirft wenn $p(X) \leq \alpha$.

Der p-Value ist das Ergebnis eines Integrals und kann vom Programm  sehr schnell berechnet werden.

3.6. Der lm-Befehl des Statistik Programms

Zur Vorbereitung auf die folgenden Kapitel, wird hier kurz der für diese Arbeit wichtige und häufig verwendete Befehle des Statistik-Programms  anhand eines sehr einfachen Beispiels erklärt. Der Befehl lautet "lm" und steht für "linear model". Als Input hat man die Responsevariable und den Prädiktor. Die Zusammenfassung (summary) dieses Befehls liefert die Regressionsgerade mit den dazugehörigen Parameterschätzungen die im Abschnitt 4.1 und 4.2 erklärt wurden. Zusätzlich erhält man zu jedem der β 's auch einen p-Value. Dieser p-Value bewertet die Nullhypothese $H_0 : \beta = 0$. Kleine p-Values deuten also auf eine Signifikanz des β 's hin. Als einfaches Beispiel wählen wir einen Response Vektor y und als Prädiktor einfach die Position der Response Beobachtung x . Das sieht wie folgt aus:

```
x = 1:20
```

```
y = y=c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)
```

Wendet man nun den eben erklärten Befehl auf diese Daten an, so liefert das folgende Ergebnis:

3. Statistische Grundlagen

Call:

lm(formula = y ~ x)

Residuals:

Min	1Q	Median	3Q	Max
-0.2820	-0.2462	0.0000	0.2462	0.2820

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.28947	0.12195	2.374	0.0289 *
x	0.49624	0.01018	48.744	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2625 on 18 degrees of freedom

Multiple R-squared: 0.9925, Adjusted R-squared: 0.9921

F-statistic: 2376 on 1 and 18 DF, p-value: < 2.2e-16

Im mittleren Bereich, unter der Kategorie "Coefficients" findet man die Parameterschätzungen. Der Wert der β 's befindet sich unterhalb von "Estimate". Die letzte Spalte der Koeffizienten besteht aus den p-Values der jeweiligen Parameter. Wenn der p-Value mit einem oder mehreren Sternchen gekennzeichnet wurde, so bedeutet dies, dass der Wert des Parameters sich signifikant von der 0 unterscheidet. Die Nullhypothese ist natürlich, dass das jeweilige β der 0 entspricht. Diese Hypothese wird zugunsten der Alternativhypothese verworfen.

4. Methoden der Beharrungserkennung

Die Beharrungserkennung (oder Steady-State Detection) ist ein bekanntes Problem und gewann über die letzten Jahrzehnte immer mehr an Bedeutung, insbesondere wenn es um eine Signalüberwachung geht. Zunächst hat man sich auf "offline" Verfahren beschränkt, welche speziell sich auf das Erkennen einer Aufwärmphase gerichtet haben. In HOAD et al. (2009) werden diese in 5 verschiedene Kategorien eingeteilt und mit verschiedenen weiteren Literaturverweisen versehen:

1. Graphische Methode:

Die Zeitreihe wird grafisch dargestellt und menschlich ausgewertet.

Beispiel: CUSUM plots SALVENDY (1982)

2. Heuristische Methode: Einfache Kriterien werden angewendet, die allerdings auf einigen Annahmen beruhen.

Beispiel: MSER SANDIKCI et al. (2006)

3. Statistische Methode:

Kriterien beziehungsweise Methoden werden anhand von statistischen Kenngrößen entwickelt.

Beispiel: ASD BAUSE et al. (2003)

4. "Befangenheitstests":

Es wird getestet ob es Verzerrungen in der Aufwärmphase gibt.

Beispiel: STS SCHRUBEN et al. (1983)

5. Hybridmethoden:

Die "Befangenheitstests" werden mit Abbruchkriterien kombiniert um Aufwärmphasen zu bestimmen.

Beispiel: PAWLIKOWSKI (1990)

4. Methoden der Beharrungserkennung

Allerdings wurde der Wunsch nach einer "online detection" immer größer und so wurden immer mehr dieser Verfahren entwickelt. Die grundlegenden Ideen hierfür sind

1. Lineare Regression über ein gleitendes Zeitfenster zu finden in WU et al. (2013)
2. Anwendung eines t-Tests auf Mittelwerte von zwei benachbarten Zeitfenstern in NARASIMHAN et al. (1987)
3. Die Überwachung der errechneten Standardabweichung in KIM et al. (1987)
4. Anwendung des F-Tests auf das Verhältnis der Varianzen in CROW et al. (1960)

4.1. Die Methode von Wu, Chen, Zhou und Li

Ein sehr aktuelles Verfahren findet man in WU et al. (2014) und wird nun hier beschrieben. Dabei wird ein Signal sequentiell durch ein stückweise lineares Modell gefittet indem man Bayesianische Inferenztechniken verwendet. Die Bayesianische Statistik wird in **hoff** beziehungsweise in **bayesian** ausgezeichnet erklärt. Einfach ausgedrückt, kann man mithilfe der Bayesianischen Statistik die eigenen Vermutung anhand von Sampling Schritten überprüfen. Dazu wird eine vermutete Dichte, auch Prioridichte genannt, gewählt. Im Anschluss werden Samplings und Beobachtungen verwendet um eine Posterioridichte zu erhalten. Je nach Ähnlichkeit beziehungsweise Unterschied der Priori- und Posterioridichte kann nun die ursprüngliche Annahme bestätigt oder verworfen werden.

Die Methode in WU et al. (2014) kann in einfachen Worten folgendermaßen zusammengefasst werden: Die Zeitreihe des Signals wird in einzelne Liniensegmente aufgeteilt. Die Stellen, an denen die Unterteilung der Zeitreihe vorgenommen wird, sind Changepoints. Das bedeutet, zwischen zwei aufeinanderfolgenden Changepoints wird eine lineare Funktion gelegt, welche das Signal beschreibt. Die Steigung der linearen Funktion wird bewertet und je nach Absolutbetrag der Steigung kann Beharrung angenommen oder verworfen werden.

Die Beharrungserkennung wird also zu einem Inferenzproblem für multiple Changepoint Modelle. Über den aktuellsten Changepoint wird dazu eine Prioridichte gelegt mit deren Hilfe die Modellparameter (das bedeutet die lineare Funktion) und die Posterioridichte geschätzt werden kann. Man kann sich das Vorgehen also folgendermaßen vorstellen: Ein Signal folgt ungefähr dem Verlauf einer linearen Steigung.

4.1. Die Methode von Wu, Chen, Zhou und Li

Jeder neue Datenpunkt wird mithilfe der Bayesianischen Statistik überprüft. Sollte dieser Datenpunkt eine Posterioridichte liefern, die der Prioridichte entspricht, so gibt es keine Änderung im Verhalten des Signals. Sollte nun allerdings die Bayesianische Statistik auf Unstimmigkeiten der Dichten stoßen, so liegt ein neuer Changepoint vor. Von diesem neuen Changepoint ausgehend wird nun wieder ein lineares Modell geschätzt und mithilfe von Samplings die neue Steigung bewertet.

Die Bayesianische Inferenz des aktuellsten Changepoints (latest change-point LCP) und die Modellparameter der stückweise linearen Funktion werden in der folgenden Grafik eines Beispiels illustriert. Die Posteriori Verteilungen werden sequentiell aktualisiert beziehungsweise neu geschätzt sobald neue Datenpunkte vorhanden sind. Im oberen Drittel der Grafik findet man das tatsächliche Signal und die Steigung die von dem aktuellsten Changepoint ausgehend berechnet wird. Im mittleren Drittel wird die Dichte über den Changepoint gelegt und im letzten Drittel erhält man eine geschätzte Verteilung des Slope Parameters.

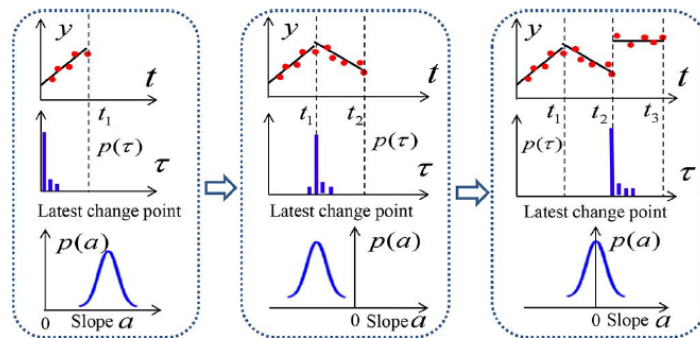


Abbildung 4.1.: Changepoints und ihre Liniensegmente aus WU et al. (2014)

Seien $\xi_t = (a_t, b_t, \sigma_t^2)$ die Modell Parameter zum Zeitpunkt t mit

- a_t die Steigung der Gerade
- b_t der Intercept der Geraden
- σ_t^2 Varianz des Rauschens zum Zeitpunkt t

Bezeichne den LCP zum Zeitpunkt t mit τ_t . Die Priori Übergangsdichte der Veränderung zum Zeitpunkt t gegeben der vorherige Changepoint τ_{t-1} lag an der Stelle i wird mit $p(\tau_t = t | \tau_{t-1} = i)$ bezeichnet.

4. Methoden der Beharrungserkennung

Die populärste und einfachste Prioridichte für den Changepoint ist der geometrische Prior. Dieser Prior entspricht einem Markov-Übergangsprozess mit konstanter Priori Übergangswahrscheinlichkeit. Man hat aus Gründen der Einfachheit angenommen, dass diese Übergangswahrscheinlichkeit konstant ist und mit p beziffert werden kann. Sei nun $x_t = (\xi_t, \tau_t)$ und bezeichne mit y_t die tatsächliche Beobachtung zum Zeitpunkt t . Zu jedem Zeitpunkt t kann nun der Übergangsprozess formuliert werden als:

$$x_t = \begin{cases} x_{t-1} & \text{mit Wahrscheinlichkeit } 1 - p \\ (\xi', t) & \text{mit Wahrscheinlichkeit } p \end{cases}$$

wobei ξ' gesampelt wurde aus der Prioriverteilung $f(\cdot)$ von ξ . Die Beobachtungen y_t werden modelliert durch $y_t = a_t \cdot t + b_t + \epsilon_t$ wobei $\epsilon_t \sim N(0, \sigma_t^2)$ weißes Rauschen. Wegen der allerdings unbekanntenen Changepoints ist dieses Zustandsraum Modell nichtlinear, was Rückschlüsse sehr schwierig macht. Particle Filtering Techniken die auf Sequentieller Wichtigkeit basieren, sind sehr effektiv für nichtlineare Zustandsraum Modelle. Deswegen wurde auch ein Partikelfiltender Algorithmus für diesen Algorithmus entwickelt.

Die Grundidee einer Filtering Technik ist das sequential importance sampling (SIS). Angenommen das Zustandsraum Modell, welches in **zustand** zu finden ist, kann geschrieben werden als

$$\begin{aligned} x_1 &\sim f(x) \\ x_t|x_{t-1} &\sim f(x_t|x_{t-1}) \\ y_t|x_t &\sim g(y_t|x_t) \end{aligned}$$

Die Posteriori Verteilung wird mit $\pi_n(x_{1:n})$ bezeichnet

$$\begin{aligned} \pi_n(x_{1:n}) &= p(x_{1:n}|y_{1:n}) = \frac{p(x_{1:n}, y_{1:n})}{p(y_{1:n})} \\ &= \frac{f(x_1) \prod_{t=2}^n f(x_t|x_{t-1}) \prod_{t=2}^n g(y_t|x_t)}{p(y_{1:n})} \end{aligned}$$

wobei $z_{i:j}$ definiert ist als der Vektor $(z_i, z_{i+1}, \dots, z_j)$. Wenn wir eine "Importance"

4.1. Die Methode von Wu, Chen, Zhou und Li

Verteilung wie folgt wählen

$$q_n(x_{1:n}) = q_1(x_1) \prod_{t=2}^n q_t(x_t|x_{1:t-1})$$

(also das Produkt von bedingter Dichten) dann können die nicht normalisierten Gewichte der Partikel $x_{1:n}$ geschrieben werden als

$$\begin{aligned} \omega_n(x_{1:n}) &= \frac{p(x_{1:n}, y_{1:n})}{q_n(x_{1:n})} \\ &= \frac{f(x_1)g(y_1|x_1)}{q_1(x_1)} \times \frac{f(x_2|x_1)g(y_2|x_2)}{q_2(x_2|x_1)} \cdots \frac{g(y_n|x_n)f(x_n|x_{n-1})}{q_n(x_n|x_{1:n-1})} \\ &= \omega_1(x_1) \prod_{t=2}^n \alpha_t \end{aligned} \quad (4.1)$$

mit

$$\begin{aligned} \omega_1(x_1) &= \frac{f(x_1)g(y_1|x_1)}{q_1(x_1)} \\ \alpha_t &= \frac{g(y_t|x_t)f(x_t|x_{t-1})}{q_t(x_t|x_{1:t-1})} \end{aligned}$$

Die Bezeichnung "Importance Sampling" wird deswegen verwendet, weil jeder generierte Partikel gewichtet wird. Gleichung 4.1 legt nahe, dass das Gewicht rekursiv beziehungsweise sequentiell berechnet werden kann. Die Erwartung einer beliebigen Funktion $\varphi_n(x_{1:n})$ in Betrachtung mit der Posteriori Wahrscheinlichkeit $\pi_n(x_{1:n})$ kann geschätzt werden durch:

$$\mathbb{E}(\varphi_n(x_{1:n})|y_{1:n}) \approx \sum_{i=1}^{N_S} W_n^{(i)} \varphi_n(x_{1:n}^{(i)})$$

$W_n^{(i)}$ ist das normalisierte Gewicht des Partikels $x_{1:n}^{(i)}$ und N_S ist die Anzahl der Partikel. Bezeichne mit $\delta_n^{\text{SIS}}(\varphi) = \sum_{i=1}^{N_S} W_n^{(i)} \varphi_n(x_{1:n}^{(i)})$ und $I_n(\varphi) = \mathbb{E}(\varphi_n(x_{1:n})|y_{1:n})$

Dann gilt

$$\sqrt{N_s} (\delta_n^{\text{SIS}}(\varphi) - I_n(\varphi)) \xrightarrow{D} N(0, \phi_{q_n}^2)$$

mit asymptotischer Varianz:

4. Methoden der Beharrungserkennung

$$\phi_{q_n}^2 = \int \frac{\pi_n^2(x_{1:n})}{q_n(x_{1:n})} (\varphi_n(x_{1:n}) - I_n(\varphi))^2 dx_{1:n}$$

Der Schätzer ist verzerrt für endliches N_S und der asymptotische bias ist:

$$\lim_{N_S \rightarrow \infty} N_S (\delta_n^{\text{SIS}}(\varphi) - I_n(\varphi)) = - \int \frac{\pi_n^2(x_{1:n})}{q_n(x_{1:n})} (\varphi_n(x_{1:n}) - I_n(\varphi)) dx_{1:n}$$

Der generische Partikel Filtering Algorithmus sieht wie folgt aus.

Algorithm 1. Sequential Monte Carlo Filtering Algorithm

At time $t = 1$,

- 1) Sample $\mathbf{x}_1^{(i)} \sim q_1(\mathbf{x}_1)$, $i = 1, 2, \dots, N_s$.
- 2) Compute the weights $\omega_1(\mathbf{x}_1^{(i)})$ and normalized weights $W_1^{(i)} = (\omega_1(\mathbf{x}_1^{(i)})) / (\sum_{i=1}^{N_s} \omega_1(\mathbf{x}_1^{(i)}))$.
- 3) Resample $\{W_1^{(i)}, \mathbf{x}_1^{(i)}\}$ according to the particle weight $W_1^{(i)}$ to obtain N_s equally weighted particles $\{1/N_s, \bar{\mathbf{x}}_1^{(i)}\}$ and set $\{W_1^{(i)}, \mathbf{x}_1^{(i)}\} \leftarrow \{1/N_s, \bar{\mathbf{x}}_1^{(i)}\}$.

At time $t \geq 2$,

- 1) Sample $\mathbf{x}_t^{(i)} \sim q_t(\mathbf{x}_t | \mathbf{x}_{1:t-1}^{(i)})$, set $\mathbf{x}_{1:t}^{(i)} \leftarrow (\mathbf{x}_{1:t-1}^{(i)}, \mathbf{x}_t^{(i)})$.
 - 2) Compute $\alpha_t(\mathbf{x}_{1:t}^{(i)})$ and $W_t^{(i)} \propto \alpha_t(\mathbf{x}_{1:t}^{(i)})$.
 - 3) Resample $\{W_t^{(i)}, \mathbf{x}_{1:t}^{(i)}\}$ to obtain particles $\{1/N_s, \bar{\mathbf{x}}_{1:t}^{(i)}\}$ and set $\{W_t^{(i)}, \mathbf{x}_{1:t}^{(i)}\} \leftarrow \{1/N_s, \bar{\mathbf{x}}_{1:t}^{(i)}\}$.
-

Abbildung 4.2.: Generischer Partikelfilter Algorithmus aus WU et al. (2014)

Eine Kurzfassung in einfachen Worten folgt nun sofort. Man geht von einem Zustandsraummodell aus und den vermuteten Prioridichten. Die Posterioridichte kann mittels der Prioridichte bezüglich des Modells, der Prioridichte der Beobachtungen und der Priori Übergangsdichte bestimmt werden. Im Anschluss wählt man sich eine passende Verteilung für die gewichte der Partikel Erzeugung. Diese Gewichte

4.1. Die Methode von Wu, Chen, Zhou und Li

werden auch normalisiert. Schlussendlich kann gezeigt werden, dass die Differenz der gewichteten Summe der Partikel und dem Erwartungswerts des Modells in Verteilung gegen eine Normalverteilung konvergiert.

Man hat allerdings eine Schwierigkeit mit der sogenannten Partikel Degeneration. Das bedeutet, dass nach wenigen Iterationen das Gewicht nur auf ein paar wenigen Partikeln konzentriert ist und die meisten anderen Partikel ein vernachlässigbares Gewicht besitzen. Eine Möglichkeit dieses Problem zu umgehen ist eine "Importance" Dichte zu nehmen die die Varianz der "Importance" Gewichte reduziert. Man kann zeigen dass die optimale "Importance" Dichte $q_t(x_t | x_{1:t-1}^{(i)}) = p(x_t | x_{t-1}^{(i)}, y_t)$ ist. In der Praxis wird allerdings

$$q_t(x_t | x_{1:t-1}^{(i)}) = f(x_t | x_{t-1}^{(i)})$$

verwendet.

Diese Wahl vereinfacht den Rechenaufwand enorm, da nach dem Sampling Schritt, die Gewichte zum Zeitpunkt t die Form $\omega_t = g(y_t | x_t^{(i)})$ haben. Es wurde allerdings $f(x_t | x_{t-1})$ als Importance Priori Dichte gewählt. Eine effiziente sampling Methode nennt sich schichtenweise (stratified) sampling. Dabei wird eine deterministische Stichprobe der Größe N_0 aus $f(x)$ gezogen, und aus den Partikeln die zum Zeitpunkt $t-1$ vorhanden waren, werden $N_1 = N_S - N_0$ Werte gezogen. Zum Zeitpunkt $t-1$ sind die Partikel ungefähr $\pi_{t-1}(x_{t-1})$ verteilt. Die Zielverteilung ist $\pi_t(x_t) = p(x_t | x_{1:t}) = \frac{p(x_t | y_{1:t-1})g(y_t | x_t)}{p(y_t | y_{1:t-1})}$ mit Importance Dichte

$$p(x_t = x | y_{1:t-1}) = pf(x) + (1 - p)\pi_{t-1}(x_{t-1} = x)$$

Dies ergibt folgenden Algorithmus.

Algorithm 2. Stratified Sampling with Deterministic Sample Size

At time step $t = 1$,

- 1) Sample $\mathbf{x}_1^{(i)} \sim f(\mathbf{x}), i = 1, 2, \dots, N_s$.
- 2) Compute $\omega_1(\mathbf{x}_1^{(i)}) = g(y_1 | \mathbf{x}_1^{(i)})$ and $W_1^{(i)}$.

At time step $t > 1$,

- 1) Sample $\mathbf{x}_{t,0}^{(i)} \sim f(\mathbf{x}), i = 1, 2, \dots, n_0$ to obtain n_0 new particles. Resample $\{W_{t-1}^{(i)}, \mathbf{x}_{t-1}^{(i)}\}$ to obtain $n_1 = N_s - n_0$ particles $\bar{\mathbf{x}}_{t-1}^{(i)}, i = 1, 2, \dots, n_1$ and set $\mathbf{x}_{t,1}^{(i)} \leftarrow \bar{\mathbf{x}}_{t-1}^{(i)}$.
 - 2) Set $\mathbf{x}_t \leftarrow \{\mathbf{x}_{t,0}, \mathbf{x}_{t,1}\}$ and compute $W_t^{(i)}$ using (16).
-

Abbildung 4.3.: Algorithmus für schichtweises Sampling aus WU et al. (2014)

Diese Methode reduziert zwar die Degenerierung, aber es tritt noch ein weiteres Problem auf; Die Vielfalt der Partikel kann reduziert werden. Man muss also dafür sorgen, dass einige verschiedene Partikel ein nicht vernachlässigbares Gewicht erhalten. Es ist also ein Balanceakt zwischen Degeneration und Vielfalt. In dem besagten Paper wird auch hierfür eine Lösung angegeben und schlussendlich bekommt man folgenden Code

Algorithm 3. Particle Filtering Algorithm

At time step $t = 1$,

- 1) Sample $\mathbf{x}_1^{(i)} \sim f(\mathbf{x}), i = 1, 2, \dots, N_s$
- 2) Compute $\omega_1(\mathbf{x}_1^{(i)}) = g(y_1 | \mathbf{x}_1^{(i)})$ and $W_1^{(i)}$
- 3) Resample $\{W_1^{(i)}, \mathbf{x}_1^{(i)}\}$ to obtain $\{1/N_s, \bar{\mathbf{x}}_1^{(i)}\}$ and set $\{W_1^{(i)}, \mathbf{x}_1^{(i)}\} \leftarrow \{(1)/(N_s), \bar{\mathbf{x}}_1^{(i)}\}$

For time step $t > 1$,

- 1) Sample $\mathbf{x}_t^{(i)}$ according to Algorithm 2.
 - 2) Compute $W_t^{(i)}$ according to Algorithm 2.
 - 3) Resample $\{W_t^{(i)}, \mathbf{x}_t^{(i)}\}$ to obtain $\{1/N_s, \bar{\mathbf{x}}_t^{(i)}\}$ and set $\{W_t^{(i)}, \mathbf{x}_t^{(i)}\} \leftarrow \{1/N_s, \bar{\mathbf{x}}_t^{(i)}\}$.
 - 4) Gibbs move: select a subset S of resampled particles such that $\sum_{j \in S} d_t^{(j)} \leq C$.
 - Sample $\bar{\boldsymbol{\beta}}_t^{(j)} \sim N(\boldsymbol{\mu}_t^{(j)}, \boldsymbol{\Sigma}_t^{(j)})$ and set $\boldsymbol{\beta}_t^{(j)} \leftarrow \bar{\boldsymbol{\beta}}_t^{(j)}$.
 - Sample $\bar{\sigma}_t^{2(j)} \sim IG(\alpha_1 + (t - \tau_t^{(j)} + 1)/2, \alpha_2 + (\|\mathbf{Y}_t - \mathbf{X}_t^{(j)} \boldsymbol{\beta}_t^{(j)}\|^2)/2)$ and set $(\sigma_t^2)^{(j)} \leftarrow \bar{\sigma}_t^{2(j)}$.
-

Abbildung 4.4.: Partikelfilter Algorithmus aus WU et al. (2014)

Um sich diese Methode nun leichter vorstellen zu können folgt eine Zusammenfassung in einfacheren Worten. Man hat zwei Fälle zu unterscheiden. Fall 1: Man hat nur einen Datenpunkt ($t = 1$), Fall 2: man hat bereits mehrere Datenpunkte ($t > 1$). Für jeden dieser Fälle gibt es nun ein Vorgehen.

- $t = 1$

- Erzeuge zufällig N_S Zustandsraumvektoren $x_1^{(i)}, i = 1, \dots, N_S$ aus der Priori-Verteilung. In WU et al. (2014) wurde $N_S = 2000$ gewählt und die Priori-Verteilung ist entweder die Normalverteilung oder die Inverse Gamma Verteilung
- Berechne die Gewichte jedes Vektors $w_1^{(i)}$ hinsichtlich der aktuellen Beobachtung y_1 und normiere diese Gewichte. Das Normieren bringt den Vorteil, dass die erhaltenen Werte als Wahrscheinlichkeiten benutzt werden können
- Führe danach wiederholte Ziehungen einer Stichprobe aus der Ausgangsstichprobe bzgl. der Gewichte durch. Im Anschluss wird $W_1^{(i)} = \frac{1}{N_S}$ gesetzt

4. Methoden der Beharrungserkennung

- $t > 1$

- Erzeuge zufällig $N_0 = N_S \cdot p$ Zustandsraumvektoren $x_t^{(i)}$, $i = 1, \dots, N_0$ aus der Priori-Verteilung und ziehe $N_1 = N_S - N_0$ Vektoren aus dem Schritt zuvor
- Berechne und normiere die Gewichte jedes Vektors $W_t^{(i)}$.
- Ziehe wiederholt aus einer Stichprobe und setze im Anschluss $W_t^{(i)} = \frac{1}{N_S}$
- Der sogenannte "Gibbs-Move"
Wähle eine Teilmenge der Vektoren und passe die Parameter (wie z.B. Steigung) an die Datensituation an.
- Berechne die Wahrscheinlichkeit, dass die aktuelle Steigung kleiner als eine bestimmte Grenze s_0 ist. Die Wahrscheinlichkeit wird berechnet als $\frac{\text{Anzahl der Vektoren die eine kleine Steigung besitzen}}{N_S}$
- Berechne die Schätzung der aktuellsten Beobachtung

Die Funktionalität und Güte dieser Methode wurde durch einige Beispiel in WU et al. (2014) bestätigt. Eines davon wird nun kurz gezeigt.

