

Masterthesis im Studiengang: Angewandte Forschung und Entwicklung

Andreas Bernatzky 818849

Untersuchung des Einflusses der Hagelabwehr auf Hagelschäden und meteorologische Daten im Landkreis Rosenheim

Erstkorrektor der TH Rosenheim:

Prof. Dr.-Ing. Peter Zentgraf, M.Sc.

Zweitkorrektor der TH Rosenheim:

Prof. Dr. Ulrich Wellisch

Abgegeben am:

ERKLÄRUNG:

Ich versichere, dass ich diese Arbeit selbständig angefertigt, nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benützt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

Rosenheim den,

Andreas Bernatzky

Danksagung:

Herrn Prof. Dr.-Ing. Peter Zentgraf, M.Sc., will ich für die Bereitstellung des Themas und der jahrelangen Betreuung, bei Projekten an der Hochschule, danken.

Mein Dank gilt auch Herrn Prof. Dr. Ulrich Wellisch, der mit seiner Begeisterung für Statistik mein Interesse für die statistische Modellbildung geweckt hat.

Dem Verein zur Erforschung der Wirksamkeit der Hagelbekämpfung e.V. und Herrn Vogl, will ich für die gute Kooperation und Bereitstellung der Flugzeugdaten danken.

Für die Bereitstellung der Daten möchte ich den beiden Versicherern, der Versicherungskammer Bayern und der Allianz, danken. Ohne die Bereitstellung der Daten wäre ein großer Teil der Masterarbeit nicht möglich gewesen.

Der Firma nowcast GmbH und Herrn Prof. Dr. Hans-Dieter Betz, will ich für die Bereitstellung der Strokedaten danken.

Zum Schluss würde ich gerne Frau Katharina Schnötzing, Herrn Maximilian Wohlschläger und Herrn Markus Feldhaus für die Durchsicht meiner Arbeit danken.

Kurzfassung

Der primäre Kern der Arbeit befasst sich mit dem Nachweis der Wirksamkeit von Hagelabwehrflügen des Hagelabwehrvereins aus dem Landkreis Rosenheim. Hierzu werden zwei unterschiedliche statistische Modelle erstellt. Das erste Modell versucht eine Aussage über die Wirksamkeit von Silberjodidimpfungen, im Hagelzentrum einer Wolke, zu treffen. Durch ein Wetterradar können Wolken räumlich geortet werden und die Intensität des Echosignals, welches von der Reflexivität und somit vom Wassergehalt, abhängt, kann bestimmt werden. Basierend auf der räumlichen Lage und der Intensität eines Echosignals, soll ein Algorithmus entwickelt werden, welcher auf statistischen Verfahren beruht und in der Lage ist, die Wirksamkeit von Hagelabwehrflügen darzulegen.

Das zweite Modell basiert auf Versicherungsschadendaten. Zu diesem Zweck wird das betrachtete Gebiet (Südostbayern) in ein von Hagelfliegern geschütztes Gebiet und ein benachbartes Gebiet unterteilt. Statistische Parameter werden zwischen diesen Gebieten mit und ohne Hagelschutz verglichen.

Abstract

The primary focus of the work is the proof of effectiveness of hail defence flights of the Hagelabwehrverein from the district of Rosenheim. Two different statistical models are developed for this purpose. The first model tries to make a statement about the effectiveness of silver iodide vaccinations in the hail centre of a cloud. By a weather radar clouds can be spatially located and the intensity of the echo signal, which depends on the reflectivity and thus on the water content, can be determined. Based on the spatial position and the intensity of an echo signal, an algorithm is to be developed, which is based on statistical procedures and is able to demonstrate the effectiveness of hail defense flights.

The second model is based on insurance loss data. For this purpose, the area under consideration (southeast Bavaria) is divided into an area protected by hail pilots and an adjacent area. Statistical parameters are compared between these areas with and without hail protection.

1 Inhalt

1	Einleitung	8
1.1	Hagelabwehr im Landkreis Rosenheim.....	8
1.2	Stand des RO-BERTA Projekts.....	9
1.3	Entstehung von Hagel	9
1.4	Ansätze zur Beurteilung der Hagelabwehr.....	10
2	Ansatz zur Beurteilung der Hagelabwehr mittels eines Weterradar.....	11
2.1	Datenbankserver Hochfellen	11
2.1.1	Flugzeugdaten	11
2.1.2	Radardaten.....	12
2.1.3	Datenstruktur Radardaten.....	13
2.2	Grundlagen der Radartechnik.....	16
2.2.1	Radargrundgleichungen.....	18
2.2.2	Radargleichung für ein Weterradar	18
2.3	Ansätze zur Analyse der Weterradardaten	21
2.3.1	Nachweis der Veränderung von Wolkenclustern	21
2.3.2	Erhebung von Rasterkennzahlen	23
2.3.3	Betrachtung einer lokalen Region um den Hagelabwehrflieger.....	24
2.3.4	Bewertung der Ansätze zur Analyse der Weterradardaten.....	27
2.4	Erdkoordinatensystem	27
2.5	Distanzberechnung aus GPS-Daten.....	28
2.5.1	Pythagoreische Distanzberechnung.....	28
2.5.2	Distanzberechnung mit abhängigem Längengrad	29
2.5.3	Distanzberechnung über Orthodrome (Großkreisbogen)	30
2.5.4	Distanzberechnung über den Referenzellipsoid WGS84	32
2.5.5	Vergleich der einzelnen Verfahren.....	33
2.5.6	Mercator Transformation.....	34
2.6	Interpolation der Radarkarte.....	36
2.6.1	1D-Interpolation (Linear)	36
2.6.2	Bilineare Interpolation.....	37
2.7	Hailshow.....	39
2.7.1	Anbindung an die Datenbank und Datenmanagement	39
2.7.2	Hailshow Exportfunktionalität	40
2.7.3	Hailshow Importfunktionalität	40
2.7.4	Visualisierung und Erfassung der Flüge.....	41
2.7.5	Manuelle Untersuchung von einzelnen Gewitterzellen.....	46

2.7.6	Algorithmus zur Erkennung von Gewitterzellen.....	47
2.7.7	Flugverlauf	51
2.7.8	Virtuelle Flüge.....	52
3	Auswertung des Radarechoansatzes.....	55
3.1.1	Zufallszahlenerkennung	57
3.1.2	Ergebnis der Auswertung des Geometrieansatzes.....	59
3.1.3	Ergebnis der Auswertung des Cloudtarget Ansatzes.....	61
3.2	Fazit zum Radarechoansatz	63
4	Ansatz zur Beurteilung der Hagelabwehr mithilfe von Versicherungsschadendaten.....	64
4.1	Datenstruktur.....	64
4.2	Übersicht über die Versicherer.....	65
4.2.1	Extremereignisse	65
4.2.2	Auswertung aller Schadenstage.....	66
4.2.3	Auswertung aller guten Flugtage ohne Extremereignisse	71
4.3	Multiple lineare Regressionsanalyse der Versicherungsdaten	76
4.4	Erweiterte Analyse einer Multiplen linearen Regression (Residuenanalyse).....	82
4.5	Analyse der Versicherungsdaten mithilfe eines generalisierten linearen Modells	87
4.5.2	Maximum-Likelihood Methode	88
4.5.3	Bestimmung der Score-Funktion für die Parameterschätzer bei GLMs.....	89
4.6	Ergebnis des GLMs für die Versicherungsdaten.....	90
4.7	Analyse der Versicherungsdaten mithilfe eines bereinigten GLM.....	93
	Verbesserung des statistischen Modells mithilfe des Akaike Informationskriterium	96
4.7.1	Verbesserung des GLM mithilfe des stepAIC.....	97
4.8	Vergleich der Versicherungsdaten mithilfe eines Scoring-Verfahrens	100
4.9	Zusammenfassung des Scoringverfahrens.....	101
4.10	Untersuchung der Schadensverteilung der einzelnen Gebiete.....	102
4.10.1	T-Test - Vergleich zweier Mittelwerte	102
4.10.2	F-Test - Vergleich auf Varianzheterogenität	102
4.10.3	Wilcox Vorzeichen-Rangsummentest (Mann-Whitney Test)	102
4.11	Ergebnis der Verteilungsanalyse zwischen den Schäden im Nachbargebiet und im Schutzgebiet, an Flugtagen.....	103
4.12	Ergebnis der Verteilungsanalyse zwischen den Schäden im Schutzgebiet an Flug und Nichtflugtagen.....	103
4.13	Ergebnis der Verteilungsanalysen	104
4.14	Möglichkeit der Generierung eines Faktors zur Klassifikation, der Unwetterstärke	104
4.14.1	Atmosphärische Impulsstrahlung	104
4.14.2	AIS-Aktivität – Absoluter Schaden Schutzgebiet und Nachbargebiet	107

4.14.3	AIS-Aktivität – Absoluter Schaden Schutzgebiet	108
4.14.4	AIS-Aktivität – Absoluter Schaden Nachbargebiet.....	108
4.14.5	AIS-Aktivität – Absoluter Schaden Schutzgebiet und Nachbargebiet bereinigt.....	109
4.14.6	AIS-Aktivität – Absoluter Schaden Schutzgebiet um Ausreißer bereinigt	109
4.14.7	AIS-Aktivität – Absoluter Schaden Nachbargebiet um Ausreißer bereinigt.....	110
4.14.8	Fazit zur AIS-Aktivität im Zusammenhang mit den zu erwartenden Schäden	110
4.15	Fazit zur Analyse der Versicherungsdaten	110
5	Zusätzliche Ergänzungen zum ROBERTA-Projekt	111
5.1	Datenaquise von <i>FlightAware</i>	111
5.2	Raspberry ADS-B Empfänger Datenbeschaffung.....	112
5.2.1	Flightradar24 Export	113
6	Fazit.....	116
7	Literaturverzeichnis.....	118
8	Anhang	121

Abkürzungsverzeichnis Radar

c_0	<i>Lichtgeschwindigkeitskonstante</i>	$\frac{m}{s}$
t_1	<i>Laufzeitsendeimpuls</i>	s
t_2	<i>Laufzeitechoimpuls</i>	s
P_r	<i>Empfangsleistung</i>	W
P_t	<i>Sendeleistung</i>	W
G_t	<i>Antennengewinn der Sendeantenne</i>	dB
G_r	<i>Antennengewinn der Empfangsantenne</i>	dB
λ	<i>Wellenlänge der Trägerfrequenz</i>	m
σ	<i>Effektive Reflexionsfläche</i>	m ²
R_t	<i>Entfernung Sendeantenne – reflektierendes Objekt</i>	m
R_r	<i>Entfernung reflektierendes Objekt – Empfangsantenne</i>	m
φ_{az}	<i>Vertikaler Öffnungswinkel des Radarsignals</i>	°
φ_{el}	<i>Horizontaler Öffnungswinkel des Radarsignals</i>	°
R	<i>Entfernung Antenne Impulsvolumen</i>	m
τ	<i>Laufzeit Sendeimpuls</i>	s
m	<i>Optischer Brechungsindex</i>	–
D	<i>Durchmesser der Niederschlags –/wolkenteilchen</i>	m
N	<i>Anzahl der Niederschlags –/wolkenteilchen</i>	–

Gebietskodierung

<i>LK9163</i>	<i>Stadt Rosenheim</i>
<i>LK9182</i>	<i>Landkreis Miesbach</i>
<i>LK9187</i>	<i>Landkreis Rosenheim</i>
<i>LK9189</i>	<i>Landkreis Traunstein</i>
<i>LK9171</i>	<i>Landkreis Altötting</i>
<i>LK9172</i>	<i>Landkreis Berchtesgadener Land</i>
<i>LK9173</i>	<i>Landkreis Bad Tölz</i>
<i>LK9175</i>	<i>Landkreis Ebersberg</i>
<i>LK9177</i>	<i>Landkreis Erding</i>
<i>LK9183</i>	<i>Landkreis Mühldorf am Inn</i>
<i>LK9184</i>	<i>Landkreis München</i>
<i>LK9162</i>	<i>Stadt München</i>

1 Einleitung

„Die heftigen Gewitter mit Sturmböen, Starkregen und Hagel am Pfingstmontag haben im Süden Bayerns Millionenschäden verursacht. Die Versicherungskammer Bayern berichtete am Donnerstag in einer ersten Schätzung von Schäden in Höhe von voraussichtlich 30 Millionen Euro. Der Gebäudeversicherer rechnet demnach mit rund 10.000 Schadensfällen durch Hagel und Starkregen. Die Hagelschneise sei weiträumig von Kaufbeuren bis München verlaufen. Betroffen waren demnach neben München und dem Münchner Umland vor allem die Landkreise Freising, Dachau, Landsberg am Lech, Fürstfeldbruck, die Regionen rund um Ammersee und Wörthsee sowie der Landkreis Kaufbeuren“ (Süddeutsche Zeitung, 2019). Um solche Ereignisse, wie am 10.06.2019, zu verhindern, betreibt der Landkreis Rosenheim ein Hagelabwehr Programm. Hier werden, zur Bekämpfung von Gewitterwolken, welche Hagel enthalten, Silberjodidteilchen (AgJ-Teilchen) in den Wolken freigesetzt. Durch die enthaltenen Kristallisationskeime kommt es zu einer Abnahme von festen Bestandteilen in den Wolken und zu einer Zunahme der flüssigeren Bestandteile. Der Hagel regnet also in Form von Wasser ab, bevor die einzelnen Hagelkörner eine kritische Größe erreichen und ihr volles Schädigungspotenzial entfalten können.

1.1 Hagelabwehr im Landkreis Rosenheim

Die Hagelabwehr besitzt im Landkreis Rosenheim, seit den 1930er Jahren, eine langjährige Tradition. Hierbei wurden anfangs Silberjodid (AgJ) gefüllte Raketen auf Gewitterwolken geschossen. Durch den zweiten Weltkrieg wurde diese Möglichkeit der Bekämpfung von Gewitterwolken, mittels Raketen, eingestellt. Erst im Jahre 1958 wurde ein zehnjähriger Freilandversuch auf Initiative des damaligen Landrats Georg Knotts gestartet. Dabei wurde versucht, mithilfe von Bodengeneratoren, welche Silberjodid am Boden verbrennen und dem Einsatz von Hagelabwehrraketen, eine Verminderung von Hagelschäden zu erreichen. Während des Versuchszeitraums wurde ein Rückgang der Hagelschläge um fast 30 % verzeichnet (Hagelabwehr Rosenheim, 2019). Durch Änderungen des Sprengstoffgesetzes konnten die Treibsätze der Hagelabwehrraketen nicht mehr gelagert werden und es musste eine Alternative, zur Einbringung von AgJ in Gewitterwolken, gefunden werden. In diesem Fall bot sich die Einbringung von AgJ mittels Flugzeugen an. Der Rosenheimer Bauingenieur Hermann Selbertinger flog das erste Einsatzflugzeug der Hagelabwehr Rosenheim, mit selbst konstruierten AgJ-Generatoren. Im Jahr 2007 startete der Hagelforschungsverein, in Zusammenarbeit mit der Hochschule Rosenheim, das Forschungsprojekt ROBERT unter der damaligen Leitung von Prof. Dr. Ernst Schneider. Es war das Ziel eine verbesserte Aufbereitung von Wetterdaten und die damit einhergehende erhöhte Impfgenaugigkeit zu erhöhen. Seit 2010 wird das Projekt mit dem Namen RO-BERTA, unter der Leitung von Prof. Dr. Peter Zentgraf, an der Technischen Hochschule Rosenheim weitergeführt.

1.2 Stand des RO-BERTA Projekts

Das RO-BERTA Projekt befindet sich mittlerweile in der dritten Projektphase. Das Projekt umfasst derzeit zwei mit Silberjodid-Generatoren ausgestattete Flieger, wovon einer zusätzlich mit Sensorik zur Positionserfassung ausgestattet ist. Dieser sendet bei einem Hagelabwehrflug seine Positionsdaten und Umgebungsmesswerte, wie zum Beispiel Temperatur und Luftfeuchte, an eine Bodenstation. Mithilfe von Radarkarten können Unwetterzentren visuell auf einem Bildschirm im Cockpit dargestellt werden und erlauben so dem Piloten eine zielgenaue Bekämpfung von Gewitterwolken. Hinsichtlich der Öffentlichkeitsarbeit existiert eine Handy App (TH-Rosenheim, 2019), welche es ermöglicht einen Hagelabwehrflug live mitzerleben. Für die dritte Projektphase sind unter anderem eine Überarbeitung und Erneuerung der kompletten Soft- und Hardware geplant.

1.3 Entstehung von Hagel

Das Naturphänomen Hagel kann bei dem Aufeinandertreffen von kalten Luftmassen mit warmer Luft entstehen. Hierbei entsteht in einer Wolke ein sogenannter Aufwindschlot, welcher kleine Regentropfen in eine Höhe von bis zu 12000 m transportieren kann. Dabei gefrieren die Regentropfen zu ersten kleinen Eiskristallen, da in solch einer Höhe Temperaturen bis zu $-80\text{ }^{\circ}\text{C}$ herrschen können (Stieglmair, 2019). Je höher die Eiskristalle getragen werden, desto geringere Aufwinde sind vorhanden. Ab einem gewissen Gewicht stürzen die noch kleinen Eiskristalle herab, um erneut vom Aufwind erfasst zu werden. Dort setzen sich, in den unteren Wolkenschichten, erneut Regentropfen an die jetzt vorhandenen Eiskristalle, welche von den Aufwinden wieder empor gerissen werden. In hohen Höhen kommt es schließlich zu einem erneuten anfrieren an dem Eiskristall, und der Prozess wiederholt sich, bis der Eiskristall zu schwer ist, um erneut vom Aufwind erfasst zu werden. Die gewachsenen Eiskristalle gehen schließlich als Hagelkörner nieder, wenn diese zu schwer sind um vom Aufwind erfasst zu werden.

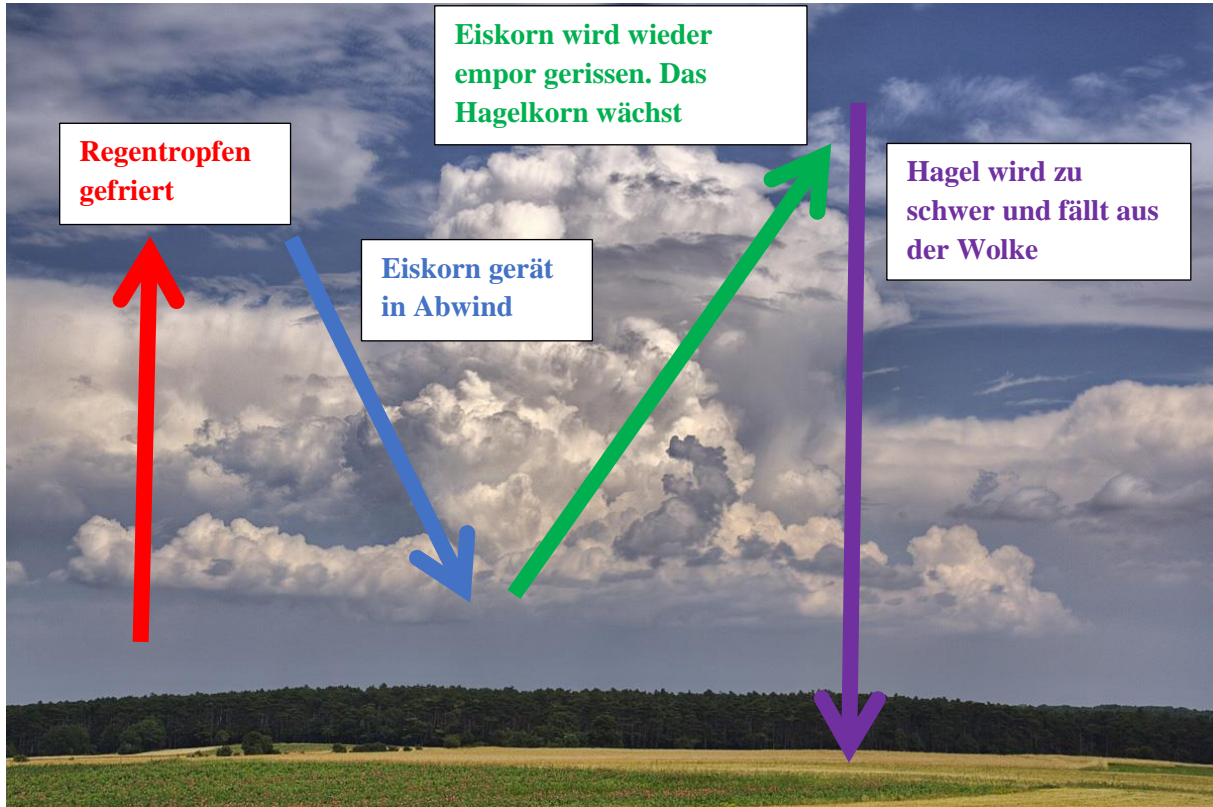


Abbildung 1: Entstehung von Hagel (Manfred Wasshuber, 2019)

Zur Bekämpfung von Hagelentstehung werden Silberjodidteilchen, mittels sogenannter AgJ-Generatoren, in den Wolken freigesetzt, bevor es zur eigentlichen Hagelbildung kommt. Durch die Kristallisationskeime, in Form von AgJ, kommt es zu einem Konglomerieren von Wassertropfen, sodass diese vorzeitig abregnen. Der Hagel regnet also in Form von Wasser ab, bevor die einzelnen Hagelkörner eine kritische Größe erreichen und ihr volles Zerstörungspotenzial entfalten können.

1.4 Ansätze zur Beurteilung der Hagelabwehr

In der Vergangenheit wurden einige Versuche unternommen, die Effektivität von Hagelabwehr nachzuweisen. Hierbei soll die Studie „Grossversuch IV“, aus dem Jahre 1985 erwähnt werden, welche die Sowjetische-Hagelabwehrmethode in der Schweiz untersuchte (B. Federer, 1986). Die Sowjetische-Hagelabwehrmethode wird in „Hail process investigation and hail suppression activities in the USSR“ beschrieben und verspricht auf den Gebieten der ukrainischen Sowjetunion und Osteuropa eine Hagelschadenreduktion von 70 bis 90 % (Burtsev, 1979). Es handelt sich um eine Methode, welche Wolken in festen Abständen mit Raketen impft. Die Studie konnte keinen statistisch signifikanten Unterschied zwischen geimpften und nicht geimpften Gewitterzellen feststellen.

Eine andere wissenschaftliche Begleituntersuchung aus dem Jahr 2015/16, welche vom Landwirtschaftsamt Stuttgart, für den Rems-Murr-Kreis, durchgeführt wurde, bestätigt eine Wirksamkeit der Hagelabwehr für einzelne Ereignisse (Landratsamt Rems-Murr-Kreis, 2015/2016). Hierbei wurden Radardaten von geimpften und nicht geimpften Gewitterzellen miteinander verglichen. Es konnte mit Messungen bestätigt werden, dass das Einbringen von Silberjodid die Bildung großer Hagelkörner verhindert.

2 Ansatz zur Beurteilung der Hagelabwehr mittels eines Wettersradar

Durch einen Wettersradar können Wolken räumlich geortet werden. Mit der räumlichen Lage und der Intensität eines Radarechosignals, soll ein, auf statistischen Verfahren beruhender, Algorithmus entwickelt werden, welcher in der Lage ist die Wirksamkeit von Hagelabwehrflügen nachzuweisen. Hierzu soll ein Programm in MATLAB geschrieben werden, dass in der Lage ist, den Erfolg von vergangenen und kommenden Hagelabwehrflügen zu beurteilen.

Im Folgenden soll jedoch zuerst auf die, aus dem ROBERTA-Projekt, zur Verfügung stehenden Daten und auf die Funktionsweise eines Radargerätes eingegangen werden, wobei es sich hier im speziellen um ein Wettersradar handelt.

2.1 Datenbankserver Hochfellen

Seit dem Start des ROBERTA-Projekts an der TH-Rosenheim speichert die Hochschule, mit einer MySQL-Datenbank, Radardaten, welche vom Deutschen Wetterdienst (DWD) zur Verfügung gestellt werden. Zusätzlich zu den Radardaten werden, bei stattfinden eines Hagelabwehrfluges, die Positionsdaten eines Hagelabwehrfliegers, an die Datenbank gesendet. Die Datenbank selbst befindet sich auf dem Hochfellen in den Chiemgauer Alpen.

Die Datenbank kann mit dem Programm *Hailshow* (siehe Kapitel 2.7) abgefragt werden. Es wurden hierfür drei SQL-Querys erstellt, welche von MATLAB ausgeführt werden können. Zu Demonstrationszwecken wurde ein Beispielskript geschrieben, welches die Konnektivität zur Datenbank mit MATLAB demonstriert (siehe Anhang A). Da das derzeitige Datenbankschema ab 2020 neu strukturiert wird, soll der alte Stand von 2019 dokumentiert werden.

2.1.1 Flugzeugdaten

Das Hagelabwehrflugzeug mit der Flugzeugkennung D-GITY, übermittelt, mithilfe eines Funksignals, die momentanen Positionsdaten, sowie Temperatur und Luftfeuchte, an die Datenbank. Eine komplette Übersicht, über die vom Flugzeug ermittelten Daten, ist im Anhang B zu finden. Bei einigen Flügen kommt es, aufgrund von terrestrischen Störungen, zu einem Abriss der Datenkommunikation. Diese Lücken in den Positionsdaten können, mithilfe von einem Mikrocomputer gesammelten Daten, ergänzt werden (siehe Kapitel 5.2). Da das zweite Hagelabwehrflugzeug D-GOGO nicht mit einem GPS-Transmitter ausgestattet ist, können dessen Hagelabwehrflüge nur über die vom Mikrocomputer gesammelten Daten analysiert werden. Diese Daten können jedoch mithilfe der Importfunktion von *Hailshow* (siehe Kapitel 2.7.3) ebenfalls analysiert werden.

Die Flugzeugdaten befinden sich auf dem Hochfellen Server in der Datenbank „hailena_airplane0“. Nachfolgende Abbildung zeigt das SQL-Query, mit dem im Programm *Hailshow* auf die Datenbank zugegriffen wird.

```
SELECT GP_ID,  
       S06_DateTime,  
       S06_LON,  
       S06_LAT,  
       S06_ALTsea  
FROM hailena_airplane0_20180705.hailenaacquire
```

Abbildung 2: SQL Query für die Flugdaten, hier beispielhaft für den Flug vom 05.07.2018

2.1.2 Radardaten

Die Radardaten werden vom Deutschen Wetterdienst über das Isen Radar zur Verfügung gestellt. Hierbei werden 2-Dimensionale (im folgenden Px) und 3-Dimensionale (im Folgenden Pz) Radardaten zur Verfügung gestellt. Die Daten können in ihrer Rohform für die Px-Daten unter folgendem Link abgerufen werden:

opendata.dwd.de/weather/radar/sites/px/isn/

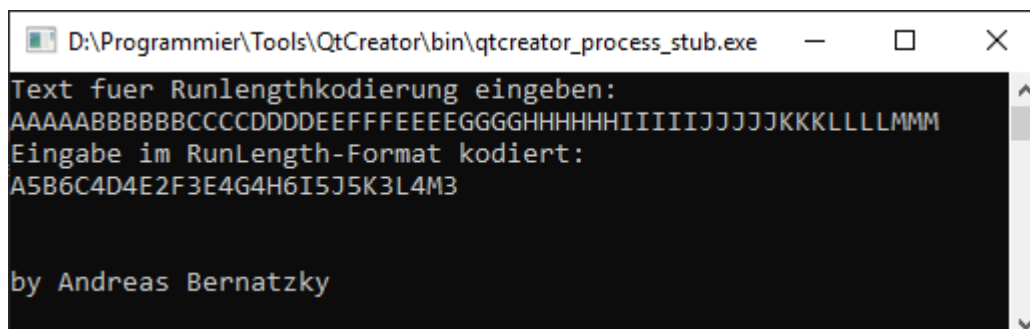
Die Pz-Daten können unter folgendem Link abgerufen werden:

opendata.dwd.de/weather/radar/sites/pz/isn/

Ein, von Herrn Martin Heigl bereitgestelltes, C++ Programm überprüft zyklisch beide Webseiten und lädt, falls ein neuer Radarscan stattfindet, diesen herunter und decodiert die binär vorliegende Datei. Anschließend werden die Informationen auf dem Hochfellnserver auf der Datenbank „hailena_radar.radardwdpxisen“ bzw. „hailena_radar.radardwdpzisen“ abgelegt.

2.1.2.1 Lauflängenkodierung

Die als Binärdatei vorliegenden Daten sind mit einer Lauflängenkodierung versehen. Eine Lauflängenkodierung ist ein Kompressionsalgorithmus, welcher sich gut für Datensätze eignet, in denen sich mehrere Zeichen aufeinanderfolgend wiederholen. Eine Implementierung, für einen eigenen Lauflängenalgorithmus, ist in einem C++ Programm im Anhang A zu finden.



```
D:\Programmier\Tools\QtCreator\bin\qtcreator_process_stub.exe
Text fuer Runlengthkodierung eingeben:
AAAAAABBBBBBCCCCDDDEEFFEEEEGGGGHHHHIIIIJJJJKKLLLLMMM
Eingabe im RunLength-Format kodiert:
A5B6C4D4E2F3E4G4H6I5J5K3L4M3
by Andreas Bernatzky
```

Abbildung 3: Beispiel einer Lauflängenkodierung

Der in Abbildung 3 gezeigte Zeichensatz benötigt 59 Byte, der komprimierte Schriftsatz hingegen nur 28 Byte. Nach dem hier aufgezeigten Prinzip sind die Radarscans vom DWD ebenfalls kodiert. Die Px-Daten stellen decodiert eine 200x200 Matrix dar, wobei jedes Feld in der Matrix für eine bestimmte GPS-Position steht. Jedes Feld besitzt einen Wert zwischen null und sechs, welches die Intensität der einzelnen Radarechos repräsentiert. Ein Radarecho der Stärke sechs steht für eine Wolke, mit einem hohen Reflexionsvermögen, was ein Indiz für Hagel ist. Es können Binärdateien mit dem, vom DWD zur Verfügung gestellten, Perl-Skript konvertiert werden, welches ebenfalls Bestandteil des Programm *Hailshow* ist (siehe hierzu Kapitel 2.7.1). In dem nachfolgenden Kapitel soll die Datenstruktur der Px- und der Pz-Daten erläutert werden.