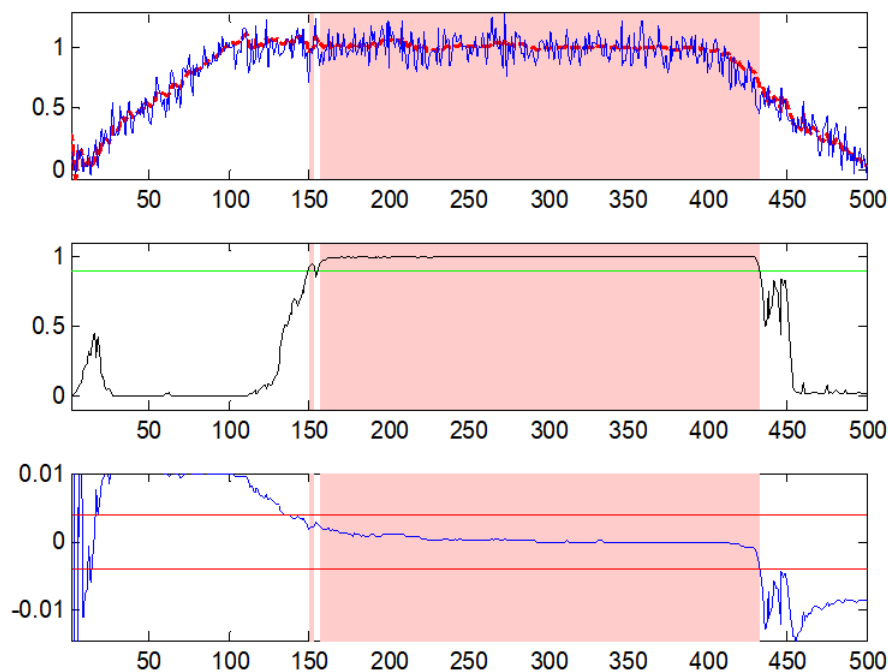


Abbildung 4.5.: Erkannte Beharrung der Datenzeitreihe

4.1. Die Methode von Wu, Chen, Zhou und Li

Im ersten Bild ist das tatsächliche Signal in Blau eingezeichnet. Der rot schattierte Bereich ist der beharrte Zustand des Signals. Im zweiten Bild wird die Wahrscheinlichkeit angegeben, dass die Vektoren die zuvor erwähnten Bedingungen erfüllen. Das letzte Bild zeigt die durchschnittliche Steigung. Wie man deutlich erkennen kann, wurde der beharrte Zustand erfolgreich erkannt. Auch bei zyklischen Daten scheint dieser Algorithmus gut zu funktionieren.

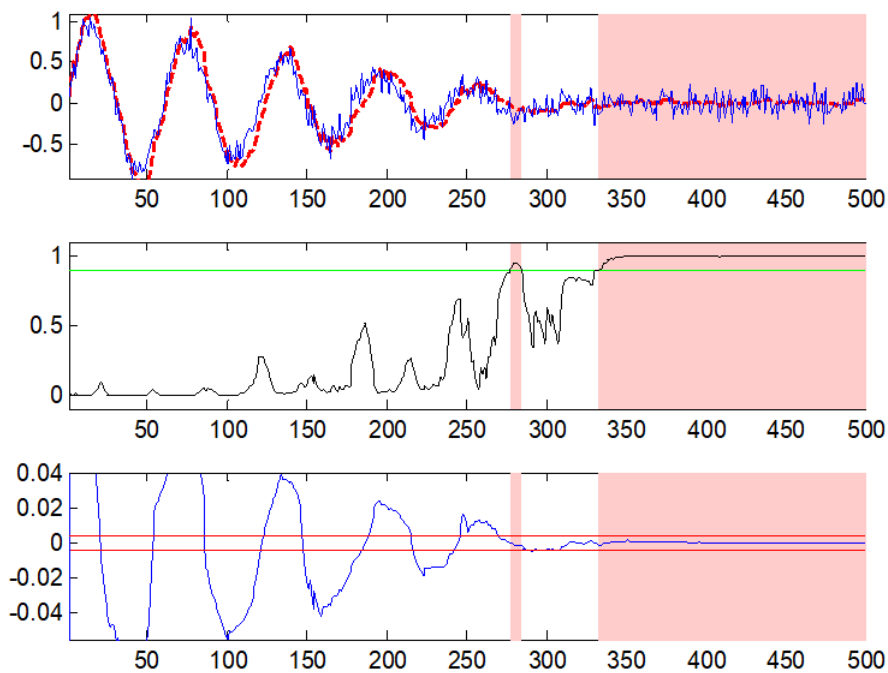


Abbildung 4.6.: Erkannte Beharrung der Datenzeitreihe

Leider ist diese Methode aus mehreren Gründen in unserem Projekt nicht anwendbar. Zum einen muss man sich eine Grenze der Steigung vorgeben. Eine Vorgabe der Steigung müsste aber für jede Messgröße eigens bestimmt werden. Des Weiteren kann diese Grenze während einer einzigen Messreihe variieren, was natürlich keine Vorgabe mehr möglich macht. In vielen Fällen ist diese Steigung sogar unbekannt, und eine falsche Grenze wäre katastrophal, denn es würde zu enormen finanziellen und zeitlichen Verlusten kommen, wenn hunderttausende Daten falsch bewertet werden. Zum anderen ist der Rechenaufwand für diese Methode enorm und benötigt einiges an Zeit die am Motorprüfstand nicht vorhanden ist. In unserer Datensituation werden pro Sekunde 5 Datenpunkte erzeugt (das bedeutet eine Datenerzeugung von 5Hz), was zu einem unvermeidlichen Rückstau bei schon einer zu

4. Methoden der Beharrungserkennung

überwachenden Messgröße führt. Außerdem treten auch weitere Probleme auf. Es wird nun ein Beispiel angegeben wo selbst diese Methode versagt:

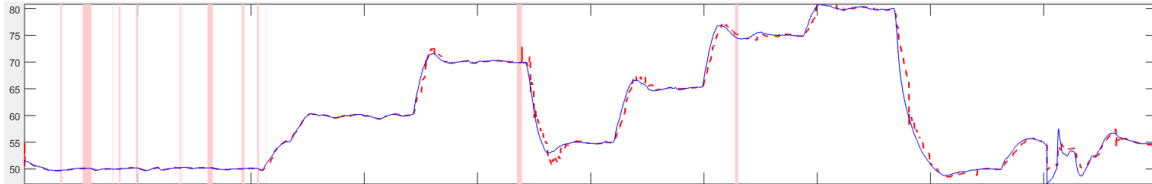


Abbildung 4.7.: Verwendete Methode auf echten Daten

Man kann deutlich erkennen, dass die gewünschten stationären Bereiche nicht erkannt werden. Dies liegt daran, dass es sich hierbei um eine diskrete Messgröße handelt. Zusätzlich zum enormen zeitlichen Aufwand ist also diese Methode nicht auf alle Datensätze anwendbar.

Daher wird nun die Methode vorgestellt welche tatsächlich am Motorprüfstand zum Einsatz kommt.

4.2. Eine neue Methode für den Motorprüfstand

Unsere Methode zur Beharrungserkennung basiert auf der Anwendung der einfachen linearen Regression in einem gleitenden Zeitfenster. In jedem Schritt wird also ein neuer Datenpunkt hinzugefügt und der letzte Datenpunkt entfernt. In jedem dieser Zeitfenster wird die Steigung der linearen Regressionsgerade berechnet und der dazugehörige p-Value zum Test der Nullhypothese $H_0 : \beta_1 = 0$ errechnet. Wenn die Steigung nicht signifikant von 0 verschieden ist, so ist der p-Value annähernd Gleichverteilt auf dem Intervall $[0,1]$. Keine Steigung bedeutet, dass die Responsevariable keine signifikante Veränderung aufweist und daher ein stationärer Zustand angenommen werden kann. Die Zeitreihe der Daten wird also in eine Zeitreihe von p-Values transformiert und diese dann ebenfalls mit einem gleitenden Zeitfenster untersucht. In jedem dieser gleitenden Zeitfenster wird nun überprüft, ob ausreichend viele p-Values groß genug sind. Weitere Einzelheiten werden im Verlauf der Arbeit beschrieben.

Zur Illustration geben wir ein einfaches Beispiel. Es wurde eine stückweise lineare Funktion erstellt. Im ersten linearen Teilstück besitzt die Funktion eine signifikante Steigung und im zweiten Teilstück ist die Steigung 0. Diese Gerade ist nun mit

4.2. Eine neue Methode für den Motorprüfstand

einem weißen Rauschen überlagert. Dazu werden 1000 Werte aus einer Normalverteilung mit konstanter Varianz generiert. Über diese Daten wird nun ein gleitendes Zeitfenster, welches 150 Datenpunkte umfasst, laufen gelassen. Nun werden die oben erwähnte Zeitreihe von p-Values generiert. (Siehe Abbildung 4.8)

Um die Datensituation am LEC zu simulieren wird das Datenfenster mit jedem Wert einzeln befüllt und die Analyse sofort durchgeführt. Das bedeutet, dass zu Beginn die lineare Regression beziehungsweise die Berechnung des p-Values mit weniger Daten durchgeführt wird, da das Datenfenster noch nicht vollständig befüllt werden kann. Somit sind die ersten p-Values noch nicht sehr aussagekräftig, da die Datenmasse noch nicht ausreichend groß war. Es ist verständlich, dass eine Regression mit sehr wenigen Datenpunkten keine zuverlässige Richtlinie ist. Allerdings gibt es am LEC ein paar Messgrößen, die keine Startphase benötigen und kaum Schwankungen aufweisen. Daher hat man sich für eine sofortige Evaluierung ausgesprochen um diese wenigen Parameter einschätzen zu können. Für die anderen Parameter ist es irrelevant, ob die ersten p-Values zuverlässig sind oder nicht, da ohnehin eine Aufwärmphase benötigt wird. Diese Aufwärmphase dauert wesentlich länger als das vollständige Befüllen des Datenfensters, denn wenige Sekunden nach Inbetriebnahme werden niemals Messungen vorgenommen. Daher spielt der anfängliche unsichere Bereich keine Rolle.

4. Methoden der Beharrungserkennung

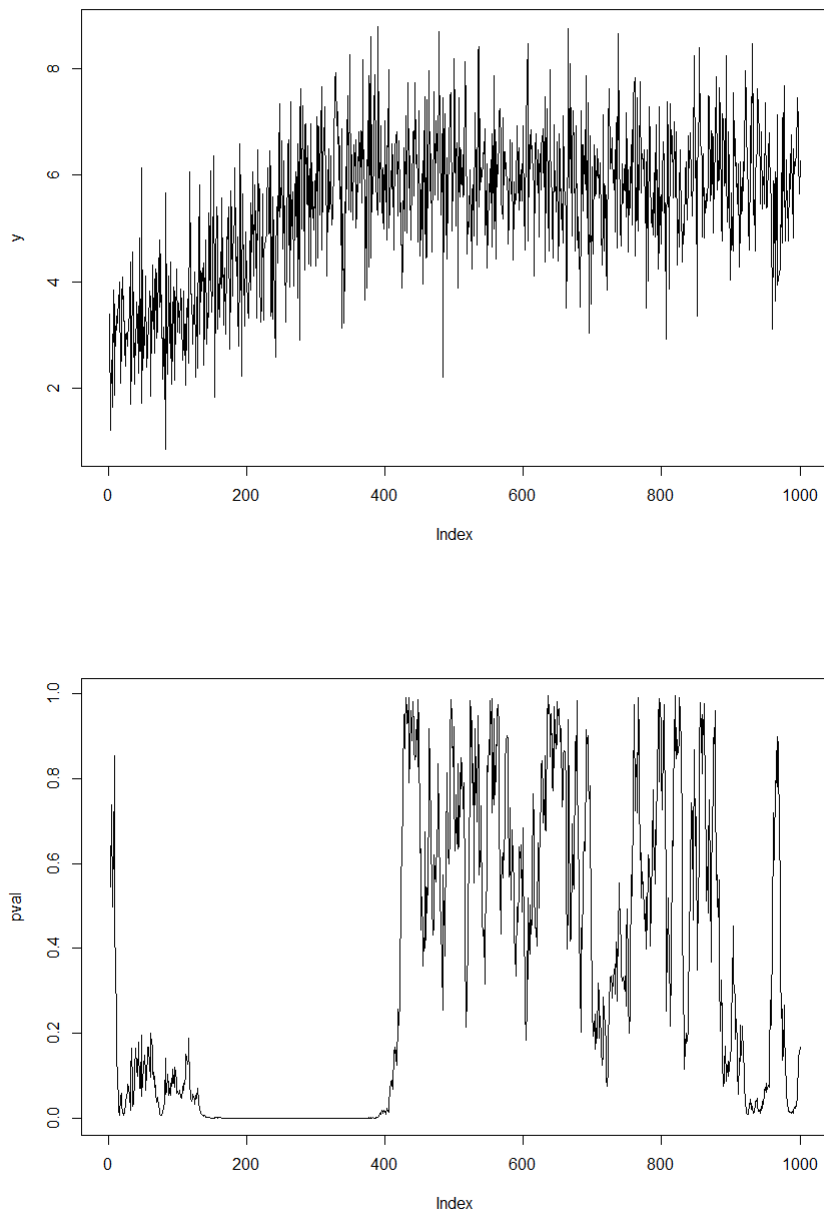


Abbildung 4.8.: Datenverlauf und die dazugehörige p-Value Zeitreihe

Nun wird die Zeitreihe der p-values untersucht. Hier verwendet man ebenfalls ein gleitendes Zeitfenster und überprüft auf ungefähre Gleichverteilung der p-Values. In diesem Fall wird einfach geprüft ob $x\%$ der Daten einen Wert von $x/100$ unterschreiten. Falls dies der Fall ist, so sind sehr viele p-values zu klein, was bedeutet, dass eine signifikante Steigung vorhanden ist und daher keine Beharrung angenom-

4.2. Eine neue Methode für den Motorprüfstand

men werden kann. Wenn allerdings sehr viele p-Values "groß" sind, ist dies ein Indiz dafür, dass die Steigung eine längere Zeit ungefähr 0 war. Diese Bereiche werden nun im Datenverlauf grau schattiert.

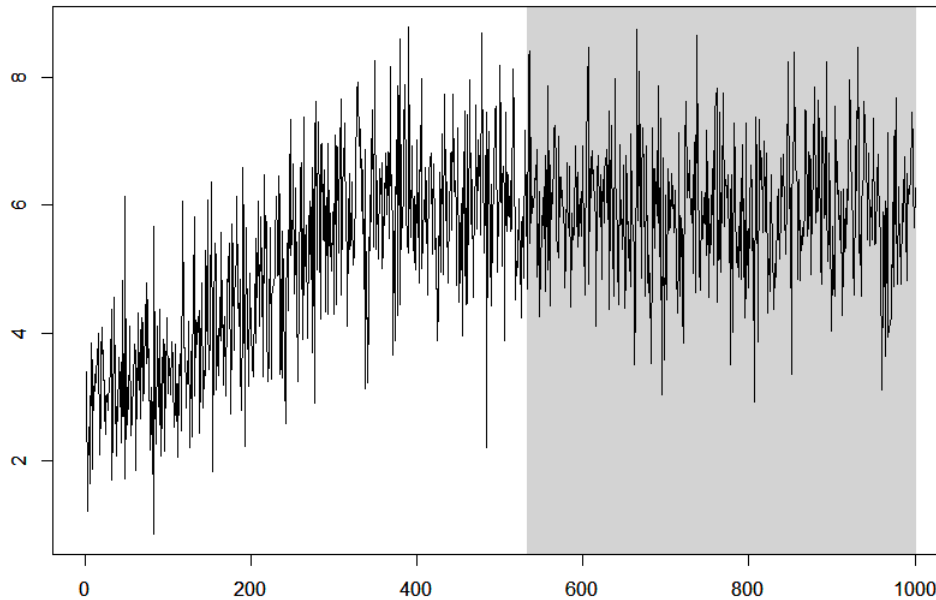


Abbildung 4.9.: Erkannte Beharrung der Datenzeitreihe

Man erkennt sofort, dass die Beharrung erst dann erkannt wurde als sie wirklich eingetreten ist. Zu Beginn war eine Steigung vorhanden und erst als diese beendet war, begann die Schattierung. Es kann nicht sofort nach Beendigung der Steigung die Beharrung entdeckt werden, da es sich beispielsweise nur um ein Wechseln von positiver Steigung zu einer negativen Steigung handeln könnte. Man muss sicherstellen, dass die Steigung ausreichend lange unverändert klein bleibt, um annehmen zu können, dass sie sich auch weiterhin so verhält. Dies wird mit der Überprüfung auf Gleichverteilung des gleitenden Zeitfensters auf der p-Value Zeitreihe ermöglicht. Denn erst, wenn ausreichend viele p-Values dieses Verhalten bestätigen kann man davon ausgehen, dass dieses Verhalten weiterhin bestehen bleibt. Man kann in diesem Beispiel mit freiem Auge die Beharrung erkennen und auch mit freiem Auge verifizieren, dass die Methode zufriedenstellende Ergebnisse liefert.

Ein zusätzlicher Vorteil dieser Herangehensweise ist, dass auch das Ende eines beharrten Zustands erkannt werden kann. Alternativ könnte man hierzu Changepoint Methoden verwenden.

4.3. Das Changepoint Problem

Das Changepoint Problem wird herangezogen um abrupte Änderungen eines stationären Prozesses zu ermitteln. Eine solche Detektion kann in unserem Fall von essentieller Bedeutung sein, da es wichtig ist so schnell wie möglich zu erkennen wann eine Messgröße nicht mehr stationär ist. Ein sehr beliebtes Beispiel, um die Problematik zu beschreiben, ist die Verwendung einer fairen und einer unfairen Münze. Angenommen man verwendet für ein Glücksspiel eine faire Münze (das bedeutet $\mathbb{P}(\text{Kopf}) = \mathbb{P}(\text{Zahl}) = 0.5$). Nach einiger Zeit hat einer der Spieler den Verdacht, dass die faire Münze ausgetauscht wurde. Wenn sich diese Annahme bestätigt, will man herausfinden, zu welchem Zeitpunkt der Austausch vorgenommen wurde.

Für diesen Spezialfall kann man in PETTITT (1980) nachlesen wie das Problem für eine Bernoulli Variable gelöst worden ist. Für den etwas allgemeineren Fall betrachte man die folgende Situation:

Sei Y_1, \dots, Y_n die Zeitreihe der Response Variablen. Jedes Y_i kann geschrieben werden als: $Y_i = \mu_i + \epsilon_i$ wobei μ_i der Erwartungswert von Y_i und $\epsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$ weißes Rauschen. Wir nehmen an, dass die ϵ_i aus einer Normalverteilung stammen. Nun gilt es noch eine gültige Hypothese aufzustellen um herauszufinden ob und wann ein Changepoint aufgetreten ist.

$$H_0 : \underbrace{\mu_1 = \mu_2 = \dots = \mu_n}_{=\mu} \quad \text{vs.} \quad H_1 : \underbrace{\mu_1 = \mu_2 = \dots = \mu_k}_{=\mu} \neq \underbrace{\mu_{k+1} = \mu_{k+2} = \dots = \mu_n}_{=\eta}$$

In der Alternativhypothese wird ein Changepoint zum Zeitpunkt k . Der klassische Ansatz verwendet nun die sogenannte CUSUM (cumulative sum) Methode. Definiere $S_i := \sum_{j=1}^i Y_j$. S_i , also die Partialsumme bis zur i -ten Response. Unter H_0 gilt

natürlich $S_i \approx i \cdot \mu$. Sei weiters $T_i := \sum_{j=1}^i \epsilon_j$.

Sei nun die Teststatistik $C_i = \frac{1}{\sqrt{n}} (S_i - \frac{i}{n} S_n)$. Es gilt $C_i \stackrel{H_0}{=} \frac{1}{\sqrt{n}} (T_i - \frac{i}{n} T_n)$. Unter H_0 gilt für C_i außerdem:

$$\begin{aligned} \bullet \quad \mathbb{E}(C_i) &= \mathbb{E} \left(\frac{1}{\sqrt{n}} \left(T_i - \frac{i}{n} T_n \right) \right) = \frac{1}{\sqrt{n}} \mathbb{E} \left((\epsilon_1 + \dots + \epsilon_i) - \frac{i}{n} (\epsilon_1 + \dots + \epsilon_n) \right) \\ &= \frac{1}{\sqrt{n}} \left(\mathbb{E}(\epsilon_1) + \dots + \mathbb{E}(\epsilon_i) + \frac{i}{n} (\mathbb{E}(\epsilon_1) + \dots + \mathbb{E}(\epsilon_n)) \right) = 0 \end{aligned}$$

$$\begin{aligned}
 \bullet \text{ Var}(C_i) &= \text{Var} \left(\frac{1}{\sqrt{n}} \left(T_i - \frac{i}{n} T_n \right) \right) \\
 &= \frac{1}{n} \text{Var} \left[\left(1 - \frac{i}{n} \right) (\epsilon_1 + \dots + \epsilon_i) - \frac{i}{n} (\epsilon_{i+1} + \dots + \epsilon_n) \right] \\
 &= \frac{1}{n} \text{Var} \left[\left(\frac{n-i}{n} \right) (\epsilon_1 + \dots + \epsilon_i) - \frac{i}{n} (\epsilon_{i+1} + \dots + \epsilon_n) \right] \\
 &= \frac{1}{n^3} \text{Var} ((n-i)(\epsilon_1 + \dots + \epsilon_i) - i(\epsilon_{i+1} + \dots + \epsilon_n)) \\
 &= \frac{1}{n^3} ((n-i)^2 (i \cdot \sigma^2) + i^2 ((n-i)\sigma^2)) = \frac{1}{n^3} ((n-i)i ((n-i)\sigma^2 + i\sigma^2)) \\
 &= \frac{(n-i)i}{n^2} \sigma^2
 \end{aligned}$$

Unter H_1 sei unser Changepoint $k = \theta \cdot n$, $\theta \in (0, 1)$. Dann gilt

$$C_k = \frac{1}{\sqrt{n}} \left(S_k - \frac{k}{n} S_n \right) = \frac{1}{\sqrt{n}} \left(\underbrace{k\mu - \frac{k}{n}(k\mu + (n-k)\eta)}_{*} + \underbrace{\left(T_k - \frac{k}{n} T_n \right)}_{\circ} \right)$$

Der Ausdruck (\circ) ist die Summe von Zufallsvariablen mit Erwartungswert 0 und kann daher vernachlässigt werden. Ersetze im Ausdruck $(*)$ das k durch $\theta \cdot n$. Dann folgt

$$(*) = \left(\theta n \mu - \frac{\theta n}{n} (\theta n \mu + (n - \theta n) \eta) \right) = \theta (n \mu - (\theta n \mu + n(1 - \theta) \eta)) := \alpha(\theta, \mu, \eta) n$$

Weiters gilt $\frac{1}{\sqrt{n}} \alpha n = \sqrt{n} \alpha \rightarrow \infty$ mit Geschwindigkeit \sqrt{n} . Wir sind also interessiert ab welchem k das C_k zu groß wird. Dazu bestimmen wir nun die Verteilung von C_i unter H_0 . Da wir zu Beginn angenommen haben, dass jedes ϵ_i aus einer Normalverteilung stammt, ist natürlich C_i als Summe ebenfalls wieder normalverteilt. Es gilt $C_i \sim N(0, \frac{(n-i)i}{n^2} \sigma^2)$. Als Teststatistik verwendet man nun $Z_n := \max_{1 \leq i \leq n} |C_i|$ und unter H_0 gilt

$$Z_n \xrightarrow{d} \sup_{t \in [0,1]} |BB(t)| \text{ wobei } BB(t) \text{ die Brown'sche Brücke ist.}$$

Da man nun aber die Verteilung von C_i unter H_0 kennt, braucht man nur das betragsmäßig größte C_i mit dem vorgegebenen α der Normalverteilung vergleichen um zu wissen wann die Nullhypothese verworfen werden kann. Genaueres und weitere interessante Aspekte können in CAO et al. (2015) nachgelesen werden.

4. Methoden der Beharrungserkennung

Ein kurzes Beispiel zur Demonstration. Es wurden 500 Werte $\sim N(0, 1)$ und 500 Werte $\sim N(5, 1)$ generiert. Auf diesen Datensatz angewandt führt man den Befehl "cpt.mean" in \mathbb{R} aus. Das Ergebnis sieht dann wie folgt aus:

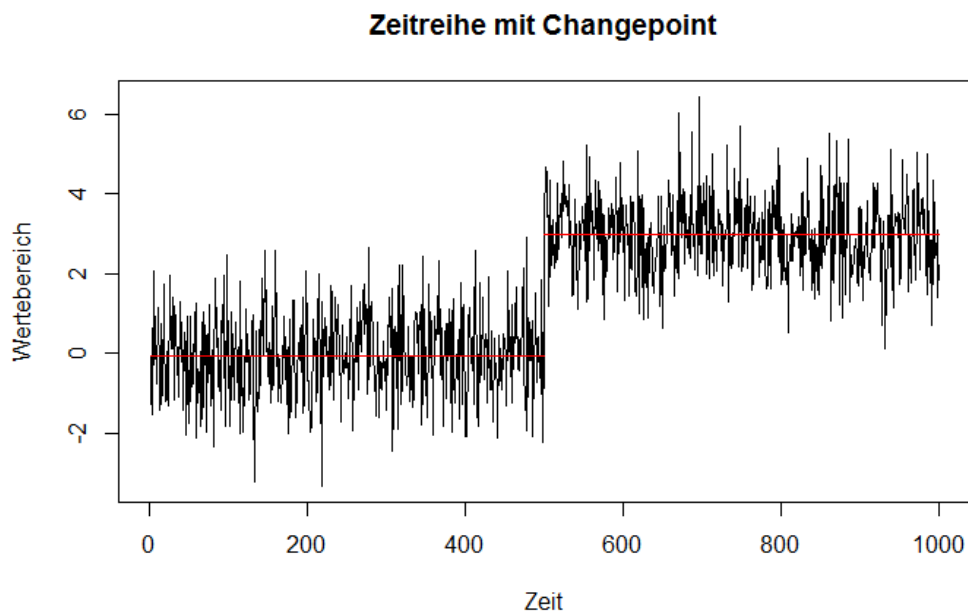


Abbildung 4.10.: Analysierter Changepoint

summary(.) :

```
Created Using changepoint version 2.2.2
Changepoint type      : Change in mean
Method of analysis    : AMOC
Test Statistic       : Normal
Type of penalty       : MBIC with value , 20.72327
Minimum Segment Length : 1
Maximum no. of cpts   : 1
Changepoint Locations : 500
```

In diesem Fall wurde der exakte Zeitpunkt erkannt, an dem sich das Verhalten geändert hat. Durch die explorative Datenanalyse konnten wir feststellen, dass die einzelnen Messwerte womöglich nicht unabhängig sind. Aus diesem Grund wird nun kurz erklärt wie unser Verfahren bei korrelierten Daten angepasst werden muss.

4.4. Korrelation & Tests

Der in Kapitel 4.2 beschriebene Ansatz basiert auf der Annahme, dass im Regressionsmodell $Y_i = \beta_0 + i\beta_1 + \epsilon_i$ die Fehler ϵ_i *iid* $N(0, \sigma^2)$ verteilt sind für $1 \leq i \leq n$. Da wir aber gesehen haben, dass die ϵ_i in unserem Fall abhängig sind, muss ein Test auf $\beta_1 = 0$ entsprechend adaptiert werden. In unserem Fall ist $(\epsilon_1, \dots, \epsilon_n)^\top = \epsilon \sim N(0, \Omega)$ mit Ω der Varianz-Kovarianz Matrix und im allgemeinen ist $\Omega \neq \sigma^2 I$. Das bedeutet insbesondere wenn die Fehler (ϵ_i) einen stationären Prozess definieren mit Kovarianzfunktion $\rho(h)$, dann gilt

$$\Omega = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \dots & \omega_{1,n} \\ \omega_{2,1} & \omega_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \omega_{n-1,n} \\ \omega_{n,1} & \dots & \omega_{n,n-1} & \omega_{n,n} \end{pmatrix} = \begin{pmatrix} \rho(0) & \rho(1) & \dots & \rho(n-1) \\ \rho(1) & \rho(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho(1) \\ \rho(n-1) & \dots & \rho(1) & \rho(0) \end{pmatrix}$$

Nun soll für $Y = \mu + \epsilon = \beta_0 \cdot \mathbb{1} + \beta_1 x + \epsilon$ bestimmt werden ob die Steigung relevant ist. Hier ist $x = (1, \dots, n)$ und $\mathbb{1} = (1, \dots, 1)$ wobei x und $\mathbb{1}$ dieselbe Dimension aufweisen.

$$H_0 : \beta_1 = 0 \quad vs. \quad H_1 : \beta_1 \neq 0.$$

Ein klassischer Weg ist die Bestimmung des Likelihood-Quotienten-Tests. Bezeichne mit $f_Y(y_1, \dots, y_n | \beta_0, \beta_1)$ die Dichte von Y . Dann ist der Likelihood-Quotienten-Test definiert durch

$$L = \frac{\sup_{\beta_1 \in \mathbb{R}} f_Y(y_1, \dots, y_n | \beta_0, 0)}{\sup_{\beta_0, \beta_1 \in \mathbb{R}} f_Y(y_1, \dots, y_n | \beta_0, \beta_1)}$$

Da wir in unserem Fall $\epsilon \sim N(0, \Omega)$ angenommen haben, gilt $Y \sim N(\beta_0 \mathbb{1} + \beta_1 x, \Omega)$ wobei $(\mathbb{1}, x) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \end{pmatrix}^\top$ und wir Ω als bekannt annehmen. Daher kann die Likelihood-Funktion geschrieben werden als

$$f_Y(\beta_0, \beta_1 | Y) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\sqrt{|\Omega|}} \exp\left(-\frac{1}{2}(Y - \mu)^\top \Omega^{-1} (Y - \mu)\right)$$

4. Methoden der Beharrungserkennung

Der Likelihood Quotient sieht nun wie folgt aus

$$\Lambda_Y = \frac{\sup_{\beta_0 \in \mathbb{R}} L(\beta_0, 0|x)}{\sup_{\beta_0, \beta_1 \in \mathbb{R}} L(\beta_0, \beta_1|x)} = \frac{\sup_{\beta_0 \in \mathbb{R}} \exp\left(-\frac{1}{2}(Y - \mathbb{1}\beta_0)^\top \Omega^{-1}(Y - \mathbb{1}\beta_0)\right)}{\sup_{\beta_0, \beta_1 \in \mathbb{R}} \exp\left(-\frac{1}{2}(Y - (\mathbb{1}\beta_0 + \beta_1 x))^\top \Omega^{-1}(Y - (\mathbb{1}\beta_0 + \beta_1 x))\right)} \quad (4.2)$$

Hier ist das Maximieren der Likelihood-Funktion äquivalent mit dem Minimieren des Exponenten. Sei dazu $\omega'_{i,j}$ der Eintrag für Ω^{-1} . Für H_0 erhält man dafür den Ausdruck

$$\sum_{i=1}^n \sum_{j=1}^n (Y_i - \beta_0) \omega'_{i,j} (Y_j - \beta_0) = \sum_{i=1}^n \sum_{j=1}^n Y_i \omega'_{i,j} Y_j - \omega'_{i,j} \beta_0 (Y_i + Y_j) + \beta_0^2 \omega'_{i,j}$$

Unter H_1 kann dieses Vorgehen ebenfalls durchgeführt werden. Dabei ist ein entscheidender Faktor, dass $x_i = i$ vorliegt.

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n (Y_i - \beta_0 - i\beta_1) \omega'_{i,j} (Y_j - \beta_0 - j\beta_1) = \\ \sum_{i=1}^n \sum_{j=1}^n Y_i \omega'_{i,j} Y_j - \beta_0 \omega'_{i,j} (Y_i + Y_j) + \beta_0^2 \omega'_{i,j} - \beta_1 \omega'_{i,j} (iY_j + jY_i) + ij \omega'_{i,j} \beta_1^2 + \beta_0 \beta_1 \omega'_{i,j} (i + j) \end{aligned} \quad (4.3)$$

Natürlich ist man nun an den Werten von β_0 und β_1 interessiert für die der Ausdruck 4.3 minimal ist. Leite dazu 4.3 nach β_0 und nach β_1 ab und erhalte die Normalengleichungen wie in Kapitel 3 $\hat{\beta}_0$ und $\hat{\beta}_1$ ergeben sich dann wie folgt.

$$\hat{\beta}_0 = \frac{\sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j} (Y_i + Y_j - \hat{\beta}_1(i+j))}{2 \sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j}}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j} (iY_j + jY_i) - \left(\frac{\sum_{i'=1}^n \sum_{j'=1}^n \omega'_{i',j'} (Y_{i'} + Y_{j'})}{2 \sum_{i'=1}^n \sum_{j'=1}^n \omega'_{i',j'}} \right) \sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j}}{2 \sum_{i=1}^n \sum_{j=1}^n ij \omega_{i,j} - \left(\frac{\sum_{i'=1}^n \sum_{j'=1}^n (i+j) \omega'_{i',j'}}{2 \sum_{i'=1}^n \sum_{j'=1}^n \omega'_{i',j'}} \right) \sum_{i=1}^n \sum_{j=1}^n \omega_{i,j} (i+j)}$$

Auf den ersten Blick wirken diese Ausdrücke möglicherweise sehr fremdartig, doch man kann sich leicht vergewissern, dass die Rechnung stimmt. Setze dafür $\Omega = I$, dann ist man im klassischen Fall der linearen Regression. Ω^{-1} ist dann natürlich ebenfalls die Einheitsmatrix und damit ist $\omega'_{i,j} = 0$ für $i \neq j$ und $\omega_{i,j} = 1$ für $i = j$. Für $\hat{\beta}_0$ ergibt sich dadurch

$$\hat{\beta}_0 = \frac{\sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j} (Y_i + Y_j - \beta_1(i+j))}{2 \sum_{i=1}^n \sum_{j=1}^n \omega'_{i,j}} = \frac{\sum_{i=1}^n 1(2Y_i + \beta_1(2i))}{2 \sum_{i=1}^n 1} = \frac{2 \sum_{i=1}^n Y_i + 2\beta_1 \sum_{i=1}^n i}{2 \sum_{i=1}^n 1}$$

$$= \frac{n\bar{Y} + \beta_1 \frac{n(n+1)}{2}}{n} = \bar{Y} + \frac{n+1}{2} \beta_1$$

Was exakt jenem kleinste Quadrate Schätzer in Kapitel 3 entspricht für $x = (1, 2, \dots, n)$. Doch nun soll bestimmt werden, wann H_0 überhaupt zutrifft. Dazu verwenden wir eine Transformation von Y . Laut unserem Modell gilt $Y = X\beta + \epsilon$. Sei nun $\tilde{Y} = \Omega^{-1/2}Y = \Omega^{-1/2}X\beta + \Omega^{-1/2}\epsilon = \tilde{X}\beta + \tilde{\epsilon}$. Durch diese Transformation entsteht $\tilde{\epsilon} \sim N(0, I_n)$ und somit ist $\tilde{Y} \sim N(\tilde{X}\beta, I_n)$. Für dieses Modell können wir nun den kleinste Quadrate Schätzer aus Kapitel 3 angeben.

$$\hat{\beta} = \left(\tilde{X}^\top \tilde{X} \right)^{-1} \tilde{X}^\top \tilde{Y} = \left(X^\top \Omega^{-1/2} \Omega^{-1/2} X \right)^{-1} X^\top \Omega^{-1/2} \Omega^{-1/2} Y$$

$$= \left(X^\top \Omega^{-1} X \right)^{-1} X^\top \Omega^{-1} X \beta + \left(X^\top \Omega^{-1} X \right)^{-1} X^\top \Omega^{-1} \epsilon = \beta + \left(X^\top \Omega^{-1} X \right)^{-1} X^\top \Omega^{-1} \epsilon$$

4. Methoden der Beharrungserkennung

Unter H_0 ist

$$\hat{\beta} = \hat{\beta}_0 = \beta_0 + (\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} \epsilon$$

Damit kann der Ausdruck in 4.2 neu geschrieben werden und zwar gilt dann

$$\Lambda_Y = \frac{\exp\left(-\frac{1}{2}(Y - \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y)^\top \Omega^{-1} (Y - \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y)\right)}{\exp\left(-\frac{1}{2}(Y - X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y)^\top \Omega^{-1} (Y - X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y)\right)}$$

Man kann den Exponenten nun zu einem großen Term zusammenfassen. Dazu multiplizieren wir ihn zuerst mit (-2) und erhalten folgenden Ausdruck für den Exponenten:

$$\begin{aligned} M &= ((Y - \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y)^\top \Omega^{-1} (Y - \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y)) - \\ &\quad \left(-\frac{1}{2}(Y - X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y)^\top \Omega^{-1} (Y - X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y) \right) \\ &= Y^\top \Omega^{-1} Y - 2Y^\top \Omega^{-1} \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y + Y^\top \Omega^{-1} \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top Y \\ &\quad - Y^\top \Omega^{-1} Y - 2Y^\top \Omega^{-1} X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y \\ &\quad + Y^\top \Omega^{-1} X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y \\ &= Y^\top \Omega^{-1} X(X^\top \Omega^{-1} X)^{-1} \Omega^{-1} X^\top Y - Y^\top \Omega^{-1} \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top \Omega^{-1} Y \end{aligned}$$

Sei nun $Z := \Omega^{-1/2} Y$ und gehen wir davon aus, dass Y zentriert ist. Dann ist $Z \sim N(0, I_n)$ und M lässt sich schreiben als

$$M = Z^\top \left[\Omega^{-1/2} \left(\underbrace{X(X^\top \Omega^{-1} X)^{-1} X^\top - \mathbb{1}(\mathbb{1}^\top \Omega^{-1} \mathbb{1})^{-1} \mathbb{1}^\top}_{:=A} \right) \Omega^{-1/2} \right] Z$$

Man kann zeigen, dass unter gewissen Bedingungen an Ω bzw Ω^{-1} die Matrix A nicht negativ definit ist. Weiters kann man $\Omega^{-1/2} A \Omega^{-1/2}$ mithilfe des Spektralsatzes darstellen als $\Omega^{-1/2} A \Omega^{-1/2} = O \Lambda O^\top$, wobei $O O^\top = I_n$ und $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ mit λ_i der i -te Eigenwert von $\Omega^{-1/2} A \Omega^{-1/2}$. Hier sind alle $\lambda_i \geq 0$. Dann ist

$$M = Z^\top \Omega^{-1/2} A \Omega^{-1/2} Z = Z^\top O \Lambda O^\top Z = W^\top \Lambda W,$$

wobei $W \sim N_n(0, O I_n O^\top) = N_n(0, I_n)$ verteilt ist und somit $w_i \stackrel{iid}{\sim} N(0, 1)$. Also ist

$$M = \sum_{i=1}^n w_i^2 \lambda_i,$$

was einer gewichteten Summe von χ_1^2 verteilten Zufallsvariablen. Leider ist eine gewichtete Summe dieser Zufallsvariablen nicht wieder χ^2 -verteilt. Allerdings kann man entweder die Verteilung schätzen oder die Wahrscheinlichkeit berechnen, dass diese gewichtete Summe einen Schwellwert nicht überschreitet. Dazu kann man in BAUSCH (2013), SOLOMON et al. (1977) oder in JEZ (2015) diverse Methoden und Ergebnisse nachschlagen.

Nun führen wir noch eine kleine Simulationsstudie durch, um weitere Erkenntnisse über diesen Test zu erlangen. Sei also $T := Y^\top \Omega^{-1} A \Omega^{-1} Y$ unsere Teststatistik. Wir simulieren dafür einen AR(1) Prozess ($X_t = \rho X_{t-1} + \epsilon_t$ wobei $\epsilon_t \stackrel{iid}{\sim} WN(0, \sigma^2)$) für verschiedene ρ 's und für verschiedene Werte von n wobei n der Anzahl der Beobachtungen entspricht. Ω hängt hier von ρ und n ab und sieht für einen AR(1) Prozess folgendermaßen aus:

$$\Omega = \frac{1}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \rho \\ \rho^{n-1} & \dots & \dots & \rho & 1 \end{pmatrix}.$$

Wir haben nun für simulierte Daten zunächst eine Tabelle von kritischen Werten für unsere Teststatistik erstellt. Dabei wurden für verschieden n und verschieden ρ Daten simuliert und das 95% Quantil bestimmt. Die Tabelle sieht nun wie folgt aus:

95% Quantil	n=100	n=200	n=300	n=500
rho = 0.1	4,6753	3,6670	3,7682	3,5077
rho = 0.4	4,8864	3,7746	3,6268	3,8798
rho = 0.7	3,0409	3,6844	3,4022	3,6743
rho = 0.9	3,5463	4,4423	4,0203	4,0199

Abbildung 4.11.: Kritische Werte der 95% Quantile

Mithilfe dieser Werte können wir nun auch einen Vergleich der beiden Tests machen. Wir wenden einmal den adaptierten Test und einmal den klassischen Test auf simulierte Daten an. Dabei wird einmal unter H_0 getestet (kein Trend vorhanden) und einmal unter H_1 (Trend vorhanden). Wir betrachten den relativen Anteil an Tests,

4. Methoden der Beharrungserkennung

die verworfen werden. Unter H_0 sollte der adaptierte Test eine Verwerfungsrate von ungefähr 5% besitzen. Die folgende Tabelle bestätigt diese Vermutung.

	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test
n	100		200		300		500	
rho=0,1	0,03	0,05	0,06	0,075	0,046666	0,09	0,054	0,07
rho=0,4	0,02	0,18	0,04	0,165	0,0466666	0,1733333	0,05	0,184
rho=0,7	0,03	0,39	0,05	0,435	0,0366667	0,393333	0,068	0,422
rho=0,9	0,04	0,97	0,02	0,64	0,046667	0,653333	0,042	0,662

Abbildung 4.12.: Anteil an fälschlicherweise verworfenen H_0

Unser adaptierter Test liefert also sowohl für hohe als auch kleine Korrelation immer das gewünschte Resultat. Der klassische Test hingegen legt ein gänzlich anderes Verhalten an den Tag. Je größer die Korrelation, desto öfter wird fälschlicherweise die Nullhypothese verworfen. Unsere Vermutungen werden also für Daten unter H_0 bestätigt. Betrachten wir nun auch noch die Ergebnisse, wenn wir wissen, dass H_1 korrekt ist. Falls nun auch hier der adaptierte Test bessere Resultate erzielt, haben wir eine ausgesprochen effektive Methode gefunden, korrelierte Daten zu bewerten. Dazu führen wir künstlich unterschiedlich große Steigungen ein. Eine Steigung von 0.005 bedeutet, dass jeder Datenpunkt im Vergleich zum Vorgänger um 0.005 angehoben wird. Die Ergebnisse für die verschiedenen n und Trends sehen wie folgt aus:

4.4. Korrelation & Tests

n=100	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test
Trend	0,005		0,01		0,02	
rho=0,1	0,16	0,21	0,68	0,75	1	1
rho=0,4	0,07	0,28	0,42	0,75	0,91	0,99
rho=0,7	0,14	0,39	0,21	0,61	0,53	0,9
rho=0,9	0,06	0,69	0,07	0,75	0,14	0,71

n=200	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test
Trend	0,005		0,01		0,02	
rho=0,1	0,97	0,97	1	1	1	1
rho=0,4	0,725	0,895	1	1	1	1
rho=0,7	0,33	0,69	0,655	0,945	1	1
rho=0,9	0,065	0,67	0,0725	0,7	0,455	0,915

n=300	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test
Trend	0,005		0,01		0,02	
rho=0,1	1	1	1	1	1	1
rho=0,4	0,9933	1	1	1	1	1
rho=0,7	0,716667	0,93	0,9933	1	1	1
rho=0,9	0,103333	0,756667	0,346667	0,886667	0,89	0,9933

n=500	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test	Adaptierter Test	Klassischer Test
Trend	0,005		0,01		0,02	
rho=0,1	1	1	1	1	1	1
rho=0,4	1	1	1	1	1	1
rho=0,7	1	1	1	1	1	1
rho=0,9	0,376	0,896	0,932	1	1	1

Abbildung 4.13.: Anteil an verworfenen H_0

Wenn H_1 korrekt ist, so liefert der klassische Test die "besseren" Resultate. Allerdings verwirft dieser auch unter H_0 teilweise mit einer Rate von über 90%. Der klassische Test ist also verzerrt, denn wenn H_0 korrekt ist, wird die Verwerfungsrate höher für großes ρ . Wenn H_1 korrekt ist, nimmt die Verwerfungsrate bei großem ρ sogar bei beiden Tests ab. Zusammenfassend kann man sagen, dass unter H_0 der adaptierte Test wesentlich bessere Resultate liefert, während der klassische Test durch seine Verzerrung unzuverlässigere Werte liefert.

5. Die Messfreigabe

Dieses große Kapitel beschreibt nun im Detail die Vorgehensweise für eine Messung am Motorprüfstand und die einzelnen Komponenten. Als erstes stellt man den Betriebspunkt ein und lässt den Motor so lange laufen bis man die Messfreigabe erhält. Die Messfreigabe legt jenen Zeitpunkt fest, ab dem eine Messung gestartet werden kann. Die Freigabe besteht aus einer Beharrungserkennung und gegebenenfalls einem Soll-Ist-Vergleich. Nachdem eine Messung durchgeführt wurde, liegt ein Messpunkt vor, den es zu bewerten gibt. Dieses Bewerten nennt man im Folgenden Messpunktdiagnose und wird im nächsten Kapitel genauer erläutert. Sollte dieser Messpunkt fehlerfrei sein, so liegt eine erfolgreiche Messung vor. Bei einer negativen Bewertung kann der Prüfstandsfahrer das Ergebnis allerdings trotzdem manuell auf positiv verändern.

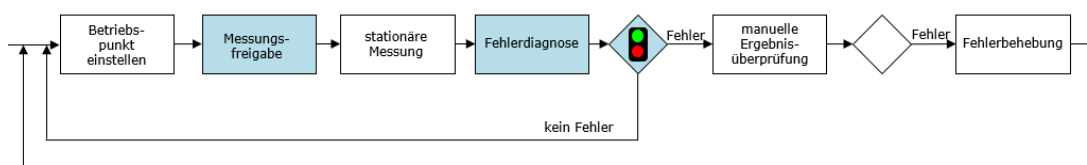


Abbildung 5.1.: Übersicht der Vorgehensweise

Die abgebildete Grafik veranschaulicht den soeben beschriebenen Ablauf. Die Felder Messungsfreigabe und Fehlerdiagnose wurden blau gefärbt da die beiden Punkte in dieser Arbeit behandelt werden, wobei Fehlerdiagnose für den Begriff Messpunktdiagnose steht. Die Grafik wurde von DI Michael Wohlthan im Zuge seiner Dissertation erstellt und freundlicherweise für diese Masterarbeit zur Verfügung gestellt. Im Folgenden Abschnitt wird zunächst kurz der Soll-Ist-Vergleich beschrieben und danach die Beharrungserkennung im Detail erklärt. Danach wird ein umfangreicher Überblick über die Messfreigabe als programmierte Funktion gegeben.

5.1. Der Soll-Ist-Vergleich

Ein entscheidendes Kriterium für die Messfreigabe ist eine Überprüfung des Wertebereichs. Wenn es Sollwerte gibt die eingehalten werden müssen, so müssen die Daten natürlich ungefähr diesem Wert entsprechen. Zu jedem gegebenen Sollwert gibt es auch einen Toleranzbereich. Im ersten Schritt wird also überprüft ob der Mittelwert der Daten mit dem vorgegebenen Wert ungefähr übereinstimmt. Der Mittelwert wird natürlich nicht von allen Datenpunkten genommen. Es wird immer nur der Mittelwert der aktuellsten Werte berechnet und mit dem Sollwert verglichen. In der Regel sind es 150 Messwerte die für diese Berechnung beziehungsweise Bewertung herangezogen werden. In der folgenden Grafik kann man den Soll-Ist-Vergleich sehr gut erkennen. In der Praxis hat man natürlich immer nur einen Sollwert der gerade angenommen werden soll. Aus Gründen der Einfachheit, haben wir angenommen, all diese Sollwerte aber schon zu kennen.

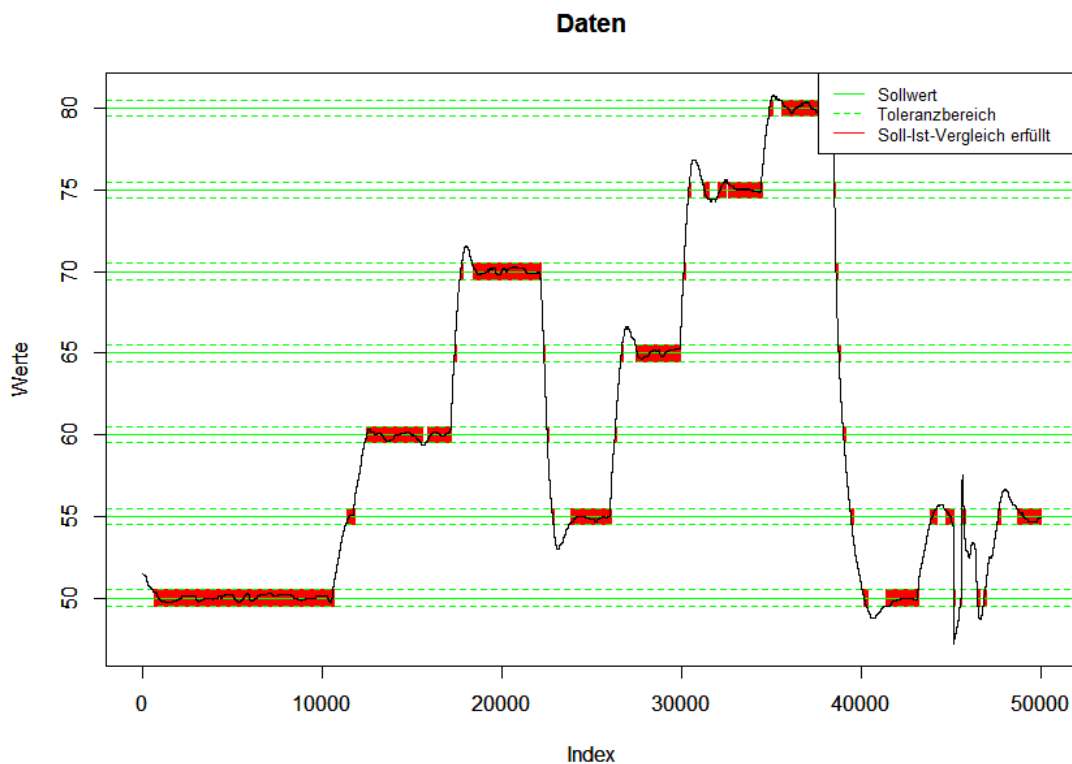


Abbildung 5.2.: Soll-Ist-Überprüfung des Datenverlaufs

Die Stellen an denen 150 Messwerte den Mittelwertsvergleich erfüllt haben wurden

rot markiert. Doch diese Überprüfung reicht natürlich bei weitem nicht aus um eine Messfreigabe zu erlauben. Man sieht speziell im Indexbereich von ca. 38.000 - 40.000 den Grund dafür. Die Messwerte fallen rapide ab und passieren dadurch die Sollwerte. Der Mittelwert tut dies natürlich ebenfalls, aber nur für sehr kurz Zeit. Hätte man in diesem Bereich zu messen begonnen, müsste man den erhaltenen Messpunkt sofort wieder verwerfen. Das ist der Grund, warum eine Beharrungserkennung notwendig ist. Des Weiteren stehen Sollwerte nicht immer zur Verfügung, was die Beharrungserkennung sogar unverzichtbar werden lässt.