2.1.3 Datenstruktur Radardaten

Die Radardaten erfassen Wolken in 2-D und 3-D Wolkenbildern (siehe Abbildung 4). Das blaue Rechteck (2D-Daten P_x) kann zu einer 200×200 Matrix dekodiert werden. Das rote Rechteck (3D-Daten P_z) wird in eine 200×2400 Matrix dekodiert, wobei 200 Reihen der Matrix eine neue Höhe kennzeichnen. Hier werden unterschiedliche Höhenschichten betrachtet. Die 3D-Daten lassen sich daher in eine $200 \times 200 \times 12$ Matrix umwandeln. Beispielsweise kann die Höhenschichtung für die 3D-Daten von 1000 bis 12000 Meter in Abbildung 5 betrachtet werden.

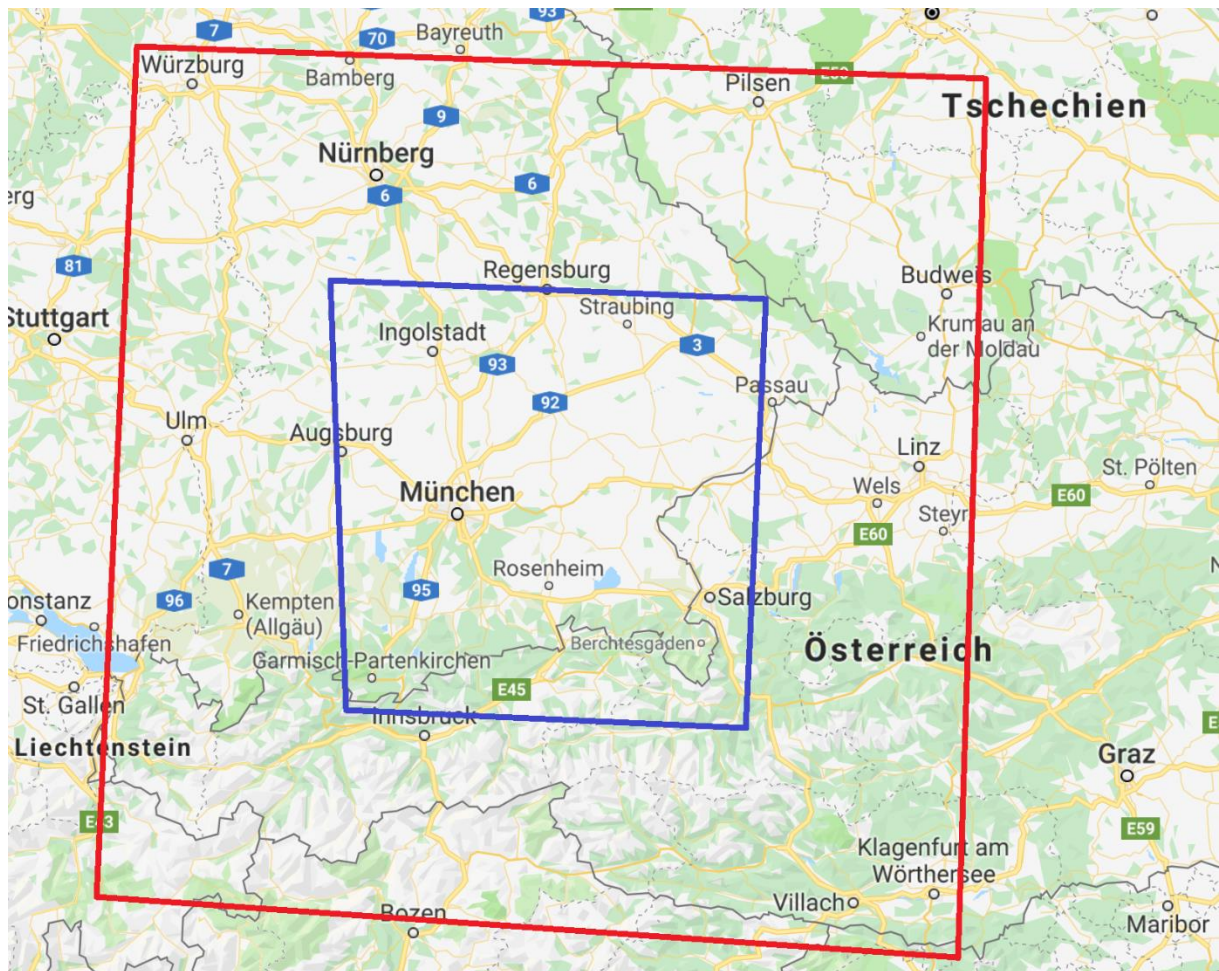


Abbildung 4: Das blaue Viereck zeigt den vom P_x -Radar abgedeckten Bereich. Das rote Viereck zeigt den vom P_z -Radar abgedeckten Bereich. (Google Inc., 2019)

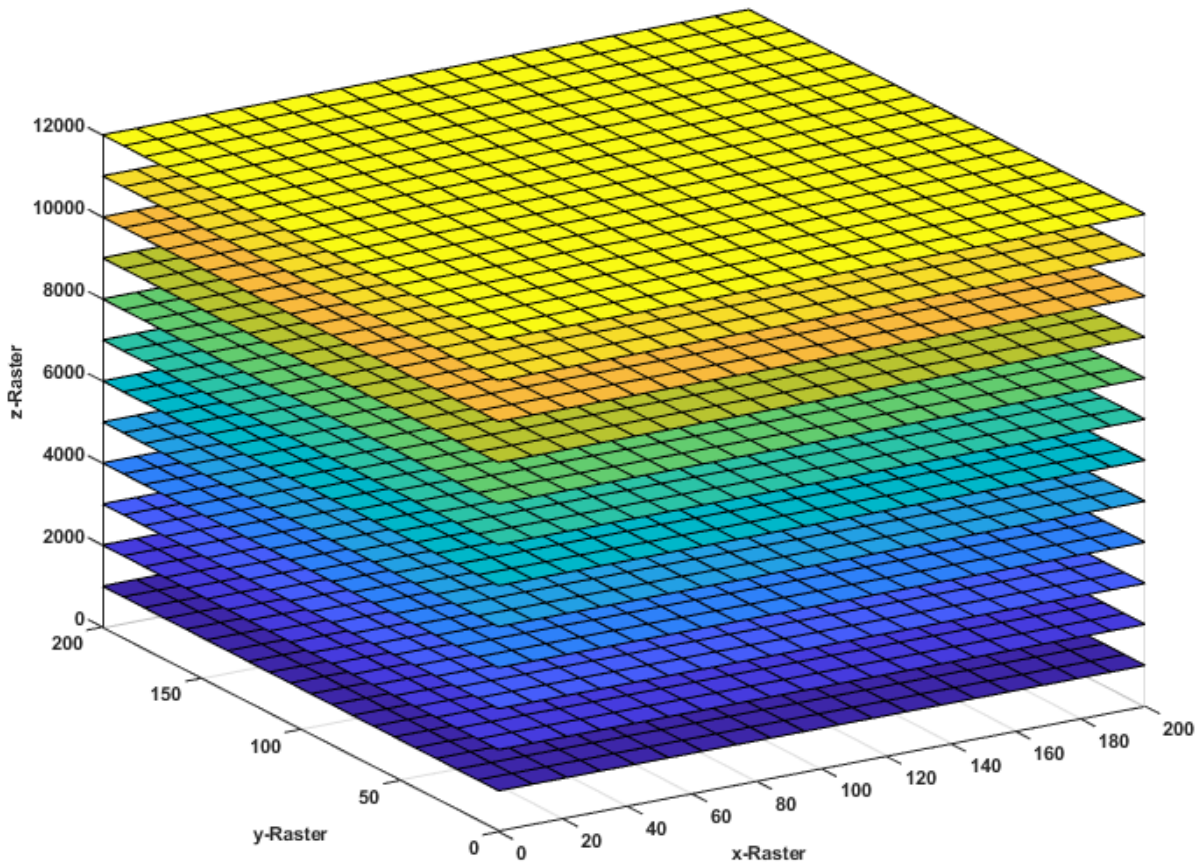


Abbildung 5: Visuelle Darstellung der Datenstruktur für die 3D-Radarbilder

Die GPS-Daten des abgedeckten Radarbereiches sind bekannt. Es wurde daher eine Lookup-Tabelle erstellt, um jedes Raster einer GPS-Koordinate zuweisen zu können (siehe Kapitel 2.6). Die Reflexion der einzelnen Echos wird in einer Skala von 1-6 (sehr leicht bis extrem) angegeben.

Reflektivitätsfaktor	
●	sehr leicht
●	leicht
●	mittel
●	schwer
●	sehr schwer
●	extrem

Abbildung 6: Farbliche Kennzeichnung und Bewertung der Reflektivitätsfaktoren

2.1.3.1 Datenstruktur der Radardaten innerhalb der Datenbank

Die Radardaten selbst liegen in der Datenbank als Binarylargeobject (BLOB) vor (siehe Abbildung 7). Der detaillierte Aufbau der Radardaten soll in diesem Kapitel erklärt werden.

GP_ID	S01_Address	S01_GenesisDate	S01_GenesisTime	S01_Class1_Size	S01_Class1_Points	S01_Class1
20849	68097	2013.04.27	16:05	4454	2227	[BLOB - 4.3 KiB]
20850	68097	2013.04.27	16:15	3872	1936	[BLOB - 3.8 KiB]

Abbildung 7: Datenbankauszug für Radardaten zu verschiedenen Uhrzeiten

Zu einer bestimmten Uhrzeit liegen sechs Reflexivitätsstufen vor, welche ausfolgenden Informationen bestehen:

S01_ClassX_Size: Gesamtmenge der Koordinaten
S01_ClassX_Points: Gesamtmenge aller vorhandenen Echos
S01_ClassX: Binarylargeobject welches die Koordinaten enthält.

S01_Class1_Size und S01_Class1_Points können zum Überprüfen der Korrektheit der Daten hinzugezogen werden.

Ein Beispiel für Class1

S01_Class1_Size: 6 6 Koordinaten gegeben (jeweils XY-Koordinate)
S01_Class1_Points: 3 Aus 6 Koordinaten können 3 XY-Pärchen gebildet werden

1.X, 1.Y, 2.X, 2.Y, 3.X, 3.Y
23, 34, 95, 118, 0, 23...

Abbildung 8: Struktur der XY-Koordinaten in einem S01_Class Object

S01_Class1: Es folgen stets XY-Koordinatenpaare abwechselnd aufeinander, wie auch Abbildung 8 zeigt. Die PX-Daten sind valide, wenn $S01_ClassX_Size/2 = S01_ClassX_Points$ ergibt. Für PZ Daten ist das Schema analog. Hier ist die Abfolge X, Y, Z und die Anzahl der Koordinatentriple muss durch die Anzahl der Koordinaten, in diesem Fall drei, dividiert werden, um die Datenintegrität zu validieren.

2.2 Grundlagen der Radartechnik

Die Nutzung und Entwicklung von Radargeräten haben ihren Ursprung im 2. Weltkrieg, um Flugzeuge und U-Boote zu orten. Die Theorie zu elektromagnetischen Wellen reicht jedoch schon bis in das Jahr 1865 zurück, wo der schottische Physiker James Clerk Maxwell seine elektromagnetische Lichttheorie aufstellte (Wolff, 2019).

Im Folgenden soll die grundlegende physikalische Funktionsweise eines Radargerätes erklärt werden und im Anschluss speziell auf Wetterradargeräte eingegangen werden. Die Erkenntnisse zur Theorie der Radartechnologie und Radarmeteorologie liegen Herrn Christian Wolff (Wolff, radartutorial, 2019) und Herrn Hagen (Martin Hagen,DLR, 2007) zugrunde. Hierzu sollen zunächst drei grundlegende physikalische Gesetzmäßigkeiten erwähnt werden, welche für Radargeräte relevant sind (Wolff, radartutorial, 2019):

1. Die **Reflexion** elektromagnetischer Wellen:
Wenn diese Wellen auf einen elektrisch leitenden Körper treffen, werden sie reflektiert. Wird die reflektierte Welle am Ursprungsort, als elektromagnetisches Echo, wieder registriert, ist das ein Beweis dafür, dass sich in Ausbreitungsrichtung ein Hindernis befindet.
2. Die **konstante Ausbreitungsgeschwindigkeit** der elektromagnetischen Wellen:
Die elektromagnetischen Wellen breiten sich mit annähernder Lichtgeschwindigkeit aus. Durch diese konstante Ausbreitungsgeschwindigkeit lässt sich die Entfernung von reflektierenden Objekten (Flugzeugen, Schiffen, Fahrzeugen) durch Laufzeitmessung der Impulse exakt bestimmen.
3. Die **geradlinige Ausbreitung** der elektromagnetischen Wellen:
Bei der Ausbreitung der elektromagnetischen Wellen im Frequenzbereich der Radargeräte geht man von einer geradlinigen Ausbreitung aus. Durch spezielle Antennen können die elektromagnetischen Wellen in eine bestimmte Richtung gebündelt werden. Somit ist es möglich, die Winkelkoordinaten (Seitenwinkel und Höhenwinkel) zu bestimmen.

Das Radargerät sendet einen gebündelten kurzen Radioimpuls mit hoher Impulsleistung in eine vorgegebene Richtung aus. Befindet sich in dieser Richtung ein Hindernis, kommt es zu einer Zerstreuung des Impulses in alle Richtungen. Ein gewisser Anteil wird wieder zum Radar reflektiert und dort von einem Detektor erfasst. Zur Bestimmung der Distanz müssen zwei Zeitpunkte gemessen werden. Zum einen der Zeitpunkt t_1 des Aussendens des Impulses und zum anderen die Registrierung des Echos t_2 am Detektor. Abbildung 9 zeigt das Prinzip der Laufzeitmessung schematisch. Folgende Formel beschreibt die Entfernung R eines Radargerätes zu einem Objekt in Abhängigkeit der Zeitpunkte t_1 und t_2 , sowie der Lichtgeschwindigkeit c_0 :

$$R = \frac{c_0 * (t_2 - t_1)}{2}$$

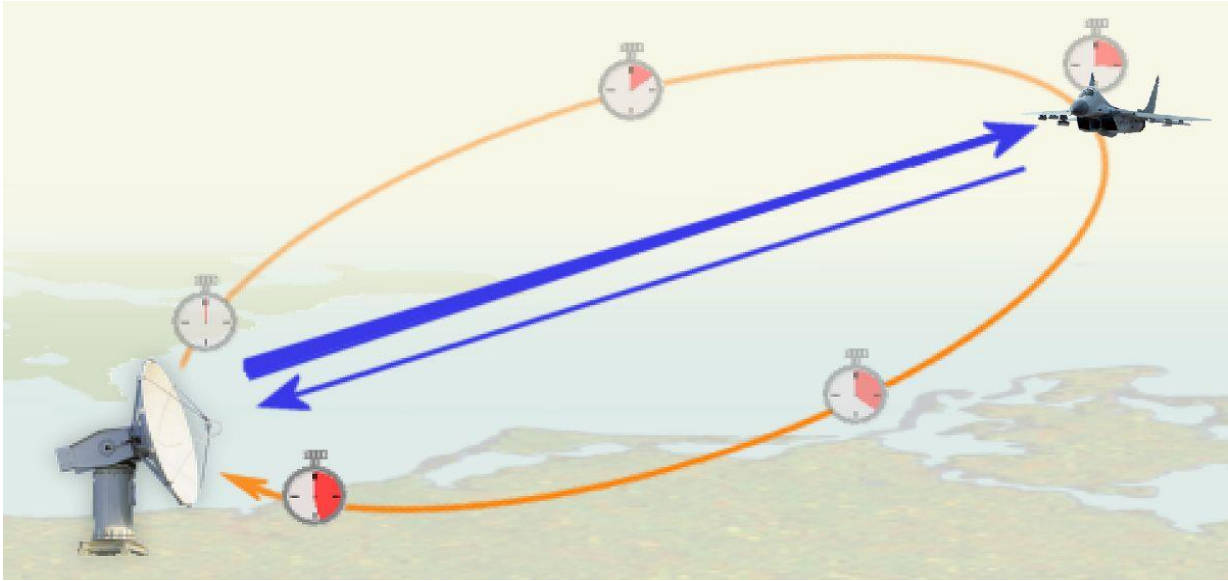


Abbildung 9: Prinzip der Laufzeitmessung (Wolff, 2019)

Der größte Unterschied zwischen Wetterradar und einem klassischen Radar zur Luftraumaufklärung besteht darin, dass bei Wetterradaren, zusätzlich zur Positionsbestimmung, die Amplitude des Echosignals mitgemessen wird (Wolff, 2019).

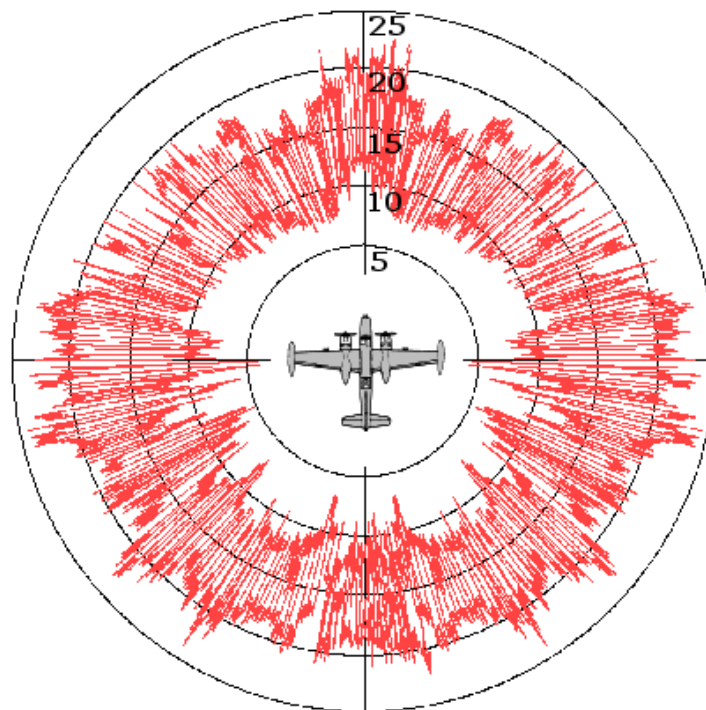


Abbildung 10: Diagramm der experimentell ermittelten, winkelabhängigen relativen Reflexionsfläche eines Flugzeuges (Wolff, 2019)

2.2.1 Radargrundgleichungen

Die Radargrundgleichung wird benutzt, um die physikalischen Zusammenhänge von der Sendeleistung, über die Wellenausbreitung, bis zum Empfang für ein Punktziel darzustellen. Mit ihr kann die Leistung P_r am Eingang des Radarempfängers, in Abhängigkeit von der Sendeleistung P_t , der Entfernung R und den Reflexionseigenschaften des Objektes σ , bestimmt werden (für ein Wetterradar würde es sich um die Reflexionseigenschaft von Regentropfen/Hagel handeln, jedoch darf eine Wolke nicht als Punktziel betrachtet werden). Der effektive Reflexionsgrad σ , oder auch oftmals Radarquerschnitt eines Objektes, gibt an, wie groß die Reflexion eines Objekts in Richtung einer Empfangsantenne ist. Abbildung 10 zeigt den Reflexionsgrad für ein Flugzeug.

Der Reflexionsgrad (Radarquerschnitt) kann nur für einfache Körper analytisch bestimmt werden, für komplexe Objekte, wie Flugzeuge, Schiffe etc. ist daher eine experimentelle Bestimmung nötig. Tabelle 1 zeigt einige typische Werte für den Reflexionsgrad:

Tabelle 1: Typische Werte für den effektiven Reflexionsgrad σ

Objekt:	Effektiver Reflexionsgrad σ:
Mensch	1 m ²
Transportflugzeug	Bis zu 100 m ²
Containerschiff	10.000-80.000 m ²

Nach (Wolff, 2019) setzt sich die Radargleichung für ein Punktziel folgendermaßen zusammen:

$$P_r = \frac{G_t * G_r * \lambda^2 * \sigma}{(4\pi)^3 * R_t^2 * R_r^2}$$

Der Sendeantennengewinn G_t und der Empfangsantennengewinn G_r ist das Produkt aus dem Richtfaktor einer Antenne und ihres Wirkungsgrades.

Die Distanz zwischen Objekt und Sender muss deutlich größer, als die Wellenlänge des Radars, sein. Allgemein gesagt, muss das Objekt sich im Fernfeld der Antenne befinden, welches wiederum von Art und Beschaffenheit der Antenne abhängig ist.

2.2.2 Radargleichung für ein Wetterradar

Im vorhergehenden Kapitel wurde das Grundprinzip und die allgemeine Radargleichung für Punktziele erklärt. Um in atmosphärischen Schichten Regen oder Hagel detektieren zu können, ist technisch betrachtet ein gewöhnliches Radargerät ausreichend. In der Theorie ändert sich hier, dass die effektive Rückstrahlfläche σ , z.B. eines Flugkörpers, sehr viel kleiner ist, als die Gesamtheit aller, sich in einer Wolke befindenden, Teilchen mit dem Durchmesser D . Problematisch ist zusätzlich, dass Wolken für Radarfrequenzen „halbdurchlässige“ Gebilde sind. Um dennoch eine Aussage über die Beschaffenheit von Wolken, welche aus winzig kleinen Regenteilchen bestehen, treffen zu können, ist es notwendig, die allgemeine Radargleichung um das Impulsvolumen V zu erweitern. Dadurch kann ein Rückschluss auf die „Menge“ an Regenteilchen, die sich in einem vom Radar abgedeckten Volumensegment befinden, erfolgen. Das Impulsvolumen ist in Abbildung 11 grafisch dargestellt.

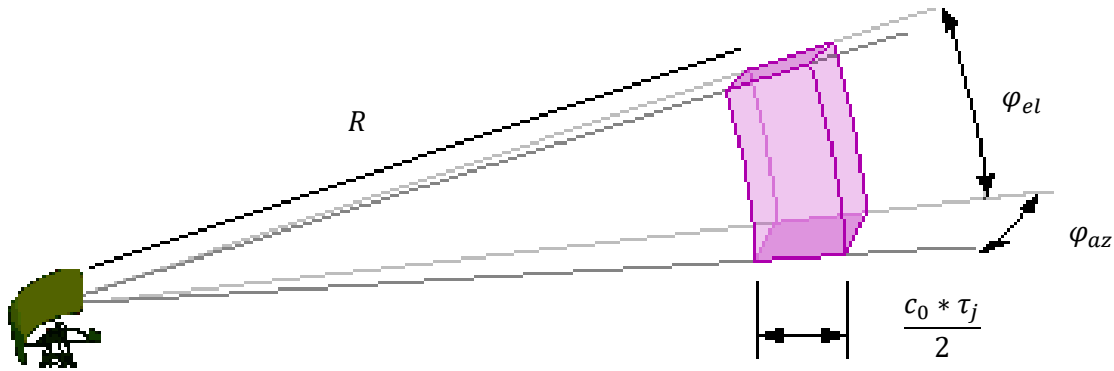


Abbildung 11: Impulsvolumen eines Radars (Wikipedia, 2019)

Das Impulsvolumen setzt sich aus dem vertikalen Öffnungswinkel φ_{el} und dem horizontalen Öffnungswinkel φ_{az} zusammen. Bei der Zeit τ_j handelt es sich um die Laufzeit des Impulses. Das Impulsvolumen bildet sich folgendermaßen:

$$V = \frac{\pi * \varphi_{el} * \varphi_{az} * R^2 * c_0 * \tau}{8}$$

Da ein Regentropfen sehr viel kleiner, als die Wellenlänge eines Radargerätes, ist, kann diese nicht mehr so einfach verwendet werden. Hierzu muss die Rückstrahlfläche nach Wolff durch die Rayleigh-Streuung bestimmt werden. Die effektive Rückstrahlfläche $\sigma_{Partikel}$ ergibt sich daher zu:

$$\sigma_{Partikel} = \frac{\pi^5}{\lambda^4} * \left| \frac{m^2 - 1}{m^2 + 2} \right|^2 * D^6 = \frac{\pi^5}{\lambda^4} * |K|^2 * D^6$$

Wenn der Teilchendurchmesser D in etwa der Radarwellenlänge entspricht, muss die Mie-Strahlung verwendet werden. Dies ist vor allem bei Hagel der Fall. Die optische Region gilt für Punktziele, wie Flugzeuge oder ähnliches. Der Brechungsindex m ist eine Materialeigenschaft aus der Optik, welche das Verhältnis der Wellenlänge des Lichts im Vakuum, zur Wellenlänge im Material wiedergibt. Abbildung 12 zeigt den Gültigkeitsbereich der Rayleighstreuung für kleine Teilchen. Der Faktor K stellt den Dielektrizitätsfaktor dar. Dieser ist abhängig von der Wellenlänge des Radarimpulses und der Temperatur (siehe Tabelle 2).

Tabelle 2: Werte für K in Abhängigkeit von der Wellenlänge

	Temperatur °C	$\lambda = 10 \text{ cm}$	$\lambda = 3 \text{ cm}$
$ K ^2$ für Wasser	20	0,9280	0,9272
$ K ^2$ für Eis	unabhängig	0,197	0,197

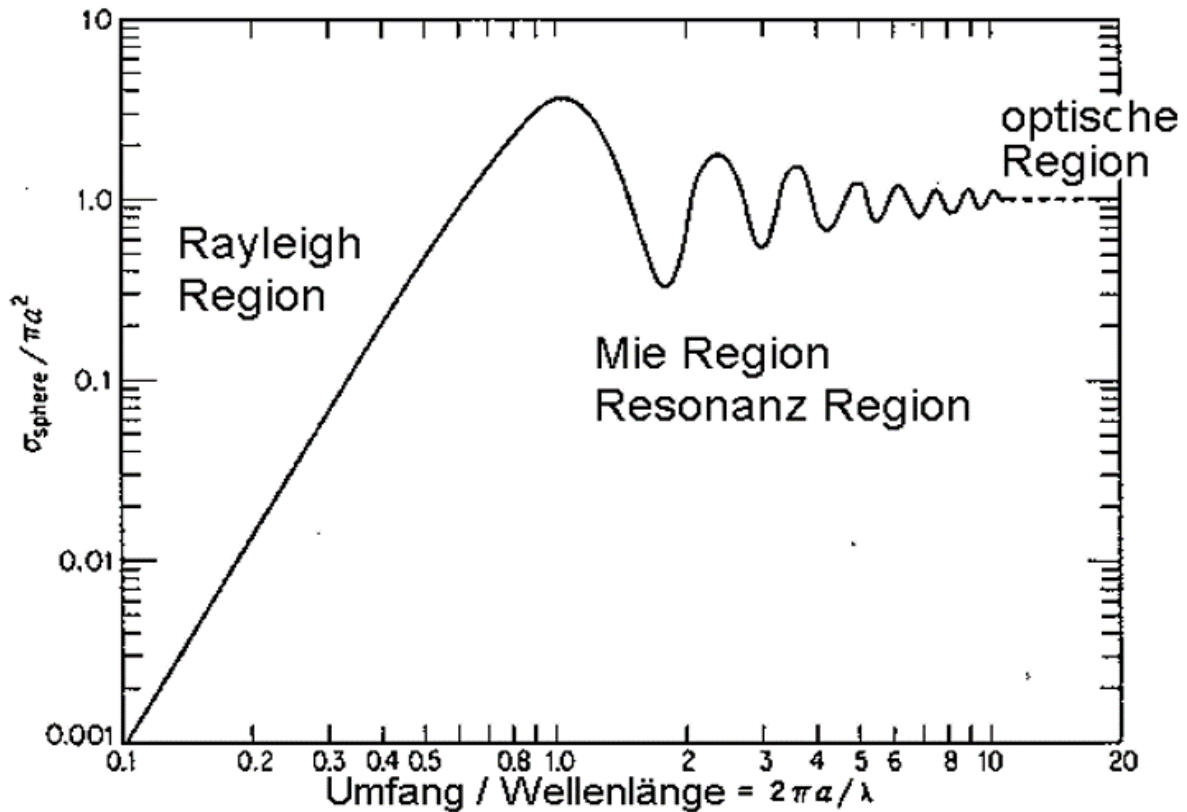


Abbildung 12: Bereich der Rayleighstreuung in Abhängigkeit der Wellenlänge λ und des Teilchendurchmessers (Martin Hagen, DLR, 2007)

Die allgemeine Radargleichung erweitert sich daher folgendermaßen:

$$P_r = \frac{P_t * G_t * G_r * \lambda^2}{(4\pi)^3 * R^4} * \sigma_{Rayleigh}$$

$$P_r = \frac{P_t * G_t * G_r * \lambda^2}{(4\pi)^3 * R^4} * V * \frac{\pi^5 * |K|^2 * D_t^6}{\lambda^4}$$

Diese kurze Einführung von Radargeräten und deren meteorologischen Nutzung, stellt nur einen Bruchteil der eigentlichen Thematik dar. Es gibt noch einiges an Störeinflüssen, welche außer Acht gelassen wurden. Als Beispiele seien hier genannt, dass sich Radarstrahlen nicht exakt parallel zur optischen Achse ausbreiten, unendlich frei im Raum emittiert werden (Freiraumdämpfung) und, dass es zu einer Krümmung durch die Erdoberfläche kommt.

2.3 Ansätze zur Analyse der Wetterradardaten

Für die Untersuchung, der Wirkung von Hagelabwehrflügen auf Gewitterzellen, muss eine geeignete Methodik entwickelt werden. Speziell für die Analyse von Wetterphänomenen mit Radardaten gibt es einige interessante Ansätze aus dem Bereich des maschinellen Lernens, wie zum Beispiel die Vorhersage von Tornados in den USA, mittels Wetterradardaten (Ellen Ferrante, 2011). Bei diesen Modellen fließen jedoch weitaus mehr meteorologische Daten mit ein, als bloße Radarbilder, wie sie hier zur Verfügung stehen. Aus diesem Grund soll eine Methodik entwickelt werden, welche sich nur auf die Daten des Wetterradars stützt.

2.3.1 Nachweis der Veränderung von Wolkenclustern

Bei der Zusammenfassung von einzelnen Radarechos, zu sogenannten Clustern, werden ab einem gewissen Abstand einzelne Echos zu einem Echocluster zusammengefasst. Ein Cluster entspricht hierbei einer großen zusammenhängenden Gewitterzelle. Anhand der zeitlichen Änderung ihrer Abstände und der Echostufen, sollen diese Cluster klassifiziert werden. Auf diesen beiden Merkmalen beruhend soll untersucht werden, ob sich geimpfte Wolken anders verhalten, als nicht geimpfte, in Bezug auf Abstand und Änderung der Echostufe.

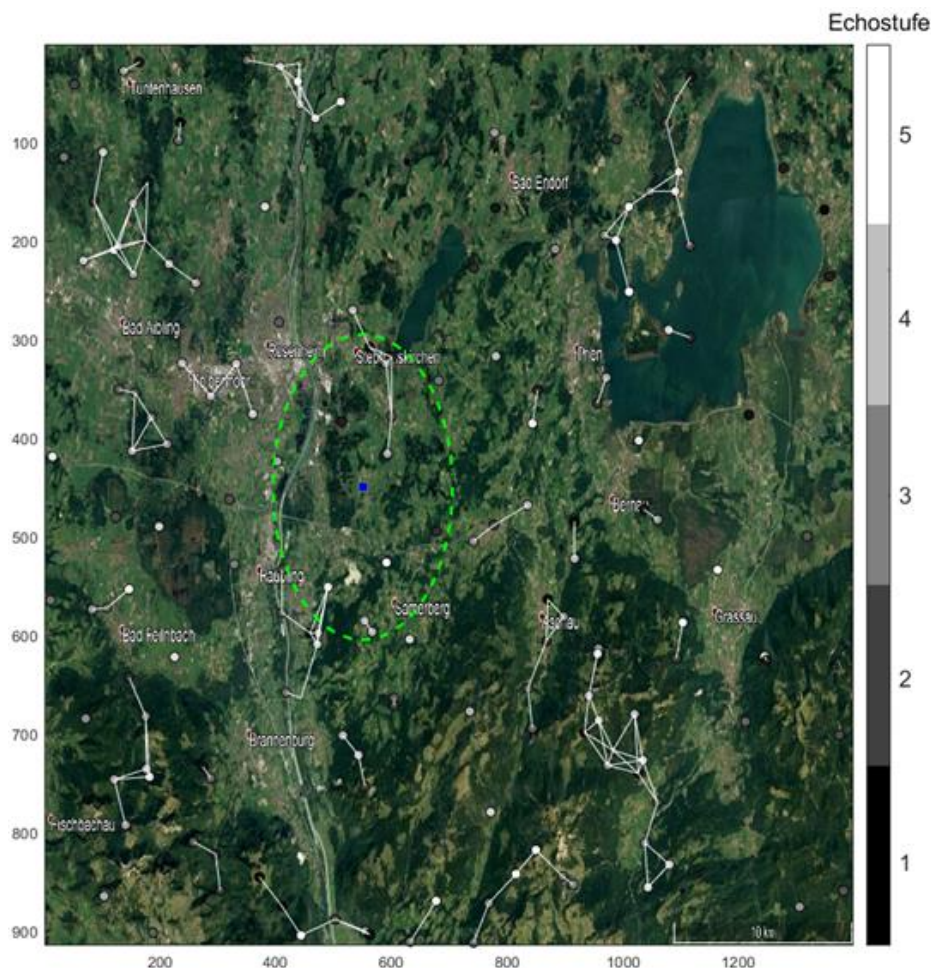


Abbildung 13: Clusterbildung (realisiert hier mit Zufallszahlen). Der Hagelflieger ist als blaues Quadrat zu sehen und der Wirkungsradius des Silberjodids als grüner Kreis.

Der Ansatz der Wolkencluster stellt sich, in Bezug auf die Verfolgbarkeit von einzelnen Radarechos, als äußerst schwierig heraus. Einzelne Radarechos unterschreiten ihren Schwellwert für die Detektierbarkeit und verschwinden somit vom Radar. Das Verfolgen von einzelnen Radarechos ist, aufgrund der relativ hohen Zeitabstände zwischen zwei Radarscans zusätzlich, äußerst schwierig. Insbesondere für Tage mit hohen Windgeschwindigkeiten, hat sich dies als schwer umsetzbar herausgestellt, da die Windrichtung und -stärke nicht bekannt sind und oftmals noch variieren. Aufgrund dieser Problematik wurde der Ansatz für den Nachweis der Wirksamkeit, von AgJ-Impfungen mittels Hagelflieger, nicht mehr weiterverfolgt. Theoretisch ließe sich mit diesem Ansatz jedoch die Bildung einer Superzelle, durch die sich verändernde Wolkenbewegung vorhersagen. Bei einer Superzelle bewegen sich die einzelnen Gewitterzellen (Radarechos) nicht mehr als einzelne kleinere Gewitterzellen, sondern als ein konvektives System, mit sehr ähnlichen Richtungsvektoren. Um eine Gewitterzelle als Superzelle zu klassifizieren zu können ist die Zugrichtung dieser nur eines von vielen Klassifizierungsmöglichkeiten (Smith, 2004).

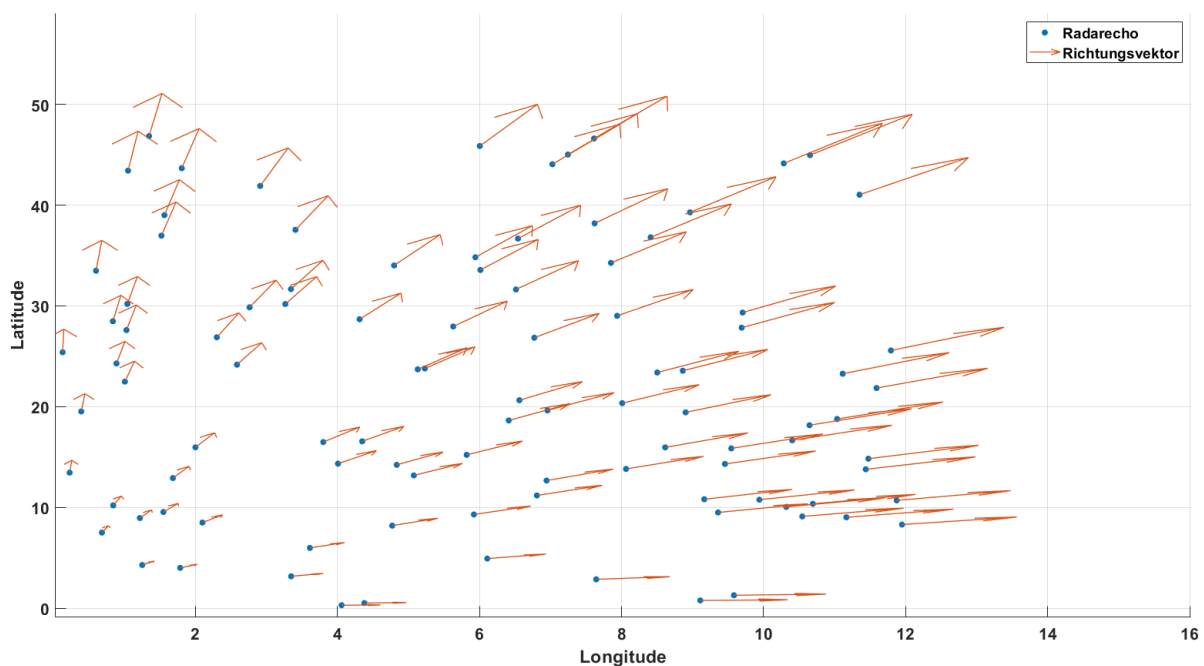


Abbildung 14: Darstellung einzelner Radarechos mit deren zugehörigen Richtungsvektoren

2.3.2 Erhebung von Rasterkennzahlen

Die Radarkarte, welche aus einer 200x200 Matrix besteht, wird in endlich kleine Raster eingeteilt, wie Abbildung 15 zeigt. Für jedes Raster wird eine Kennzahl (Rasterkennzahl) eingeführt, welche sich aus der Summe aller vorhandenen Echostufen zusammensetzt. Anhand der Flugdaten wird ein Raster als geimpft markiert, wenn der Hagelflieger hindurchgeflogen ist. Es kann im Anschluss untersucht werden, ob ein geimpftes Raster sich anders verhält, als ein nicht geimpftes. Die einzelnen Rasterkennzahlen RKZ können über die Menge der Echos einer bestimmten Stufe ME gebildet werden.

$$RKZ = ME1 * 1 + ME2 * 2 + ME3 * 3 \dots$$

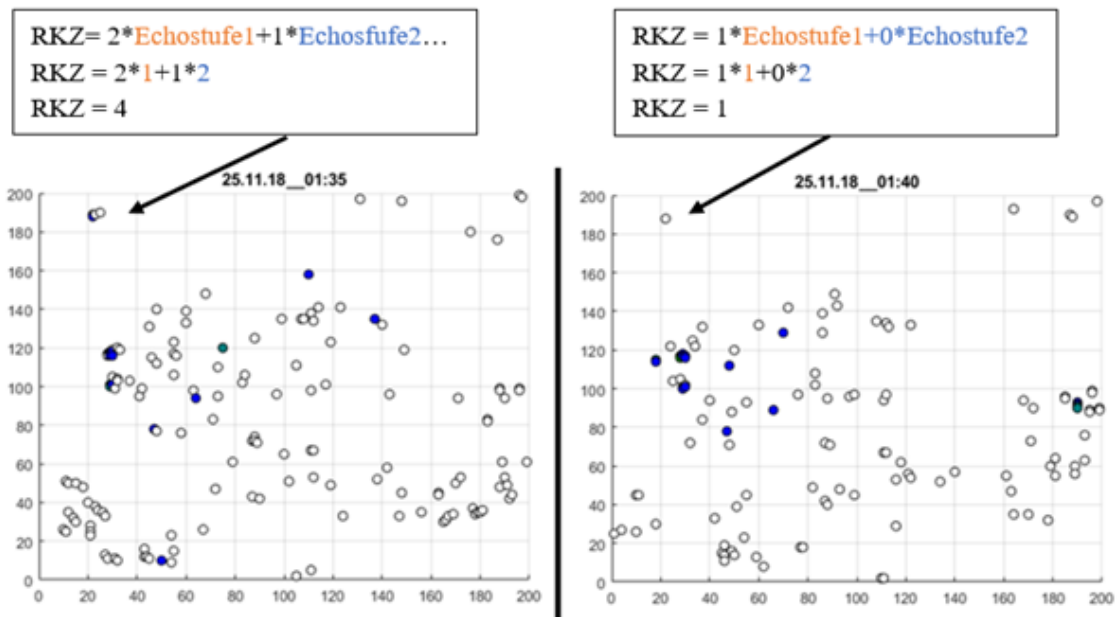


Abbildung 15: Beispielhafte Berechnung einer Rasterkennzahl

Der Vorteil bei der Erfassung von Rasterkennzahlen, gegenüber der Veränderung von Wolken-Clustern, ist, dass hier die einzelnen Radarechos nicht mehr verfolgt werden müssen. Es ist jedoch schwierig zu bestimmen, in welchen Größen die Raster festgelegt werden sollen. Für die dreidimensionalen Daten besteht zusätzlich das Problem, dass zwischen den einzelnen Schichten ein Abstand von 1000 m existiert.

2.3.3 Betrachtung einer lokalen Region um den Hagelabwehrflieger

Es existieren keine Daten zur Verteilung von AgJ, wenn dieses in Hagelabwehrflügen mithilfe von Generatoren ausgebracht wird. Am wahrscheinlichsten ist jedoch, dass sich das AgJ in Abhängigkeit von der momentanen Windrichtung und Flugrichtung verteilt. Es wurden Versuche unternommen die Ausbreitung von AgJ zu simulieren. Als erster Ansatz wurde versucht, die Ausbreitung von AgJ so zu simulieren, wie sich elektrisch geladene Teilchen im elektrischen Feld verhalten. Hierbei wirkt auf ein geladenes Teilchen eine Kraft, die zur Beschleunigung des Teilchens führt. Die Bahnkurve des Teilchens ergibt sich aus seiner Anfangsgeschwindigkeit und seiner Richtung. Um eine qualitative Aussage über die AgJ-Verteilung in den Wolken zu treffen, wird dieser Ansatz als zu einfach eingeschätzt. Für eine realistische Simulation, von der Verteilung von AgJ, wird eine CFD-Simulation empfohlen, welche auf Navier-Stokes-Gleichungen beruht (Zuo, 2019). Es wurde daher nur eine sehr simple Simulation durchgeführt, welche eine quantitative Aussage, über die Verteilung der Ausbreitung von AgJ in unmittelbarer Nähe des Flugzeugs, zulässt, wie Abbildung 16 zeigt.

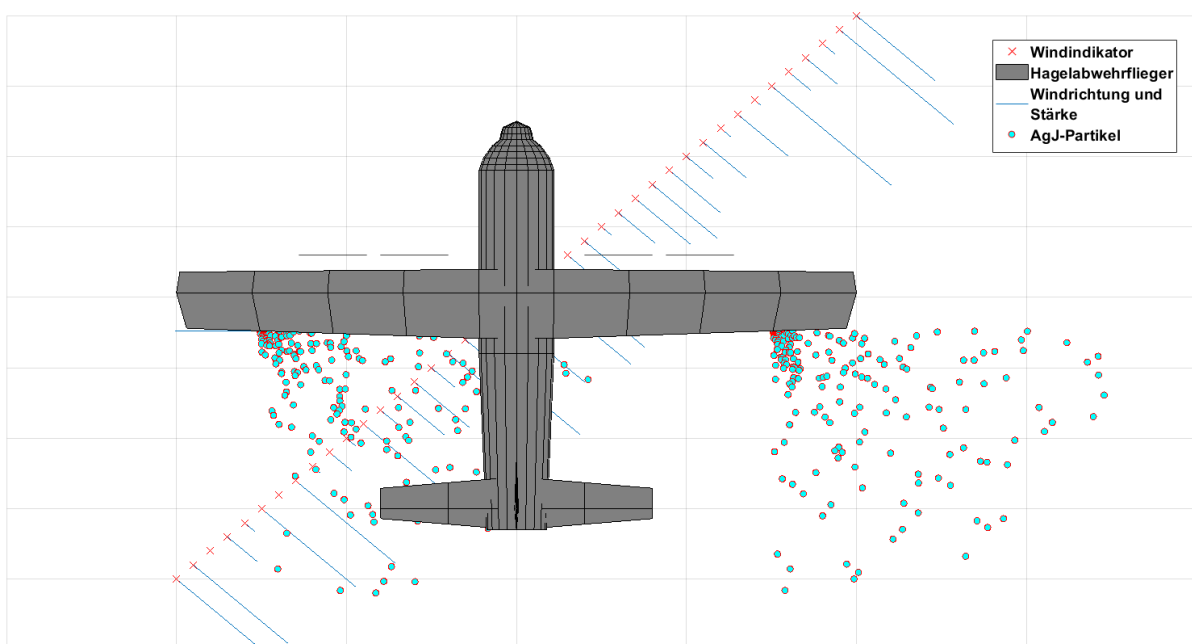


Abbildung 16: Quantitative Simulation der Ausbreitung von AgJ

Es kann eine Simulation von Partikeln (AgJ), mit unterschiedlichen Flugzeuggeschwindigkeiten, Windgeschwindigkeiten und Richtungen, im Programm *Particleflow-Simulation* erstellt werden (siehe Anhang A). Das Programm erlaubt eine quantitative Analyse der Verteilung, von als masselos angenommenen Partikeln, welche aus einem Flugzeug ausgebracht werden, in Abhängigkeit von den drei Einflussparametern Flugzeuggeschwindigkeit, Windgeschwindigkeit und Windrichtung. Die hier erstellte Simulation beruht auf stark vereinfachten physikalischen Tatsachen und bezieht keinerlei physikalische Effekte, wie die Reynoldszahl (Turbolenzverhalten der Strömung) oder die Düsenöffnung des AgJ-Generators, mit ein. Die Bewegung der Partikel beruht daher nur auf der Berechnung von Geschwindigkeitsvektoren. Abbildung 17 zeigt die Berechnung eines Partikels mit zwei Vektoren. Der blaue U-Vektor stellt die Geschwindigkeit eines ausgestoßenen Partikels dar, der grüne V-Vektor die Windgeschwindigkeit. Der Geschwindigkeitsvektor W ist die resultierende Richtung und Geschwindigkeit eines Partikels, durch die Abfälschung des ursprünglichen Geschwindigkeitsvektors U durch den Wind V.

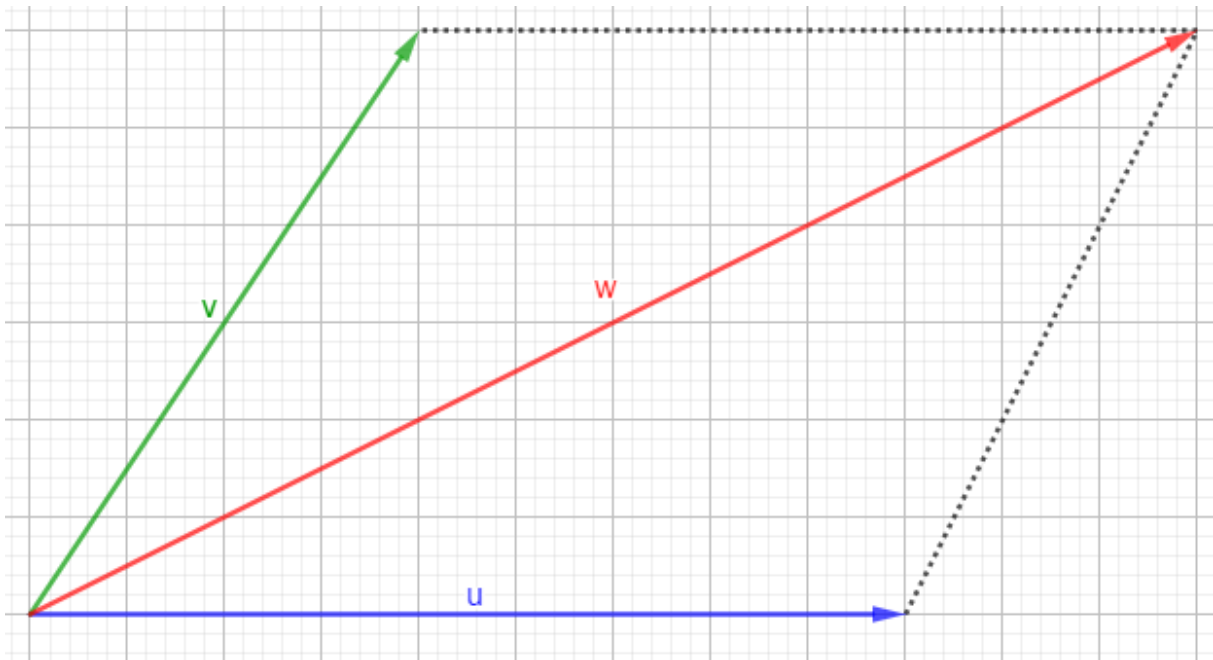


Abbildung 17: Berechnung der Partikelgeschwindigkeit

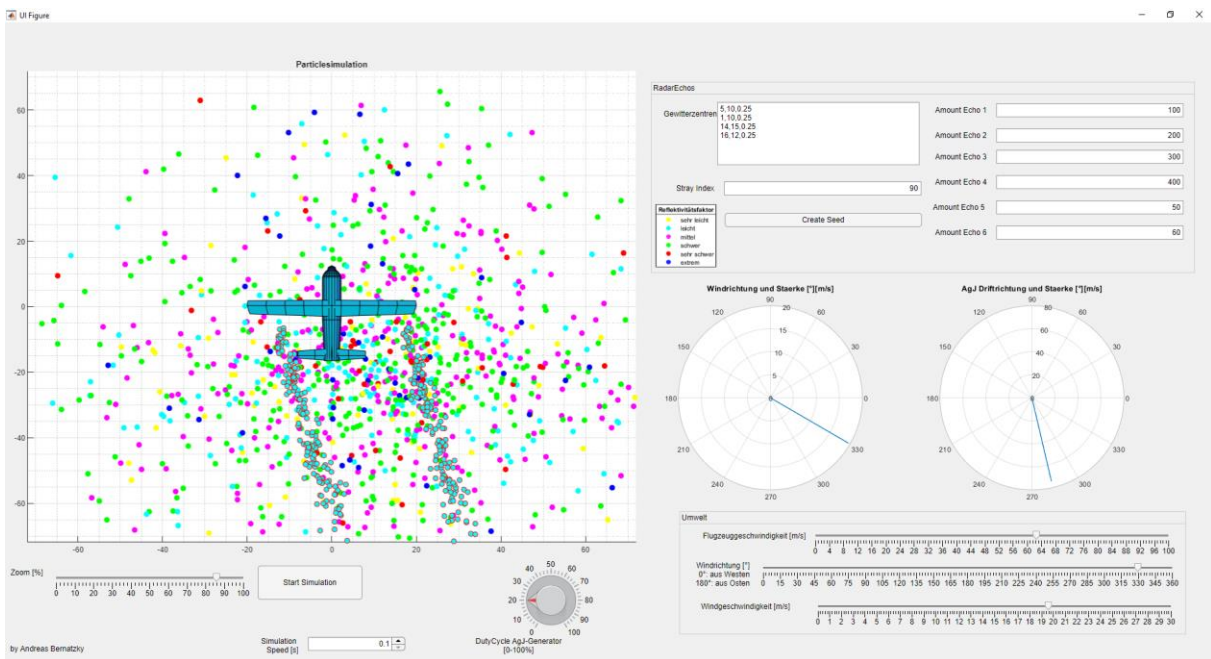


Abbildung 18: Simulation der AgJ-Ausbreitung in einer Gewitterzelle. Die AgJ-Partikel sind als türkise Punkte mit rotem Rand gekennzeichnet.

Da die Simulation derzeit nicht validiert werden kann, bleibt das Wissen, über das Ausbreitungsverhalten von Silberjodid, unbekannt. Daher wurde beschlossen, diese Unsicherheit, mithilfe eines möglichst großen Betrachtungsraum um das Flugzeug, abzuschwächen (siehe Abbildung 19).

Hierfür wurde eine sogenannte Impfgeometrie und eine Beobachtungsgeometrie eingeführt. Die Impfgeometrie kennzeichnet hierbei einen Bereich, von dem angenommen wird, dass alle Radarechos innerhalb der Impfgeometrie geimpft werden. Die Beobachtungsgeometrie stellt den Bereich dar, welcher lokal die Radarechos um das Flugzeug betrachtet. Die Geometrien werden für die 2D- und 3D-Daten erstellt.

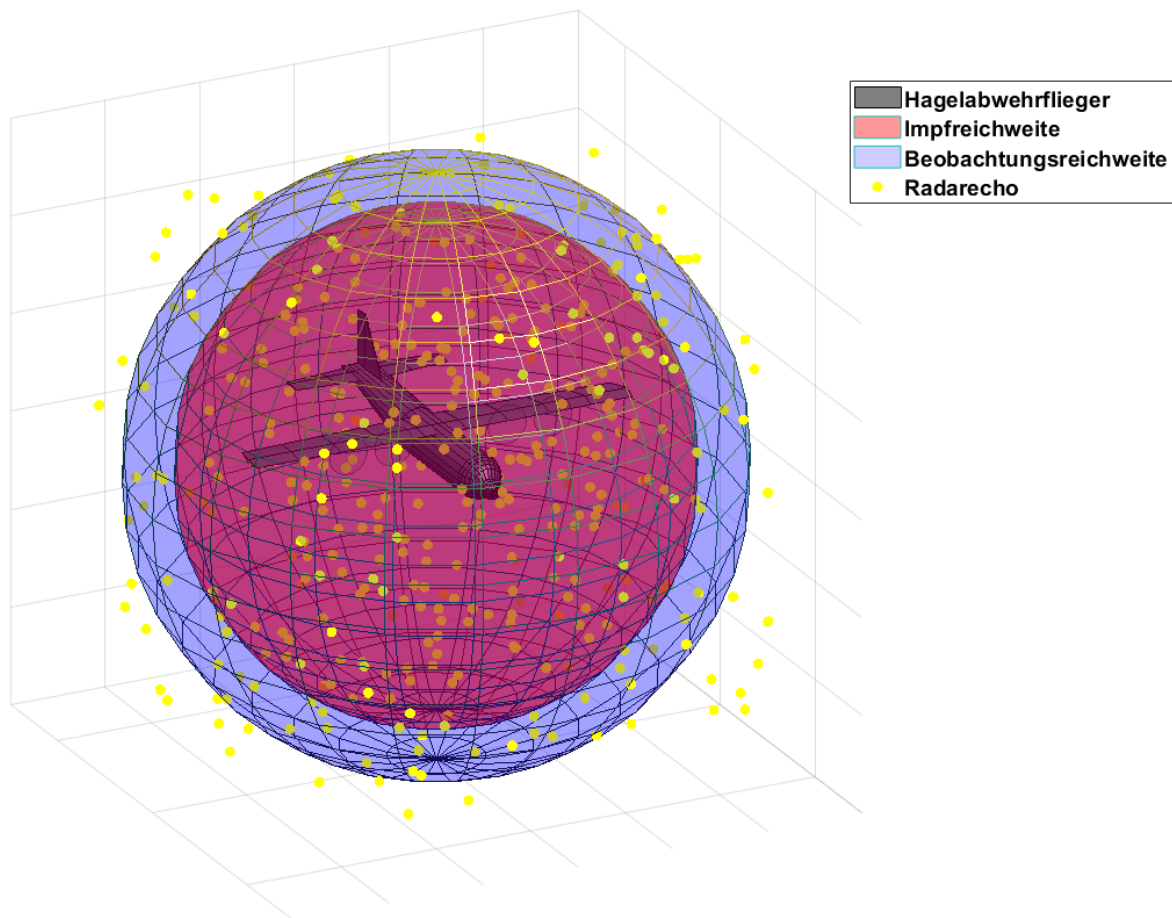


Abbildung 19: Ermittlung aller Radarechos, welche sich innerhalb der Impfkugel und der Beobachtungskugel befinden

2.3.4 Bewertung der Ansätze zur Analyse der Wetterradardaten

Die Entwicklung der Methodik, für die Auswertung der Radarechos, umfasste mehrere Stufen. Zuerst wurde versucht die Radarechos in Abhängigkeit von ihrem Abstand zu Gewitterzellenclustern zusammenzufassen. Dieser Ansatz scheiterte an der Tatsache, dass Wolken im Laufe der Zeit ihre Reflexivität ändern können oder sich ganz auflösen. Auch die Aktualisierungsrate der Radardaten, welche zwischen fünf und zehn Minuten liegt, wurde als zu groß eingeschätzt, um eine effektive Verfolgung von einzelnen Radarechos zu gewährleisten. Der Ansatz für die Erhebung von Rasterkennzahlen, wurde für die 2-D Daten für gut befunden. Für die 3-D Daten hätte sich anstelle der Unterteilung von Raster, die Unterteilung in Würfel angeboten. Da die maximale Rasterauflösung nur 200x200 betragen kann (Datenstruktur), wurde dies als zu gering auflösend erachtet, da der Hagelflieger, abhängig vom jeweiligen Flug, somit nur sehr wenige Raster durchflogen hätte. Ein durchflogenes Raster, wenn auch nur teilweise durchflogen, wäre somit als geimpft gewertet worden. Schlussendlich wurde der Ansatz mit den Impfgeometrien gewählt, da dieser zwei anpassbaren Bereiche um den Hagelflieger erfasst und gleichzeitig viele verschiedene Geometrien zulässt.

2.4 Erdkoordinatensystem

Um die Ansätze mit den Radarechos realisieren zu können, ist ein fundamentales Wissen im Bereich der Geodäsie nötig. Da die Erdoberfläche einer Kugel ähnelt, ist es notwendig diese, wie in Abbildung 20, in Längen- und Breitengrade einzuteilen, um hierauf Punkte identifizieren zu können.

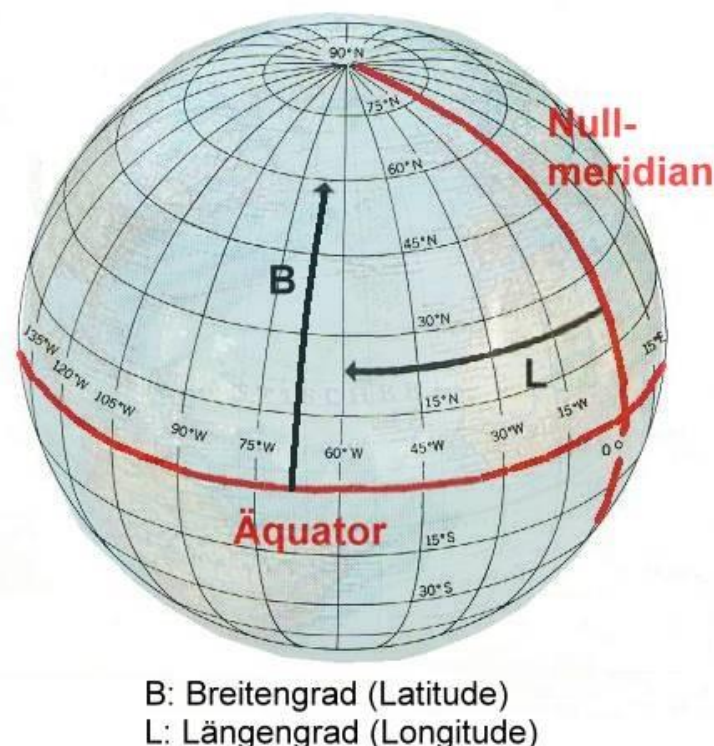


Abbildung 20: Weltkugel mit eingezeichneten Längen und Breitengraden (Liebau, kartenkunde-leichtgemacht.de, 2019)

Die Längengrade laufen in Östliche Richtung von 0 bis 180 ° O und in Westliche Richtung von 0° bis 180° W. Die geografische Länge bezeichnet den Winkel, der sich zwischen Erdmittelpunkt, einem gesuchten Punkt und dem Nullmeridian ergibt. Die Breitengrade (engl. Latitude) laufen von 0° (Äquator) bis zu 90° N (Nordpol) und 0° bis zu 90° S (Südpol). Der Abstand zwischen den Breitengraden beträgt pro Grad 111 km. Der Abstand zwischen zwei Längengraden (engl. Longitude) ist abhängig von der momentanen geographischen Breite. Für den Ansatz der Radarechos, war es

erforderlich, die Positionen der Radarechos und des Hagelabwehrfliegers korrekt abbilden zu können. Insbesondere ist es erforderlich, die Distanz zwischen zwei GPS-Koordinaten korrekt zu ermitteln. Die verschiedenen Verfahren sollen im Kapitel 2.5 vorgestellt werden (Kompf, 2019).