5.2. Die Beharrungserkennung

In diesem Abschnitt wird nun die bereits vorgestellte Methode angewendet. Es gab allerdings auch diverse prozessspezifische Probleme und daher musste die Methode mit einigen Tunings verbessert werden. Die aufgetretenen Schwierigkeiten und die Lösungsvorschläge werden nun vorgestellt und präsentiert. Schlussendlich wird alles zu einem großen Konstrukt zusammengefasst und die Beharrungserkennung als algorithmischer Pseudocode dargestellt. Es gibt auch eine Fallunterscheidung, auf die allerdings erst am Ende eingegangen wird da sich die einzelnen Fälle nur geringfügig voneinander unterscheiden.

5.2.1. Probleme dieser Methode

Zu Beginn hat man durch die explorative Datenanalyse festgestellt, dass die Daten verschiedenste Eigenschaften aufweisen. Die zuvor beschriebene Herangehensweise funktioniert nicht bei allen Daten gleich gut und man muss viele Spezialfälle berücksichtigen. Für diese Spezialfälle sollte natürlich die Methode, sofern möglich, angepasst werden um ein einheitliches Programm entwickeln zu können. Die aufgetretenen Probleme werden in diesem Abschnitt erläutert.

5.2.1.1. Fehlinterpretation der p-Values

Das größte Problem liegt in einer "Fehlinterpretation" des p-Values. Es wird durchgehend dieser Begriff unter Anführungszeichen gesetzt, da die Aussage des p-Values immer richtig ist. Allerdings gibt es Fälle, in denen das Verhalten der Messgrößen eine Fehlinterpretation zulassen können. Anders ausgedrückt: Die Datensituation wird durch den p-Value falsch bewertet. Das bedeutet, dass entweder Beharrung angezeigt wird obwohl keine vorhanden ist, oder aber keine Beharrung angezeigt wird

5. Die Messfreigabe

obwohl diese mit freiem Auge ersichtlich ist.

Ein solcher Fall tritt zum Beispiel ein, wenn Daten eine sehr geringe Streuung aufweisen. Der errechnete Slope Parameter ist in diesen Fällen quasi 0, was eindeutig für eine Beharrung spricht. Der p-Value hingegen bleibt trotzdem ausgesprochen klein und verwirft die Nullhypothese der Beharrung. Eine soeben erklärte "Fehlinterpretation" kann im folgenden Beispiel sehr gut beobachtet werden. Dieses Beispiel ist ein Ausschnitt einer Messgröße aus einer echten Messreihe.

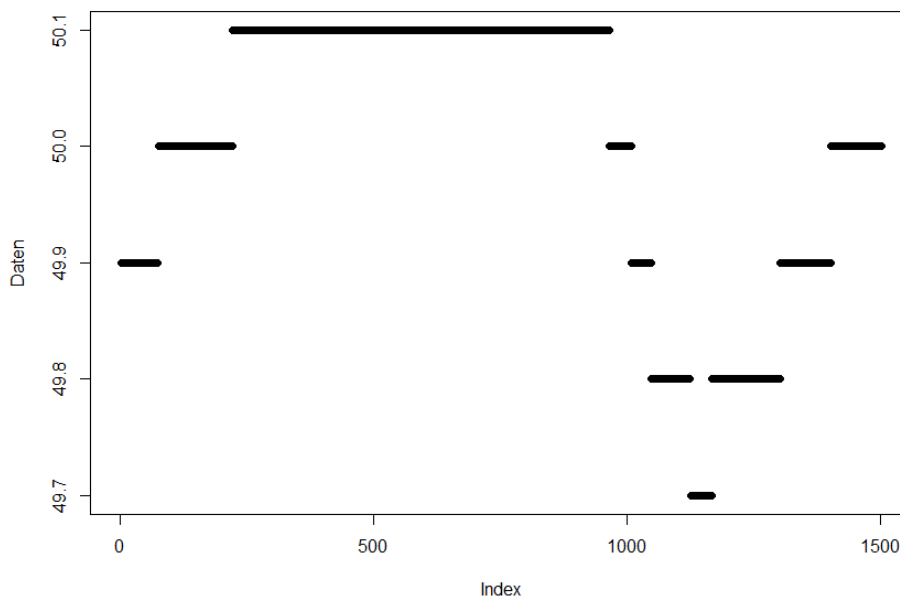


Abbildung 5.3.: Datenzeitreihe die zu einer Fehlinterpretation führt

```
> summary(lm(Daten ~ Index))$coefficients [2, 1]
[1] -0.0001304309
```

```
> summary(lm(Daten ~ Index))$coefficients [2, 4]
[1] 1.207648e-82
```

Hier tritt genau das Problem der starken Abhängigkeit auf, welches im vorherigen Kapitel besprochen wurde. Die Steigung in diesem großen Datenfenster ist ausgesprochen klein. Sie liegt ungefähr bei einem 10.000-stel, was eindeutig für eine Beharrung spricht. Der p-Value jedoch ist noch viel kleiner und würde daher eine

Beharrung ablehnen. Man weiß allerdings (in diesem speziellen Fall), dass die Beharrung vorliegt. Um die Methode weiterhin anwenden zu können, muss also der absolute Wert der Steigung mitberücksichtigt werden. Dies gestaltet sich durch die großen Unterschiede zwischen den Parametern als sehr schwierig.

5.2.1.2. Kontinuierlicher schwacher Trend

Ein weiteres Problem liegt darin, dass ein schwacher, aber anhaltender Trend, zuweilen nicht erkannt wird. Angenommen die Steigung, beziehungsweise die p-Values, legen eine Beharrung nahe, allerdings ist die Steigung konstant über eine große Anzahl an Datenfenstern. Das bedeutet, jedes einzelne Fenster bestätigt die Beharrung aufs Neue. Auf lange Sicht gesehen, kann man aber keine Beharrung annehmen, da die Mittelwerte des Datenfenster immer größer oder kleiner werden. Solch eine Situation ist in Abbildung 5.4. illustriert.

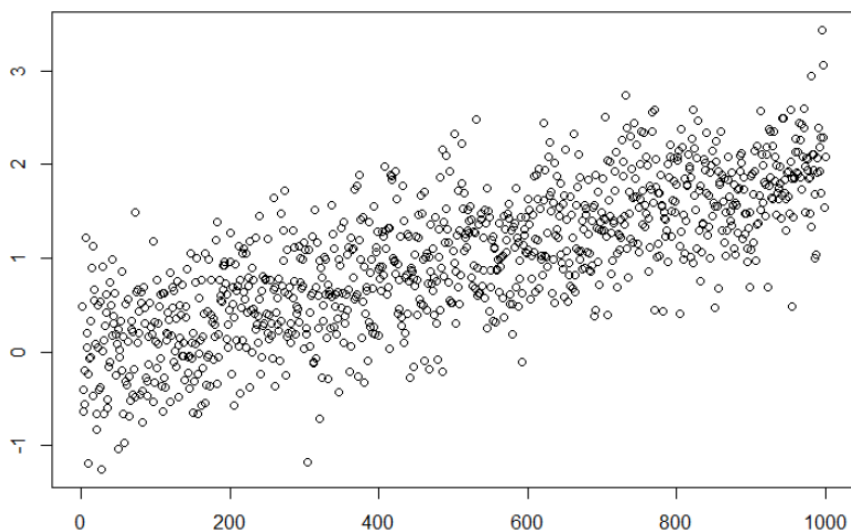


Abbildung 5.4.: Schwacher aber anhaltender Trend

Wie man erkennen kann, ist ein leichter Anstieg der Daten vorhanden. Dieser Anstieg ist allerdings in jedem einzelnen Datenfenster so gering, dass trotzdem Beharrung errechnet werden würde. Wenn dieser Trend weiter anhält, unterscheiden sich die Mittelwerte der Daten in jedem Fenster immer mehr. Nach ausreichender Dauer kann in keinem Fall Beharrung mehr angenommen werden.

5. Die Messfreigabe

5.2.1.3. Zyklische Daten

Daten, welche einen zyklischen Verlauf aufweisen, sind ebenso schwierig zu bewerten. Als Simulations- und Illustrationsbeispiel wählen wir eine Sinusfunktion mit weißem Rauschen, siehe Abbildung 5.5.

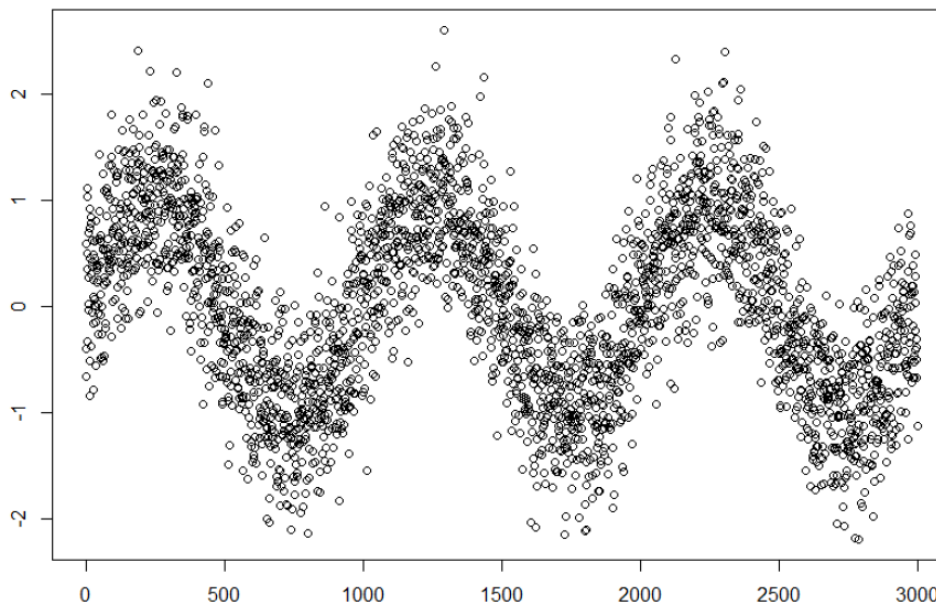


Abbildung 5.5.: Zyklischer Datenverlauf mit weißem Rauschen

Die Wahl der Fensterbreite für die Regressionsgerade ist hier ein wichtiges Kriterium. Wenn die Fensterbreite sehr gering ist, so wird die Steigung nur an den Extremstellen für kurze Zeit nahe bei 0 sein. Beharrung wird allerdings nicht angenommen, da diese Zeit zu kurz ist und nur sehr wenige p-Values groß genug sind.

Wenn die Fensterbreite sehr groß gewählt wird (hier im speziellen größer als die Länge eines Zyklus) so wird durchgehend Beharrung angenommen. Dies könnte durchaus wünschenswert sein, allerdings könnte die Schwankung der Daten zu groß beziehungsweise die Extremstellen des Zyklus zu weit voneinander entfernt sein. Darin liegt das zweite Problem. Wir müssen bewerten können, ob die Schwankung der Daten (beziehungsweise die Differenz zwischen Minimum und Maximum) im Datenfenster ausreichend klein ist.

Die Fensterbreite zu variieren kommt leider nicht in Frage, da man im Vorfeld nicht

wissen kann ob man einen zyklischen Datensatz vorliegen hat oder nicht. Bestehende Methoden für Periodizitätstests können beispielsweise in ANSMANN (2015) oder AMARIEI et al. (2014). Des Weiteren birgt eine große Fensterbreite weitere Gefahren, auf die später noch eingegangen wird. Aus der explorativen Datenanalyse weiß man, dass manche Messgrößen ein zyklisches Verhalten besitzen, welches manchmal toleriert und manchmal als fehlerhaft gedeutet werden muss.

5.2.1.4. Variation der Streuung

Dieser letzte Punkt verursacht eines der schwierigsten Probleme der Beharrungserkennung und tritt speziell bei jenen Parametern auf, welche Poissonverteilte Eigenschaften aufweisen. Ein Beispiel hierfür ist die Messgröße "AbgNox05mg". Die Grafik dazu findet man in Abschnitt 2.1. Man muss also in der Lage sein, auch dann eine Beharrung zu erkennen, wenn sich die Streuung immer wieder verändert, abhängig davon in welchem Wertebereich sich die Messgröße momentan befindet.

5.2.2. Lösung der Probleme

Zu jedem der oben aufgeführten Probleme wird nun eine Lösung erarbeitet, beziehungsweise eine Lösungsidee vorgeschlagen, die auf der gesammelten Erfahrung der Simulationen und echten Daten beruht. Jede Lösung wird in Folge erläutert und gegebenenfalls mit Beispielen illustriert.

5.2.2.1. Fehlinterpretation der p-Values

Wie zuvor beschrieben können die p-Values sehr klein sein, obwohl die der Slopeparameter im Regressionsmodell deutlich für eine Beharrung spricht. Der Lösungsansatz liegt in einem Vergleich des p-Values mit einem Wert α und einem Vergleich der Steigung mit der Streuung. Dieses α entspricht einem Signifikanzniveau aus der Statistik.

Man wählt ein klassisches α ($\{0.01,0.05,0.1\}$) und überprüft ob der errechnete p-Value kleiner als das gewählte α ist. Sollte dies der Fall sein, muss überprüft werden wie groß der Slope Parameter ist. Die erste Idee wäre gewesen, den Slope Parameter des Datenfensters mit der Streuung des Datenfensters zu vergleichen. Sollte die Steigung kleiner als diese Streuung sein, so war der p-Value fälschlicherweise zu klein und muss erhöht werden. Dieser Ansatz muss allerdings kritisch betrachtet werden. Wenn ein Parameter variiert wird, so steigt oder sinkt er von einem Wert

5. Die Messfreigabe

zum anderen. In genau dieser Zeit kann die Streuung nicht beurteilt werden.

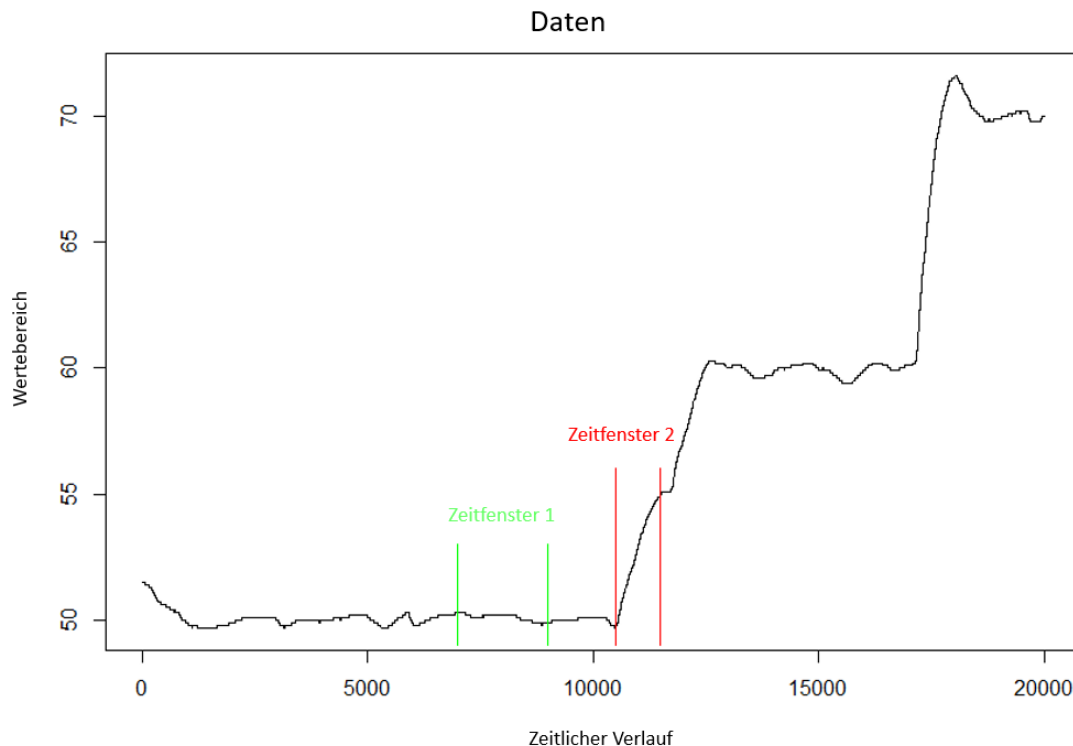


Abbildung 5.6.: Anstieg der Daten führt zu falsch berechneter Streuung

Die empirische Streuung im grünen Datenfenster ist jene die wirklich auf den Parameter zutrifft. Betrachtet man allerdings die empirische Streuung im roten Datenfenster, so ist diese im Vergleich zur tatsächlichen Streuung um ein Vielfaches höher. Der Grund ist aber der Anstieg der zugrunde liegenden Mittelwertsfunktion, der bewusst durchgeführt wurde.

In diesem Fall führt ein Vergleich nur mit der errechneten Abweichung zu falschen Ergebnissen. Daher wird davor die Streuung des aktuellen Datenfensters (im Folgenden mit Streuung 1 bezeichnet) und die Streuung der bisherigen Daten (im Folgenden mit Streuung 2 bezeichnet) miteinander verglichen und das Minimum der beiden ausgewählt. Mit diesem Vorgehen können solche Fälle korrekt erkannt und daher die Steigung mit einem stets adäquaten Wert verglichen werden. Die Errechnung der Streuung der bisherigen Daten kann auch ersetzt werden durch den Verlauf der Streuung der einzelnen Datenfenster. Es werden dabei die errechneten Werte gespeichert und aus der Historie ein Mittelwert gebildet. Bei einem Anstieg der Daten,

beziehungsweise einem Wechsel des Wertebereichs, erhöht sich der Mittelwert der Streuung nur langsam, da die neuen Streuungen den Mittelwert nicht zu sehr beeinflussen. Sobald sich die Messgröße wieder eingependelt hat, enthält das aktuelle Datenfenster wieder die zutreffende Streuung. Es gibt zu diesem Thema einige Vorgehensweisen, allerdings wird hier nicht weiter darauf eingegangen.

Simulationen haben allerdings noch zusätzlich gezeigt, dass die Abweichung selbst nach unten und nach oben beschränkt werden sollte. Als Obergrenze wurde $1/\alpha$ und als Untergrenze α gewählt. Der p-Value wird also künstlich angehoben, wenn folgende zwei Bedingungen zutreffen

1. $p\text{-Value} < \alpha$
2. $\text{Steigung} < \max(\min(\min(\text{Streuung } 1, \text{Streuung } 2), 1/\alpha), \alpha)$

Nun stellt sich noch die Frage, wie weit der p-Value angehoben wird. Bei der Überprüfung auf Gleichverteilung wird errechnet ob $x\%$ der Daten kleiner als $1/x$ sind. Setzt man $x = \alpha$, so muss der p-Value echt größer als α werden. Der doppelte Wert von α ist eine gute Wahl.

5.2.2.2. Kontinuierlicher schwacher Trend

Angenommen man hat einen schwachen Trend der nicht erkannt wird. Wenn man sich die Bedingung für die Beharrung nochmals vor Augen führt, findet man eine Lösung zu diesem Problem. Sobald Beharrung eingetreten ist, müssen folgende Bedingungen erfüllt sein:

1. Der p-Value, beziehungsweise die Steigung, müssen ausreichend groß beziehungsweise ausreichend klein sein.
2. Die p-Value Zeitreihe muss in einem gewissen Abschnitt ungefähr Gleichverteilt sein.

Dieser Abschnitt der p-Value Zeitreihe ist ausschlaggebend für die Problemlösung. Sobald Beharrung eintritt bedeutet dies, dass die aktuellsten p-Values annähernd gleichverteilt sind. Dies bedeutet nun wiederum, dass die aktuellsten Datenpunkte sich in einem beharrten Zustand befinden. Nun berechnet man den Mittelwert exakt jener soeben genannten Daten und nimmt diesen Wert als Vergleichswert. Wenn der schwache Trend bestehen bleibt, so werden sich zeitlich nachfolgende Daten von dem

5. Die Messfreigabe

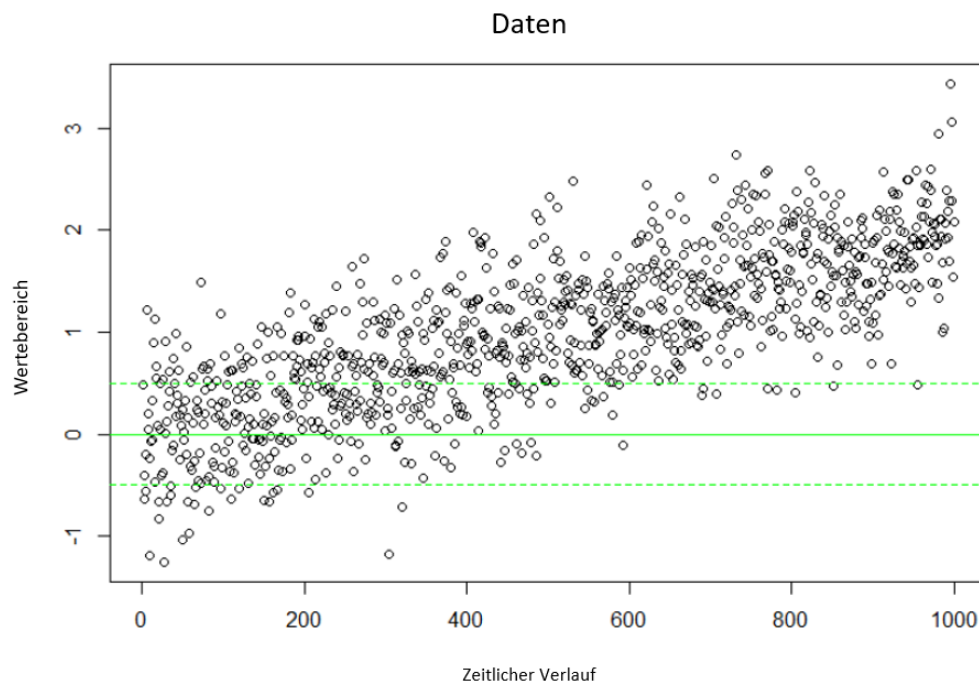


Abbildung 5.7.: Anstieg der Datenwerte mit Vergleichswert

errechneten Vergleichswert immer weiter entfernen. Sobald sie weit genug entfernt sind, wird die Beharrung beendet.

Die mittlere grüne Linie repräsentiert den errechneten Mittelwert und die beiden strichlierten Linien stehen für die Streuung um den Mittelwert. Sobald nun der Mittelwert zukünftiger Daten den Streubereich des errechneten Vergleichswertes verlässt, ist dies ein Zeichen für einen Trend und die Beharrung wird beendet. Sehr bald danach könnte wieder Beharrung eintreten, allerdings ist es auffällig, wenn ständig Beharrung errechnet und kurz darauf wieder beendet wird.

5.2.2.3. Zyklische Daten

Die Lösung des Trendproblems hilft auch bei der Lösung des zyklischen Problems. Man kann durchaus die Fensterbreite für die Regression größer wählen um Beharrung bei Daten mit kurzem Zyklus zu erkennen. Denn sollte der Zyklus eine langsame Frequenz aufweisen, ist die Steigung beim Anstieg beziehungsweise Abfall sehr klein und man ist wieder im vorherigen Fall des schwachen Trends. Die Beharrung wird beibehalten, wenn der Trend kurz genug ist. In anderen Worten, wenn die Streuung der Daten im Zyklus ausreichend klein ist, so pendeln die Daten um den errechneten

5.2. Die Beharrungserkennung

Vergleichswert und daher kann Beharrung durchgehend angenommen werden. Sollte der Trend zu lange anhalten, was bedeutet, dass die Streuung größer ist als sie sein sollte, so wird die Beharrung regelmäßig verworfen.

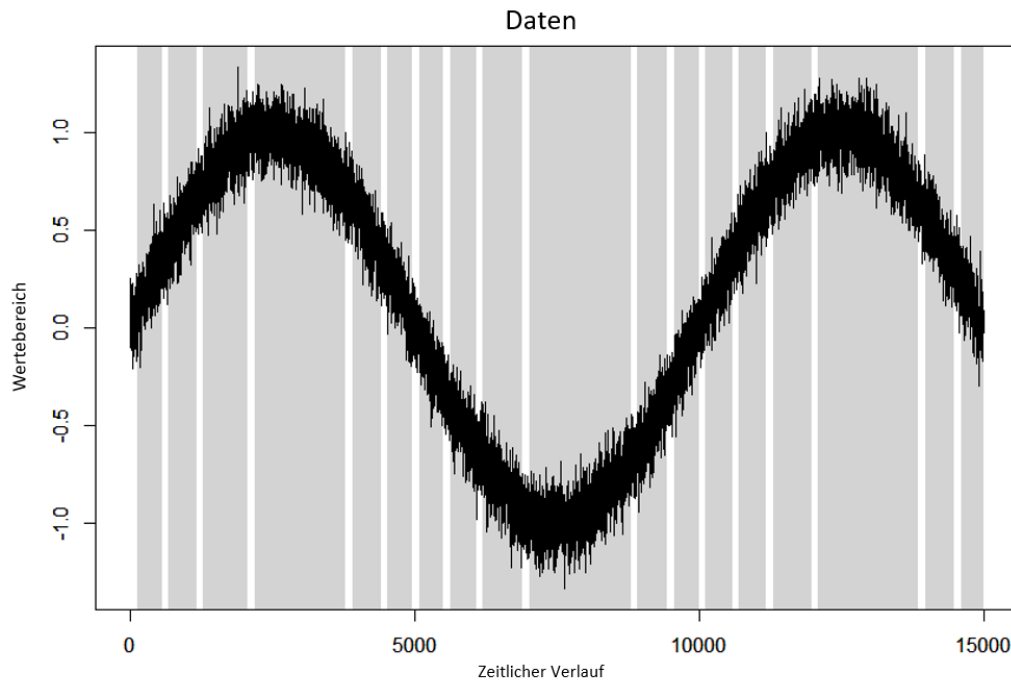


Abbildung 5.8.: Zyklische Daten ohne durchgehender Beharrung

In dieser Grafik wurde die erlaubte Streuung sehr klein gewählt. Der grau hinterlegte Bereich steht für den beharrten Zustand, der in regelmäßigen Abständen verworfen wird weil der Trend zu lange angehalten hat. Wenn die erlaubte Streuung größer ist, so würde man durchgehend Beharrung erhalten und hätte das folgende Bild.