2.5 Distanzberechnung aus GPS-Daten

Die Algorithmen zur Längenberechnung basieren auf den Veröffentlichungen von Herrn Martin Kompf und Herrn Dr. Louis Strous. Diese Algorithmen wurden mit der Softwarebibliothek proj4 (proj4, 2019) validiert. Gebietsweise kommt es zu Abweichungen von wenigen Metern, da die Erde keine perfekte Kugel ist, sondern mit unterschiedlichen Ellipsoiden angenähert werden kann. In Mitteleuropa, und speziell in Deutschland, wird beispielsweise das Bessel-Ellipsoid verwendet. Auf Gebieten der ehemaligen Sowjetunion erfolgt eine Annäherung der Erdoberfläche mit dem Krassowski-Ellipsoid.

Es soll, mit drei verschiedenen Verfahren, die Distanz zwischen der TH-Rosenheim und dem Max-Josef-Platz (Rosenheim) und die Distanz zwischen der TH-Rosenheim und Kairo errechnet werden (siehe Abbildung 31).

Abbildung 3: Längen- und Breitengrade der zu vermessenden Punkte

Ort	Longitude	Latitude
TH-Rosenheim	12,1095889	47,8669244
Max-Josef-Platz	12,1273618	47,8558643
Kairo	31,24967	30,06263

2.5.1 Pythagoreische Distanzberechnung

Die simpelste Art der Entfernungsberechnung geschieht über den Satz des Pythagoras. Hierbei wird angenommen, dass Längen- und Breitengrade Rechtecke aufspannen. Dieser Ansatz sollte nur für sehr kleine Gebiete verwendet werden und stellt eine starke Vereinfachung dar.

Die Berechnungsformel setzt sich für die pythagoreische Distanzberechnung folgendermaßen zusammen:

$$\begin{aligned}\Delta X &= lon1 - lon2 \\ \Delta Y &= lat1 - lat2 \\ Distanz_{pythagor\ddot{a}isch} &= \sqrt{distLon * \Delta X^2 + 111,3 * \Delta Y^2}\end{aligned}$$

Die Konstante 111,3 km ist der Abstand zwischen zwei Breitengraden. Die Variable *distLon* bezeichnet die Distanz zwischen zwei Längengraden und ist abhängig von der geographischen Breite. Diese muss zuerst für jeden Breitengrad individuell ermittelt werden. Für Distanzberechnungen innerhalb kleiner Gebiete, z.B. Städten, ist dieser Ansatz absolut ausreichend.

2.5.2 Distanzberechnung mit abhängigem Längengrad

Der Abstand zwischen zwei Längengraden variiert in Abhängigkeit vom Breitenkreis. Am Äquator beträgt der Abstand zwischen zwei Längengraden 111,3 km und ist somit identisch zum Abstand zwischen zwei Breitenkreisen. Im Gegensatz zur pythagoreischen Distanzberechnung (siehe Kapitel 2.5.1), kann nun nicht mehr von Rechtecken ausgegangen werden, wie auch Abbildung 21 zeigt.

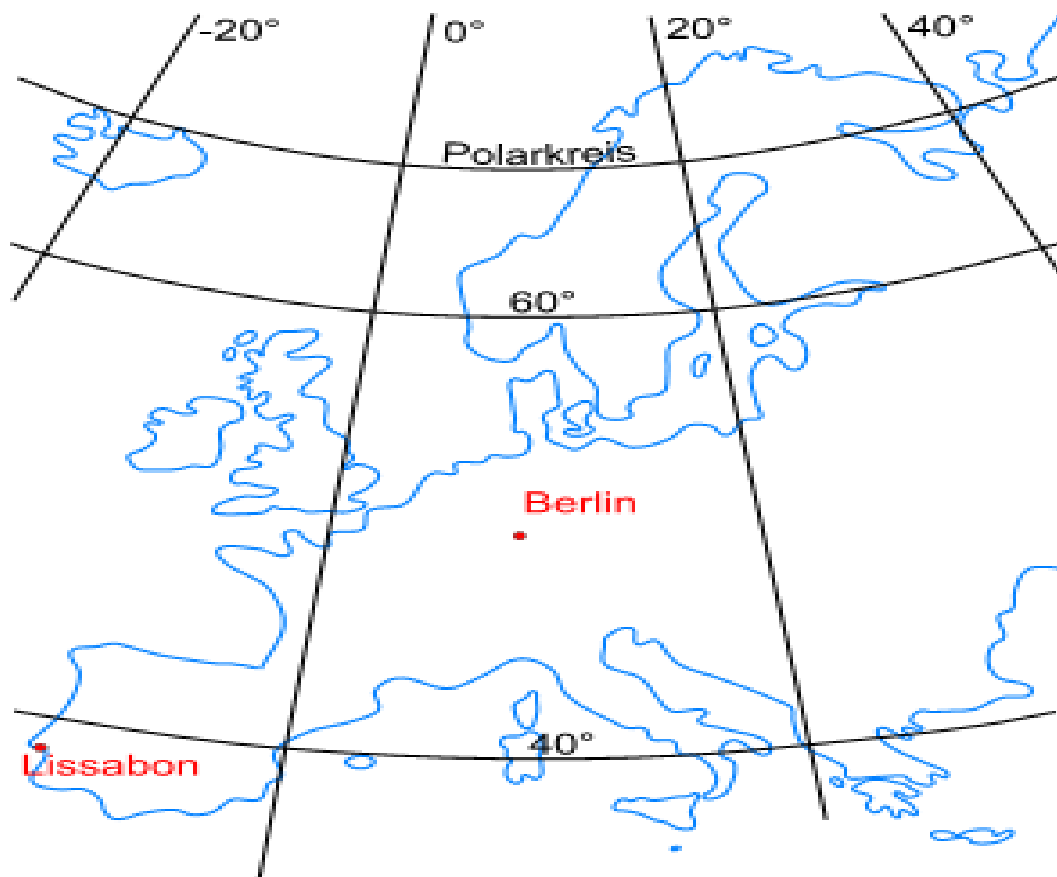


Abbildung 21: Mit zunehmendem Breitengrad steigt die Krümmung dieser. Es werden keine Rechtecke mehr zwischen Längen- und Breitengrade aufgespannt (Kompf, 2019)

Der Abstand zwischen zwei GPS-Koordinaten, in Abhängigkeit zwischen zwei Längengraden, errechnete sich nachfolgender Formel:

$$lat_{Med} = \frac{lat1 + lat2}{2} * \frac{\pi}{180}$$

$$\Delta X = 111,3 * \cos(lat_{Med}) * (lon1 - lon2)$$

$$\Delta Y = 113,3 * (lat1 - lat2)$$

$$Distanz_{abhängig} = \sqrt{\Delta X^2 + \Delta Y^2}$$

2.5.3 Distanzberechnung über Orthodrome (Großkreisbogen)

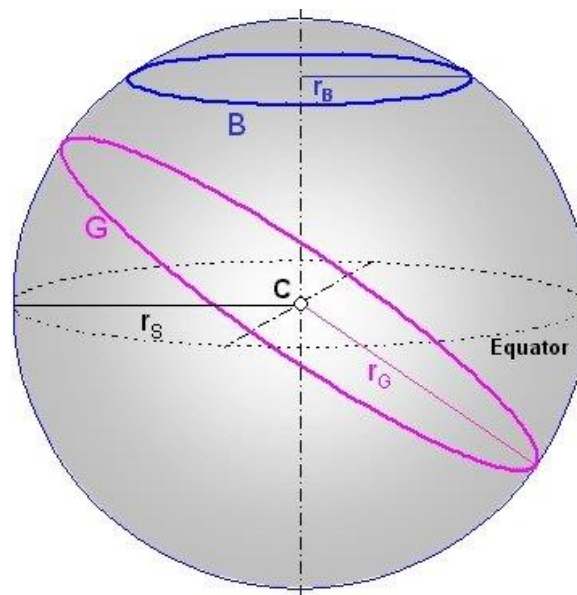


Abbildung 22: Eine Orthodrome entspricht immer dem Kugelradius (Frassek, 2019)

Die kürzeste Verbindung zwischen zwei Punkten auf einer Ebene lässt sich durch eine Gerade darstellen. Auf einer Kugeloberfläche ist die kürzeste Verbindung zwischen zwei Punkten ein Teilstück einer Orthodromen (griech. Orthos für „gerade“, dromos für „Lauf“). Eine Orthodrome wird umgangssprachlich auch als „Luftlinie“ zwischen zwei Punkten verstanden.

Abbildung 22 zeigt solch eine Orthodrome G, auf einer Kugel mit dem Radius r_s . Sein Mittelpunkt liegt immer im Kugelmittelpunkt C, der Radius der Orthodromen r_G entspricht immer dem Kugelradius r_s . Der Kreis B mit dem Radius r_B kennzeichnet einen umlaufenden Breitengrad.

Abbildung 22 zeigt die Erde als Kugel, nun aber mit den beiden eingezeichneten Punkten P1 und P2, deren Abstand bestimmt werden soll. Diese (kürzeste) Entfernung wird auch als Großkreisbogen bezeichnet. Verbindet man drei, nicht alle auf einem Großkreis liegende, Punkte der Kugeloberfläche mit drei Großkreisbögen, so erhält man acht Kugeldreiecke. Die Punkte P1 und P2 bilden zusammen mit dem Nordpol ein Kugeldreieck. Zwei Seiten dieses Dreiecks und der von ihnen umschlossene Winkel sind bekannt: Die Seitenlänge entspricht der Entfernung des Punktes vom Nordpol, d.h. 90 Grad minus seiner geographischen Breite. Der eingeschlossene Winkel wird aus der Differenz zwischen der geographischen Länge der beiden Punkte berechnet.

Unter Zuhilfenahme des Seitenkosinus-Satzes lässt sich, für den gesuchten Großkreisbogen g und den beiden geographischen Punkten P1 und P2 aus Abbildung 23, folgende Formel aufstellen:

$$\cos(g) = \cos(90^\circ - \text{lat}_1) * \cos(90^\circ - \text{lat}_2) + \sin(90^\circ - \text{lat}_1) * \sin(90^\circ - \text{lat}_2) * \cos(\text{lon}_2 - \text{lon}_1)$$

Da $\cos(90^\circ - a) = \sin(a)$ und $\sin(90^\circ - a) = \cos(a)$ gilt, vereinfacht sich obige Formel zu:

$$\cos(g) = \sin(\text{lat}_1) * \sin(\text{lat}_2) + \cos(\text{lat}_1) * \cos(\text{lat}_2) * \cos(\text{lon}_2 - \text{lon}_1)$$

Um die Entfernung in Kilometern zu erhalten, muss der Arkuskosinus gebildet werden und das Ergebnis mit dem Erdradius multipliziert werden.

$$g = r_{\text{Erde}} * \text{acos}(\sin(\text{lat}_1) * \sin(\text{lat}_2) + \cos(\text{lat}_1) * \cos(\text{lat}_2) * \cos(\text{lon}_2 - \text{lon}_1))$$

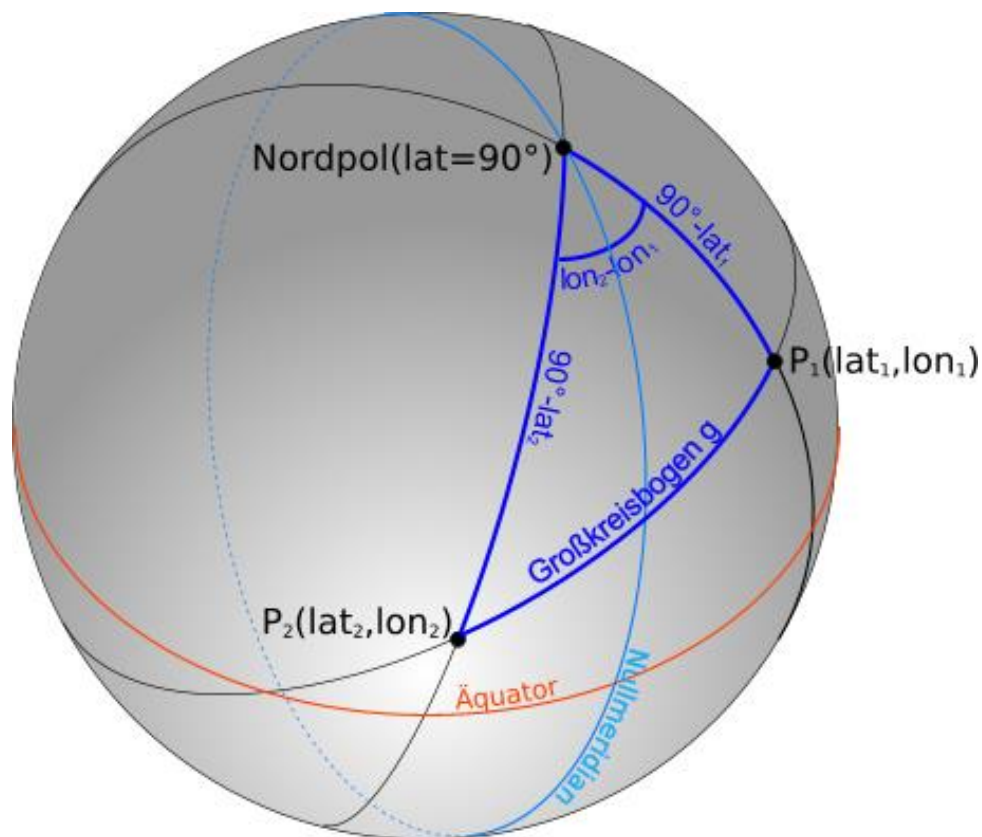


Abbildung 23: Berechnung des Großkreisbogens zwischen zwei Punkten auf einer Kugel (Kopf, 2019)

2.5.4 Distanzberechnung über den Referenzellipsoid WGS84

In Kapitel 2.5.3 wurde ein Verfahren vorgestellt, um die Distanz zwischen zwei Punkten auf einer perfekten Kugel zu berechnen. In der Realität gleicht die Erde aber keiner Kugel, sondern einem Ellipsoid, welcher an den Polen abgeflacht ist. Die Abweichung von der perfekten Kugelform ist durch die Erdrotation bedingt. Um die Geometrie der Erde dennoch global beschreiben zu können, existieren für globale Näherungen Referenzellipsoide und für lokale Näherungen Geoide. Geoide stellen die Fläche der Erde auf Meeresebene dar. Das Referenzellipsoid ist gegenüber dem Geoid global gültig und bezieht die Topographie nicht mit ein. Referenzellipsoiden können, der Vorstellung halber, als genäherte Ausgleichsellipsoid durch eine Punktwolke betrachtet werden.

Speziell der *WGS84* Referenzellipsoid wurde durch Beschluss der ICAO 1989 für die Luftfahrt übernommen. Dieser ist ein weltweit einheitliches Referenzsystem, welches für die Positionsangaben auf der Erde und im erdnahen Weltraum genutzt wird (Eurocontrol & IfEN, 1998). Daher soll dieser, in Anbetracht auf die Verwendung der Hagelfliegerposition, speziell behandelt werden.

Für den *WGS84* werden folgende Konstanten benötigt:

$$\text{Abplattung der Erde: } f = \frac{1}{298,257223563}$$

$$\text{Äquatorradius der Erde: } a = 6378,137 \text{ km}$$

Um mit anderen Referenzellipsoiden zu arbeiten, sind der Äquatorradius und die Abplattung der Erde speziellen Tabellen für die einzelnen Referenzellipsoide zu entnehmen. Weitere Ellipsoide wären, speziell für Eurasien, der *Bessel 1841* und für Asien beispielsweise der *Krassowski* oder *Clarke* Ellipsoid.

$$F = \frac{\text{lat1} + \text{lat2}}{2}, G = \frac{\text{lat1} - \text{lat2}}{2}, l = \frac{\text{lon1} - \text{lon2}}{2}$$

Ermittlung des groben Abstandes D:

$$S = \sin(G)^2 * \cos(l)^2 + \cos(F)^2 * \sin(l)^2$$

$$C = \cos(G)^2 * \cos(l)^2 + \sin(F)^2 * \sin(l)^2$$

$$w = \text{atan}\left(\sqrt{\frac{S}{C}}\right), D = 2 * w * a$$

Ermittlung der beiden Korrekturfaktoren H_1 und H_2 :

$$T = \frac{\sqrt{S * C}}{w}$$

$$H_1 = \frac{3 * T - 1}{2 * C} \text{ und } H_2 = \frac{3 * T + 1}{2 * S}$$

$$\text{Distanz}_{WGS84} = D * (1 + f * H_1 * \sin(F)^2 * \cos(G)^2 - f * H_2 * \cos(F)^2 * \sin(G)^2)$$

2.5.5 Vergleich der einzelnen Verfahren

Nachfolgende Tabelle 4 zeigt die Ergebnisse der Entfernungsberechnung, zwischen den in Kapitel 2.5 erwähnten Punkten. Um die Daten zu validieren wurde openstreetmaps.org verwendet.

Tabelle 4: Ergebnis der einzelnen Distanzberechnungsverfahren

Verfahren	Distanz: TH-Rosenheim-Max-Josef-Platz [km]	Distanz: TH-Rosenheim- Kairo[km]
<i>Distanz_{pythagoräisch}</i>	1,7911	2426,8831
<i>Distanz_{abhängig}</i>	1,8103	2582,8170
<i>Distanz_{Orthodrome}</i>	1,8106	2568,5117
<i>Distanz_{WGS84}</i>	1,8113	2565,6247
<i>Haversine (Validierung)</i>	1,8102	2566,3421

Wie aus Tabelle 4 hervorgeht, bestehen zwischen den einzelnen Verfahren Unterschiede. Die herangezogene Validierung von openstreetmaps.org basiert auf der Haversine-Formel, welches ein Verfahren zur Berechnung des Großkreisbogens (Orthodrome) auf einer Kugel darstellt. Es wird auch hier von einer perfekten Kugel ausgegangen. Die Validierung eignet sich daher nicht für eine qualitative Aussage um die einzelnen Verfahren beurteilen zu können, sondern soll lediglich ein unabhängiges Verfahren darstellen, um die eigenen Berechnungen, bezüglich der Größenordnung, zu validieren. Es soll noch die Vincenty-Formel erwähnt werden, welche aufgrund ihrer Komplexität nicht implementiert wurde, da sie als zu rechenaufwändig erachtet wird. Die Vincenty-Formel geht von einem abgeflachten Rotationsellipsoid aus und berechnet die Distanz zwischen zwei Punkten, mittels eines iterativen konvergierenden Verfahrens (Movable, 2019). Es wird sich daher schlussendlich für die Berechnung auf Basis des WGS84-Ellipsoids entschieden, da dieser einen weltweiten Standard abdeckt. Wenn eine möglichst exakte Positionsbestimmung notwendig ist, wird empfohlen, anstelle von Referenzellipsoiden auf Geoiden auszuweichen. Der Vorteil bei Geoiden besteht insbesondere darin, dass hier eine topographische Abbildung der Höhe mitberücksichtigt wird. Speziell für Deutschland werden verschieden Geoiden für verschiedene Anwendungsfälle verwendet. Für die Positionsbestimmung können hier der *ETRS89* und der *DREF91* verwendet werden (Bundesamt für Kartographie und Geodäsie, 2019). Die Skripte zur Distanzberechnung zwischen Geokoordinaten sind im Anhang A zu finden.

2.5.6 Mercator Transformation

Mit dem Aufkommen der Seeschifffahrt im 16. Jahrhundert wurden exakte und lückenlose Karten immer wichtiger. Die Mercator-Projektion ist eine winkeltreue Kartenprojektion, die besonders zur Navigation in der Schifffahrt eingesetzt wird. Die konforme (=winkeltreue) Mercator-Projektion wurde in Europa zuerst von Etzlaub (1511) und Mercator (1569) beschrieben und angewendet. Der Geoinformatik-Service der Universität Rostock (Uni-Rostock, 2019) und die Internetseite von Herrn Liebau bieten hier eine gute Einführung (Liebau, kartenkunde-leichtgemacht, 2019).

Um die Hagelabwehrflüge verzerrungsfrei 2-dimensional abbilden zu können, ist es erforderlich, diese in ein kartesisches-Koordinatensystem zu überführen. Für eine bildliche Vorstellung dieser Projektion, kann sich die Erde wieder vereinfacht als perfekte Kugel vorgestellt werden. Um diese wird ein Zylinder gelegt (Zylinderprojektion), der die Kugel am Äquator berührt. Vom Projektionszentrum im Erdmittelpunkt werden nun alle Punkte vom Globus auf den Zylinder übertragen. Rollt man danach den Zylinder ab, so erhält man eine winkeltreue, plane Karte der Erdoberfläche (Liebau, kartenkunde-leichtgemacht, 2019).

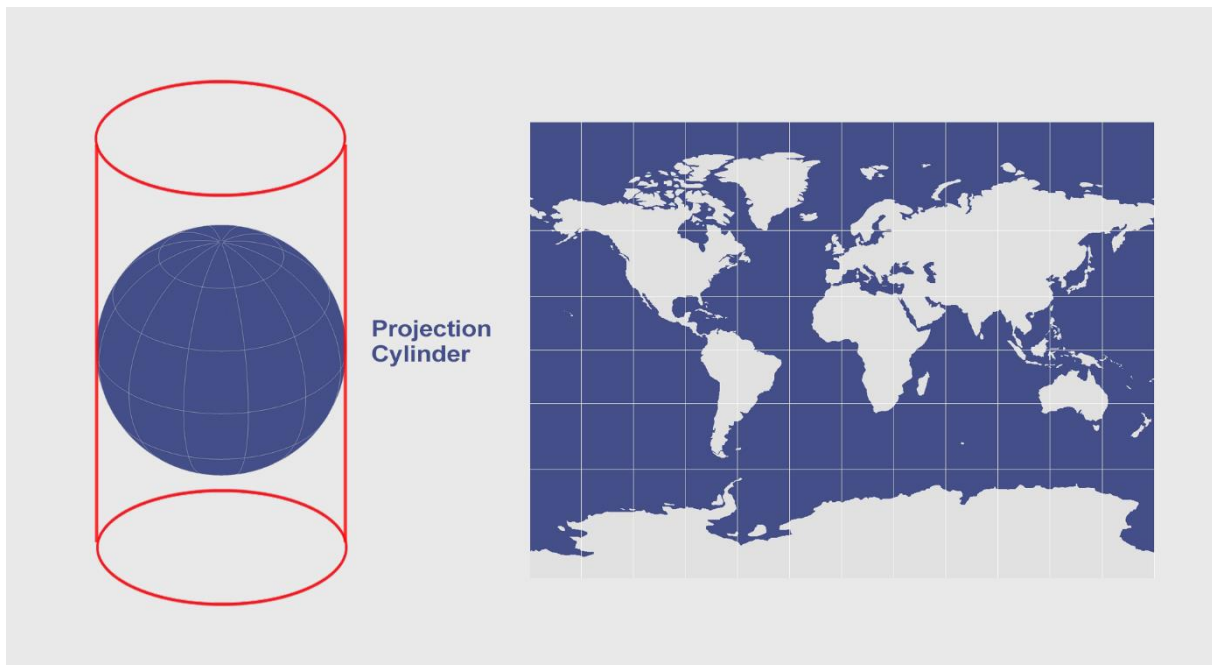


Abbildung 24: Bildhafte Darstellung der Mercatortransformation. Jeder Punkt auf der Weltkugel wird auf einen Zylinder projiziert, welcher anschließend abgerollt wird. (GIS Geography, 2019)

Da der Äquator die Berührungslinie von Zylinder und Globus ist, wird dieser längentreu abgebildet. Zum Nord- und Südpol hin werden die Verzerrungen immer größer. Dadurch ist die Insel Grönland (2,2 Mio. km²) in dieser Kartenprojektion fast so groß dargestellt wie der Kontinent Afrika (30,3 Mio. km²). Die Mercator-Projektion ist also nicht flächentreu. Der Nord- und der Südpol können nicht dargestellt werden, da der projizierte Punkt im Unendlichen liegen würde (Liebau, kartenkunde-leichtgemacht, 2019).

Der große Vorteil der Mercator-Projektion ist die Möglichkeit, Kurslinien als gerade Linien auch über weite Distanzen darzustellen. Die Kurslinie schneidet die Meridiane immer im gleichen Winkel. Die Karte ist daher winkeltreu, aber nicht flächentreu. Gerade wegen dieser Eigenschaft revolutionierte die Mercator-Karte die Navigation und markierte den Wandel der Kartographie von der Kunst zur mathematisch genauen Wissenschaft, der Kartographie. Die meisten Seekarten sind auch heute noch Mercator-Karten (Liebau, kartenkunde-leichtgemacht, 2019).

Die Überführung in kartesische Koordinaten erfolgt folgendermaßen:

$$\begin{aligned}x &= FE + R * (\lambda - \lambda_0) \\y &= FN + R * \ln\left(\tan\left(\frac{\pi}{4} + \frac{\varphi}{2}\right)\right)\end{aligned}$$

λ_0 stellt hierbei der Längengrad im Kartenzentrum dar (Längen-Null). FN und FE sind False Northing und False Easting. False Easting und False Northing sind hierbei Offsets in X- bzw. Y-Richtung, gemessen vom Nullmeridian aus. Dies wird verwendet um negative x- und y-Werte für die Mercatortransformierte Karte zu vermeiden. R ist der Radius der Erdkugel in Meter. Die hier gezeigte Formel sollte nicht für eine Latitude von $\varphi > 88^\circ$ genutzt werden.

Um Mercatortransformierte Koordinaten wieder in GPS-Koordinaten überführen zu können, wird die Gudermannfunktion verwendet:

$$D = \frac{FN - X}{R}$$
$$\lambda = \left(y - \frac{FE}{R}\right) + \lambda_0 \quad (12)$$
$$\varphi = \frac{\pi}{2} - 2 * \operatorname{atan}(e^D) \quad (13)$$

Neben der Mercatortransformation existieren noch etliche weitere Kartentransformationen, welche Kartenausschnitte in einer Ebene darstellen können. Die Mercatortransformation liegt als Matlabskript und als C++ Programm im Anhang A vor.

2.6 Interpolation der Radarkarte

Das Radargerät löst den in Abbildung 4 gezeigten Bereich in Form einer 200x200 Matrix auf. Von dem abgedeckten Radarbereich sind nur die vier Eckpunkte bekannt. Um die Flüge, zusammen mit den Radarechos, darstellen zu können, muss von diesen Eckpunkten aus die Radarkarte interpoliert werden. Es muss hier eine 2-dimensionale Interpolation, sowohl für die geographische Länge, als auch für die geographische Breite erfolgen. Im Folgenden sollen Interpolationen zunächst an einem einfachen 1D-Interpolationsverfahren und im Anschluss das benötigte 2D-Interpolationsverfahren erklärt werden.

2.6.1 1D-Interpolation (Linear)

Die 1-dimensionale Interpolation ist speziell für das vervollständigen zwischen zwei gegebenen diskreten Datenpunkten geeignet. Für die 1D-Interpolation muss zwischen einem Startpunkt $P_0(x_0, y_0)$ und einem Endpunkt $P_1(x_1, y_1)$ eine Sekante berechnet werden. Für diese ist die Steigung m und der Ordinatenwert t durch zwei gegebene Punkt zu berechnen.

$$m = \frac{y_1 - y_0}{x_1 - x_0}$$
$$t = y_0 - m * x_0$$

Der interpolierte Punkt P_{interp} , welcher auf der Sekante liegt und sich zwischen x_0 und x_1 befindet, errechnet sich folgendermaßen:

$$x_{interp} = \frac{x_1 - x_0}{2} + x_0$$
$$y_{interp} = m * x_{interp} + t$$

Das Skript zur 1D-Interpolation findet sich in Anhang A.

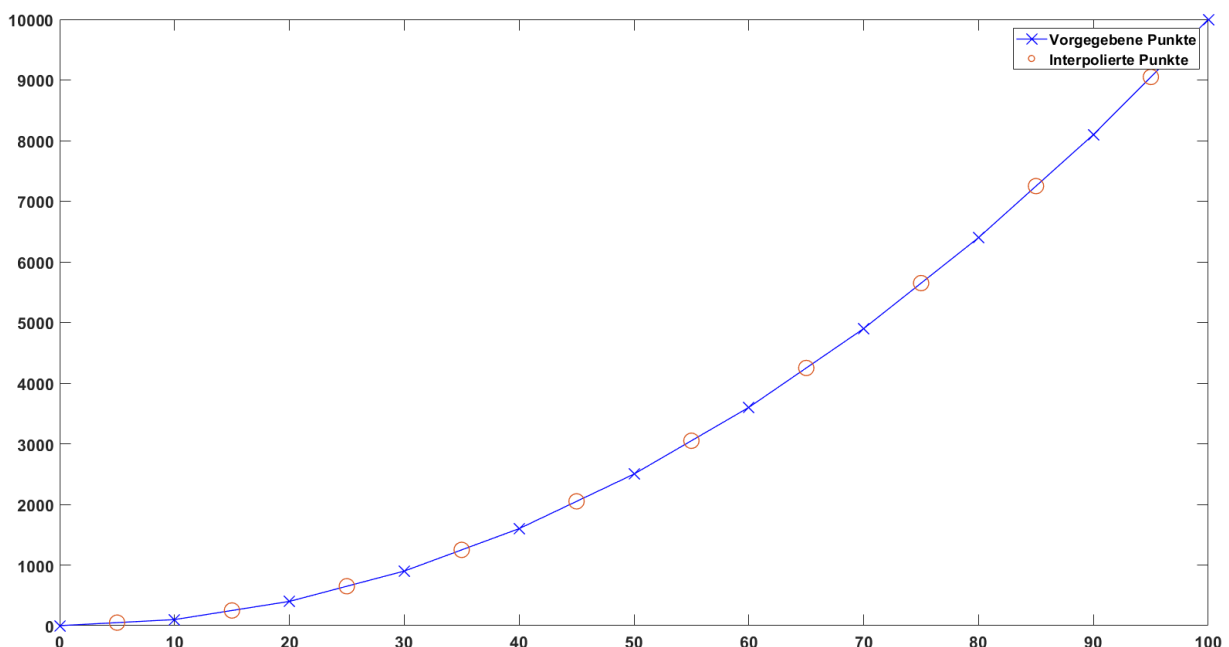


Abbildung 25: Ergebnis der 1D-Interpolation

2.6.2 Bilineare Interpolation

Die bilineare Interpolation stellt eine Erweiterung der linearen Interpolation dar, um Werte innerhalb einer 2D-Matrix zu berechnen. 2D-Interpolationen sind gängige Anwendungen in der Bildverarbeitung um Bilder, z.B. auf eine höhere Auflösung, zu skalieren. Hier existieren noch weitere Algorithmen, wie die Bikubische Interpolation, welche eine Interpolation mit kubischen Splines ausführt, oder die Nearest Neighbor Interpolation, wo jeder hinzu gewonnene Punkt den Wert des Nachbarpunktes annimmt (Matuschek, 2019). Für den Anwendungsfall der zu interpolierenden Radarkarte, wurde sich für den bilinearen Interpolationsalgorithmus entschieden, welcher in MATLAB programmiert wurde. Hierbei wird zuerst eine gewichtete 1D-Interpolation in X-Richtung ausgeführt und anschließend eine gewichtete 1D-Interpolation in Y-Richtung. Die vom Deutschen Wetterdienst zur Verfügung gestellten Radardaten lösen den erfassten Bereich der Px und Pz Daten (siehe Kapitel 2.1.2) auf eine 200x200 Matrix auf, von welcher nur die vier äußersten Eckpunkte bekannt sind. Die bilineare Interpolation soll an dem Beispiel in Abbildung 26 gezeigt werden.

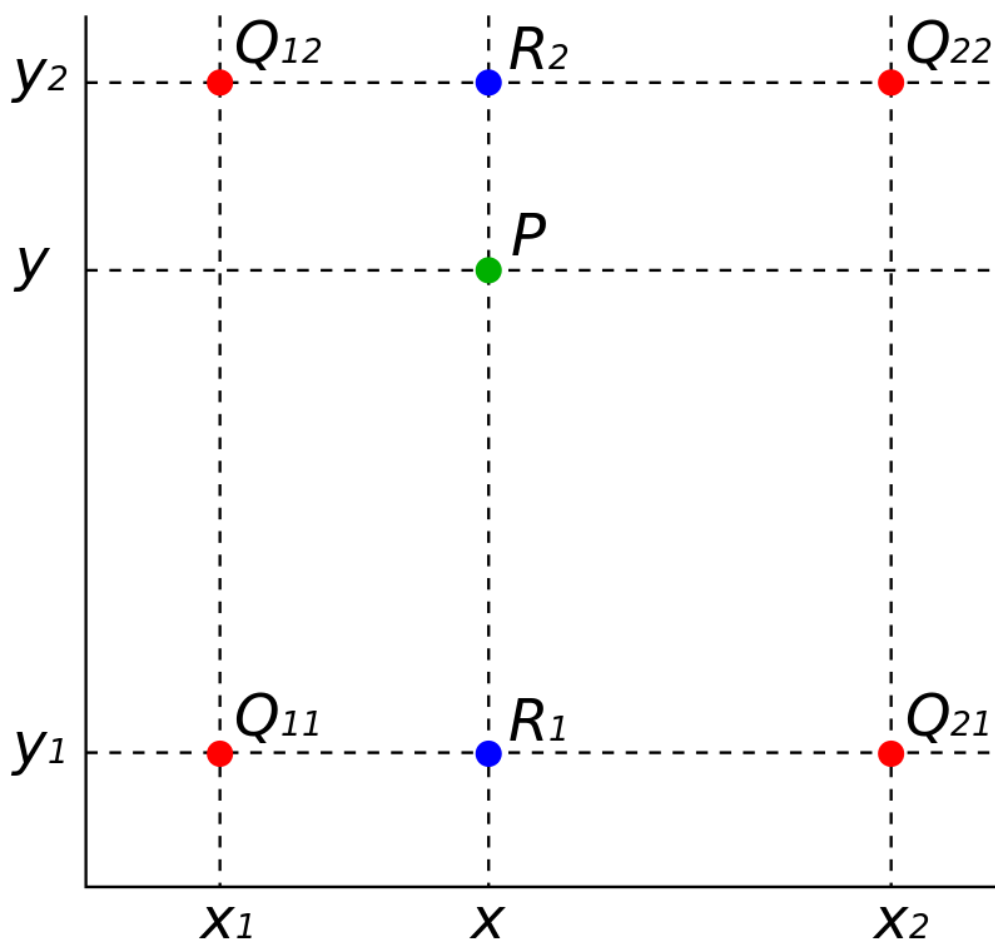


Abbildung 26: Bilineare Interpolation des Punktes P (Wikipedia, 2019)

Es sind die Punkte Q_{11} , Q_{12} , Q_{22} und Q_{21} gegeben. Es soll der Punkt P bestimmt werden. Hierzu wird eine gewichtete 1D-Interpolation in x-Richtung zwischen Q_{11} und Q_{21} ausgeführt um den Punkt R_1 zu bestimmen. Für R_2 wird eine 1D-Interpolation in x-Richtung zwischen Q_{12} und Q_{22} durchgeführt. Dabei wird zuerst die Gewichtung w zwischen den beiden Punkten berechnet.

$$w = \frac{(x_2 - x)}{x_2 - x_1}$$

Die Punkte R_1 und R_2 errechnen sich hierbei folgendermaßen:

$$R = w * x_1 + (1 - w) * x_2$$

Nachdem eine 1D-Interpolation in x-Richtung ausgeführt wurde kann nach gleichem Schema in die y-Richtung interpoliert werden.

$$w = \frac{(y_2 - y)}{y_2 - y_1}$$

$$P = w * y_1 + (1 - y) * y_2$$

Ein beispielhaftes Ergebnis, der bilinearen Interpolation, für die folgende 200x200 Matrix M , mit vorgegebenen Eckpunkten, ist in Abbildung 27 zu sehen.

$$M = \begin{matrix} & 1 & \dots & 2 \\ & \vdots & \ddots & \vdots \\ & 3 & \dots & 4 \end{matrix}$$

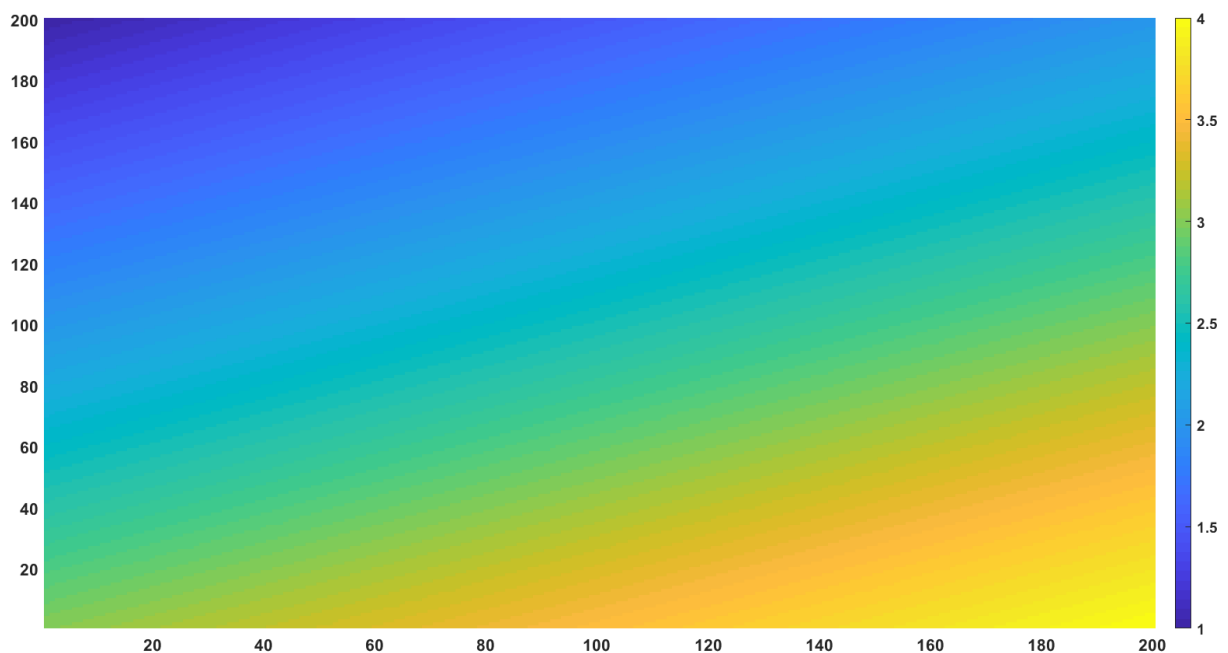


Abbildung 27: Ergebnis der bilinearen Interpolation für die Matrix M

Das Skript zur bilinearen Interpolation wurde in MATLAB geschrieben und ist im Anhang A zu finden.

2.7 Hailshow

Das Programm *Hailshow* ist eine, unter der MATLAB Version R2019a, entwickelte Applikation. *HailShow* ermöglicht einen bequemen Import von Flug und Radardaten (2D und 3D), von der Hochfelln Datenbank, und eine Visualisierung der absolvierten Hagelabwehrflüge. Im Anschluss an den visualisierten Flug, kann eine statistische Auswertung zu dem Hagelabwehrflug erfolgen. Das Programm gliedert sich hier in drei Hauptteile, welche im Folgenden kurz erläutert werden sollen, um ein besseres Verständnis über die Arbeitsweise und Bedienung zu geben. Um das Programm *Hailshow* voll nutzen zu können, muss der ausführende Rechner sich im Hochschulnetzwerk befinden, oder es muss ein VPN-Client genutzt werden. Das Programm *Hailshow* kann im Anhang A gefunden werden.

2.7.1 Anbindung an die Datenbank und Datenmanagement

Die, in Kapitel 2.1 beschriebene, Datenbank enthält alle absolvierten Hagelabwehrflüge und Radarbilder (2D und 3D). Die Anbindung an die Datenbank wurde in MATLAB, mithilfe der Database Toolbox™ (MathWorks, 2019), erstellt. Die Database Toolbox™ stellt hierbei Schnittstellen und Funktionen zur Verfügung, um Daten von relationalen und nicht relationalen Datenbanken zu lesen und zu schreiben. Die Datenbankanbindung wird hierbei von einem JDBC-Connector durchgeführt (Oracle, 2019). Es wurde eine Installationsroutine geschrieben, welche den JDBC-Connector korrekt auf jedem Windows 10 Rechner installiert.

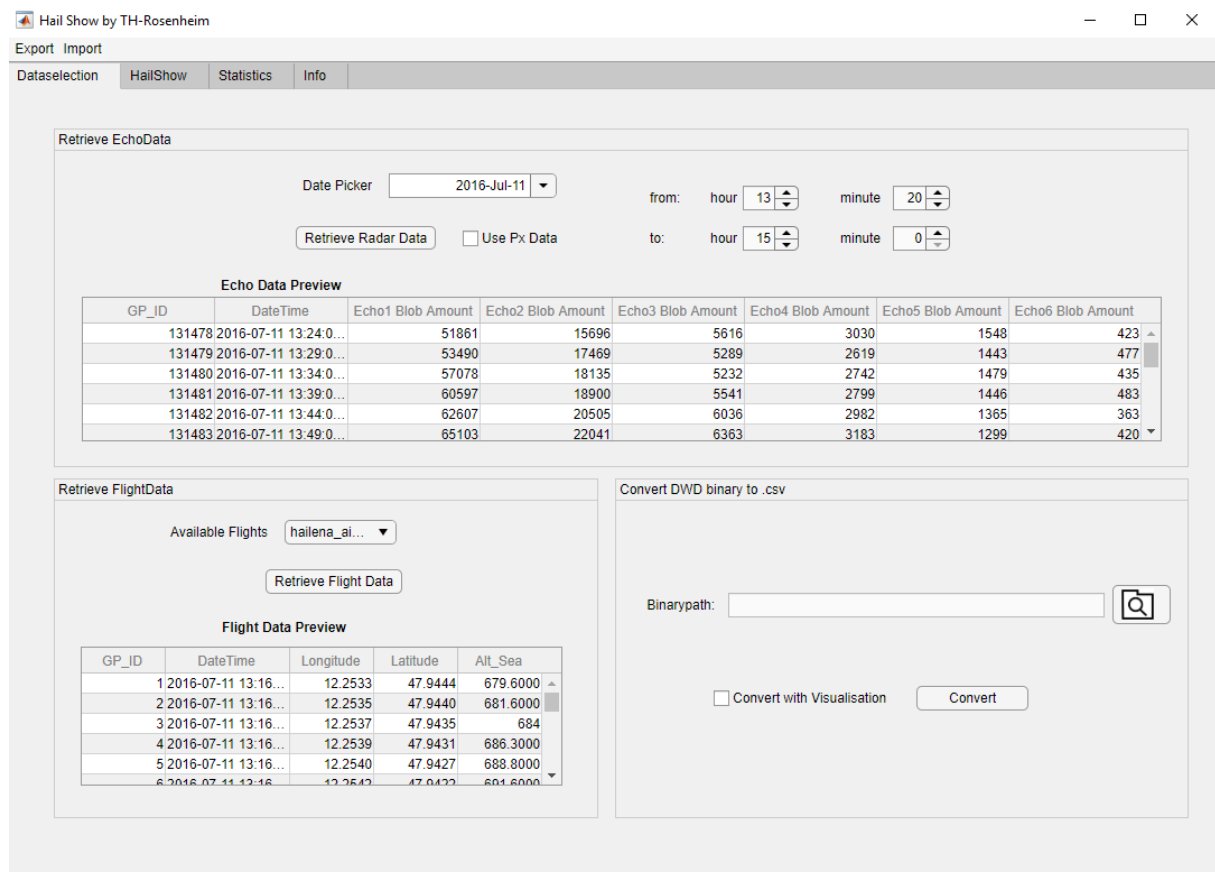


Abbildung 28: Dataselection Oberfläche von Hailshow

Die Kommunikation mit der Datenbank ist im Tab *Dataselection* hinterlegt, wie Abbildung 28 zeigt. Es können hier die einzelnen Flüge über eine Dropdown-Listenfeld selektiert und über den Button *Retrieve Flight Data* abgerufen werden. Die Radarechos können für einen spezifischen Tag und einen

festgelegten Zeitraum von der Datenbank angefordert werden. Das Programm bietet die Auswahl, ob Px-Daten (2-D) oder Pz-Daten (3-D) angefordert werden sollen.

Es wurde ebenfalls die Funktion implementiert, eine im *Runlength*-Format kodierte Binärdatei, in ein lesbares CSV Format umzuwandeln. Hierbei kann bei Bedarf eine Visualisierung der Daten erfolgen. Die Konvertierung von binär zu CSV wird durch ein, vom DWD zur Verfügung gestelltes, Perl-Skript übernommen. Es muss hierbei nicht zwischen Px und Pz Daten unterschieden werden. Die CSV Datei wird in das gleiche Verzeichnis gespeichert in dem die Hailshow.exe ausgeführt wird.

2.7.2 Hailshow Exportfunktionalität

Die von der Datenbank angeforderten Flüge können über die Schaltfläche *Export* als CSV exportiert werden. Ebenso können die Radardaten sowohl für Px, als auch für Pz, für ganze Tage bzw. Zeiträume exportiert werden. Es sollte jedoch darauf geachtet werden, dass mit größeren Zeiträumen erhebliche Datenmengen anfallen, welche entsprechend lange zum Exportieren benötigen. Daher sollte darauf verzichtet werden, einen ganzen Tag (24h) von Pz Daten zu exportieren, da dies eine XLS-Datei mit 288 Sheets, wobei jedes Sheet eine 2400x200 Matrix beinhaltet, zur Folge hat. Die 2400x200 Matrix kommt für Pz-Daten zu Stande, indem die einzelnen Schichten untereinander in eine XLS-Datei geschrieben werden. Die erste 200x200 Matrix kennzeichnet die unterste der 12 Schichten.

2.7.3 Hailshow Importfunktionalität

Der Hagelabwehrverein Rosenheim ist im Besitz von zwei Flugzeugen, wobei nur eines seine Position an den Hochfellnserver übertragen kann. An einigen Flugtagen ist es auch zu Ausfällen des Systems gekommen, weshalb eine Importschnittstelle für Flüge von Flightradar24 geschaffen wurde (siehe Kapitel 5.2.1). Selbst generierte Flüge, sogenannte Virtuelle Flüge, können ebenfalls importiert werden (siehe Kapitel 2.7.8). Es wurde die Möglichkeit vorgesehen 2D Radarechos selbst zu erstellen und einzulesen. Für die Erstellung von eigenen Echos wurden Ansätze unternommen, Wolkenformationen zu simulieren. Jedoch hat sich die Thematik, über die Zu- und Abnahme von Wolken, als zu komplex herausgestellt, daher wurde dieser Ansatz verworfen. Eigene 2D Wolkenformationen können über eine einzulesende XLS-Datei vorgegeben werden (siehe Abbildung 29). Als anknüpfendes Projekt könnte ein Programm erstellt werden, welches dem Benutzer ermöglicht, eigene Wolkenformationen mithilfe einer grafischen Bedienoberfläche zu erstellen.

GP_ID	DateTime	Class01	Class02	Class03	Class04	Class05	Class06
1	2019-04-24 12:00:00.0	80 100 100 80 120 100 100 120					
2	2019-04-24 12:05:00.0		80 100 100 80 120 100 100 120				
3	2019-04-24 12:10:00.0			80 100 100 80 120 100 100 120			
4	2019-04-24 12:15:00.0				80 100 100 80 120 100 100 120		
5	2019-04-24 12:20:00.0					80 100 100 80 120 100 100	
6	2019-04-24 12:25:00.0						80 100 100 8

Abbildung 29: Datenformat für eigene Wolkenformationen

2.7.4 Visualisierung und Erfassung der Flüge

Um vergangene Flüge betrachten zu können, werden diese im Reiter *Hailshow* visualisiert. Flüge mit Px-Daten werden 2-dimensional dargestellt und Flüge mit Pz-Daten 3-dimensional. Der Ansatz der lokalen Geometrien um den Hagelabwehrflieger, aus Kapitel 2.3.3, wurde hier ebenfalls implementiert. Für jeden Zeitpunkt, einer Positionsänderung des Flugzeugs, wird die entsprechende Anzahl aller Radarechos der einzelnen Klassen, innerhalb einer Impf- und Beobachtungsgeometrie, erfasst.

2.7.4.1 Geometrien PX-Daten

Für die Px-Daten besteht die Wahl zwischen einem Kreis und einem Kreissektor. Durch eine einfache Distanzberechnung kann festgestellt werden, ob sich ein Radarecho innerhalb eines Kreises befindet. Wenn der Abstand zwischen Flugzeug und Radarecho geringer ist, als der Radius des Kreises, so befindet sich das Radarecho innerhalb einer Geometrie. Für die Verwendung der Kreisgeometrie muss im Reiter *Hailshow* lediglich ein Radius, für den Impf- und den Beobachtungskreis, vorgegeben werden.

Für die Verwendung des Kreissektors, muss ein Radius und ein Zentriwinkel vorgegeben werden, welcher den Kreissektor aufspannt (siehe Abbildung 30). Der Kreissektor selbst ist immer entgegen der Flugrichtung ausgerichtet. Das Konzept zur Annahme eines Kreissektors stammt aus dem Versuch, die AgJ-Ausbreitung, mit dem Programm *Particleflow-Simulation* (siehe Kapitel 2.3.3), zu simulieren. Es wird vermutet, dass der Kreissektor die Ausbreitung des AgJ am besten wiedergibt. Jedoch sind die Parameter, welche einen realen Ausbreitungsbereich der Impfung wiedergeben, nicht bekannt. Abbildung 31 zeigt einen beispielhaften Kreissektor für eine Impf- und Beobachtungsgeometrie.

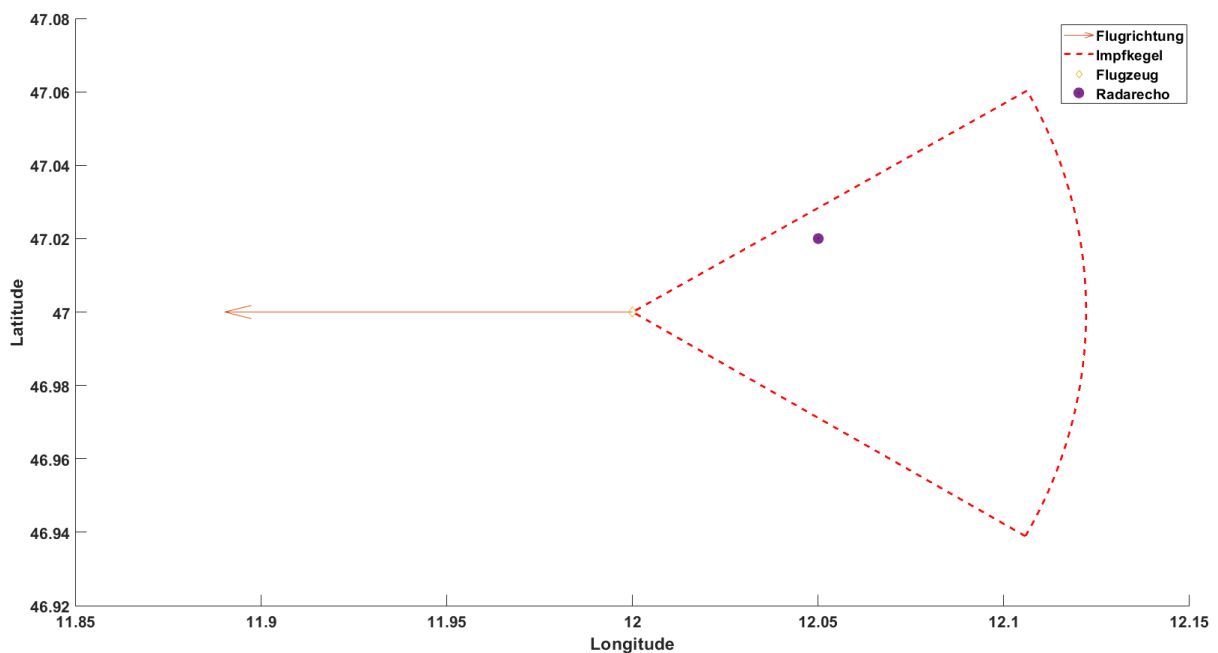


Abbildung 30: Radarecho welches sich innerhalb der Kreissektorgeometrie befindet. Der Kreissektor zeigt immer entgegengesetzt zur Flugrichtung, um den Impfbereich möglichst realistisch abzubilden

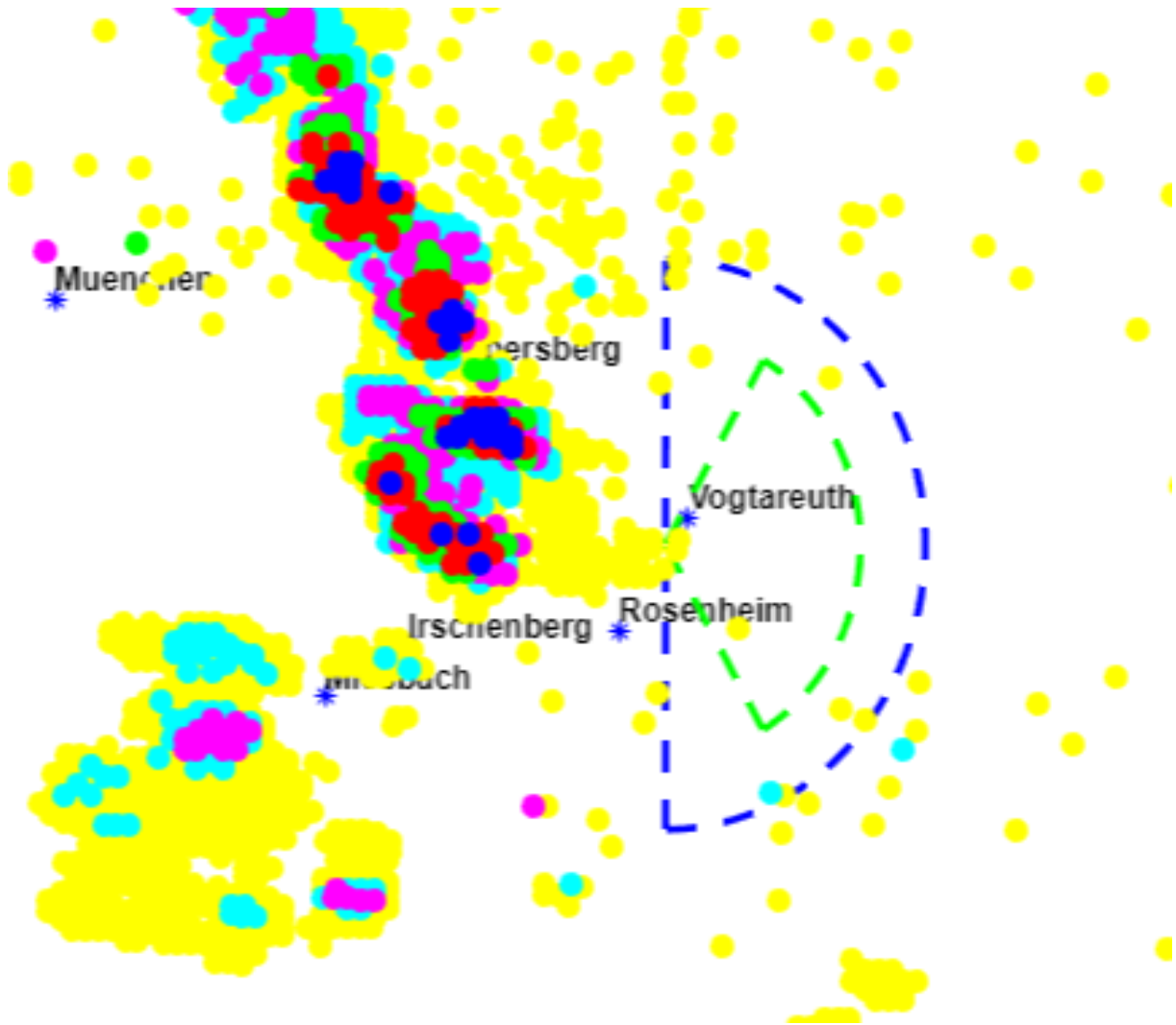


Abbildung 31: Beispielflug mit einem Kreissektor, bei welchem der blaue Kreissektor die Beobachtungsgeometrie und der grüne Kreissektor die Impfgeometrie darstellt.

2.7.4.2 Geometrien PZ Daten

Für die 3-dimensionalen Flüge kann eine Impfkugel und eine Beobachtungskugel mit einem beliebigen Radius um das Flugzeug vorgegeben werden (siehe Abbildung 32). Die Visualisierung der 3-D Radardaten stellte sich im Laufe der Arbeit als komplex heraus, da diese Performance Probleme bereitete. Daraufhin wurden diese Daten mit Methoden der objektorientierten Programmierung gekapselt und hinsichtlich ressourcenschonender Programmierung optimiert. Für Radardaten, welche dennoch enorm viele Echos abbilden, ist dies nicht immer ausreichend. Daher wurden hier Versuche unternommen, mithilfe der Parallel Computing Toolbox™ (MathWorks, 2019) die Visualisierung, der Echos auf der GPU¹ anstelle der CPU² auszuführen. Jedoch würde dies die Verwendung von *Hailshow* nur auf Computer beschränken, welche eine entsprechend leistungsstarke Hardware besitzen. Die Parallel Computing Toolbox™ unterstützt in Abhängigkeit von dem verwendeten MATLAB Release auch nicht alle GPU-Architekturen (MathWorks, 2019). Aufgrund der enormen Einschränkungen, welche eine Berechnung der Visualisierung auf der GPU mithilfe der Parallel Computing Toolbox™ mit sich bringt, wurde auf eine Verwendung dieser verzichtet. Daher wurde die Visualisierung von der Methodik vollständig entkoppelt. Es ist daher die Möglichkeit gegeben, eine Erfassung aller Echos für einen Hagelabwehrflug und eine vorgegebene Geometrie, auch ohne Visualisierung, durchzuführen. Für ein Folgeprojekt könnte an einer reinen Visualisierung von 3-dimensionalen Flügen, mithilfe von MATLAB und der Virtual Reality Modeling Language (VRLM), gearbeitet werden (siehe Abbildung 33).

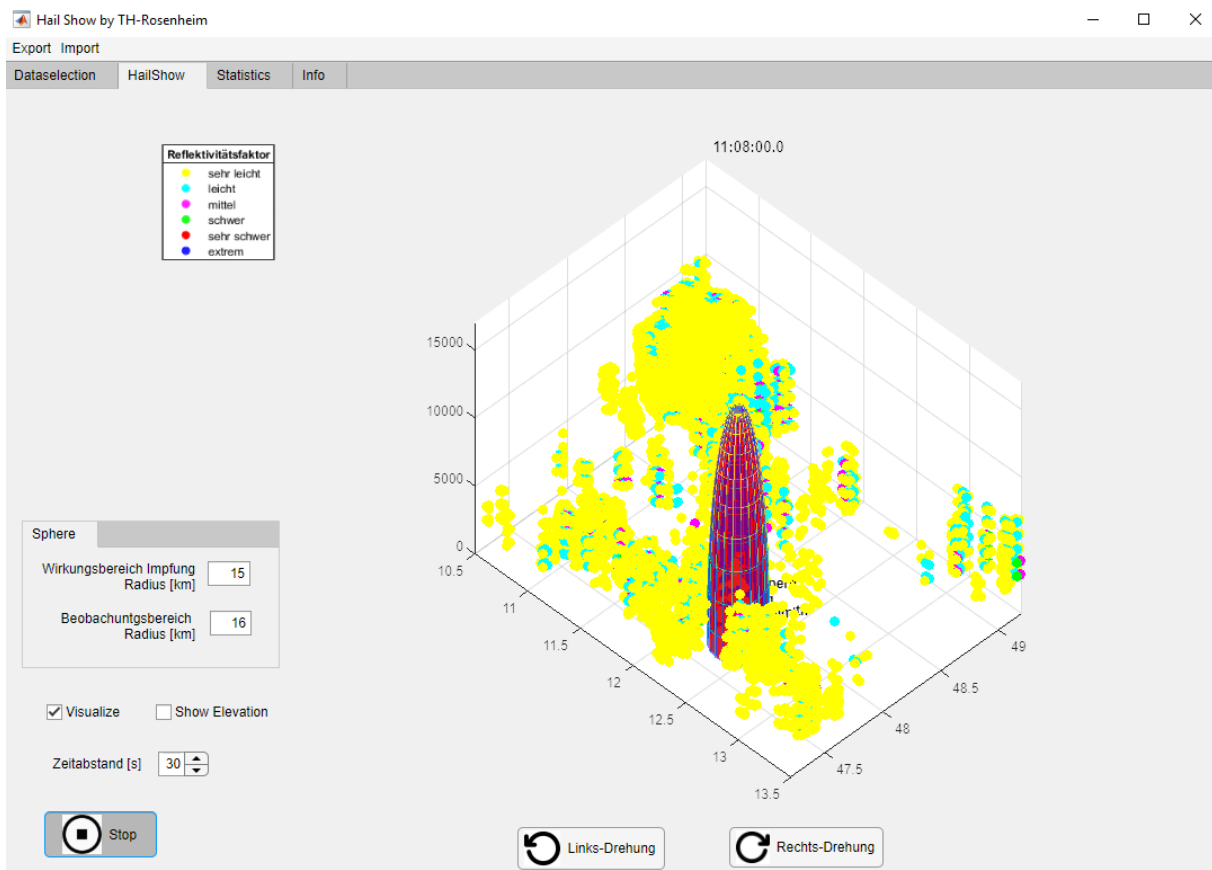


Abbildung 32: Visualisierung der PZ-Daten für eine Kugelgeometrie. Die Beobachtungsgeometrie und die Impfgeometrie überdecken sich hierbei.