5. Die Messfreigabe

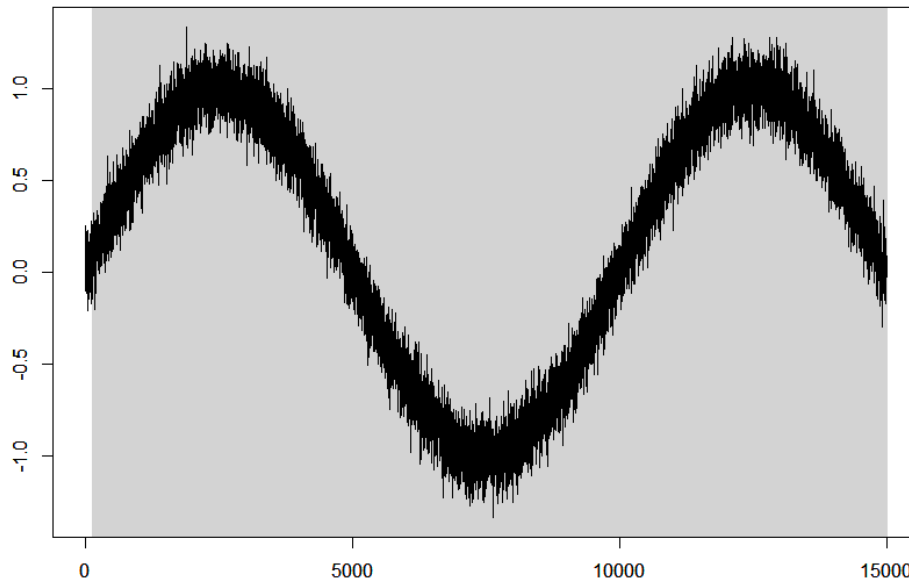


Abbildung 5.9.: Zyklische Daten mit durchgehender Beharrung

In diesem Fall ist der Bereich des Zyklus ausreichend klein um den beharrten Zustand niemals zu verwerfen. Nun taucht die berechtigte Frage auf: "Was ist überhaupt die erlaubte Streuung und wie groß darf diese sein?". Diese Frage wird am Ende des Kapitels beantwortet, denn die Antwort betrifft jeden Unterpunkt dieses Abschnitts.

5.2.2.4. Variation der Streuung

Um das Problem der variablen Streuung zu lösen waren einige Simulationen notwendig, damit die notwendigen Erfahrungen gesammelt werden konnten. Es hat sich gezeigt, dass bei der Wahl von $\alpha = 0.01$ Beharrung bei großer Streuung der Daten sehr gut erkannt wurde. Je kleiner die Streuung wurde, desto schlechtere Resultate erhielt man. Wenn man das α allerdings auf 0.1 erhöht hat, so wurde der Effekt umgedreht. Man erhielt bessere Resultate bei kleinerer Streuung, jedoch schlechtere bei großer. Da man in manchen Fällen alle möglichen Arten der Abweichung vorfindet, kann man im Vorfeld nicht festlegen welches α die besten Resultate liefert. Würde man wissen wie sich die Messgröße in der Zukunft verhalten wird, könnte man sich das "beste" α aussuchen. Des Weiteren gibt es auch keine optimale Wahl von α , wenn alle möglichen Arten der Streuung vorliegen, beispielsweise bei einer Variation einer Messgröße die sich annähernd wie eine poissonverteilte Zufallsgröße verhält.

Doch selbst wenn man nicht weiß, welche Wahl die beste wäre, gibt es eine Lösung. Sie besteht darin, gleichzeitig unterschiedliche α 's zu verwenden um die verschiedenen Streuungen besser abdecken zu können. In den folgenden Grafiken wurden 4 α 's gewählt. Jede Farbe steht für ein α . Sollte nun ein α Beharrung melden, so wird jener zeitliche Bereich in der entsprechenden Farbe markiert. Dazu wird der Datenbereich der Y -Achse in gleich große Teile unterteilt und jedes α erhält einen Bereich. Im unteren 1/4 ist jenes α welches für die große Streuung optimal ist (also das kleinste α) und im oberen 1/4 jenes α welches die besten Resultate bei kleiner Streuung liefert (also das größte α). In der ersten der drei folgenden Grafiken sieht man, dass mit dieser Idee selbst eine Variation einer annähernd poissonverteilten Messgröße sehr gut bewertet werden kann.

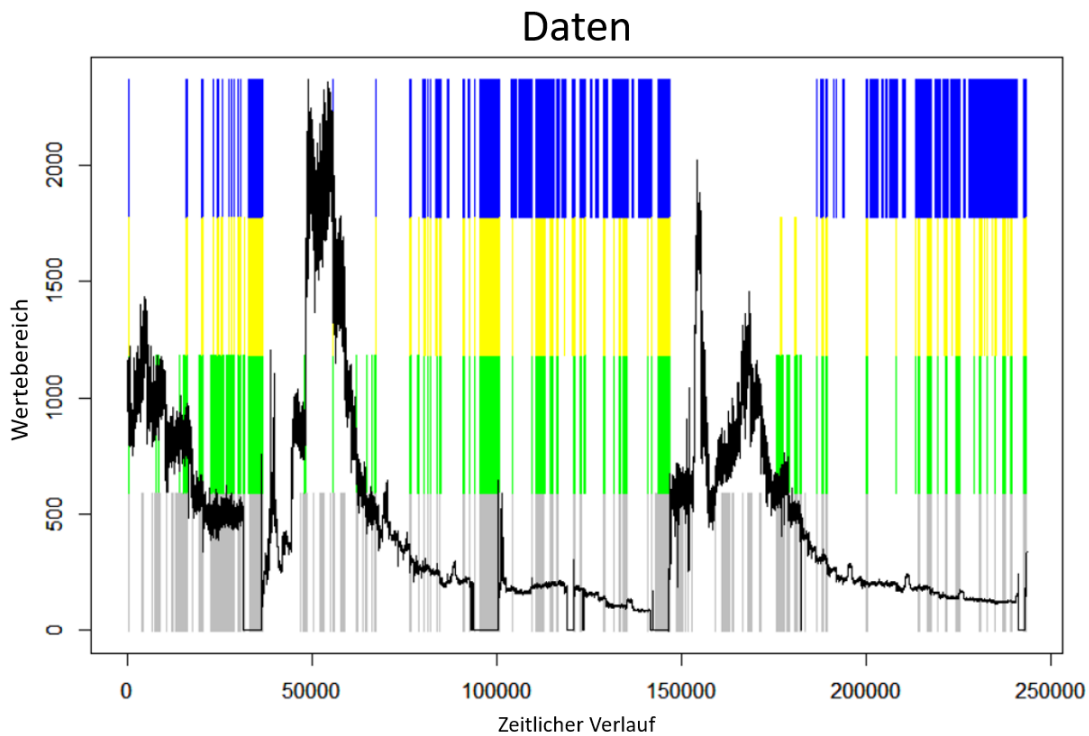


Abbildung 5.10.: Beharrungserkennung der einzelnen α 's

Man kann erkennen, dass die blaue Farbe, welche für das α der geringen Streuung steht, am präsentesten in den Bereichen der kleinen Abweichung ist. Die graue Farbe hingegen ist am stärksten bei großer Abweichung vertreten. Besonders gut erkennt man dies im letzten Bereich der Grafik und im Indexbereich von ca. 50.000. Am

5. Die Messfreigabe

Ende der Zeitreihe war die Streuung ausgesprochen klein und wird von der blauen Farbe korrekt detektiert. In dem genannten Indexbereich hingegen ist beinahe nur die graue Farbe zu erkennen. Zusammengefasst wird eine Beharrung auch dann erkannt, wenn alle Problemsituationen in einem Datensatz auftreten. Das bedeutet, sobald eines der α 's eine Beharrung meldet so ist auch gesamt Beharrung eingetreten.

Ein weiterer Vorteil ist ebenfalls ersichtlich. Selbst wenn man eine konstante Streuung hat, und im Vorfeld nicht weiß von welcher Art diese ist, lässt man alle 4 α 's die Beharrung überprüfen. Mit dieser Methode ist es völlig egal welche Situation vorliegt, denn eines der α 's ist passend. Die folgenden Grafiken belegen die Effizienz dieser Methode.

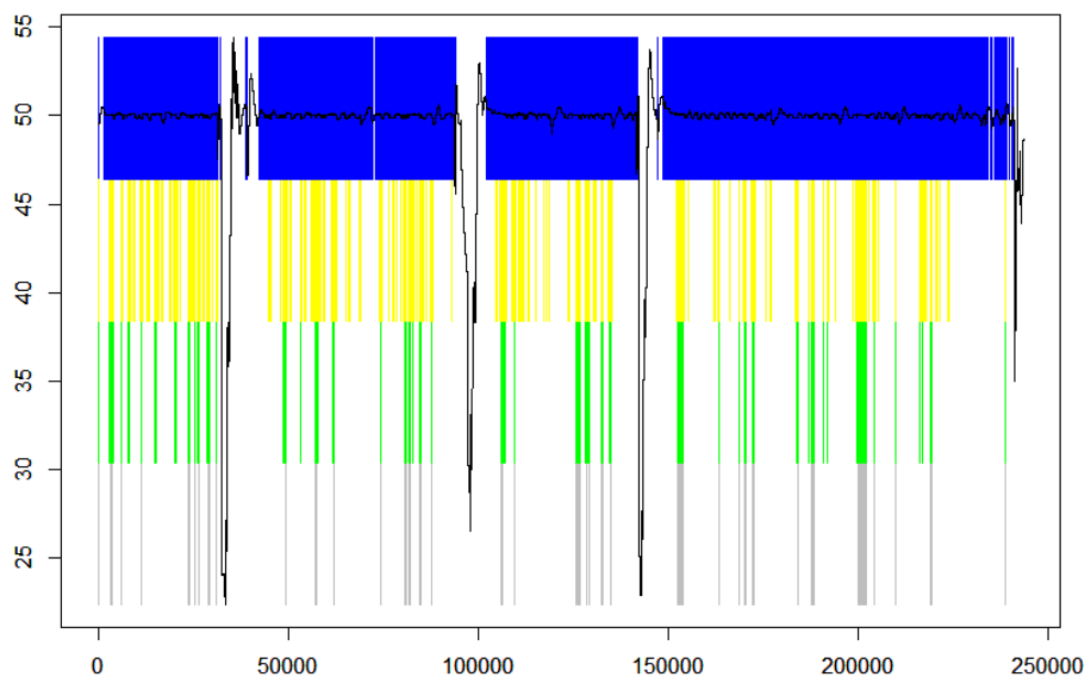
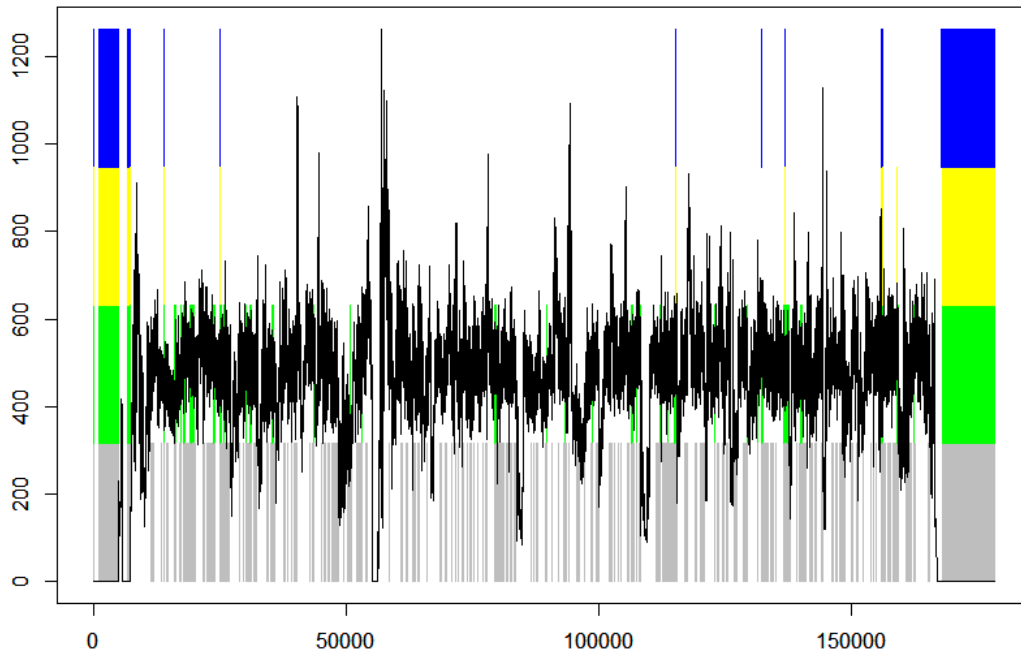


Abbildung 5.11.: Verhalten der α 's bei konstanter geringer Streuung

Abbildung 5.12.: Verhalten der α 's bei großer Streuung

In Abbildung 6.13 ist die Streuung sehr gering. Man hat sieht sofort, dass auch die blaue Farbe (zuständig für kleine Varianz) mit Abstand die größte Präsenz hat. Die graue Farbe (zuständig für große Varianz) hingegen ist kaum vorhanden.

In Abbildung 6.14 ist die Varianz sehr groß. Die graue Farbe ist hier beinahe alleine für die gesamte Beharrungserkennung verantwortlich.

5.2.3. Vorgegebene Streuung

Wie schon erwähnt, gibt es eine sogenannte vorgegebene Streuung. Beinahe jede Messgröße besitzt eine maximal erlaubte Abweichung vom Mittelwert. Diese Abweichungen sind dem Prüfstandsfahrer bekannt. Man unterscheidet hierbei zwischen einer absoluten und einer relativen Abweichung. Bei einer absoluten Abweichung darf der Parameter immer dieselbe maximale Abweichung haben, völlig egal welche Werte angenommen werden. Die relative Abweichung ist speziell für Messgrößen gedacht, die bei einer Variation verschiedene Streuungen annehmen, je nachdem wie groß der tatsächliche Messwert ist. Die Angaben erfordern also die Information einer absoluten Abweichung inklusive der absoluten Abweichungsgröße, oder einer relativen Abweichung inklusive der prozentualen Streuung.

5. Die Messfreigabe

Diese Information erlaubt uns, die Methode noch weiter zu verfeinern und zu verbessern. Zum ersten wird das größte α (also jenes für die kleine Schwankung) anders gewählt nämlich der Mittelwert der üblichen α 's. Das ergibt den Wert $0.533333 = (0.01 + 0.5 + 0.1)/3$. Doch nun wird die Steigung bezüglich dieses α 's nicht mehr mit der empirischen Streuung des aktuellen Datenfensters verglichen, sondern mit eben jener vorgegebenen Toleranzabweichung. Ist diese absolut, so kann man die Steigung direkt mit der angegebenen Abweichung vergleichen. Bei einer relativen Abweichung muss man sich den Mittelwert des Datenfensters ausrechnen, um eine absolute Abweichung zu erlangen, mit der man den Slope Parameter vergleichen kann. Für das Trendproblem gibt diese Vorgabe exakt jenen Toleranzbereich an, den man für die Überprüfung, ob die zukünftigen Mittelwerte vom Vergleichswert ausreichend weit entfernt sind, braucht.

Ohne diese Information müsste man die maximal erlaubte Abweichung durch die empirische Standardabweichung schätzen. Allerdings gibt es für diese Schätzung auch einen Spezialfall, der ein Problem bereitet hätte. Dieser Spezialfall hätte wie folgt ausgesehen: jeder zweite Datenpunkt liegt weit über dem errechneten Mittelwert und jeder andere zweite Datenpunkt liegt weit unter dem errechneten Mittelwert. Die Steigung liegt in dieser Situation fast bei 0 und die Abweichung wäre ausgesprochen groß. Mit der Vorgabe der Abweichung kann allerdings entschieden werden ob die Daten für eine Beharrung ausreichend weit vom Mittelwert entfernt sind.

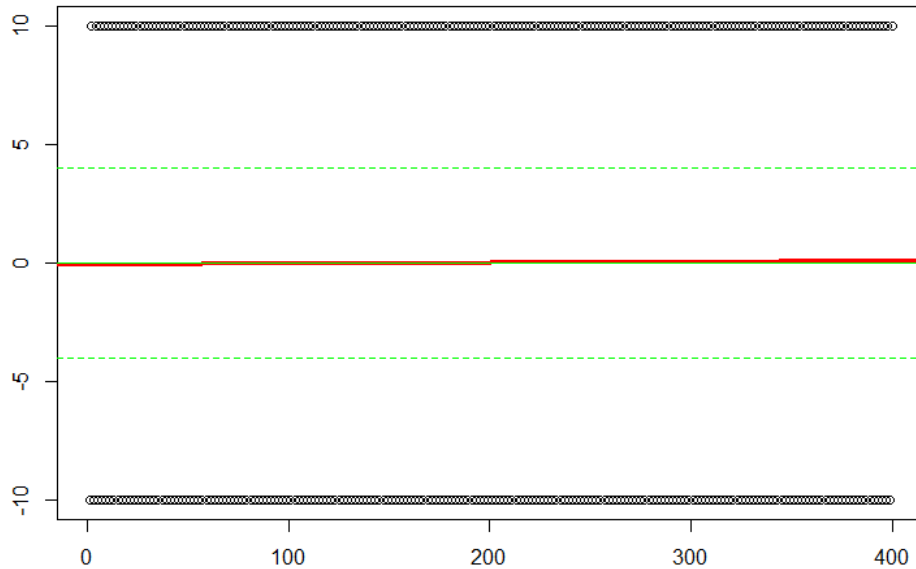


Abbildung 5.13.: Datensituation: Spezialfall erkannt

In dem Beispiel kann man aufgrund der Toleranz erkennen, dass kein einzelner Datenpunkt auch nur annähernd im vorgegebenen Bereich liegt. Solche Situationen werden also erkannt und dementsprechend angezeigt. In den Daten des LEC gab es zwar noch nie eine Datensituation die so ausgesehen hat, doch möglicherweise können solche Fälle zu späteren Zeitpunkten auftreten.

Auch beim zyklische Problem ist die Kenntnis über die maximal erlaubte Abweichung essentiell. Denn diese Information zeigt sehr schnell ob die Schwankung des Zyklus ausreichend klein ist um durchgehend in Beharrung zu sein, oder ob der Zyklus den Toleranzbereich regelmäßig verlässt. Ähnlich wie beim Trendproblem hätte man andernfalls die Streuung schätzen müssen, was zu einer Fehlinterpretation der Daten führen könnte.

5.2.4. Ausreißer Detektion

All die bisher gesammelten Informationen und Erkenntnisse liefern eine sehr gute Methode die Beharrung zu erkennen und gegebenenfalls auch wieder zu verwerfen. Es wird nun die Forderung erhoben, sogenannte Ausreißer zu erkennen, welche die Beharrung nicht beenden. Das bedeutet, man will Werte erkennen die weit von der breiten Datenmasse entfernt, aber die nicht ausreichend groß sind um die Beharrung

5. Die Messfreigabe

zu verwerfen. Mithilfe der erlaubten Abweichung allerdings lässt sich ein simples aber effizientes Verfahren entwickeln, welches Ausreißer sofort detektiert.

Wir erinnern dazu, dass $\text{Var}\left(\frac{1}{n}(X_1 + \dots + X_n)\right) = \frac{1}{n}\text{Var}(X_1)$ gilt. Also, dass die Varianz des Mittelwertes nur einem n-tel der Varianz einer einzelnen Beobachtung entspricht. In unserem Fall kennen wir die die Varianz des Mittelwertes und wollen nun die Varianz eines jeden einzelnen Punktes wissen. Bezeichne mit $\tilde{\sigma}$ die maximal erlaubte Abweichung des Mittelwerts, die vom Prüfstandsfahrer vorgegeben wird. Dieses $\tilde{\sigma}$ entspricht ungefähr 2σ für die soeben gemachte Erinnerung. Der Grund dafür ist einfach erklärt: Bei einer Normalverteilung befinden sich über 95% der Werte im Bereich $\mu \pm 2\sigma$. Unser $\tilde{\sigma}$ soll genau dieselbe Eigenschaft haben und daher wählen wir $\tilde{\sigma} = 2\sigma$. Diese Aussage gilt zwar nur für identisch und unabhängig verteilte Zufallsvariablen, allerdings verwenden wir dieses Resultat als approximative Richtlinie für nicht unabhängig verteilte Zufallsvariablen.

Angewandt auf die Problemstellung erhält man durch diese Beobachtung eine maximal erlaubte Abweichung jedes einzelnen Wertes. Sobald nun ein einzelner Messwert die errechnete Grenze überschreitet wird eben dieser Wert als Ausreißer erkannt. In der folgenden Grafik ist der Verlauf einer Messgröße zu sehen. Die grünen horizontalen Linien stehen für drei zu erreichende Mittelwerte und deren maximale erlaubte Abweichung davon. Die zwei strichlierten horizontalen blauen Linien repräsentieren die mithilfe des Satzes neu errechnete Grenze eines jeden einzelnen Messwertes, bezüglich des mittleren der drei Mittelwerte. Der grau hinterlegte Bereich ist jener, in dem Beharrung angenommen werden konnte und auch der Mittelwert der Daten mit dem Soll-Mittelwert übereinstimmt. Die vertikalen roten Liniensegmente signalisieren jene Stellen, wo die Messwerte nun außerhalb der neu errechneten Toleranz liegen.

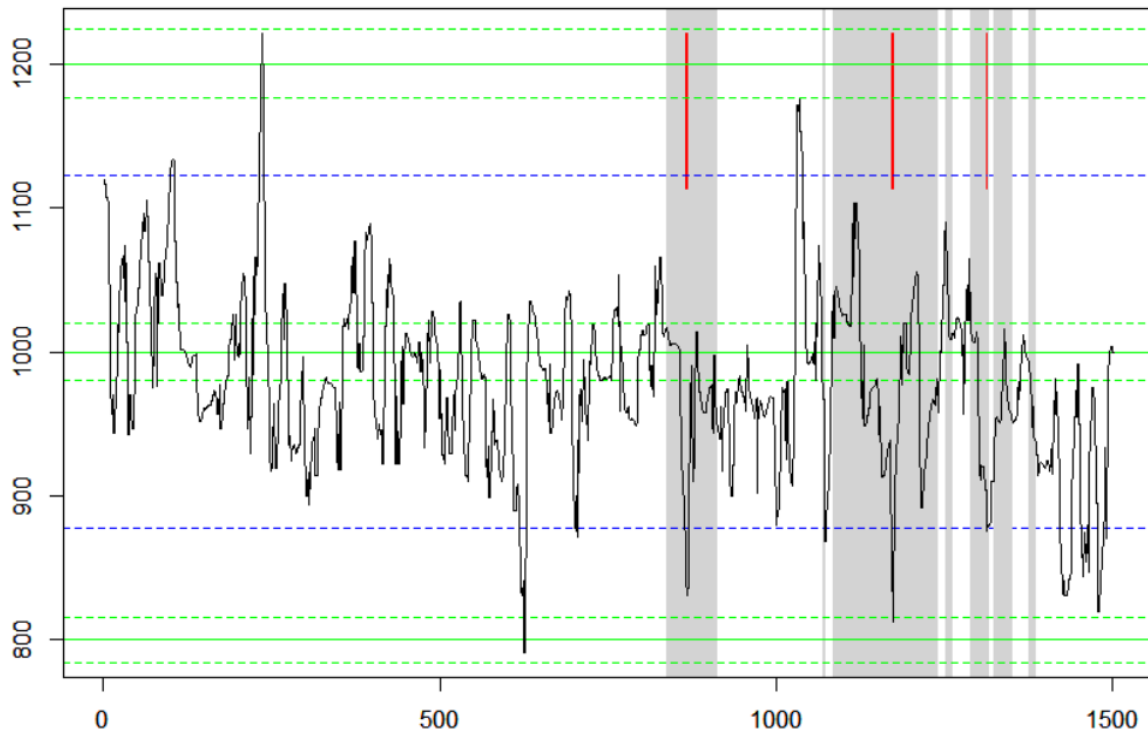


Abbildung 5.14.: Beispiel für eine Ausreißer Detektion

Man kann die Ausreißer erkennen und durch die Hilfslinien sofort mit freiem Auge verifizieren, dass es sich tatsächlich um Datenpunkte weit außerhalb der gewünschten Toleranz handelt.

5.3. Die programmierte Beharrungserkennung

Ziel ist es nun, all die bisherigen Resultate auch praktisch umzusetzen. Dies kann mittels einer neu entwickelnden Funktion sehr gut bewerkstelligt werden. Die Inputparameter der Funktion beinhalten folgende Größen.

1. Daten : Ein Spaltenvektor der jeden einzelnen Messwert beinhaltet
2. Abweichung 1: Die Art der Abweichung, entweder absolut oder relativ
3. Abweichung 2: Die erlaubte Abweichung, entweder ein fester Absolutwert oder ein Prozentsatz
4. Sollwert : Eine optionale Eingabe von einem oder mehreren Sollwerten falls diese im vorhinein schon bekannt sind

5. Die Messfreigabe

5. Fenster 1 : Die Anzahl der Daten die für die einfache lineare Regressionsanalyse herangezogen werden
6. Fenster 2 : Die Anzahl der p-Values die auf Gleichverteilung überprüft werden
7. Fenster 3 : Die Anzahl der Daten die für die Berechnung der Mittelwerte herangezogen werden.