¹ Graphics processing unit, Grafikkarte des Computers

² Central processing unit, Matlab führt standardweise alle Berechnungen auf der CPU aus

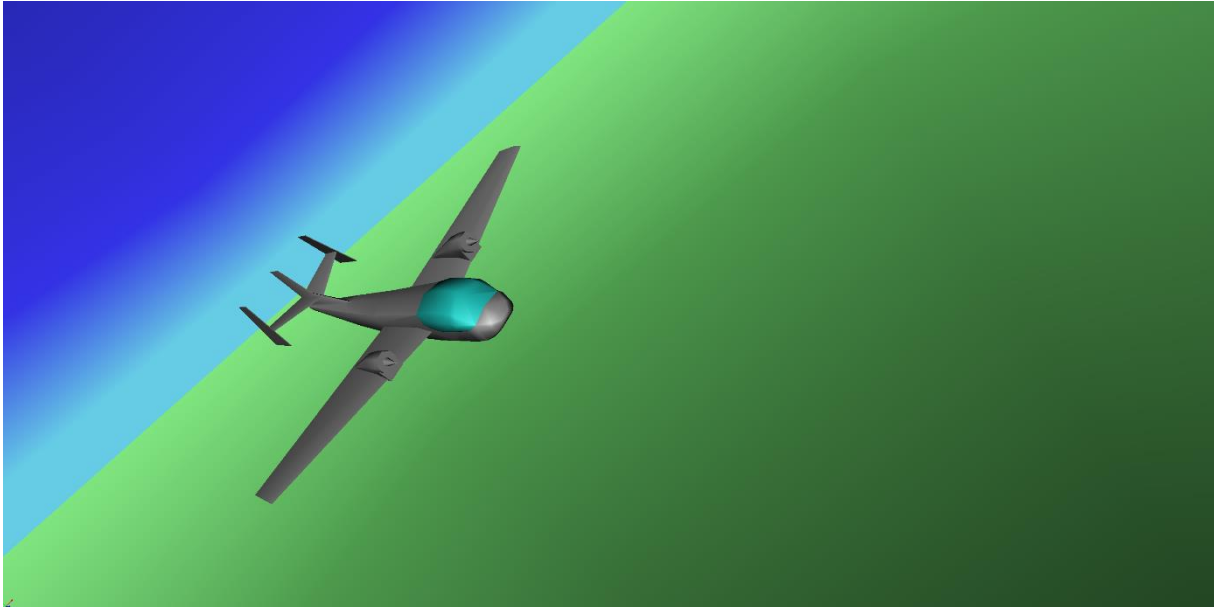


Abbildung 33: Beispielszene für eine Visualisierung der 3-dimensionalen Flüge in VRLM

2.7.4.3 Höhenprofil der Radarkarten

Die, in Kapitel 2.6, interpolierte Radarkarte wurde zusätzlich um Höhendaten ergänzt. Hierzu wurden, zu den einzelnen GPS-Punkten der Radarkarte mithilfe von digitalen Höhenmodellen, die entsprechenden Höhen über den Meeresspiegel angefordert (NASA, 2019). Die Höhenmodelle können für die Visualisierung mitgenutzt werden (siehe Kapitel 2.7.4.1 und 2.7.4.2). Abbildung 34 und Abbildung 35 zeigen eine 2-dimensionale Darstellung der Höhenprofile für die interpolierte Radarkarte. Die Pz-Karte stellt lediglich eine Erweiterung des Kartenbereichs der Px-Karte dar, welche jedoch gegenüber der Px-Karte eine geringere Auflösung besitzt.

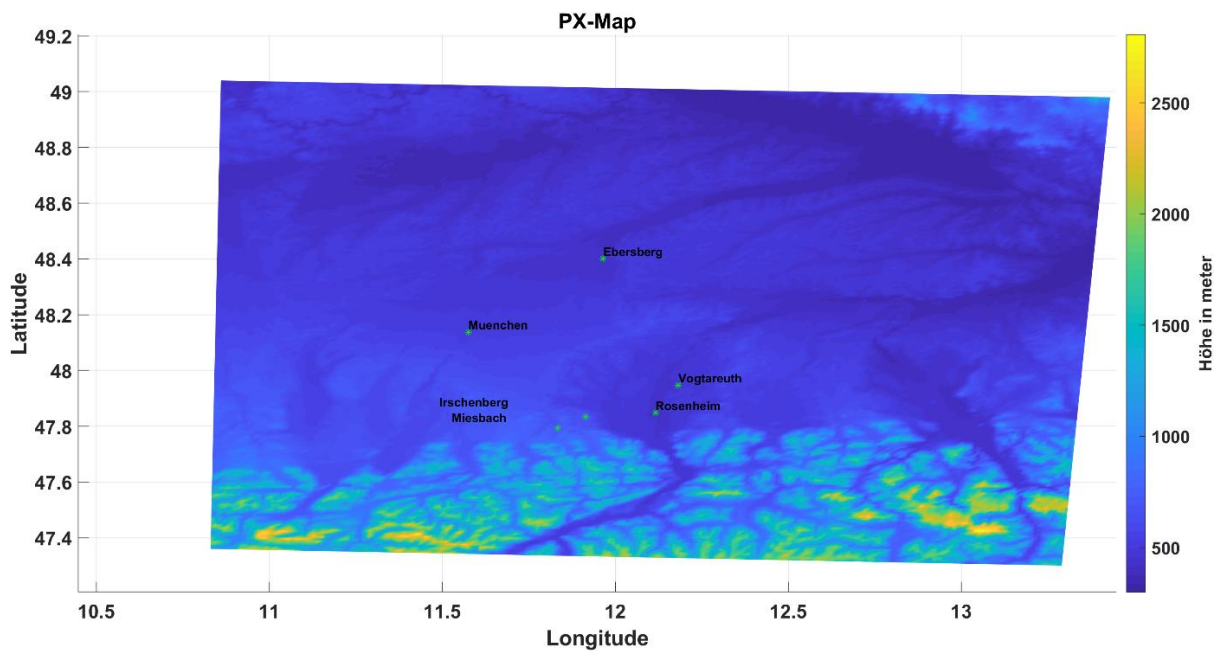


Abbildung 34: Abgedeckter Radarbereich für die Px-Daten

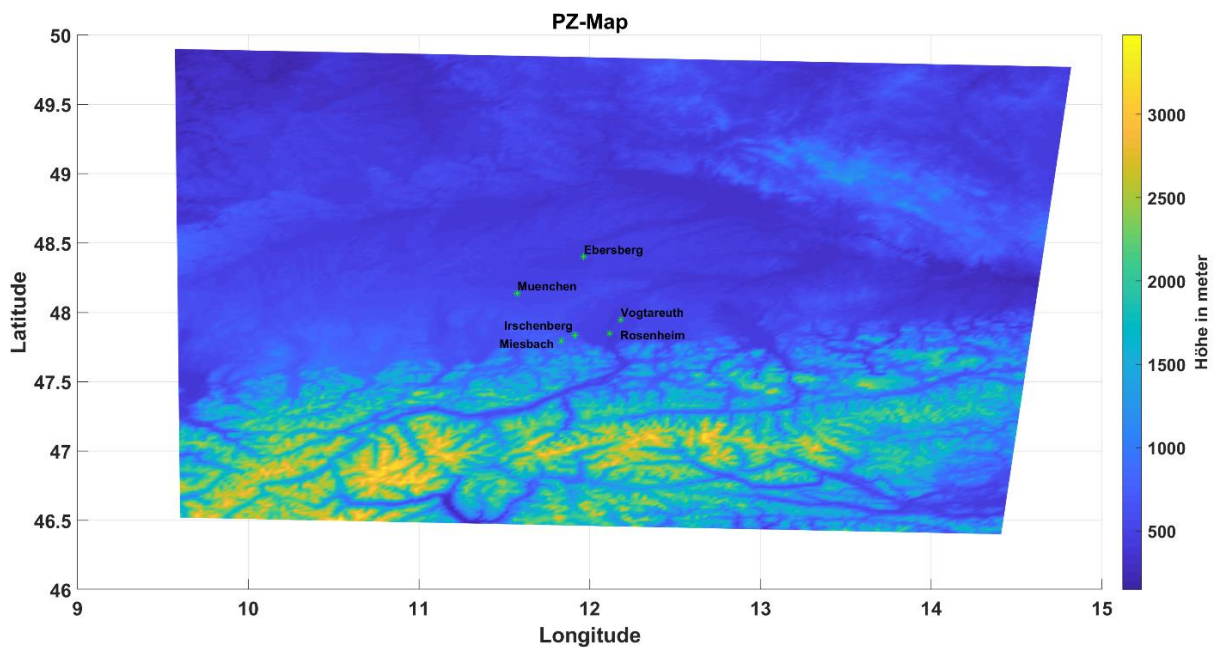


Abbildung 35: Abgedeckter Radarbereich für die Pz-Daten

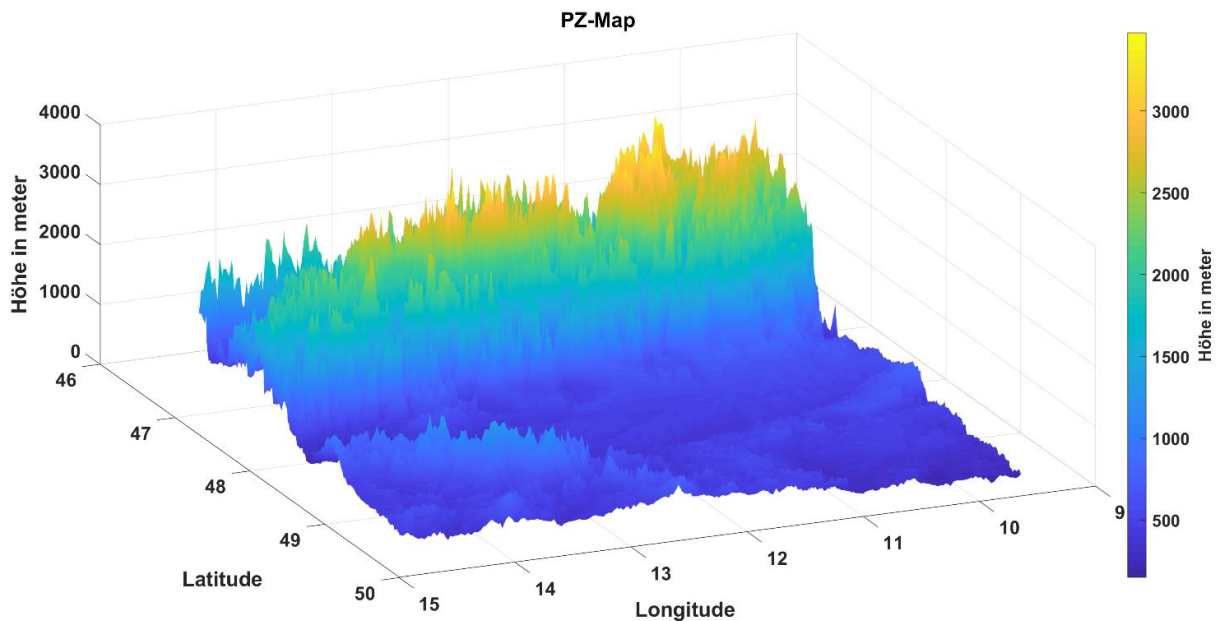


Abbildung 36: 3-dimensionale Darstellung des Höhenprofils der Pz-Daten

2.7.5 Manuelle Untersuchung von einzelnen Gewitterzellen

Im Programm *Hailshow* lässt die Methodik aus Kapitel 2.3.3 nur eine statische Betrachtung um das Flugzeug mit verschiedenen Geometrien zu. Je nach Wahl der Größe und der Impfgeometrie wird die geimpfte Gewitterzelle unterschiedlich gut erfasst. Insbesondere die Tatsache, dass der Hagelabwehrflieger der Gewitterzelle unterschiedlich nahekommen kann, begünstigt die Verfälschung des Ergebnisses der sich in den Geometrien befindlichen Echos. Daher wurde das Programm *Hailshow* um den *Cloudtarget*-Modus erweitert, welcher eine manuelle Selektion eines Bereiches einer Gewitterzelle zulässt. Hierzu wurde die MATLAB eigene Funktion `ginput` modifiziert. Im *Cloudtarget*-Modus muss für jedes Update der Radarkarte ein neuer Bereich einer zu erfassenden Gewitterzelle ausgewählt werden. Hierbei werden um den selektierten Punkt alle Echos, innerhalb eines davor festgelegten Radius, erfasst. Dies ermöglicht eine punktgenaue Untersuchung einer Gewitterzelle, wenn diese geimpft wird. Die manuelle Untersuchung von einzelnen Gewitterzellen kann für PX-Daten mit der Checkbox *Cloudtarget* aktiviert werden. Über die Vorgabe eines Radius können nun, mithilfe des Mauscurors in der MATLAB Figure, die entsprechenden Gewitterzellen selektiert werden (siehe Abbildung 37). Für ein Radarbild kann immer eine Gewitterzelle selektiert werden, hierbei werden alle Echos innerhalb des zuvor vorgegebenen Radius erfasst. Der hier verwendete Radius entspricht von der Funktionalität dem Beobachtungsradius aus Kapitel 2.7.4.1.

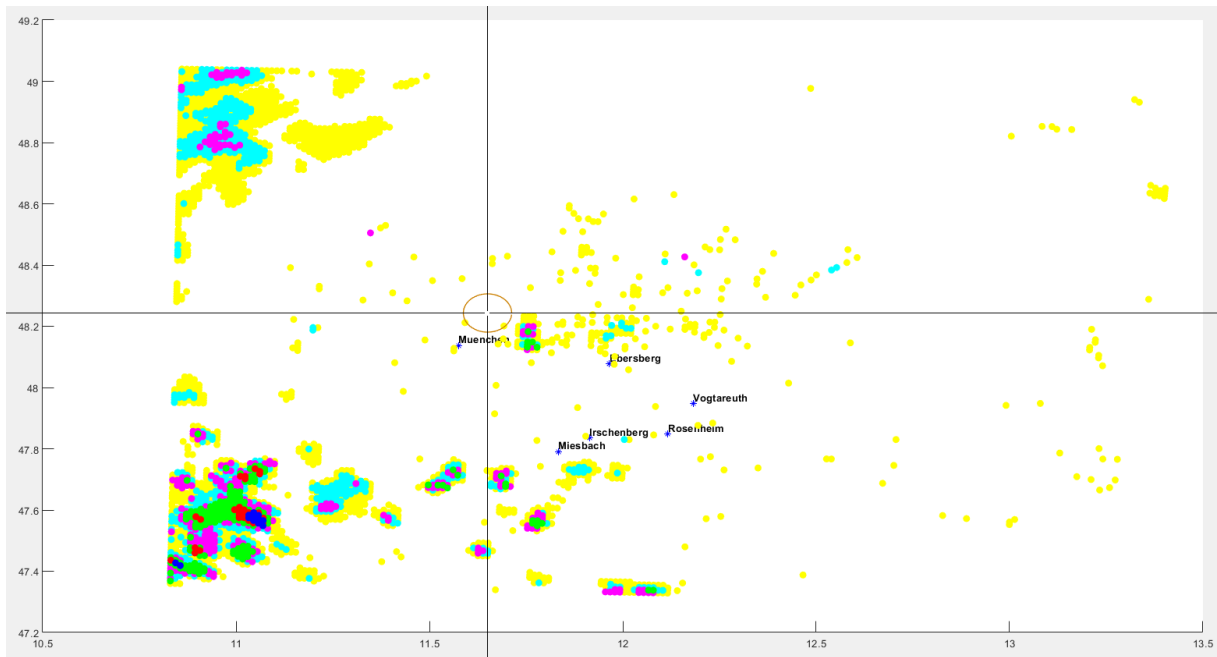


Abbildung 37: Mithilfe eines Fadenkreuzes kann eine bestimmte Gewitterzelle innerhalb eines vorgegebenen Radius analysiert werden.

2.7.6 Algorithmus zur Erkennung von Gewitterzellen

Das, in Kapitel 2.7.5 vorgestellte, Verfahren zur manuellen Untersuchung von Gewitterzellen, stellt eine Lösung für die gezielte Untersuchung von Gewitterzellen dar. In Kapitel 2.7.5 kann, mithilfe des *Cloudtarget*-Modus, nur ein Teil einer Gewitterzelle erfasst werden. Für große Untersuchungsreihen, welche mehrere Gewitterzellen umfassen, stellt dies vor allem auch einen mühsamen Prozess dar. Es wurden daher Überlegungen, ähnlich zu Kapitel 2.3.1 unternommen, um einen automatischen Algorithmus zu entwickeln, welcher zusammenhängende Echos zu einer Gewitterzelle identifiziert. Daher wird ein Algorithmus gesucht, welcher in der Lage ist, eine Punktwolke aus Radarechos zu einem Objekt zusammenzufassen.

2.7.6.1 Bildung einer Konvexen Hülle um eine Punktwolke

Das Vorhaben, einzelne Radarechos als ein zusammenhängendes Objekt zu erfassen, lässt sich auf ein Problem aus der algorithmischen Geometrie übertragen. Die Radarechos werden im Folgenden nur noch als Punkte in einer Ebene bezeichnet.

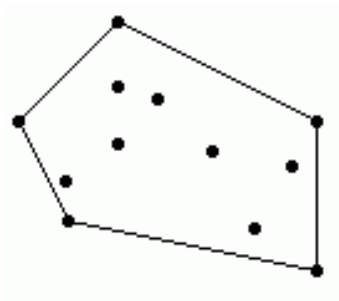


Abbildung 38: Konvexes Polygon um eine Punktwolke (Hochschule Flensburg, 2019)

In einer Ebene ist eine endliche Menge von Punkten P gegeben, um welche ein konvexes (linksgekrümmtes) Polygon gelegt werden soll (siehe Abbildung 38). Für diese Problemstellung existiert eine Vielzahl von Algorithmen, wie z.B. der *Jarvis-March-Algorithmus*, *Quick Hull-Algorithmus* und der *Graham-Scan-Algorithmus* (Hochschule Flensburg, 2019). Der für die

Problemstellung ausgewählte *Graham-Scan-Algorithmus* wurde in MATLAB und C++ programmiert und soll hier erläutert werden.

Aus der Menge P wird ein zentraler Punkt Z errechnet, welcher sich im Zentrum der Punktwolke befindet. Die Koordinaten des zentralen Punktes Z errechnen sich aus den minimalen und maximalen x - und y -Koordinaten der Punktmenge P .

$$Z_x = P_{x_{min}} + \frac{P_{x_{max}} - P_{x_{min}}}{2}$$

$$Z_y = P_{y_{min}} + \frac{P_{y_{max}} - P_{y_{min}}}{2}$$

Zwischen jedem Punkt der Menge P und dem Punkt Z wird anschließend der Winkel φ berechnet. Die Winkel und die dazugehörigen Punkte der Menge P werden nun aufsteigend sortiert.

Um einen konvexen Polygonzug um die Menge P generieren zu können ist es wichtig, nur Polygonstützpunkte für das Hüllpolygon zuzulassen, welche einen konvexen Verlauf des Polygons erzeugen. Vereinfacht gesagt, ist ein konvexes Polygon ein Polygon welches, im mathematisch positiven Drehsinn durchlaufen (Gegen den Uhrzeigersinn), nur aus linksgekrümmten Polygonzügen besteht. Ein Polygonzug besteht aus drei Punkten P_i, P_{i+1}, P_{i+2} . Ein solcher Polygonzug ist genau dann konvex, wenn die Summe der Vektorkomponenten des Kreuzprodukts aus den Vektoren $\overrightarrow{P_i P_{i+1}}$ und $\overrightarrow{P_i P_{i+2}}$ positiv ist.

$$\sum \overrightarrow{P_i P_{i+1}} \times \overrightarrow{P_i P_{i+2}}$$

Hierzu werden die, nach ihrem Winkel φ aufsteigend sortierten, Punkte auf Konvexität geprüft (siehe Abbildung 39). Falls ein Polygonzug einen konkaven (rechtsgekrümmten) Verlauf verursacht, wird der Punkt P_{i+1} entfernt. Es werden solange die Punkte P_i, P_{i+1}, P_{i+2} überprüft bis nur noch konvexe Punkte vorhanden sind.

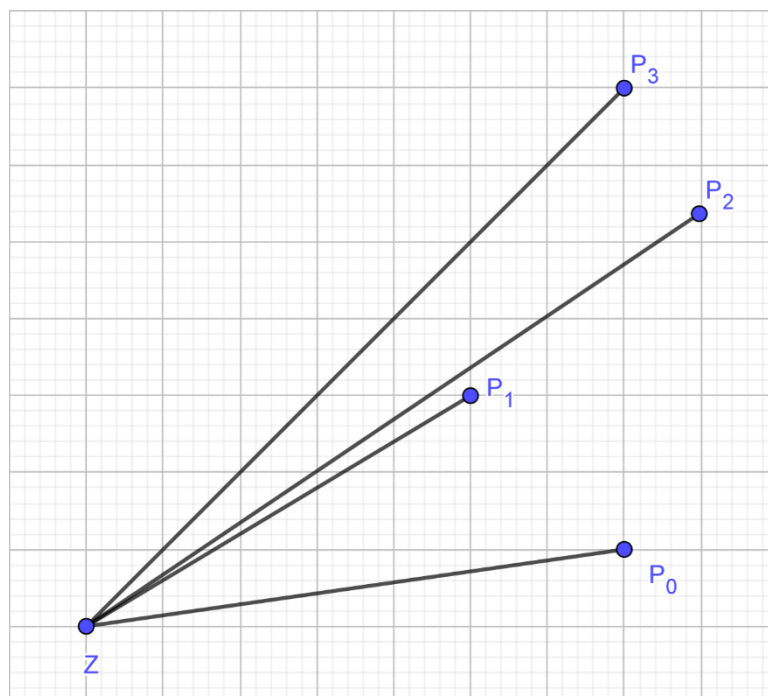


Abbildung 39: Punkte mit nach Winkeln sortierten Indizes

Abbildung 40 zeigt die schrittweise Identifizierung eines konvexen Polygons:

- Zeigt die Anfangslage der Punktwolke
- P_1 verursacht einen konkaven Verlauf und kommt daher nicht mehr als Stützpunkt für das konvexe Polygon in Frage. P_1 wird daher entfernt.
- P_2 verursacht einen konvexen Verlauf und kommt daher als Stützpunkt für das konvexe Polygon in Frage.

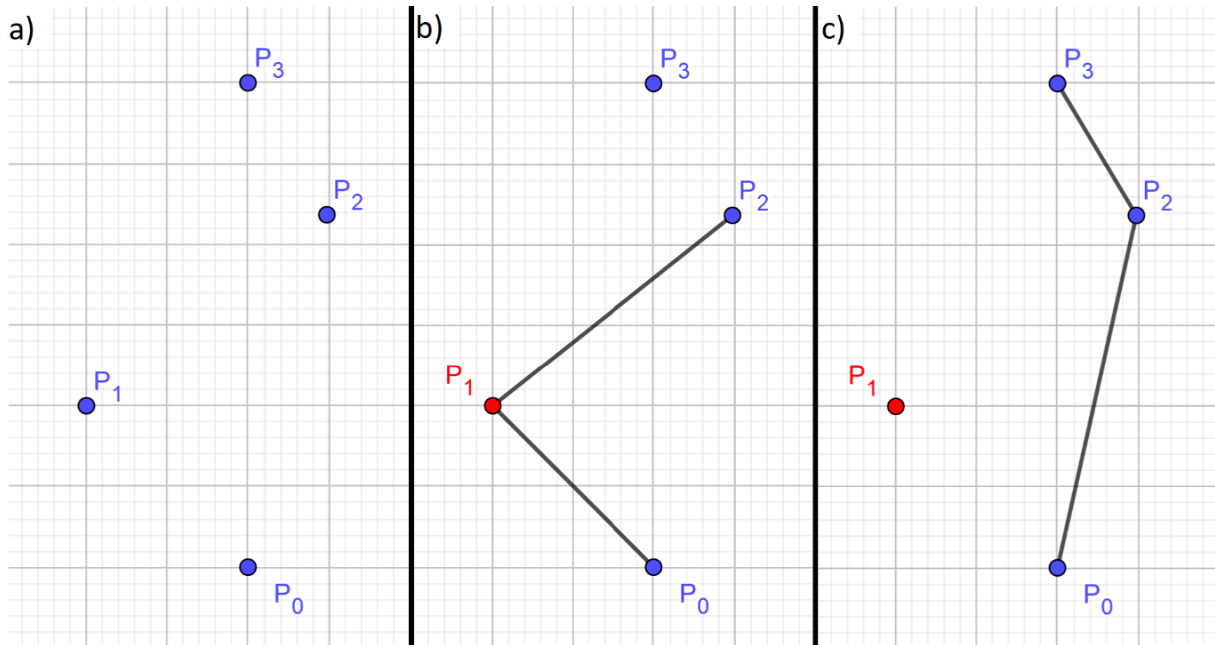


Abbildung 40: Arbeitsweise des Graham-Scan Algorithmus

Abbildung 41 und Abbildung 42 zeigen beispielhafte Ergebnisse für den *Graham-Scan-Algorithmus*.

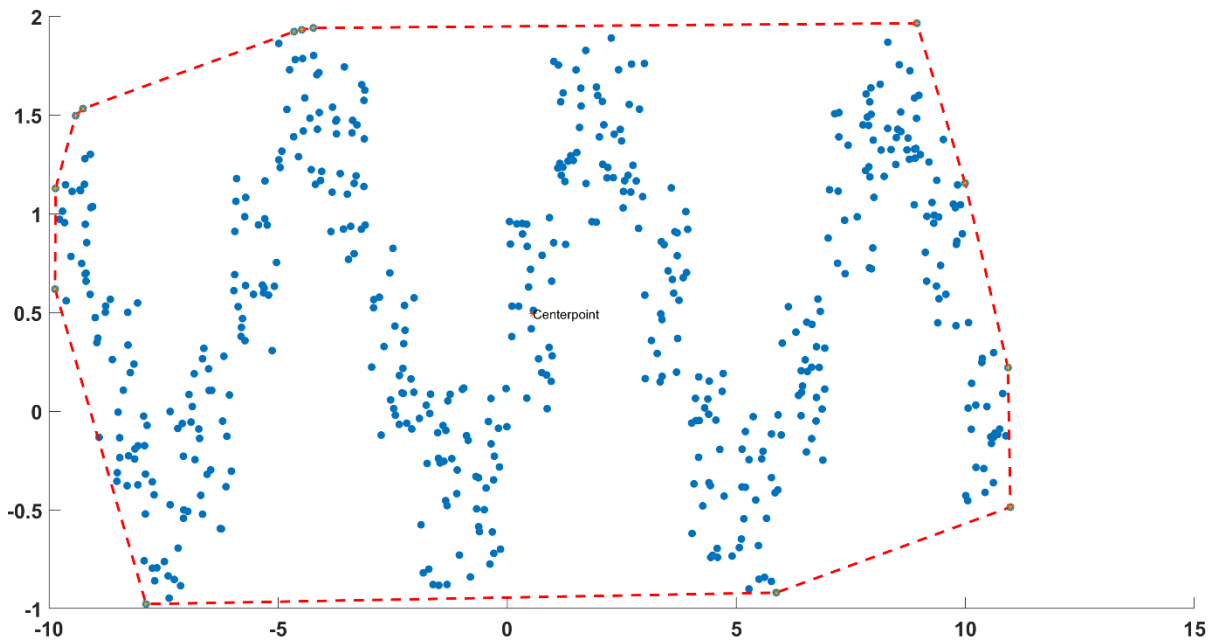


Abbildung 41: Einhüllendes konvexes Polygon um eine Punktwolke, welche einen Sinus beschreibt

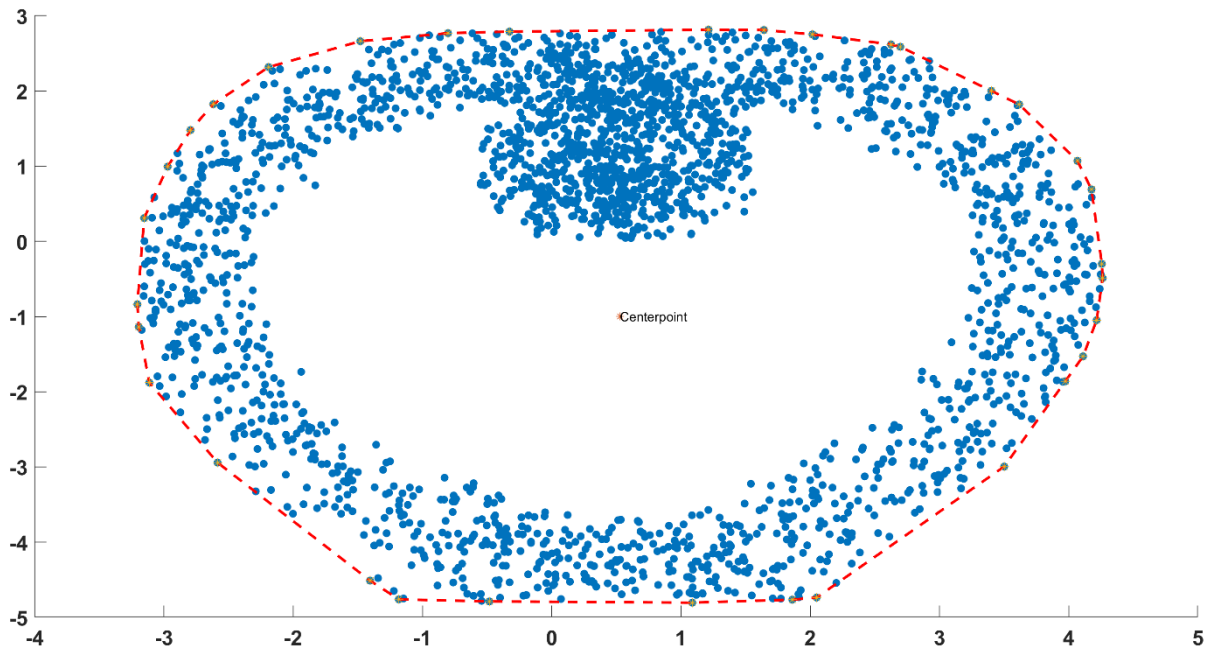


Abbildung 42: Einhüllendes Konvexes Polygon einer Punktwolke entlang einer Kardioide

Der, in diesem Kapitel gezeigte, Algorithmus kann mit dem Ansatz aus Kapitel 2.3.1 zur Segmentierung der einzelnen Gewitterzellen verbunden werden. Der *Graham-Scan-Algorithmus* ist in Anhang A zu finden. Der Algorithmus wurde sowohl in MATLAB, als auch in C++ geschrieben (siehe Abbildung 43). Für die Hochfelln Datenbank kann der *Graham-Scan-Algorithmus* zur Datenreduktion genutzt werden. Hierzu würde das geschriebene C++ Programm auf dem Server ausgeführt werden und alle neu vom DWD zur Verfügung gestellten Radarkarten in eine beschreibende Polygonstruktur zerlegen. Diese Datenreduktion würde viel Speicher auf der Datenbank einsparen und den Server bei Anfragen maßgeblich entlasten, da die beschreibende Polygonstruktur weniger Speicherkapazitäten benötigt, als jedes einzelne Echo. Im Rahmen dieser Masterarbeit wurde der *Graham-Scan Algorithmus* ausgearbeitet, jedoch aus zeitlichen Gründen nicht in das Programm *Hailshow* implementiert.

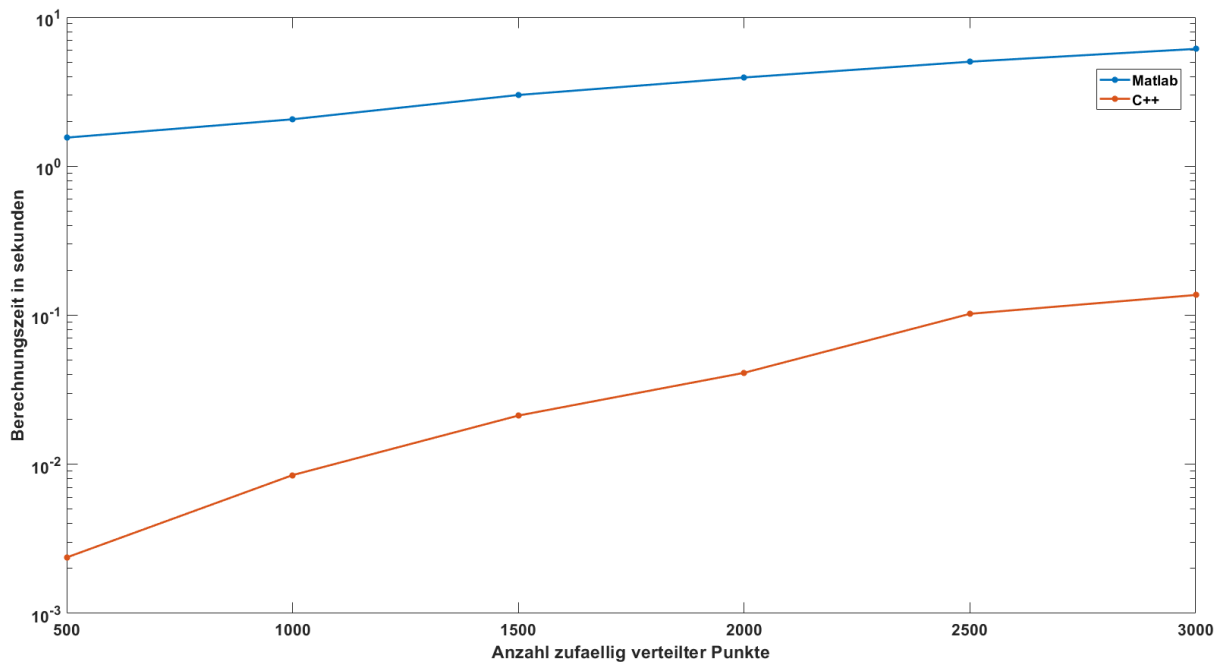


Abbildung 43: Der Graham-Scan Algorithmus in C++ zeigt eindeutig klare Performancevorteile, was für eine Implementierung auf dem Hochfellnserver spricht

2.7.7 Flugverlauf

Nachdem im Programm *Hailshow* ein Flug ausgewertet wurde, kann unter der Sektion *Flugverlauf* der zeitliche Verlauf der einzelnen Echos betrachtet werden. Es wird hierbei die Gesamtmenge aller Echos gezeigt, welche insgesamt vom Radar erfasst werden und die Menge aller Echos in den jeweiligen Geometrien (siehe Abbildung 44). Die eigentliche statistische Auswertung wird in „R“ durchgeführt, daher stellt diese Sektion nur einen Bruchteil der eigentlichen statistischen Auswertung dar.

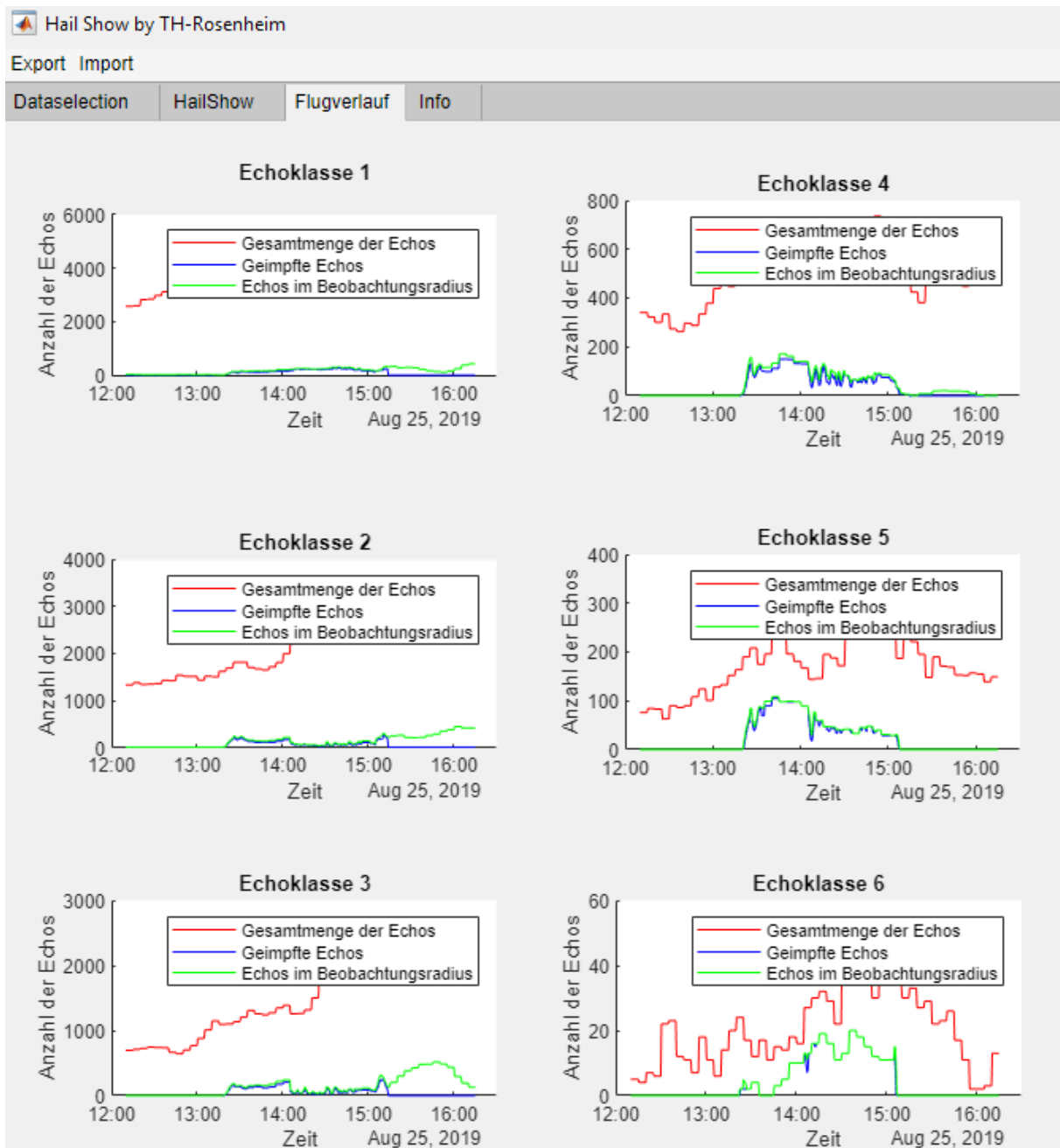


Abbildung 44: Verlauf der einzelnen Echos für einen analysierten Hagelabwehrflug

2.7.8 Virtuelle Flüge

An einigen Tagen der letzten Jahre konnte, aufgrund von einem defekten AgJ-Generator oder einem technischen Defekt am Flugzeug selbst, kein Hagelabwehrflug stattfinden. Um diese Tage dennoch untersuchen zu können, wurde in MATLAB R2019a das Programm *Virtflightcreator* geschrieben. Dieses Programm ermöglicht virtuelle Flüge darzustellen und diese in *Hailshow* für einen bestimmten Tag auszuwerten. Insbesondere kann hier untersucht werden, ob bei den virtuellen Flügen andere Ergebnisse für die Auswertung entstehen, als bei tatsächlich stattgefundenen Flügen.

2.7.8.1 Programmaufbau

Die ersten Entwicklungsschritte für die Erstellung der virtuellen Flüge sollte eine Methode umfassen, welche eine, in Microsoft Paint gezeichnete, Route in einem 200x200 Pixel JPEG mit der tatsächlichen Radarkarte verknüpft (siehe Kapitel 2.6).

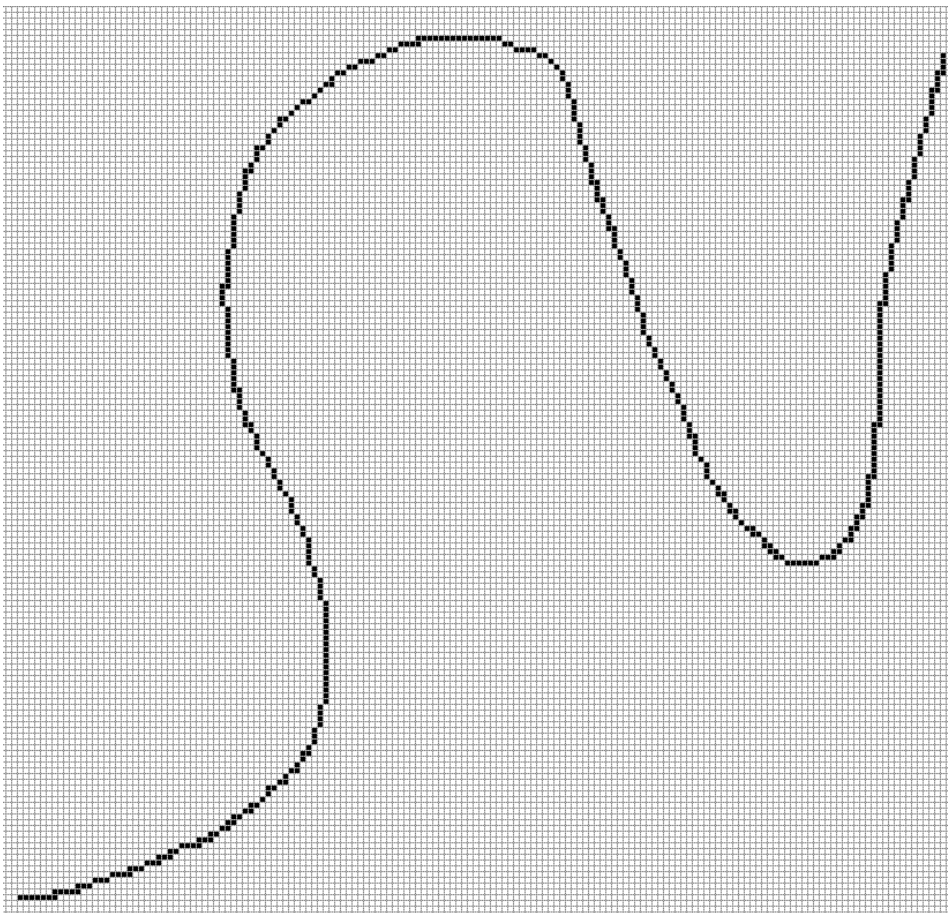


Abbildung 45: Eine in MS-Paint gezeichnete virtuelle Flugroute

Abbildung 45 zeigt eine virtuelle Flugroute, welcher aufgrund der 200x200 Pixelauflösung einen äußerst unnatürlichen „zackigen“ Flugverlauf aufzeigt. Dies konnte mit einer morphologischen Filterung der Erosion, welche aus der Bildverarbeitung stammt, behoben werden (Wagner, 2018). Jedoch hat sich die Generierung der Flugroute aus der 200x200 Pixelmatrix, als äußerst komplex herausgestellt, da eine logische Aneinanderreihung der einzelnen Pixelpunkte und damit der Flugroute nicht einfach lösbar ist. Aufgrund des theoretischen Aufwandes dieses Problem zu beheben, wurde dieser Ansatz verworfen.

In der finalen Version des Programms *VirtFlightCreator* wird die Erstellung der virtuellen Flugroute über das Setzen von sogenannten Waypoints (Wegpunkte) übernommen (siehe Abbildung 46). Zu Beginn der Programmausführung wird eine XLSX oder XLS-Datei ausgewählt, welche eine 200x200 Matrix enthält und somit ein beliebiges Radarbild aus den Px-Daten repräsentiert. Diese Datei kann aus dem Programm *Hailshow* exportiert werden (siehe Kapitel 2.7).

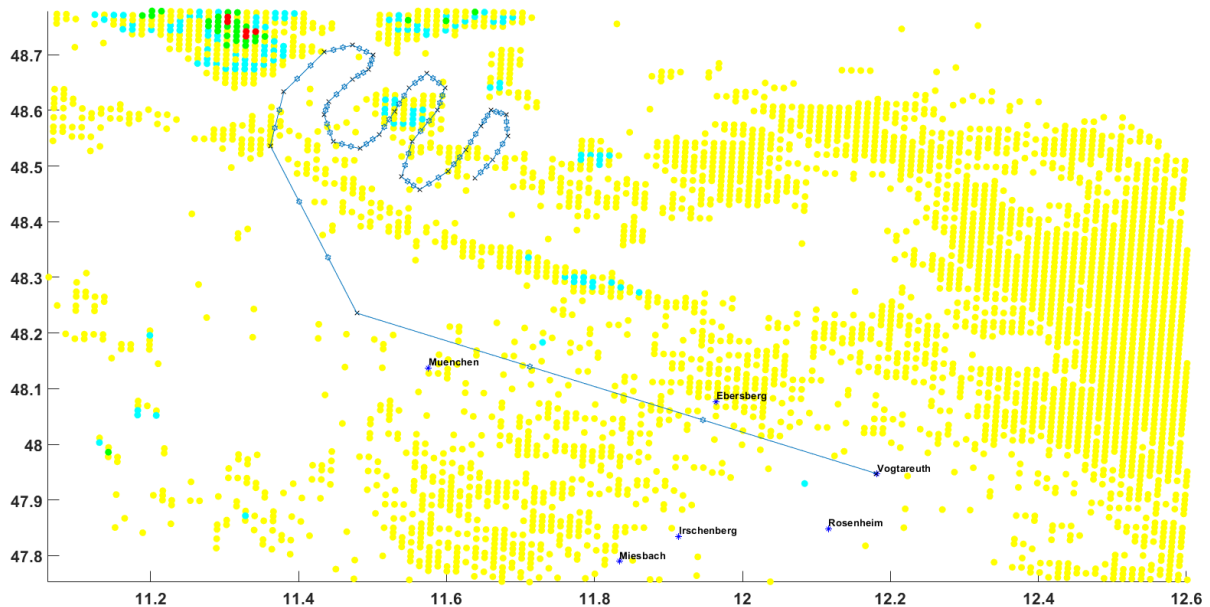


Abbildung 46: Gesetzte Wegpunkte für einen virtuellen Hagelabwehrflug. Zwischen den Wegpunkten kann eine Interpolation erfolgen

Für die manuell festgelegten Wegpunkte kann eine Interpolation zwischen den Wegpunkten durchgeführt werden. Es kann vom Nutzer spezifiziert werden, wie viele Interpolationen durchgeführt werden sollen. Hierbei handelt es sich um eine einfache lineare Interpolation zwischen zwei Punkten (siehe Abbildung 47).

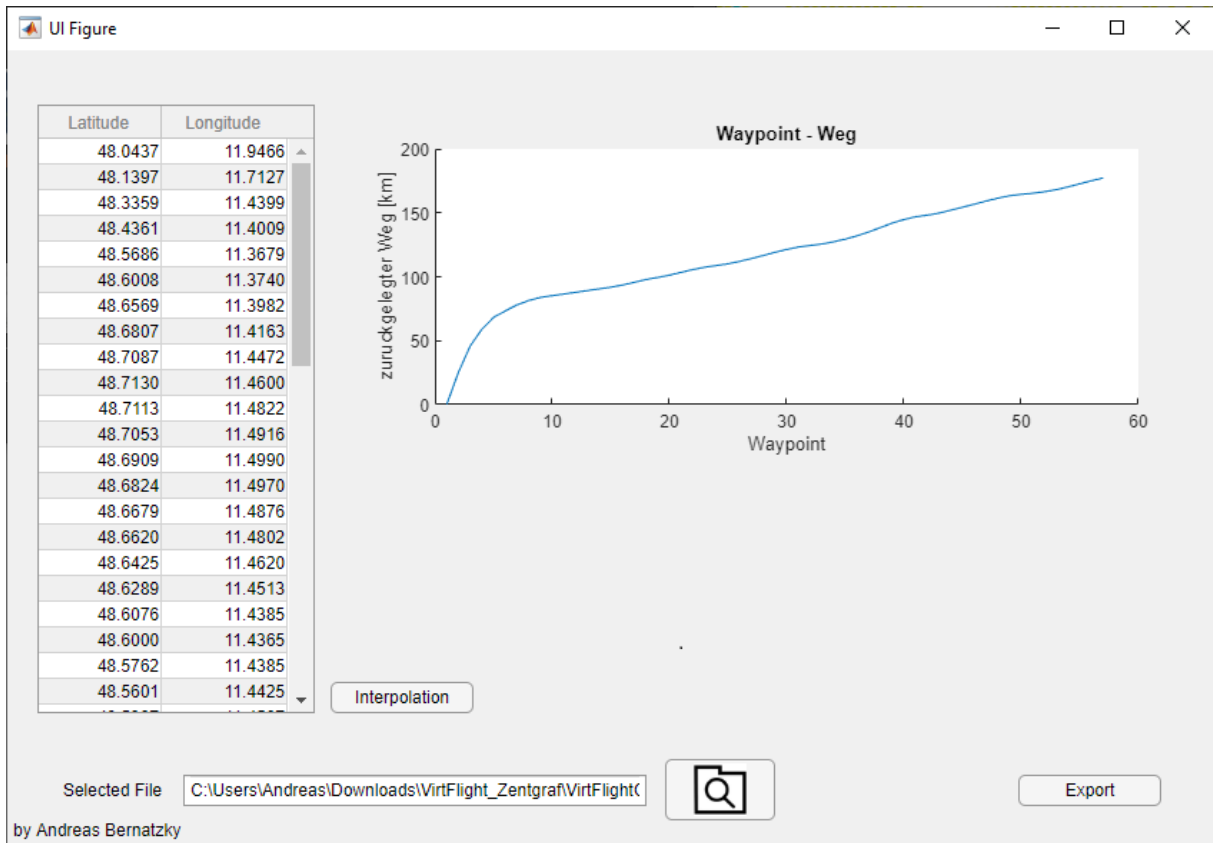


Abbildung 47: Bedienoberfläche des Programms Virtflightcreator

Die virtuell generierte Flugroute kann anschließend so exportiert werden, dass diese in *Hailshow* verwendet werden kann. Für den Export sind Geschwindigkeit des Hagelabwehrfliegers und Startzeitpunkt anzugeben. Die Geschwindigkeit wird zur Vereinfachung über den kompletten Flug als konstant angenommen. Über die Geschwindigkeit und den Startzeitpunkt werden die Zeitpunkte berechnet, an welchen der Hagelabwehrflieger die einzelnen Wegpunkte theoretisch passiert hat. Mithilfe der virtuellen Flüge können theoretische Flüge für Gewitter erstellt werden, an denen kein Hagelabwehrflug stattgefunden hat. Hierbei kann untersucht werden, ob sich solche virtuellen Flüge bezüglich der Radarechos genauso verhalten, wie tatsächlich stattgefundenene Flüge.

3 Auswertung des Radarechoansatzes

Die Auswertung des Radarechoansatzes stützt sich auf alle Daten, welche durch das Wetterradar vom DWD zur Verfügung gestellt werden. Hierbei wurde die beschriebene Methodik aus Kapitel 2.3.3 angewandt. Die Radarechodaten wurden für einzelne Intensitätsklassen untersucht, ob diese einer Zu- oder Abnahme unterliegen.

Das in MATLAB geschriebene Programm bietet die Möglichkeit, eine gezielte Analyse von einzelnen Gewitterzellen durchzuführen (siehe Kapitel 2.7.5). Hierzu wurden, für die Auswertung der Hagelabwehrflüge, jeweils eine Stunde vor dem Start und eine Stunde nach dem Start, Daten aufgenommen. Dies soll sicherstellen, dass der mögliche Effekt des Silberjodids (AgJ), welcher vermutlich nicht sofort eintritt, noch nach Möglichkeit nachgewiesen werden kann. Für jeden betrachteten Flug wird eine Gewitterzelle mit einem Radius von fünf km analysiert.

Es wurden folgende zehn Flüge ausgewertet:

08.06.18
10.06.18
11.06.18
17.06.18
21.06.18
05.07.18
18.07.18
28.07.18
08.08.18
13.08.18

Für die Datenerfassung eines Hagelabwehrfluges werden, für einen festgelegten Zeitabstand, alle 10 Sekunden folgende Daten für jede Radarechoklasse erfasst (siehe auch Abbildung 48):

- Gesamtmenge aller Radarechos einer Klasse.
- Menge aller Radarechos einer Klasse, innerhalb des Impfssektors.
- Menge aller Radarechos einer Klasse, innerhalb des Cloudtargets-Radius (wird für jedes neue Radarbild erfasst).

Zusätzlich gibt es zu jedem Zeitpunkt noch eine boolesche Variabel, ob der Hagelflieger zu einem gewissen Zeitpunkt geflogen ist.

Time	GesamtmengeEcho1	ImpfmengeEcho1	BeobachtungsmengeEcho1	GesamtmengeEcho2	ImpfmengeEcho2	BeobachtungsmengeEcho2
11:40:00	1764	0	34	225	0	0
11:40:10	1764	0	34	225	0	0
11:40:20	1764	0	34	225	0	0
11:40:30	1764	0	34	225	0	0
11:40:40	1764	0	34	225	0	0
11:40:50	1764	0	34	225	0	0
11:41:00	1764	0	34	225	0	0
11:41:10	1764	0	34	225	0	0
11:41:20	1764	0	34	225	0	0
11:41:30	1764	0	34	225	0	0

Abbildung 48: Beispielhafter Auszug für einen erfassten Flug

Zur Auswertung wird die Anzahl der Echoklassen 4 bis 6 betrachtet, da diese mit hoher Wahrscheinlichkeit Hagel enthalten. Die betrachteten Echoklassen werden in drei Unterklassen unterteilt:

- Anzahl der Echoklassen 4 bis 6, in der zu impfenden Gewitterzelle vor dem Flug
- Anzahl der Echoklassen 4 bis 6, in der zu impfenden Gewitterzelle während dem Flug
- Anzahl der Echoklassen 4 bis 6, in der zu impfenden Gewitterzelle nach dem Flug

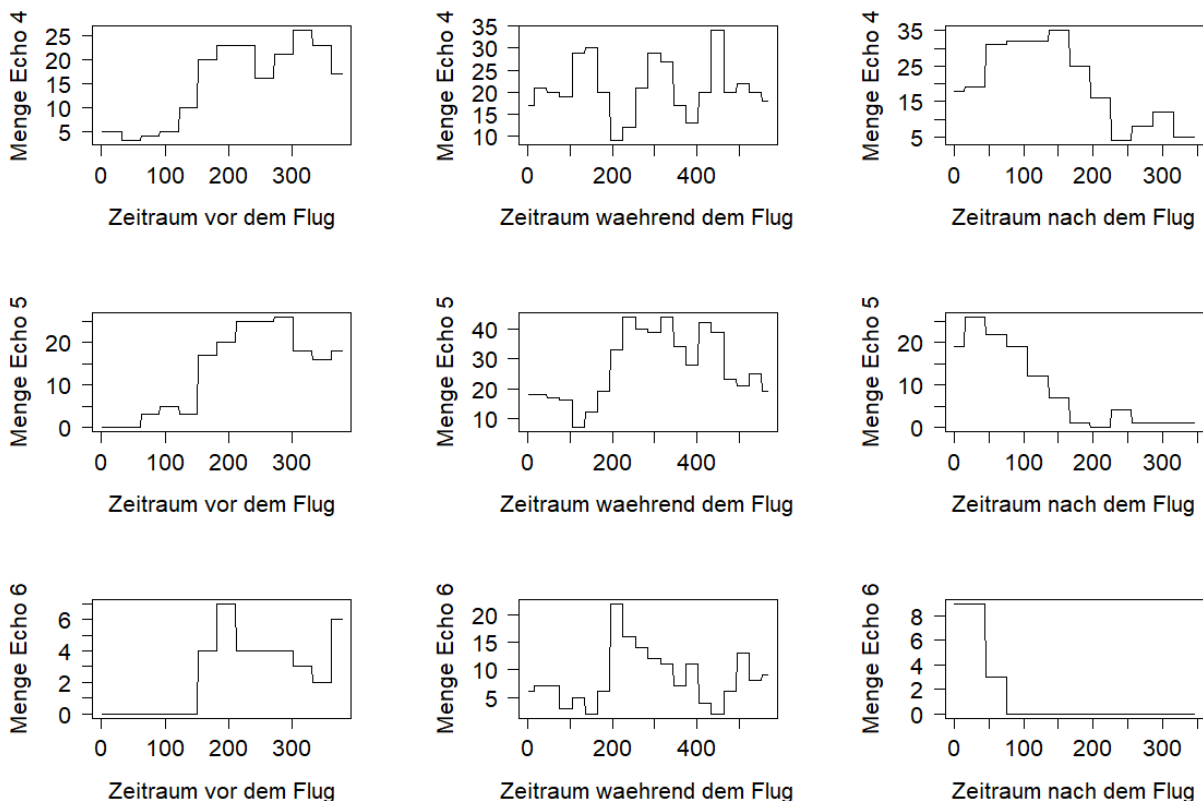


Abbildung 49: Verlauf der Gewitterzellen in den einzelnen Flugphasen für die Echoklassen 4 bis 6

Abbildung 49 stellt den Verlauf einer Gewitterzelle, für die Echoklassen fünf bis sechs, zu den genannten Zeiträumen dar. Für jeden Zeitraum einer Echoklasse wird über den Flug, mithilfe einer linearen Regressionsanalyse, der Trend β berechnet. Dieser Trend kann als Steigung einer Geraden interpretiert werden, welcher eine Aussage über den Verlauf der Menge der einzelnen Echoklassen, während den einzelnen Flugphasen, zulässt. Ein Indiz für die Wirksamkeit der Hagelabwehr wäre ein negativer (fallender) Trend für die Echos der Klassen 4, 5 und 6 für den Zeitraum nach dem Flug. Allerdings könnte die Abnahme der Echostufen auch damit erklärt werden, dass die Gewitterzelle sich abgehagelt hat und sich auflöst. Deswegen soll der Verlauf der Echos auf Zufälligkeit untersucht werden, wobei angenommen wird, dass die Zu- und Abnahme von Wolken nicht systematisch verläuft. Durch das Einbringen von Silberjodid wird erwartet, dass die zufällige Zu- und Abnahme der Echos, bedingt durch die Hagelabwehr, einen systematischen Verlauf annimmt. Im Folgenden soll die Methodik zur Erkennung von Zufallszahlenreihen erklärt werden.