Die drei Eingaben der Fenstergrößen werden Default-mäßig belegt und müssen nur dann eingegeben werden, wenn die Default Einstellung nicht verwendet werden soll. Nach Aufrufen der Funktion wird zunächst überprüft ob alle relevanten Inputgrößen (Daten, Abweichung 1 und Abweichung 2) vorhanden sind. Erst wenn dies der Fall ist wird die Funktion weiter ausgeführt und zwischen 4 Fällen unterschieden

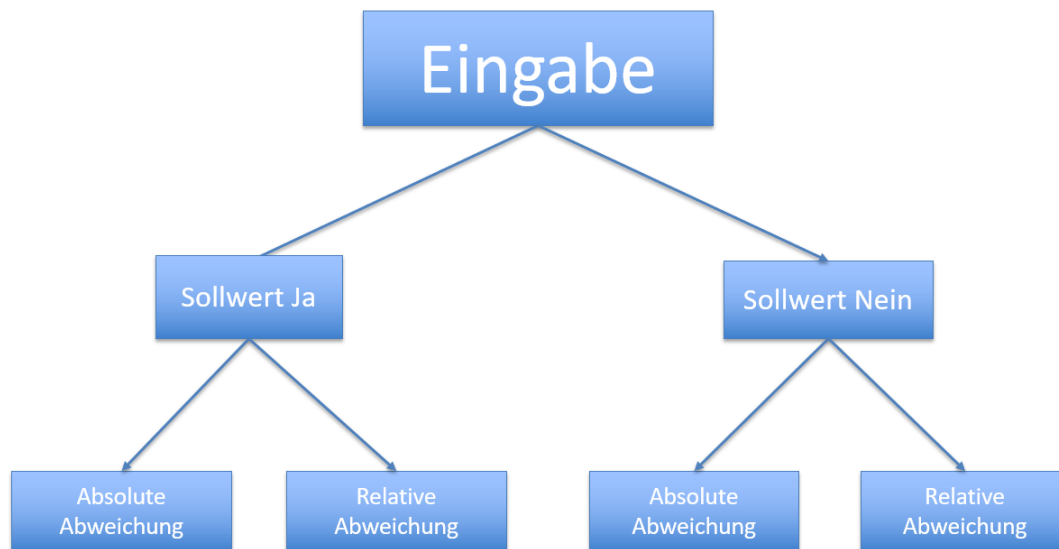


Abbildung 5.15.: Schematischer Überblick der 4 möglichen Fälle

Die Grundstruktur jeder dieser Fälle ist ident, daher wird zuerst das Vorgehen beschrieben und im Anschluss erklärt worin die kleinen Unterschiede bestehen.

Zu Beginn werden 3 Initialisierungen vorgenommen

1. 4 α 's : $\alpha_1, \dots, \alpha_4$ wie besprochen im Abschnitt 5.3.4
2. 4 p-Value Vektoren der Länge "Fenster 2" befüllt mit Nullen
3. Ein Vergleichswert der noch keinen Wert angenommen hat und mit "Sollneu" bezeichnet wird

5.3. Die programmierte Beharrungserkennung

Die α 's sind feste Werte die sich niemals ändern wohingegen die p-Value Vektoren und der Vergleichswert im Laufe der Funktion verschiedenste Werte annehmen.

Nun beginnt der eigentliche Schleifendurchlauf. Im ersten Schritt wird ein zweiter Vergleichswert definiert, mit demselben Wert wie "Sollneu". Dieser wird mit "Sollalt" bezeichnet. Der Grund dafür ist, dass sich "Sollneu" am Ende eines Schleifendurchlaufs möglicherweise verändert hat. Daher muss überprüft werden ob der Wert am Ende derselbe ist wie zu Beginn.

Des Weiteren wird auf die p-Value Vektoren ein Shiftoperator angewendet. Durch einen Shiftoperator werden die Positionen des Vektors verändert. Hier wird der p-Value auf der letzten Position entfernt und die erste Position des Vektors ist leer und kann neu befüllt werden.

Nun wird der neue Datenpunkt herangezogen und mittels der linearen Regression eine Steigung berechnet. Der dazugehörige p-Value wird (nach eventueller Anpassung wie im Abschnitt 5.3.1) in die nun freie Stelle des p-Value Vektors gespeichert.

Danach werden die aktualisierten Vektoren auf Gleichverteilung überprüft wie es im Abschnitt 5.1 erklärt wurde. Sobald bei einem dieser Vektoren Gleichverteilung angenommen werden konnte, ist eine Beharrung vorliegend. Dadurch kann angenommen werden, dass die aktuellsten Daten bereits in einem beharrten Zustand waren. Mit diesen aktuellen Daten kann der Mittelwert als Referenzwert herangezogen werden. Sollte also "Sollneu" noch keinen Wert angenommen haben so wird er nun auf den eben errechneten Mittelwert gesetzt.

Im weiteren Schleifenverlauf werden nun die Beharrungsbedingungen überprüft. In jedem Schritt wird auf Gleichverteilung der p-Value Vektoren getestet und ob der Mittelwert der aktuellsten Daten mit dem Mittelwert übereinstimmt, der zu Beginn der Beharrung errechnet wurde. Solange diese Bedingungen erfüllt sind, bleibt die Beharrung bestehen. Sobald allerdings die Gleichverteilung aller Vektoren verloren geht oder der Mittelwert sich zu weit entfernt, wird die Beharrung beendet. In diesem Fall wird zuerst der Referenzwert "Sollneu" wieder auf Initialisierungswert und jeder p-Value auf Null gesetzt.

Ein sehr großer Vorteil dieser Methode ist die geringe Komplexität. Es wird nur eine einzige Schleife benötigt in der ein paar wenige Abfragen und Berechnungen stattfinden. Auch der Speicherbedarf ist sehr gering, da man nur die aktuellsten Messwerte und p-Values benötigt. Die Kurzzusammenfassung in Form eines Algorithmus wird

5. Die Messfreigabe

hier angegeben. Eine ausführlichere Version inklusive Erklärung der verwendeten Funktion findet man im Appendix.

Algorithm 1: Beharrungserkennung als Algorithmus

```
Data: Funktion(Inputparameter)
Initialisierung:  $\alpha$ , Referenzwert;
for  $i = 1, \dots, n$  do
    1. Aktuellen p-Value berechnen
    2. Überprüfung auf Gleichverteilung der aktuellsten p-Values;
    if Gleichverteilung then
        1. Referenzwert = Mittelwert
           (falls Referenzwert = Initialwert war)
        2. Beginn der Beharrung
           (falls Beharrung nicht schon aktiv);
    end
    if Gleichverteilung & Mittelwert im Toleranzbereich then
        | Beharrung aufrecht erhalten;
    else
        | 1. Referenzwert = Initialwert
        | 2. Die aktuellsten p-Values werden auf 0 gesetzt
    end
end
```

Nun fehlen nur mehr die Unterschiede der 4 Fälle. Im Fall der absoluten Abweichung haben einige Simulationen gezeigt, dass von den 4 α 's nur 2 benötigt werden, nämlich α_1 und α_4 . Die Schwankung des Mittelwerts ist immer konstant und muss daher auch nicht jedes Mal neu berechnet werden wie bei der relativen Abweichung. Im Falle einer relativen Abweichung muss für jeden einzelnen Referenzwert eine eigene Abweichung bestimmt werden, da jeder Wert seinen eigenen Toleranzbereich besitzt.

Der Unterschied zwischen einem vorhandenen und nicht vorhandenen Sollwert ist sehr gering. Ist kein Sollwert, so führt man den Algorithmus aus wie in der vorherigen Grafik aufgelistet. Sobald nun Beharrung eintritt, ist dies äquivalent mit der Messfreigabe. Wenn man einen oder mehrere Sollwerte zur Verfügung hat so wird separat noch überprüft ob der Mittelwert der aktuellsten Daten im Toleranzbereich

5.3. Die programmierte Beharrungserkennung

der Sollwerte liegt. Die Beharrung selbst bleibt allerdings völlig isoliert von diesem Ergebnis und wird auch nicht verworfen, wenn der Soll-Ist-Vergleich nicht erfüllt ist.

Die Fallunterscheidung und ihre Eigenheiten kann in Kurzform als Grafik wie folgt dargestellt werden.

Fallunterscheidung	Sollwert gegeben	Sollwert nicht gegeben
Absolute Abweichung	1. 2/4 Alpha's 2. Soll-Ist-Wert Überprüfung	1. 2/4 Alpha's 2. Keine Soll-Ist-Wert Überprüfung
Relative Abweichung	1. Toleranzbereich für jeden Referenzwert und Sollwert bestimmen 2. Soll-Ist-Wert Überprüfung	1. Toleranzbereich für jeden Referenzwert und Sollwert bestimmen 2. Keine Soll-Ist-Wert Überprüfung

Abbildung 5.16.: Die 4 Fälle mit ihren Eigenschaften

Nun werden für jeden Fall Ergebnisse gezeigt. Dazu wählt man 2 Messgrößen aus, nämlich T2strich und AbgNox05mg. Diese beiden Parameter könnten nicht unterschiedlicher sein und sind daher ideal für eine Demonstration. T2strich besitzt eine sehr geringe aber konstante Streuung. AbgNox05mg dagegen hat eine relative Abweichung, die noch dazu im Allgemeinen sehr groß ist. Für beide Größen gibt es Sollwerte, aber um die Ergebnisse zu vergleichen nimmt man zuerst an, dass diese unbekannt sind. Des Weiteren sind diese 2 Messgrößen variiert worden um zu erkennen ob und wie gut die Funktion auf diese Variation eingehen kann.

Trotz der unterschiedlichen Verhalten liefert die Funktion sehr gute Ergebnisse, sowohl mit als auch ohne Sollwerte. Speziell beim Parameter T2strich unterscheiden sich die Zeiten der Messfreigabe mit oder ohne Sollwerte kaum. Wenn Sollwerte gegeben sind, so werden diese mit grünen durchgezogenen Linien angezeigt. Die strichlierten Linien ober- und unterhalb des Sollwerts zeigen die erlaubte Abweichung des Mittelwerts.

5. Die Messfreigabe

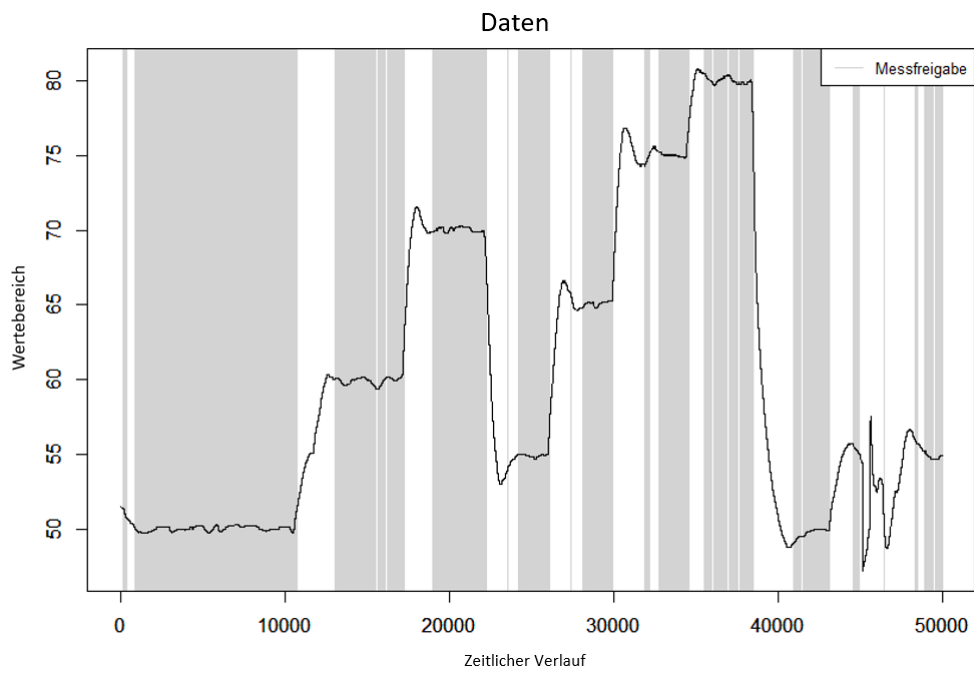


Abbildung 5.17.: Parameter T2strich ohne Sollwert

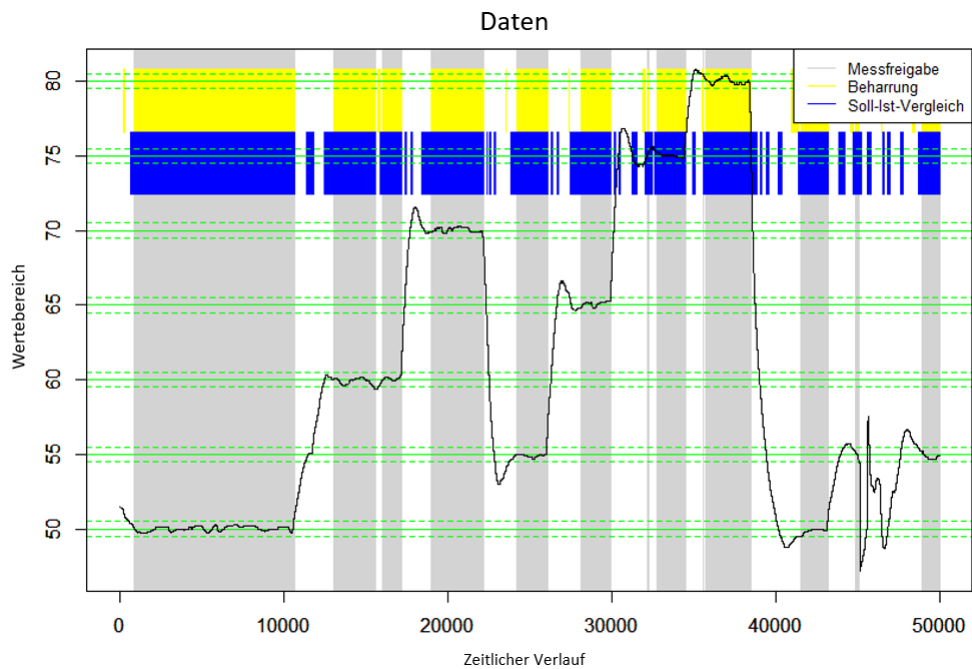


Abbildung 5.18.: Parameter T2strich mit Sollwert

5.3. Die programmierte Beharrungserkennung

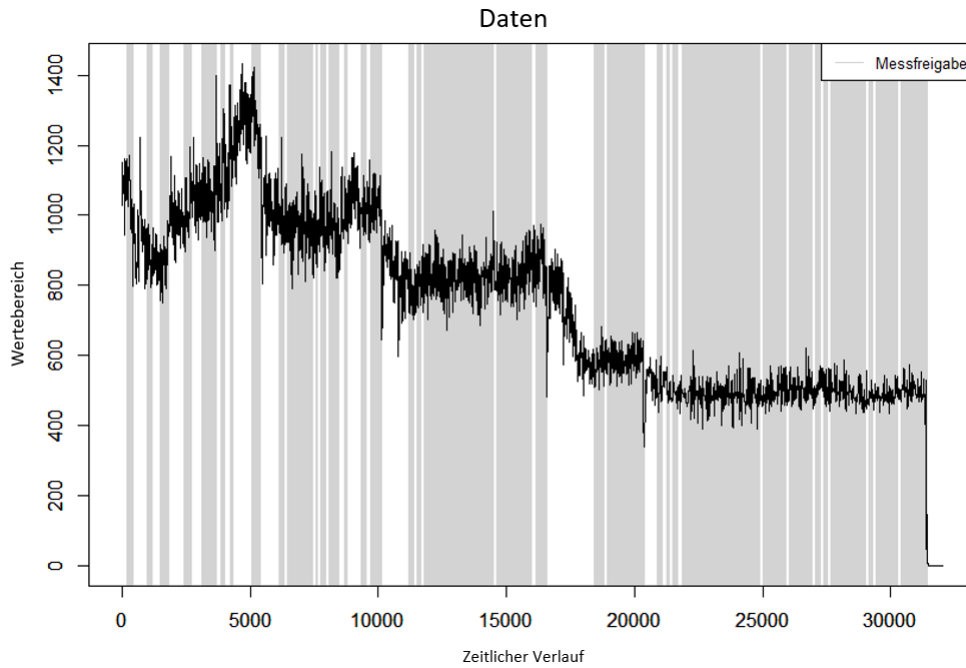


Abbildung 5.19.: Parameter AbgNox05mg ohne Sollwert

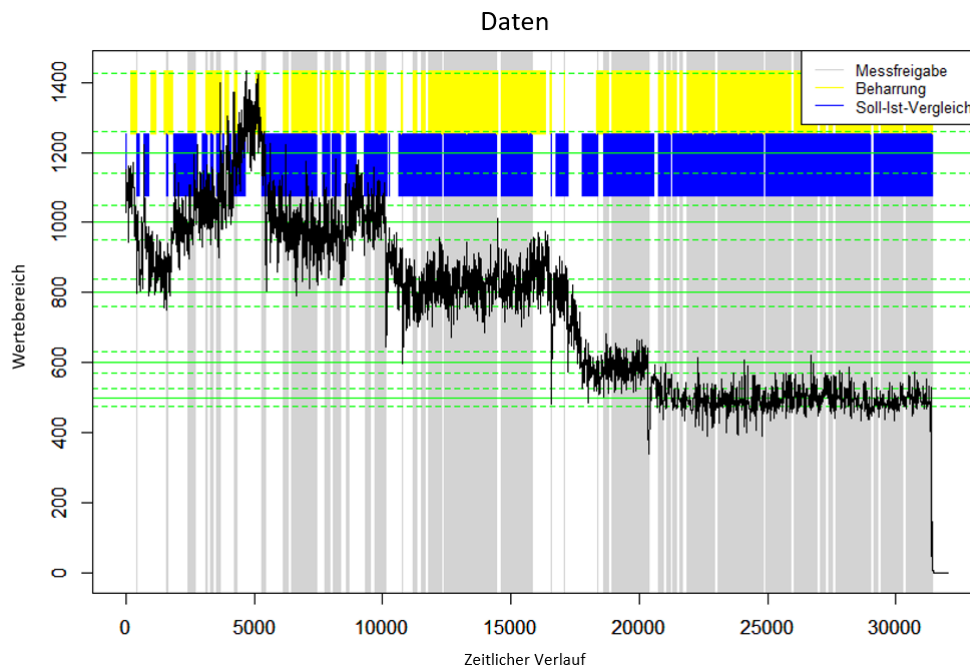


Abbildung 5.20.: Parameter AbgNox05mg mit Sollwert

6. Die Messpunktdiagnose

Sobald nun eine Messfreigabe vorhanden ist, wird eine Messung vorgenommen. Üblicherweise besteht ein Messpunkt dabei aus 150 Messwerten. Es gilt nun, diesen Messpunkt zu bewerten. Dabei kann die Methode zur Messfreigabe nicht ausschließlich herangezogen werden, denn in einigen Fällen haben Prüfstandsfahrer ihre Expertise genutzt und erraten wie sich manche Messgrößen in naher Zukunft verhalten werden, selbst wenn zum aktuellen Zeitpunkt keine Beharrung beziehungsweise Messfreigabe vorhanden war.

Gesucht ist also nun eine Bewertung, die ohne weitere Vorkenntnisse beurteilen kann, ob die aktuellsten Messwerte akzeptabel sind oder der Messpunkt verworfen werden muss. Die einzige Vorkenntnis, die man annehmen kann, ist die tolerierte Abweichung des Mittelwertes. Dieses Kapitel ist in 2 große Abschnitte unterteilt. Im ersten Teil werden bekannte statistische Möglichkeiten erklärt und anhand von Beispielen illustriert, wann und warum diese Herangehensweisen versagen. Im zweiten wird die tatsächliche Methode, die schlussendlich verwendet wurde, erläutert.

6.1. Möglichkeiten zur Bewertung allgemeiner Daten

In diesem Abschnitt werden 2 Möglichkeiten aufgelistet und erklärt wie man im Allgemeinen Daten ohne Vorkenntnisse bewerten könnte. Allerdings werden auch Beispiele und Spezialfälle angegeben, wo eine Anwendung im Fall der Messdaten des LEC nicht möglich ist.

6.1.1. Die Varianzanalyse - ANOVA

Der Grundaufbau einer Varianzanalyse, oder ANOVA (**A**Nalysis **O**f **V**ariance) besteht im Vergleich der Mittelwerte von mehreren Gruppen. Sei $X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n_i})$ die i -te Gruppe und μ_i der Erwartungswert dieser Gruppe, $1 \leq i \leq m$. Nun ist man interessiert ob die Mittelwerte der einzelnen Gruppen gleich sind, oder ob eine Grup-

6. Die Messpunktdiagnose

pe sich im Mittel deutlich von den anderen unterscheidet. Die Hypothese lautet wie folgt:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_m \quad \text{vs.} \quad H_1 : \exists i, j : \mu_i \neq \mu_j$$

Unter H_0 gilt für zwei beliebige Gruppen i, j

$$F = \frac{\sum_{i=1}^m n_i (\bar{X}_{i.} - \bar{X})^2 / (m - 1)}{\sum_{i=1}^{n_i h} \sum_{j=1}^m (X_{i,j} - \bar{X}_{i.})^2 / (n - m)} \sim F_{m-1, n-m}$$

Diese Formel findet man in sämtlichen Literaturwerken Die Anwendung für einen konkreten Messpunkt wäre wie folgt:

1. Unterteile den Messpunkt in gleich große Abschnitte. Da man 150 Messwerte hat, sind 3 oder 5 Abschnitte zu je 50 beziehungsweise 30 Messwerten eine passende Wahl.
2. Wende ANOVA auf die einzelnen Abschnitte an, um eine Aussage über die Unterschiede des Mittelwerts zu haben.

Um eine ANOVA anwenden zu können müssen gewisse Voraussetzungen erfüllt sein:

- Der statistische Fehler beziehungsweise das weiße Rauschen jeder Gruppe muss annähernd normalverteilt sein.
- Es muss Homoskedastizität vorliegen. Das bedeutet die einzelnen Gruppenvarianzen müssen homogen sein.
- Die einzelnen Beobachtungen jeder Gruppe müssen unabhängig sein.

Zur Illustration werden nun 2 Beispiele angeführt.

6.1. Möglichkeiten zur Bewertung allgemeiner Daten

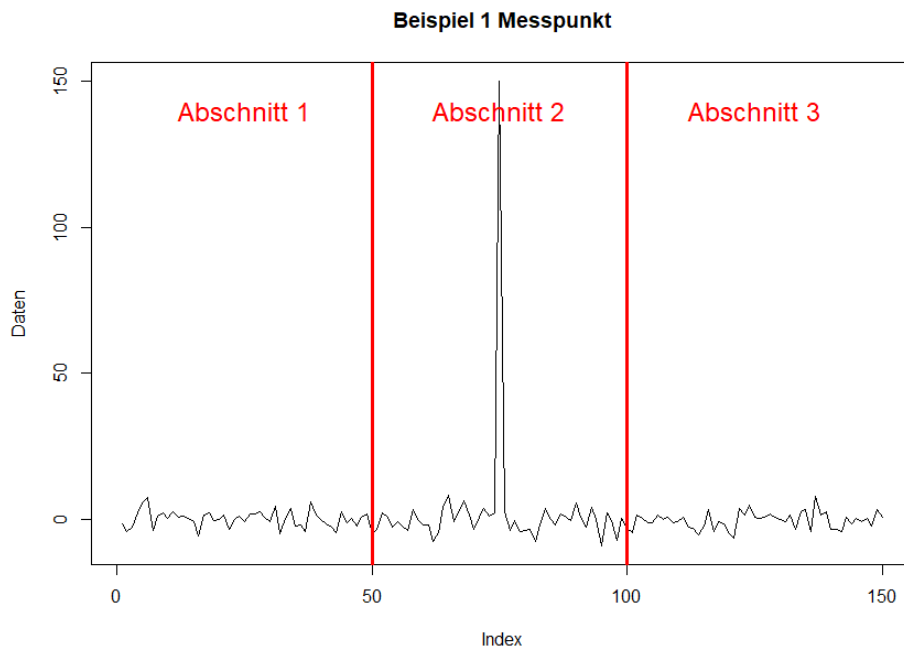


Abbildung 6.1.: Beispiel eines Messpunkts mit Ausreißer

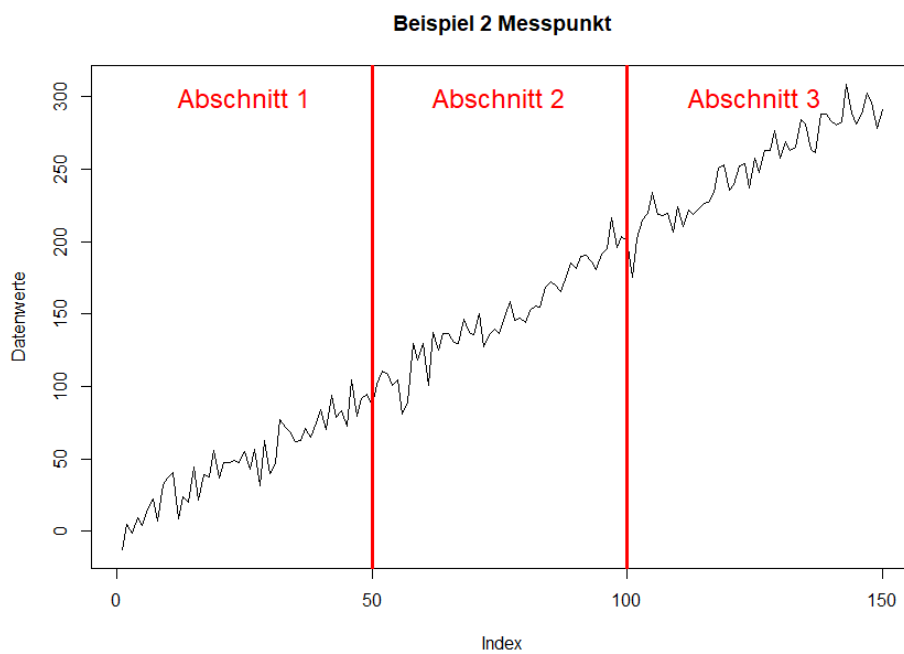


Abbildung 6.2.: Beispiel eines Messpunkts mit kontinuierlichem Anstieg

6. Die Messpunktdiagnose

Für das zweite Beispiel kann man in \mathbb{R} auch mit dem richtigen Befehl die zugehörigen Teststatistik aufrufen

```
> summary(aov(Datenwerte ~ Abschnitt))
              Df Sum Sq Mean Sq F value Pr(>F)
Abschnitt      2 1032773   516387   548.4 <2e-16 ***
Residuals    147  138427     942
-----
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Der p-Value ist sehr klein, daher kann die Nullhypothese, welche die Gleichheit der Mittelwerte angenommen hat, verworfen werden. Die ANOVA hätte also erfolgreich einen Messpunkt detektiert, den man verwerfen müsste.

Diese beiden Beispiele wurden generiert und erfüllen (größtenteils) die notwendigen Voraussetzungen für eine zuverlässige Aussage der ANOVA. Leider gibt es bei den echten Daten auch Situationen wo wie diese Methode nicht zur Anwendung kommen kann. Das folgende Beispiel wird dies erläutern.

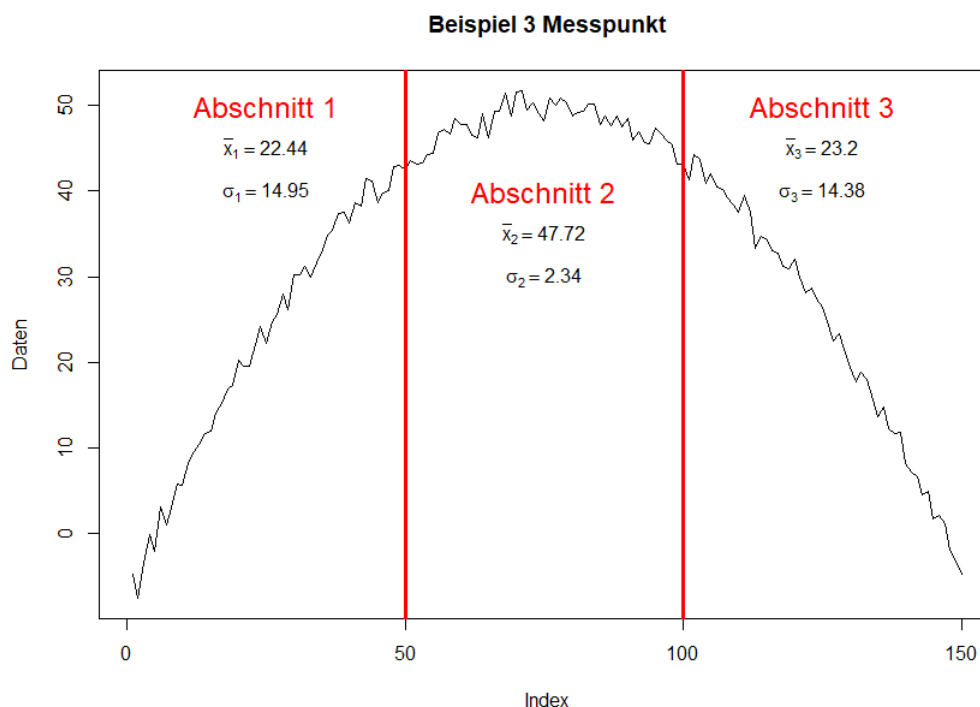


Abbildung 6.3.: Beispiel für Versagen von ANOVA

6.1. Möglichkeiten zur Bewertung allgemeiner Daten

Dies ist ein simulierter Verlauf der Messgröße "AbgNox05mg". Es gibt immer wieder Messpunkte dieser Größe die dem generierten Messpunkt sehr ähnlich sind. Da diese Messgröße eine sehr große Streuung aufweist, ist so ein Messpunkt nichts Ungewöhnliches und daher auch akzeptabel. Die ANOVA Zerlegung ist aber hier nicht anwendbar, da man schnell erkennt, dass die Varianzen in den Abschnitten zu unterschiedlich sind. Wenn man dies außer Acht lässt und trotzdem eine Analyse macht erhält man folgendes Ergebnis:

```
> summary(aov(Datenwert ~ Abschnitt))
              Df Sum Sq Mean Sq F value Pr(>F)
Abschnitt      2  20679   10340   71.19 <2e-16 ***
Residuals    147  21351     145
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Man hätte also einen sehr guten Messpunkt verworfen, weil ANOVA in diesem Fall versagt hätte. Leider gibt es Messgrößen die ein ähnliches Verhalten vorweisen und daher kann diese Methode nicht verwendet werden.

6.1.2. Residuenanalyse - Jackknife Residuen

In der Regression ist ein Residuum r_i ist die Abweichung der geschätzten i -ten Beobachtung $\hat{\mu}_i$ von der tatsächlichen i -ten Beobachtung Y_i . Das bedeutet $r_i = Y_i - \hat{\mu}_i$, beziehungsweise $r = (I - H)Y$ wobei $H = X(X^T X)^{-1}X$. Die für diese Arbeit wichtigste Aussage ist wohl, dass man mithilfe der Residuen einzelne Datenpunkte sehr gut bewerten kann. Besonders relevant ist die Frage: "Wie verändert sich das Residuum von y_i , wenn die Beobachtung x_i gar nicht verwendet wird?" Es stellt sich also die Frage, welchen Einfluss hat die i -te Beobachtung auf die Schätzungen der Modellparameter und somit auf das Residuum.

Diese Frage kann mithilfe des sogenannten "Jackknife" Residuums beantwortet werden. Sei dazu $X_{(i)}$ die Prädiktoren Matrix ohne die i -te Beobachtung und x_i die i -te Beobachtung. Der Schätzer $\hat{\beta}_{(i)}$ kann geschrieben werden als

$$\hat{\beta}_{(i)} = (X_{(i)}^T X_{(i)})^{-1} X_{(i)}^T Y_{(i)} = \hat{\beta} - \frac{r_i}{1 - h_{i,i}} (X_{(i)}^T X_{(i)})^{-1} x_i$$

6. Die Messpunktdiagnose

Die Jackknife Residuen sind folgendermaßen definiert

$$t_i = \frac{y_i - x_i^\top \hat{\beta}_{(i)}}{S_{(i)} \sqrt{1 + x_i^\top (X_{(i)}^\top X_{(i)})^{-1} x_i}} = \frac{r_i}{S_{(i)} \sqrt{1 - h_{i,i}}}$$

wobei $S_{(i)}$ die geschätzte Standardabweichung ohne der i -ten Beobachtung ist. Unter der Annahme der Normalverteilung ist $t_i^2 \sim F_{1,n-1-p}$ verteilt wobei n der Anzahl der Beobachtungen und p der Anzahl der Prädiktoren entspricht. Für detailliertere Resultate und Erkenntnisse kann BELSLEY et al. (2004) herangezogen werden. Es folgt nun wieder ein Beispiel um sich grafisch die eben beschriebene Situation vorstellen zu können.

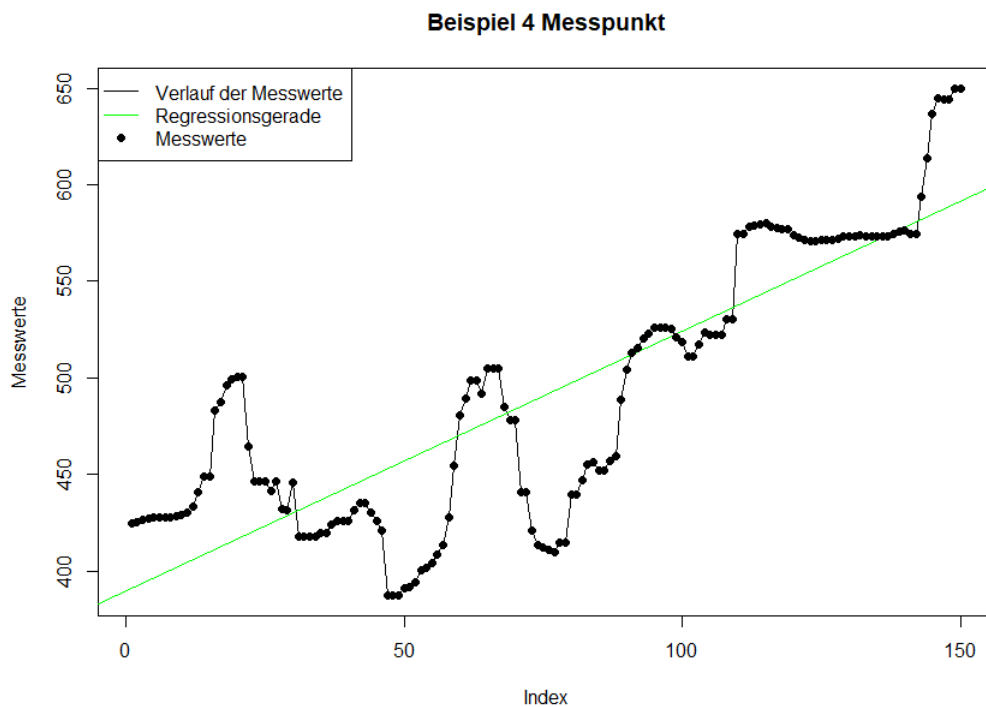


Abbildung 6.4.: Beispiel eines Messpunkts für Residuenanalyse

Die Messwerte wurden mithilfe einer einfachen linearen Regression geschätzt. An welcher Stelle sich die Messwerte befinden, ist mit schwarzen Kugeln gekennzeichnet. Dies dient zur leichteren Erkennung der Residuen in der nächsten Grafik.

6.1. Möglichkeiten zur Bewertung allgemeiner Daten

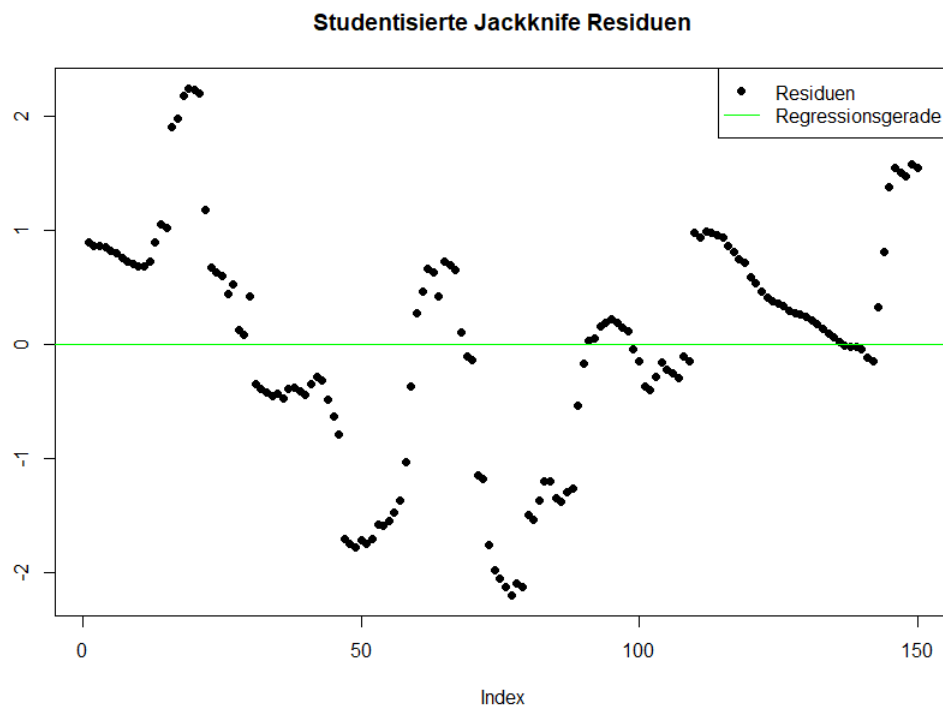


Abbildung 6.5.: Studentisierte Jackknife Residuen des vorherigen Beispiels

Wie deutlich zu sehen ist, verhalten sich die Residuen nicht auffällig. Das deutet darauf hin, dass dieser Messpunkt in Ordnung ist. Und tatsächlich war dies ein Beispiel mit echten Daten und der Messpunkt wurde von Prüfstandsfahrern akzeptiert und als gut bewertet.

Doch wie zuvor gibt es auch hier Situationen wo diese Methode versagt. Dieses Beispiel wird nun hier angegeben und beinhaltet zwar keine echten Daten, aber simulierte Daten die in genau dieser Form bei der Messgröße "T2strich" auftreten. Eine Eigenschaft dieses Parameters ist die Tatsache, dass die Nachkommastellen gerundet werden. Man hat nur eine Nachkommastelle zur Verfügung was das "sprunghafte" Verhalten erklärt. Die Messgröße springt also immer nur zwischen einigen wenigen Werten hin und her.

6. Die Messpunktdiagnose

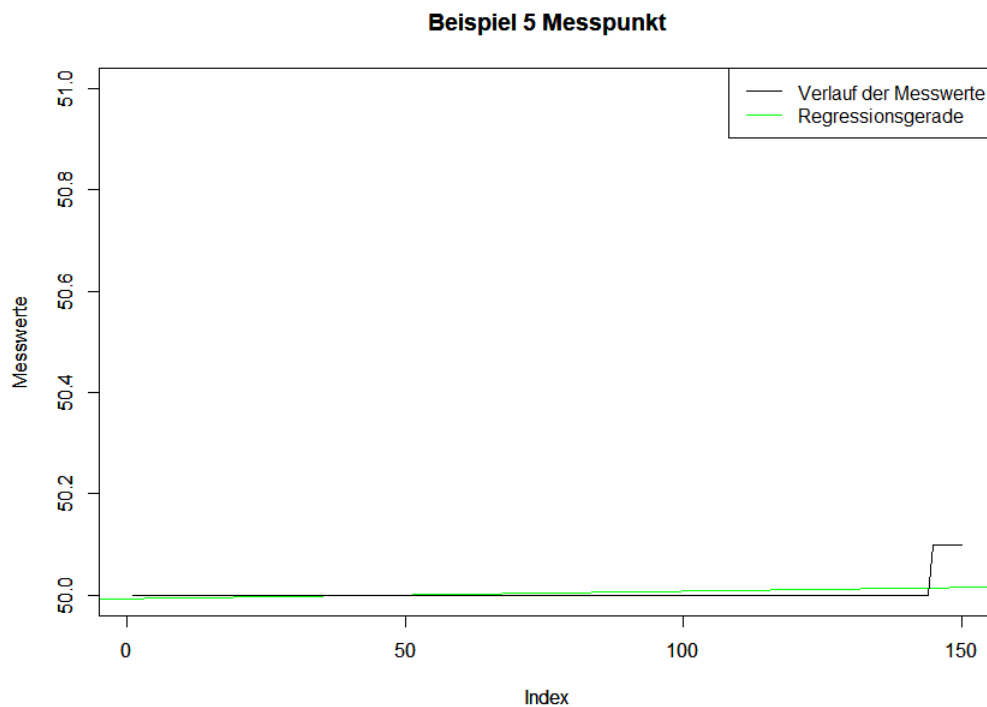


Abbildung 6.6.: Beispiel eines Messpunkts für Residuenanalyse

Wie deutlich zu sehen ist, verändern sich die ersten Werte überhaupt nicht. Nur die letzten 6 Datenpunkte sind um nur 0.1 Einheiten erhöht worden. Dies ist aber, wie schon zuvor bemerkt, ein völlig normales Verhalten dieses Parameters, da diese Messgröße diskretisierte Werte annimmt. Da dieser Strukturbruch die erlaubte Abweichung nicht überschreitet ist dies also ein ausgezeichneter Messpunkt. Als wünschenswertes Ergebnis wäre nun, dass sich die Residuen ähnlich unauffällig verhalten wie im Beispiel zuvor. Wendet man auf diesen generierten Messpunkt die Residuenanalyse an, kommt es allerdings zu einem völlig anderen Ergebnis.

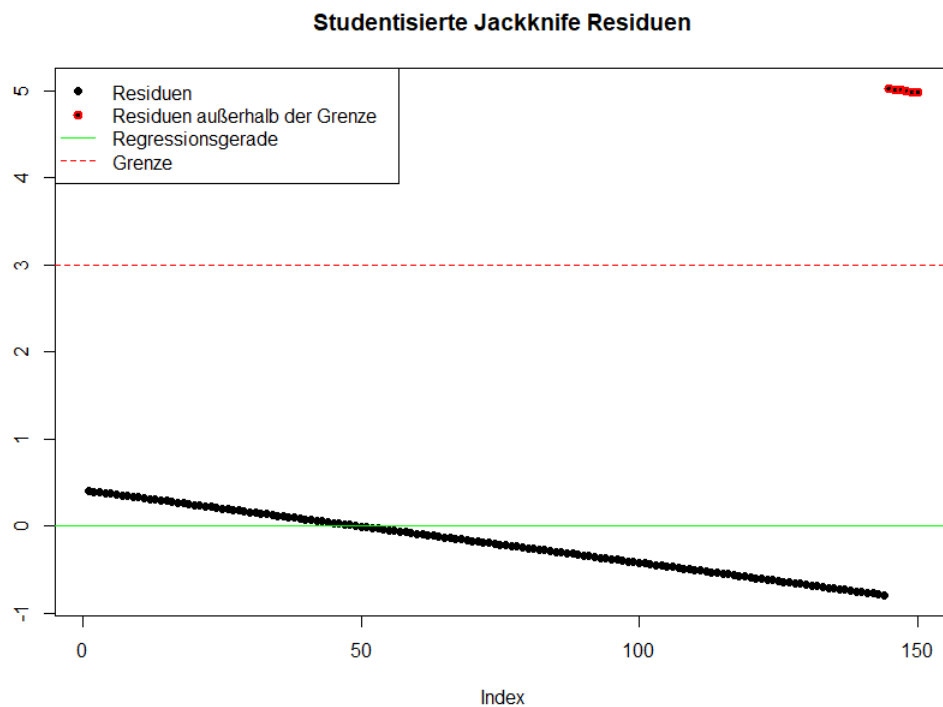


Abbildung 6.7.: Studentisierte Jackknife Residuen des vorherigen Beispiels

Leider verhalten sich die Residuen sehr auffällig, denn es wurde die empirische Varianz von abhängigen und diskretisierten Werten genommen. Man würde aufgrund dieser Grafik sofort den Schluss ziehen, dass man in diesem Messpunkt Ausreißer hat und man den Messpunkt verwerfen muss. Dieser Trugschluss macht es auch unmöglich diese Methode anzuwenden.

6.2. Bewertungsschema eines Messpunkts

Diese beiden vorgestellten Methoden haben einen Schwachpunkt. Die Varianzanalyse scheitert an quadratischen Verläufen, aber auch an der "falschen" Berechnung der Abweichung in den einzelnen Abschnitten. Es gibt sogar noch weitere Fälle in denen diese Methode nicht anwendbar wäre. Auch für die Jackknife Residuen muss die Streuung geschätzt werden aufgrund des Datensatzes den man zu Verfügung hat. Da aber diese Schätzungen in sehr vielen Fällen nicht der tatsächlichen beziehungsweise der erlaubten Abweichung entsprechen, kommt es sehr häufig zu falschen Ergebnissen.

6. Die Messpunktdiagnose

Um dieses Problem zu umgehen verwendet man ein 3-Schritte Verfahren, welches einem ähnlichen Vorgehen wie bei der Beharrungserkennung entspricht. Diese 3 Schritte werden in diesem Abschnitt aufgelistet. Die ersten beiden Punkte besitzen jeweils einen Schwachpunkt beziehungsweise Spezialfall, der mit dem darauffolgenden Schritt allerdings behoben wird.

6.2.1. Schritt 1: Soll-Ist-Vergleich

Das natürlich einleuchtendste Kriterium ist ein Vergleich des Mittelwerts mit einem gegebenen Sollwert. Ein Messpunkt, der im Mittel von dem erlaubten Wert zu stark abweicht ist natürlich unbrauchbar. Daher wird im ersten Schritt der Mittelwert berechnet und überprüft ob sich das Ergebnis stark von der Vorgabe unterscheidet. Ist dies der Fall kann der Messpunkt sofort verworfen werden. Allerdings ist ein Vergleich bei weitem nicht ausreichend um einen guten von einem schlechten Messpunkt zu unterscheiden. Ein Spezialfall wäre die folgende Grafik

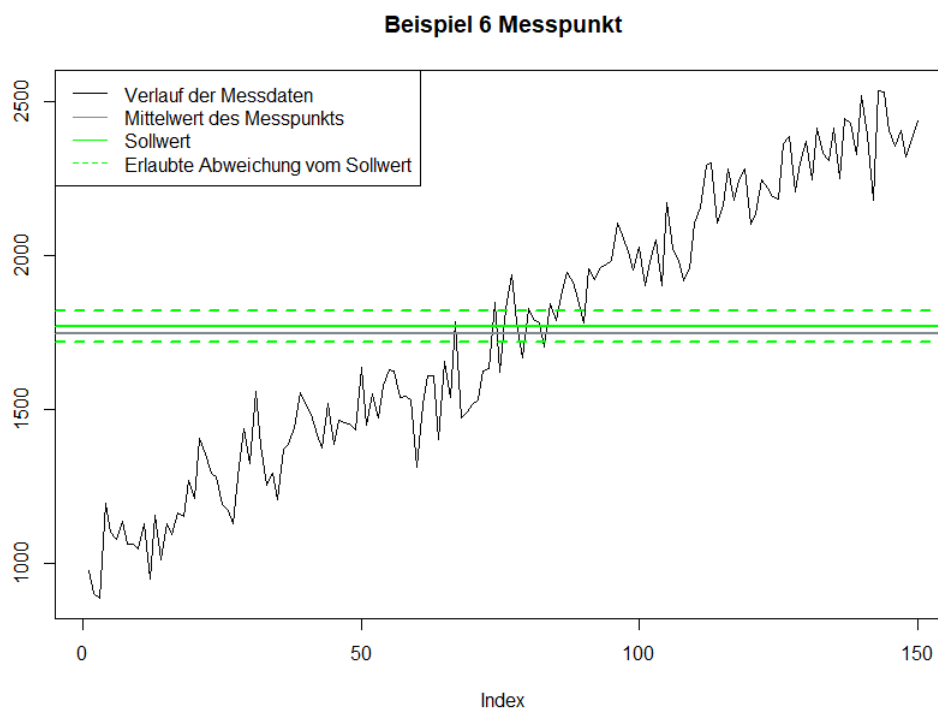


Abbildung 6.8.: Beispiel für Versagen eines reinen Soll-Ist-Vergleichs

Wie man erkennen kann, ist der Mittelwert des Messpunkts im Toleranzbereich. Es herrscht allerdings eine viel zu steile Steigung um den Messpunkt als brauchbar zu bewerten. Der Mittelwert ist nur deswegen im Toleranzbereich, weil der Sollwert dem ungefähren Zentrum der "Gerade" entspricht. Aus diesem Grund ist es unerlässlich auch den Trend des Messpunktes zu bewerten.

6.2.2. Schritt 2: Überprüfung der Steigung

Wie man gerade eben gesehen hat, muss auch das Verhalten der Messgröße berücksichtigt werden. Dies wird durch das Schätzen einer linearen Funktion bewerkstelligt. Man nimmt also wie in der Beharrungserkennung ein lineares Modell für die Messwerte an und bewertet die Steigung des Modells. Falls diese zu groß ist so muss der Messpunkt verworfen werden. Die Bewertung zu große Steigung beziehungsweise ausreichend kleine Steigung wird durch die die tolerierte Abweichung bewerkstelligt. Nimmt man die Situation von zuvor, so ergibt sich folgendes Bild:

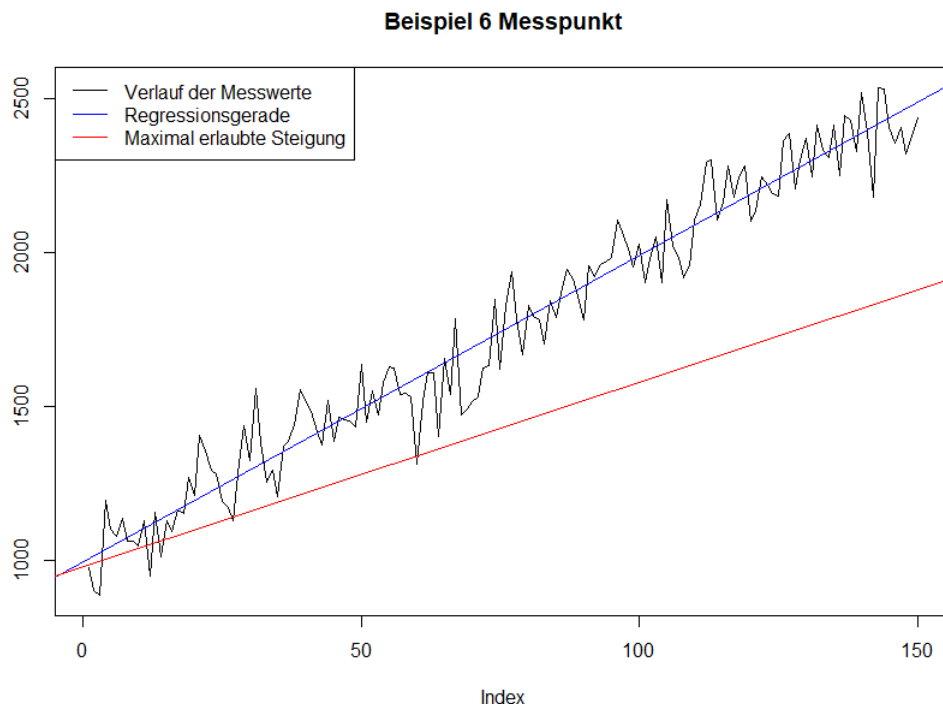


Abbildung 6.9.: Bewertung der Steigung

Die rote Linie zeigt an wie steil die Regressionsgerade maximal sein dürfte. Die

6. Die Messpunktdiagnose

tatsächliche Regressionsgerade (in blau) überschreitet offensichtlich dieses Kriterium deutlich. Daher muss dieser Messpunkt verworfen werden.

Doch wie im Schritt 1 gibt es auch hier wieder Fälle wo man erkennt, dass die bisherigen 2 Schritte nicht ausreichen um ein zuverlässiges Ergebnis zu erbringen. In diesem Fall sind es jene Messpunkte die einem quadratischen Verlauf folgen.

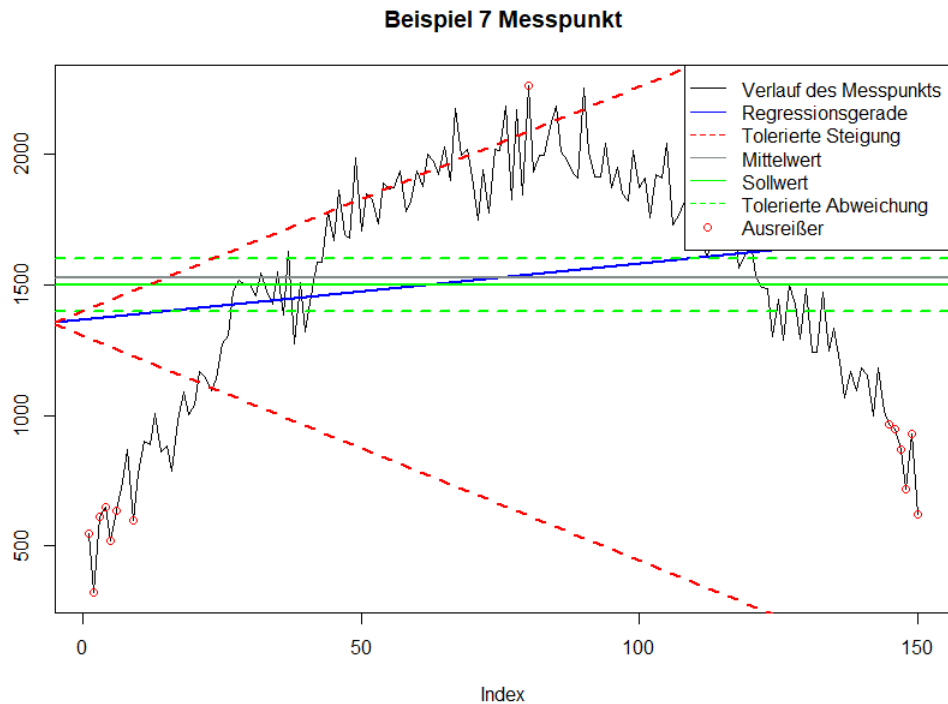


Abbildung 6.10.: Steigung und Soll-Ist Vergleich ist nicht ausreichend

Wie man deutlich erkennen kann ist die Sollwertbedingung erfüllt und auch die Steigung der Regressionsgerade ist ausreichend klein. Allerdings gibt es zu Beginn und am Ende ein paar Datenpunkte die man als Ausreißer bewerten würde, die aber bis dahin noch nicht berücksichtigt werden konnten. Eine solche Ausreißer Detektion ist der dritte und letzte Schritt.

6.2.3. Schritt 3: Ausreißer Detektion

Nun müssen nur noch die eben gezeigten Ausreißer gefunden werden beziehungsweise eine gute Richtlinie für Ausreißer ausgedacht werden. Hier hilft Abschnitt über die Ausreißer Detektion von zuvor, denn genau auf die gleiche Weise wird um den

6.2. Bewertungsschema eines Messpunkts

Mittelwert ein weiterer Vertrauensbereich gebildet. Dieser Vertrauensbereich wird um das lineare Modell gelegt.

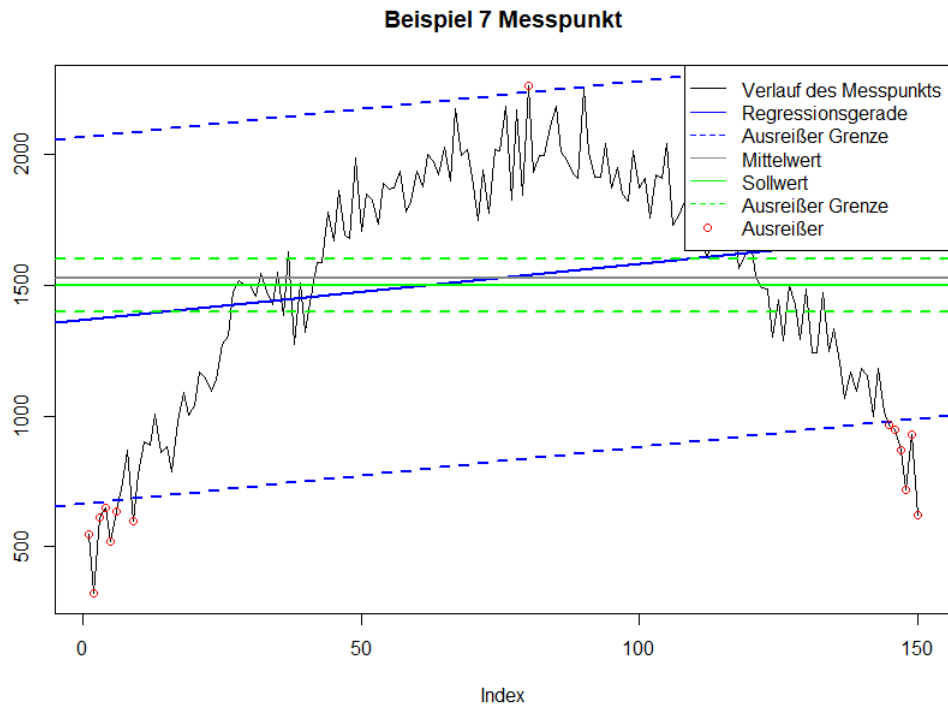


Abbildung 6.11.: Gefundene Ausreißer

Zusammengenommen ergeben diese drei Schritte beziehungsweise Überprüfungen eine gute Methode um Messpunkte zu bewerten. Wiederum hat man als Vorinformation nur die erlaubte Abweichung und die Sollwerte.

Nun zur Illustration noch ein paar Ergebnisse. Es wurden zu diesem Zweck 12 Messgrößen ausgesucht und diverse Messpunkte analysiert. Links sieht man den Verlauf der einzelnen Messgrößen kurz vor und während des Messpunkts. Der Messpunkt selbst ist jener Bereich zwischen den vertikalen grünen Linien. Die horizontale blaue Linie ist der Mittelwert des Messpunkts und falls eine grüne horizontale Linie vorhanden ist, so handelt es sich hierbei um den Sollwert der Messgröße. Falls nun ein Messpunkt in Ordnung war, so wird dies mit dem Text "Messpunkt in Ordnung" bestätigt. Falls der Messpunkt fehlerhaft war, so wird genauestens angezeigt welcher Parameter fehlerhaft war und was der Grund für den Fehler war.

6. Die Messpunkt diagnose

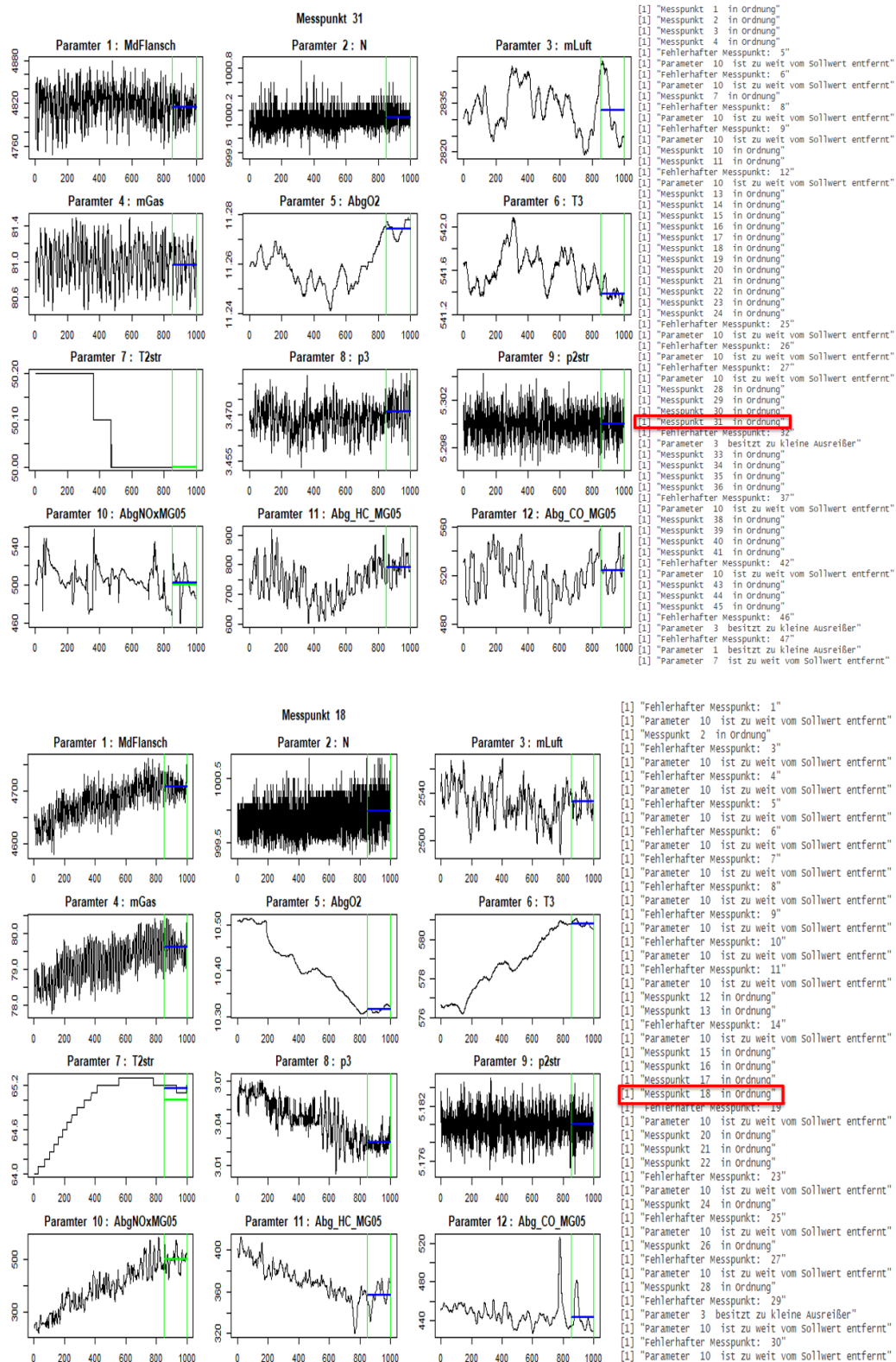


Abbildung 6.12.: Beispiele für fehlerfreie Messpunkte

6.2. Bewertungsschema eines Messpunkts

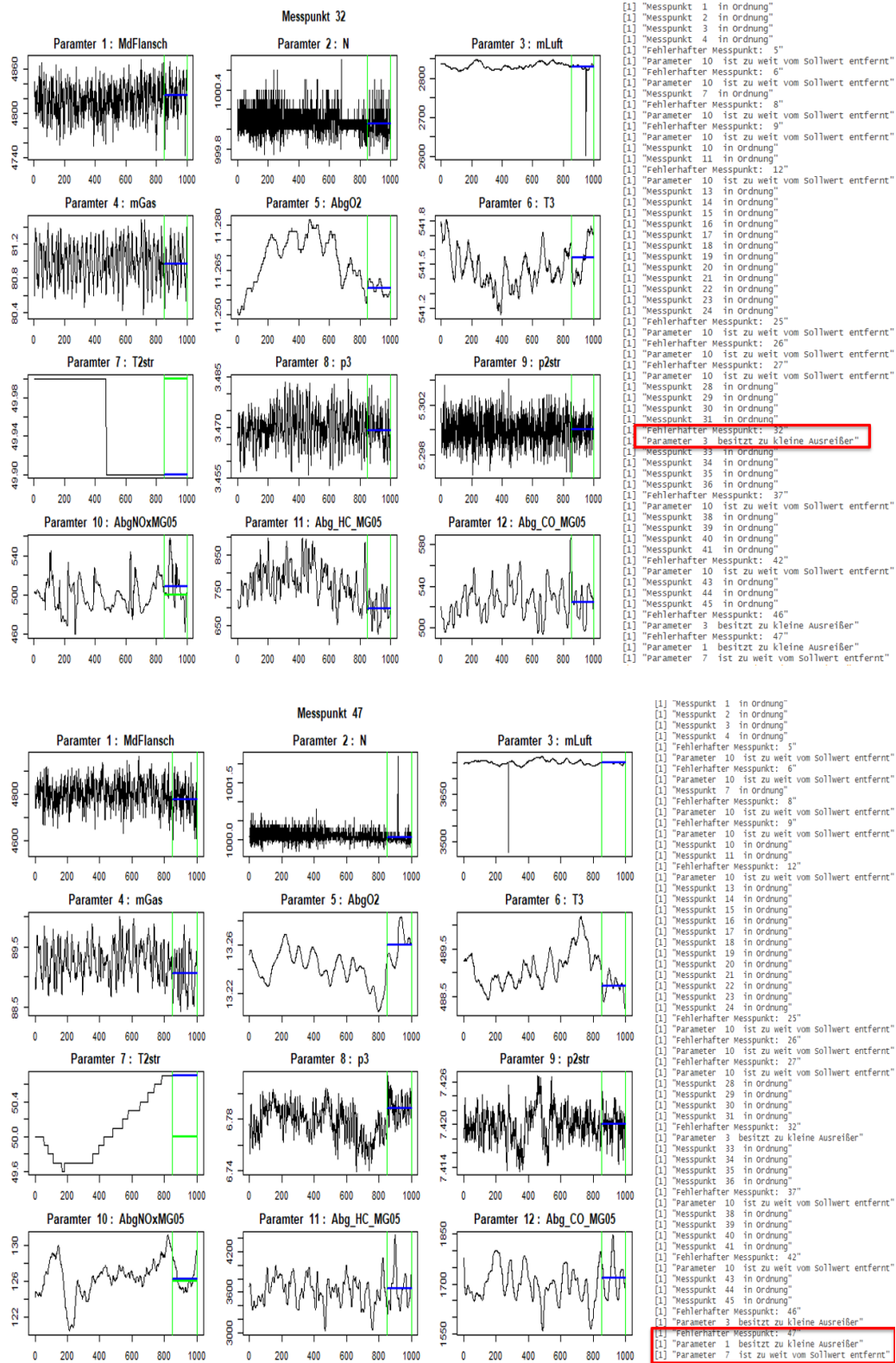


Abbildung 6.13.: Beispiele für fehlerhafte Messpunkte

7. Resümee und Ausblick

Datenanalyse: Wie sich gezeigt hat weisen die Messgrößen die unterschiedlichsten Eigenschaften auf. Diese Unterschiede reichen vom Wertebereich bis hin zum Variationsverhalten und konnten mithilfe von Histogrammen, Boxplots und dem zeitlichen Verlauf erkannt und beschrieben werden. Es wurde ebenfalls eine Korrelation einiger Messgrößen erkannt und bewertet. Aufgrund dieser breiten Variation an Verhaltensmustern der Parameter musste ein Verfahren entwickelt werden, welches all diese Erkenntnisse berücksichtigt.

Methode zur Messfreigabe: Es wurde ausführlich erläutert welche Herangehensweise praktiziert wird und welche Probleme bei der Anwendung auf konkrete Datensätze aufgetreten sind. Zu jeder aufgetretenen Schwierigkeit wird eine Lösung vorgeschlagen. Die größte Herausforderung besteht in der Rücksichtnahme aller Eigenschaften, die in den Abschnitten davor besprochen wurden. Ein großer Vorteil dieser Methode ist die geringe Komplexität sowohl in Laufzeit als auch in Speicherbedarf, da nur die aktuellsten Daten benötigt werden und mit jedem neuen Datenpunkt nur 2 Größen errechnet werden müssen.

Methode zur Messpunktanalyse: Die Schwierigkeit bei diesem Punkt lag im Versagen herkömmlicher Analysen. Es wurden Beispiele erbracht, die die echten Daten simulieren und in denen die bekannten Methoden zu falschen Ergebnissen führen. Daher musste wieder eine Reihe von Bedingungen erstellt werden die es zu überprüfen gilt. Wiederum wurden diese so gewählt, dass die Komplexität minimal ist. Sowohl die Messfreigabe als auch die Messpunktanalyse wurden in dieser Arbeit nur für den Fall einer einzig überwachten Messgröße erläutert und mit Beispielen veranschaulicht. In der Praxis müssen aber viele Parameter gleichzeitig untersucht werden. Die beschriebene Methode kann aber auch darauf eingehen, indem sie auf jede Messgröße einzeln angewendet wird und alle Einzelergebnisse zu einem großen

7. Resümee und Ausblick

Gesamtergebnis kombiniert werden.

Ausblick: Die vorgestellte Methode beruht zu einem großen Teil auf der Vorinformation der tolerierten Abweichung. Man könnte versuchen dieses Vorwissen durch eine längere Anlaufzeit der Messgrößen zu ersetzen und die Abweichungen in dieser Zeit zu schätzen. Allerdings müsste man sicherstellen, dass in dieser Anlaufzeit der Parameter sich auch wirklich so verhält wie es sein sollte. Eine weitere Möglichkeit wäre auch eine Filtermethode zu verwenden, die das Rauschen eliminieren könnte. Doch auch hier müsste man Einstellungen finden, die bei allen Messgrößen zufriedenstellende Lösungen liefern. Man könnte vielleicht auch noch weitere Kriterien für eine Messpunktanalyse finden. Eventuell kann auch ein neuer Ansatz für diese Problemstellung erprobt werden. Es gibt also viele weitere interessante Aspekte, die noch diskutiert und behandelt werden könnten.

A. Pseudocode & Funktionen

Algorithm 2: Beharrungserkennung & Messfreigabe

Data: Funktion(Daten,Sollwert,Abweichung1,Abweichung2,Fenster1,Fenster2)

Initialisierung: $\alpha_1, \dots, \alpha_4$, Sollneu(Referenzwert) = NA,

p-Value Vektoren p_1, \dots, p_4 der Länge "Fenster 2" befüllt mit Nullen

Überprüfe: Art der Abweichung

Überprüfe: Vorhandener Sollwert;

for $i = 1, \dots, n$ **do**

1. Shift der p-Value Vektoren mit dem Befehl "shift"
2. Betrachte Daten des neuen Zeitfensters und berechne die Seigung und den dazugehörigen p-Value
3. Vergleiche den p-Value mit den α 's und die Steigung
4. Befülle die frei geworden Position der Vektoren mit dem neuen, möglicherweise manuell veränderten p-Value
5. Überprüfe die Vektoren auf Gleichverteilung;

if *Gleichverteilung bei mind. 1 der Vektoren & Sollneu = NA* **then**

1. Sollneu = Mittelwert der aktuellsten Daten
 - if** *Abweichung1="relativ"* **then**
 - | Berechne zum aktuellen Mittelwert die Toleranzgrenze
- end**
2. Beginn der Beharrung;

end

if *Gleichverteilung & Mittelwert im Toleranzbereich* **then**

- Beharrung aufrecht erhalten;
- if** *Sollwert vorhanden* **then**
1. Überprüfe ob Mittelwert dem Sollwert entspricht
 2. Wenn ja, erteile Messfreigabe
- else**
- | Erteile Messfreigabe (Da Beharrung und Messfreigabe hier äquivalent)
- end**

else

1. Sollneu = NA
2. Die p-Value Vektoren werden auf 0 Vektoren gesetzt

end

end

Es gibt zwei Befehle in \mathbb{R} die hilfreich für die Programmierung sind.

1. "shift(x, n)"

Dieser Befehl führt einen Positionswechsel durch. Dabei wird jede Position des Inputvektors x um n Stellen verschoben

2. "data.table"

Mit diesem Befehl können Daten in verschiedensten Arten überprüft werden. Verwendet wird dieser Befehl um zu erfahren, ob der Mittelwert zwischen den Toleranzgrenzen liegt. Die vollständige Verwendung sieht wie folgt aus:
data.table::between(Mittelwert, Untere Grenze, Obere Grenze)

Die Funktion beziehungsweise der Algorithmus wird immer nur für eine einzige Messgröße ausgeführt. Wenn man mehrere Messgrößen zu überwachen hat, so wird diese Funktion mehrere Male aufgerufen und ein Gesamtergebnis präsentiert.

Dieselben Befehle werden für die Messpunktanalyse verwendet. Der Pseudocode für die Bewertung eines Messpunktes einer Messgröße ist unten angegeben. Wie im Fall der Beharrungserkennung kann wiederum ein Gesamtergebnis erbracht werden

A. Pseudocode & Funktionen

Algorithm 2: Messpunktbewertung

Data: Funktion(Messpunkt,Sollwert,Abweichung1,Abweichung2,Fenster1)

if *Sollwert vorhanden* **then**

| Führe Schritt 1 - 3 aus

else

| Führe Schritt 2 - 3 aus

end

Schritt 1:

if *Mittelwert stimmt mit Sollwert überein* **then**

| Gehe zu Schritt 2

else

| Abbruch und Messpunktverwerfung

end

Schritt 2:

if *Steigung der Regressionsgerade ist ausreichend klein* **then**

| Gehe zu Schritt 3

else

| Abbruch und Messpunktverwerfung

end

Schritt 3:

if *Es gibt keine Ausreißer* **then**

| Messpunkt in Ordnung

else

| Messpunkt fehlerhaft

end

Literatur

- AMARIEI, C; TOMITA, M; MURRAY, D: *Quantifying periodicity in omics data*, in: Cell and Developmental Biology, Frontiers 2014.
- ANSMANN, G: *A highly specific test for periodicity*, in: Journal of Nonlinear Science: Vol 25, No11 2015.
- BASSHUYSEN, R; SCHÄFER, F: *Handbuch Verbrennungsmotor*, Springer, 2015.
- BAUSCH, J: *On the efficient calculation of a linear combination of chi-square random variables with an application in counting string vacua*, in: Journal of Physics A: Mathematical and Theoretical 2013.
- BAUSE, F; EICKHOFF, C; MAH, R: *Truncation point estimation using multiple replications in parallel*, in: Winter Simulation Conference 2003.
- BELSLEY, A; KUH, E; WELSCH, E: *Regression Diagnostics: identifying influential data and sources of collinearity*, Wiley-Interscience, Interscience, 2004.
- BÖGE, A; BÖGE, W: *Handbuch Maschinenbau*, Springer, 2015.
- CAO, Y; THOMPSON, A; WANG, M; XIE, Yao: *Sketching for sequential change-point detection*, in: Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on 2015.
- CASELLA, G; BERGER, R: *Statistical Inference*, 2006.
- CROW, E; DAVIS, F; MAXFIELD, M: *Statistics Manual: With Examples Taken from Ordnance Development*, Dover, 1960.
- HARTMANN, P: *Mathematik für Informatiker*, Vieweg + Teubner — Springer, 2012.

Literatur

- HOAD, K; ROBINSON, S; ZHOU, S; DAVIS, R: *Automating Warm-up Length Estimation*, in: Journal of the Operational Research Society 2009.
- JEZ, P: *A note on the distribution and density of the weighted sum of independent χ^2 distributed random variables*, Techn. Ber., Universität Salzburg, 2015.
- KIM, M; YOON, S; DOMANSKI, P; PAYNE, W: *Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner*, in: International Journal of Refrigeration 1987.
- NARASIMHAN, S; KAO, M: *Detecting changes of steady states using the mathematical theory of evidence*, in: AIChE Journal 1987.
- PAWLIKOWSKI, K: *Steady-state simulation of queueing processes: survey of problems and solutions*, in: ACM Computing Surveys (CSUR) 1990.
- PETTITT, A.N: *A Simple Cumulative Sum Type Statistic for the Change-Point Problem with Zero-One Observations*, in: Biometrika 1980.
- SALVENDY, G: *Handbook of Industrial Engineering: Technology and Operations Management*, Institute of Industrial Engineers, 1982, Kap. Statistical Analysis of Simulation Results.
- SANDIKCI, B; SABUNCUOGLU, I: *Analysis of the behavior of the transient period in non-terminating simulations*, in: European Journal of Operational Research 2006.
- SCHADLER, D: *Modellbasierte Methoden zur Fehlerdiagnose an Motorprüfständen*, diploma thesis, Technische Universität Graz, 2016.
- SCHRUBEN, L; SINGH, H; TIERNEY, L: *Optimal Tests for Initialization Bias in Simulation Output*, in: Operations Research 1983.
- SOLOMON, H; STEPHENS, M: *Distribution of a Sum of Weighted Chi-Square Variables*, in: Journal of the American Statistical Association 1977.
- WIMMER, A: *Motormesstechnik*, Vorlesungsskriptum, Technische Universität Graz: Institut für Verbrennungskraftmaschinen und Thermodynamik, 2014.

WU, J; ZHOU, S; LI, X: *Acoustic emission monitoring for ultrasonic cavitation based dispersion process*, in: Journal of Manufacturing Science and Engineering 2013.

WU, J; CHEN, Y; ZHOU, S; LI, X: *Online Steady-State Detection for Process Control Using Multiple Change-Point Models and Particle Filters*, in: IEEE Transactions on Automation Science and Engineering 2014.