3.1.1 Zufallszahlenerkennung

Zufallszahlen nehmen in unserem alltäglichen Leben eine bedeutende Rolle ein. Die Hauptanwendungsgebiete sind die Informationstechnologie, die Kryptographie zur Verschlüsselung von Nachrichten oder die Überprüfung von wissenschaftlichen Thesen. Speziell bei Letzterem kann überprüft werden, ob beispielsweise ein neuartiges Verfahren zur Voraussage von Wartungsintervallen bei Maschinen, zu geringeren Maschinenstehtzeiten führt, als zufällig ausgeführte Wartungsintervalle. Für die Wirksamkeit der Hagelabwehr wird die Hypothese aufgestellt, dass die Zu- und Abnahme von Gewitterzellen zufällig geschieht. Die Ausbringung von AgJ hat in der Theorie einen systematischen Einfluss auf die Gewitterzellen, die Zu- und Abnahme wäre also nicht mehr zufällig, wenn dieses Eingebracht wird. Dies wäre ein Indiz dafür, dass die Hagelabwehr wirksam ist. Um die Zufälligkeit der Zu- oder Abnahme zu überprüfen, werden verschiedene Tests zur Überprüfung auf Zufälligkeit vorgestellt.

3.1.1.1 Turning-point-test

Der *Turning-point-test*, oder auch *Zackentest* betrachtet, anhand einer Zeitreihe von Y_n Werten, die Änderung ihres Steigungsverhaltens. Eine Zacke Y_t zum Zeitpunkt t , tritt genau dann auf, wenn die benachbarten Werte Y_{t-1} und Y_{t+1} beide größer oder beide kleiner als Y_t sind (Pruscha, Statistisches Methodenbuch, 2006).

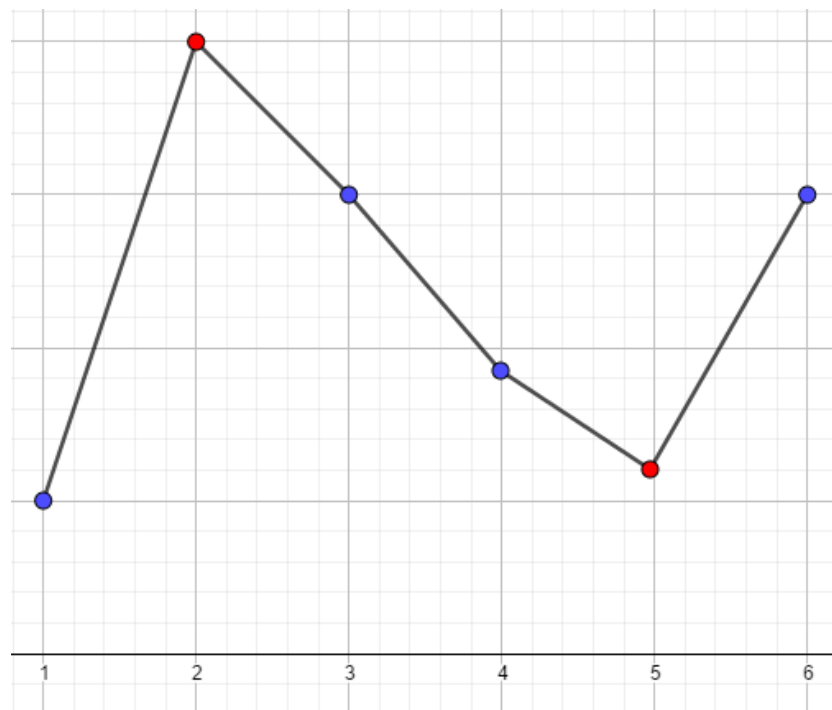


Abbildung 50: Beispielhafte Zahlenreihe, die rotmarkierten Punkte kennzeichnen Zacken

Um zu bestimmen, ob eine Zahlenreihe tatsächlich zufällig ist, wird die Teststatistik Z errechnet. n kennzeichnet dabei die Gesamtanzahl der Zahlen in der zu prüfenden Zahlenreihe. N bestimmt die Anzahl der Zacken. Die Teststatistiken berechnen sich folgendermaßen mit den beiden Hilfsgrößen Var_0 und E_0 (Pruscha, Statistisches Methodenbuch, 2006, S. 320):

$$E_0(N) = \frac{2}{3} * (n - 2) \qquad Var_0(N) = \frac{1}{90} * (16n - 29)$$

$$Z = \frac{N - E_0(N)}{\sqrt{Var_0(N)}}$$

Die Annahme einer Zufallsreihe, wird dann verworfen, falls $|Z| > u_{1-\alpha/2}$. Das $(1 - \alpha/2)$ -Quantil entstammt hierbei der $N(0,1)$ -Verteilung. Im Anhang A befindet sich ein MATLAB Skript, welches den Algorithmus des *Turning-point-tests* für eine beliebige Zahlenreihe erläutert.

3.1.1.2 Bartels rank test

Der *Rangtest nach Bartel* ist ein Test, welcher auf die Verhältnisprüfung auf Zufälligkeit von Neumann zurückgeht (Caeiro, 2015 und Bartels, 1982). Die Teststatistik RVN errechnet sich folgendermaßen (Bartels, 1982):

$$RVN = \frac{\sum_{i=1}^{n-1} (R_i - R_{i+1})^2}{\sum_{i=1}^n (R_i - (n+1)/2)^2}$$

$\sum R_i$ kennzeichnet hierbei die Summe der Rangzahlen, der zu untersuchenden Zufallszahlenreihe und n die Anzahl der Zahlen, in der zu untersuchenden Zufallszahlenreihe. Die Teststatistik RVN wird mit dem Tabellenwerk *Kritische Werte für die Verhältnisprüfung von Neumann* verglichen. Hierbei wird die Annahme von Zufallszahlen verworfen, wenn $RVN < c_0$. Der kritische Wert c_0 stammt aus dem *Neumannschen Tabellenwerk* für kritische Werte (Bartels, 1982, S. 42). Für Stichprobengrößen größer als 10 wird das Tabellenwerk mithilfe der β -Verteilung genähert (Bartels, 1982). Für kleine Stichprobenmengen ($n \leq 10$) wird empfohlen, nur die Teststatistik $NM = \sum_{i=1}^{n-1} (R_i - R_{i+1})^2$ zu nutzen. Der *Rangtest nach Bartel* soll, anhand eines kurzen Beispiels, demonstriert werden (Bartels, 1982):

Wenn Unternehmen gut geführt werden und zukünftige Aufträge genau zu erwarten sind, ist davon auszugehen, dass die jährlichen Veränderungen der Lagerbestände zufällig sind. Die jährlichen Lagerbestände für ein Produkt sind wie folgt:

$$528, 348, 264, -20, -167, 575, 410, -4, 430, -122$$

Die Rangfolge R_i der einzelnen Lagerbestände kann folgendermaßen beschrieben werden:

$$9, 6, 5, 3, 1, 10, 7, 4, 8, 2$$

Die Teststatistik $NM = \sum_{i=1}^{n-1} (R_i - R_{i+1})^2$ ergibt sich hiernach zu $NM = 169$. Dem, in Anhang B zu findenden, Tabellenwerk ist für eine Stichprobengröße von $T=10$ ein c_0 von 51, für ein Signifikanzniveau von 5 % zu entnehmen. Es ist daher davon auszugehen, dass die Lagerbestände zufällig sind.

3.1.2 Ergebnis der Auswertung des Geometrieansatzes

Die Trends für die Echoklassen 4 bis 6 wurden für alle Flüge, über die einzelnen Flugabschnitte, aufgetragen. Die einzelnen Flugabschnitte sind entsprechend Abbildung 51 gekennzeichnet. Ein negativer Trend kennzeichnet eine Abnahme der jeweiligen Echoklasse, ein positiver hingegen eine Zunahme. Ein Trend von 0 zeigt keine Veränderung an.

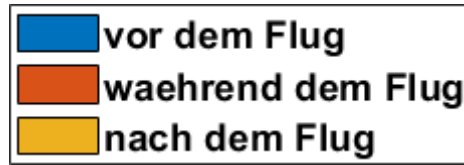


Abbildung 51: Legende der Flugabschnitte

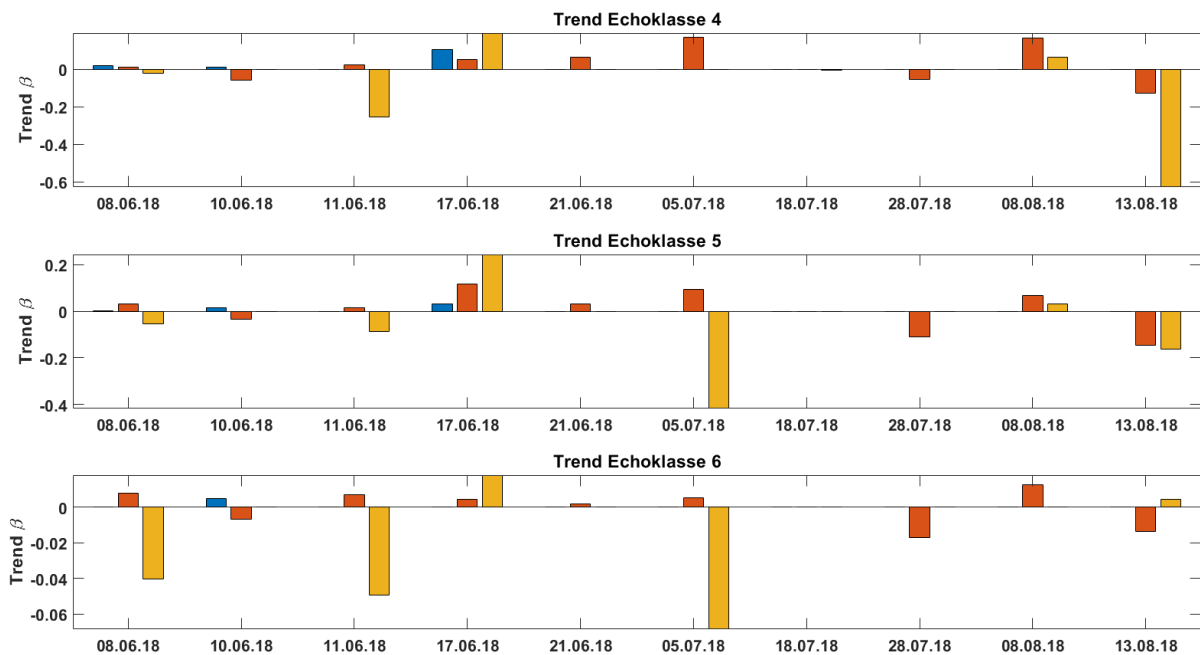


Abbildung 52: Trends der einzelnen Flugphasen für unterschiedliche Flugtage, innerhalb der Beobachtungsgeometrie

Abbildung 52 zeigt, dass für den, in Kapitel 2.3.3 beschriebenen, Ansatz keine verwertbaren Ergebnisse zu sehen sind. Dies liegt vor allem daran, dass der Trend für viele Flugphasen bei null liegt. Da die Piloten nicht immer gleich weit von einer Gewitterwolke entfernt sind kommt es vor, dass die als statisch angenommene Impf- bzw. Beobachtungsgeometrie oftmals Bereiche erfasst, welche ohne Echos sind. Der Test auf Zufälligkeit konnte daher nicht ausgeführt werden.

Um die einzelnen Flüge aus Abbildung 52, mithilfe eines numerischen Koeffizienten, ausdrücken zu können, soll eine Kennzahl eingeführt werden. Hierzu wird für jede Echoklasse die Differenz, der einzelnen, dimensionslosen Trends zwischen den Flugphasen, berechnet, um eine Aussage über die Gesamtabnahme treffen zu können. Die Driftkennzahl $\Delta\beta_x$ errechnet sich für jede Echoklasse, aus den jeweiligen Flugphasen, wobei sich das tiefgestellte x auf die entsprechende Klasse bezieht. Bei einer negativen Driftkennzahl kann eine Abnahme, der jeweiligen Echoklasse, über alle Flugphasen verzeichnet werden.

$$\Delta\beta_x = (\beta_{Flug_x} - \beta_{vorFlug_x}) + (\beta_{nachFlug_x} - \beta_{Flug_x})$$

Die Driftkennzahl $\Delta\beta_x$ lässt sich auch für eine zu betrachtende Flugreihe, bestehend aus n Flügen, für eine Echoklasse normiert darstellen:

$$\Delta\beta_{x_{norm}} = \frac{\sum_{i=1}^n \Delta\beta_{x_i}}{n}$$

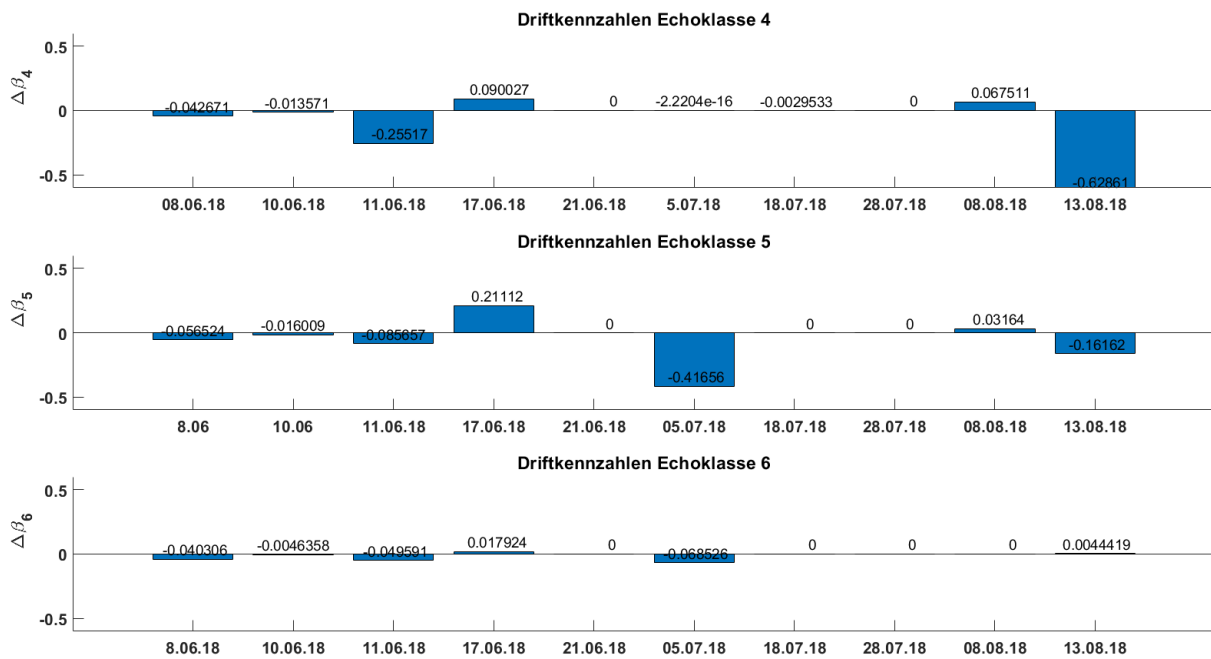


Abbildung 53: Driftkennzahl der Flüge aus dem Geometrieansatz

Abbildung 53 zeigt einige Flüge mit den Driftkennzahlen von 0, was darauf zurückzuführen ist, dass die Beobachtungsgeometrie als statisch angenommen wird.

Die normierten Driftkennzahlen zeigt Tabelle 5.

Tabelle 5: Ergebnis der normierten Driftkennzahlen

$\Delta\beta_{4_{norm}}$	$\Delta\beta_{5_{norm}}$	$\Delta\beta_{6_{norm}}$
-0,0785	-0,0494	-0,0141

3.1.3 Ergebnis der Auswertung des Clouddtarget Ansatzes

Für den Clouddtarget Ansatz wurden ebenfalls die Echoklassen 4 bis 6 für alle untersuchten Flüge, über die einzelnen Flugabschnitte, aufgetragen. Die einzelnen Flugabschnitte sind entsprechend Abbildung 51 gekennzeichnet. Ein negativer Trend kennzeichnet wieder eine Abnahme der jeweiligen Echoklasse, ein positiver hingegen eine Zunahme. Ein Trend von 0 zeigt keine Veränderung an.

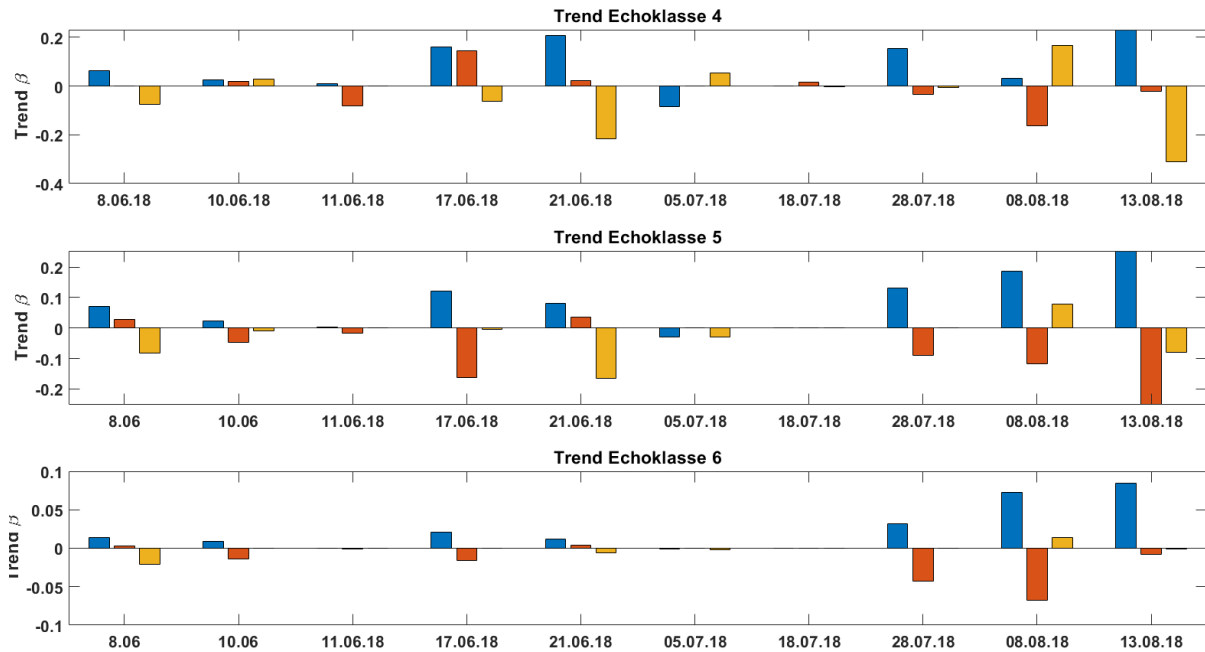


Abbildung 54: Trends der einzelnen Flugphasen für den Clouddtarget Ansatzes

Für die ausgewählten Flüge zeigen alle Trends eine Abnahme zwischen *vor dem Flug* und *während dem Flug*. Einzig der Flug vom 08.08.18 weist für *nach dem Flug*, in allen betrachteten Echoklassen eine Zunahme auf. Der Ansatz über die Zufallszahlen hat sich für den *Turning-point-test* als ebenso nicht durchführbar erwiesen, da für viele Flüge die Änderung, zwischen den einzelnen Zeitpunkten, nicht groß genug ist, was auf eine ungenügende Anzahl von *Zacken* zurückzuführen ist. Der *Rangtest nach Bartel* hat alle Flüge, für alle Zeiträume als nicht zufällig ($p\text{-Value} < 2e-16$) gewertet. Auch hier wird vermutet, dass die Änderungsrate nicht hoch genug ist, um einen Test auf Zufallszahlen durchzuführen. Interessant wäre es hier die einzelnen Gewitterzellen nicht, mithilfe von Schwellwertstufen (Echoklasse 1 – 6) zu identifizieren, sondern mit den gemessenen Werten, der relativen Reflexivität. Diese Werte werden jedoch nicht vom ROBERTA-Projekt erfasst und stehen nur dem DWD, beziehungsweise dem Betreiber des Radars, zur Verfügung.

Zu den einzelnen Flügen aus Abbildung 54 lassen sich folgende Driftkennzahlen bestimmen:

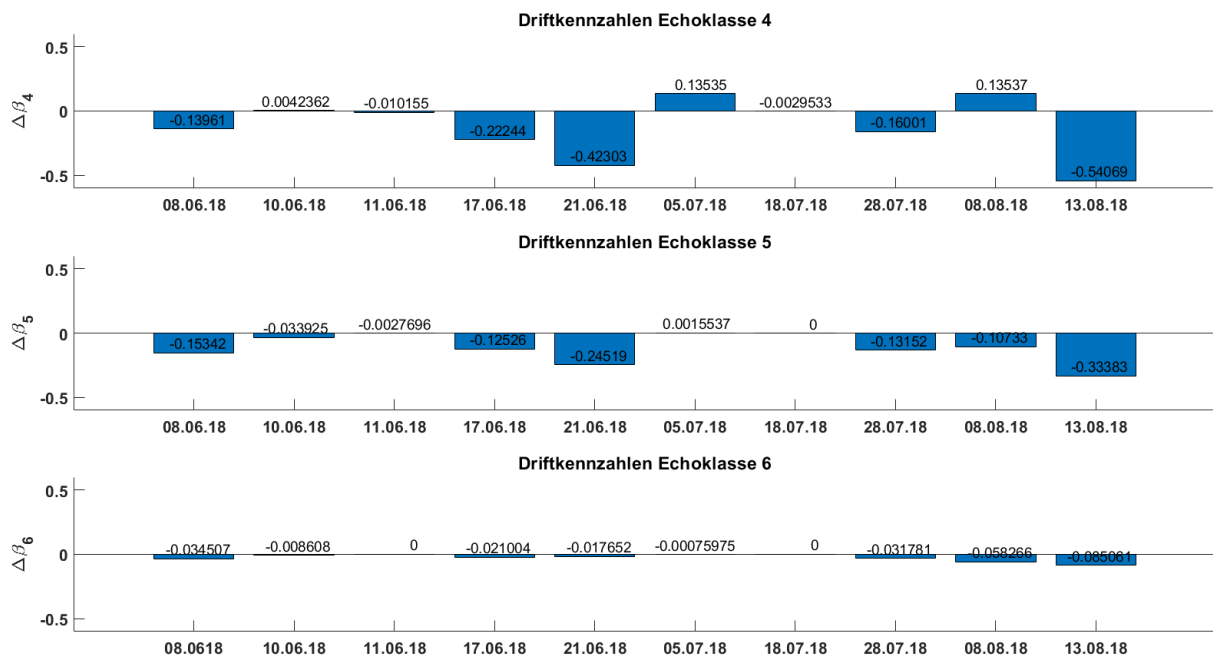


Abbildung 55: Driftkennzahlen der einzelnen Flüge

Die normierten Driftkennzahlen $\Delta\beta_{x_{norm}}$ sind in Tabelle 6 zu sehen. Da alle normierten Driftkennzahlen negativ sind, kann diese beobachtete Flugreihe als Erfolg für die Hagelabwehr gewertet werden. Um eine Saison der Hagelabwehr bewerten zu können, könnte die normierte Driftkennzahl für die kommenden Jahre erfasst werden.

Tabelle 6: Ergebnis der normierten Driftkennzahlen

$\Delta\beta_{4_{norm}}$	$\Delta\beta_{5_{norm}}$	$\Delta\beta_{6_{norm}}$
-0,1224	-0,1132	-0,0258

3.2 Fazit zum Radarechoansatz

Der Ansatz aus Kapitel 3.1.3, die einzelnen Gewitterzellen lokal zu erfassen, zeigt, dass die einzelnen normierten Trends alle negativ sind. Allerdings sollte die hier vorgestellte Methodik noch verfeinert werden. Aufgrund der Mühsamkeit, eine große Anzahl von Flügen auszuwerten, wurde hier nur eine kleine Stichprobe von zehn Flügen untersucht. Von diesen zehn Flügen weisen neun eine deutliche Abnahme der Trends auf. Es müsste daher ein Algorithmus, ähnlich zu Kapitel 2.3.1 und 2.7.6 entwickelt werden, welcher Gewitterzellen automatisch lokalisiert und bewertet. Dies bringt vor allem den Vorteil, gegenüber der momentanen Methodik, welche manuell ausgeführt wird, dass Ungenauigkeiten verhindert werden (oftmals ist der Kern einer Gewitterzelle manuell schwer zu lokalisieren). Gleichzeitig könnten andere, ungeimpfte Gewitterzellen, ebenfalls analysiert werden, ob diese sich von geimpften Gewitterzellen, bezüglich ihrer Driftkennzahl $\Delta\beta_x$, unterscheiden. Die Driftkennzahl stellt eine Möglichkeit dar, die Effektivität von Hagelabwehrflügen in eine messbare Größe zu verwandeln. Die Driftkennzahlformel kann für zukünftige Beobachtungen noch weiterentwickelt werden, sodass die betrachteten Zeiträume *vor dem Flug*, *während dem Flug* und *nach dem Flug* für die Berechnung auch berücksichtigt werden. Der Radius oder die Fläche, der in Abbildung 37 gezeigten Geometrie, sollte hierbei auch miteinbezogen werden.

Der Ansatz aus Kapitel 3.1.1, zur Zufallszahlenerkennung, kann für die Echodaten, in der momentan vorliegenden Form, verworfen werden. Ein Grund für das Scheitern der Zufallszahlenerkennung kann sein, dass die Radarechodaten zu stark durch die Intensitätsstufen quantisiert werden und deswegen kein Nachweis der Zufälligkeit erfolgen kann. Eine weitere Begründung kann sein, dass die Intensitätsänderung einer Gewitterwolke tatsächlich nicht echt zufälligen Gesetzmäßigkeiten unterliegt. Der Ansatz mit den Impf- bzw. Beobachtungsgeometrien, aus Kapitel 2.3.3, wird als zu ungenügend erachtet, da hier die Geometrie auf die Flugzeugposition und nicht auf das Gewitterzentrum bezogen wird.

4 Ansatz zur Beurteilung der Hagelabwehr mithilfe von Versicherungsschadendaten

Für den Ansatz, zur Überprüfung der Wirksamkeit von Hagelabwehrflügen über die Versicherungsdaten, haben sich die Allianz Versicherung und die Versicherungskammer Bayern freundlicherweise dazu bereit erklärt, Daten zur Verfügung zu stellen. Aus Gründen der Geheimhaltung werden die Daten, der beiden Versicherer miteinander vermischt und eine gemeinsame Datenbasis gebildet. Es kann so kein Rückschluss auf die Versicherer erfolgen.

4.1 Datenstruktur

Die Versicherungsdatensätze der Allianz und VKB liegen in folgender Form vor:

	1 Jahr	2 Datum	3 LandkreisNr	4 geschaedigteFlaeche	5 Fruchtschluessel
1	2006	21.07.2006	9161	3.0200	Ölfrüchte
2	2006	21.07.2006	9161	1.0600	Ölfrüchte
3	2006	25.06.2006	9171	2.3900	Mais
4	2006	25.06.2006	9171	2.2600	Mais
5	2006	25.06.2006	9171	2.2600	Mais
6	2006	25.06.2006	9171	2.3900	Mais
7	2006	14.07.2006	9171	1.6700	Ölfrüchte

Abbildung 56: Datenstruktur der vorliegenden Daten

Die Daten wurden so vorbereitet, dass alle Schäden in einem Landkreis und an einem Tag zusammengefasst werden. Aus Abbildung 56 geht für den 21.07.2006 folgende Zusammenfassung hervor:

21.07.2006 9161 4.08 Ölfrüchte

4.2 Übersicht über die Versicherer

Die, von der Allianz und der VKB zur Verfügung gestellten, Daten erlauben ein genaues Ermitteln der versicherten Fläche, im Zeitraum 2006 bis 2018. Der Vollständigkeit halber wurde der Faktor F gebildet, was dem Quotienten aus der Versicherten Fläche im Nachbargebiet und der Versicherten Fläche im Schutzgebiet entspricht (siehe Abbildung 57). Das rasante Einbrechen des Faktors F , im Jahre 2015, ist darauf zurückzuführen, dass hier einer der beiden Versicherer keine Datensätze zur Verfügung stellen konnte. Die folgenden Kapitel 4.2.2 und 4.2.3 sollen eine Übersicht, über die Datenlage gewähren.

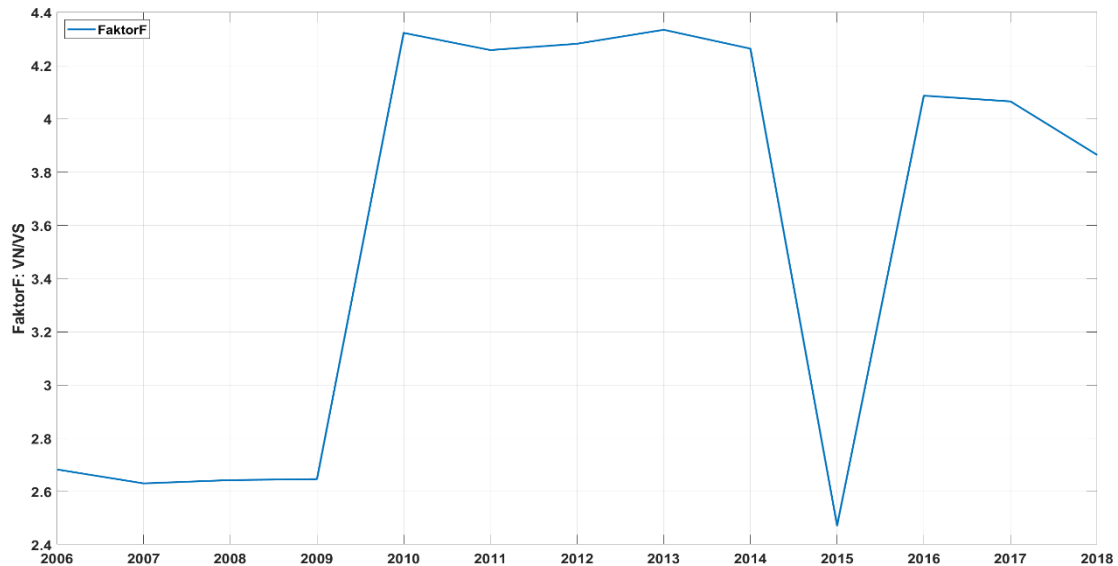


Abbildung 57: Faktor F für die gemeinsame Datenbasis

4.2.1 Extremereignisse

Die Extremereignisse wurden einer älteren Auswertung³ entnommen, welche sich mit der Wirksamkeit von Hagelabwehrflügen befasst hat. Ein Extremereignis ist hierbei ein Tag, an dem 50 % des Gesamtschadens auftritt. Diese Extremereignisse wurden im Datensatz bereinigt. Zu Extremereignisse kam es an folgenden Tagen:

15.06.2007

22.08.2008

23.07.2009

22.06.2011

10.07.2011

06.08.2013

09.06.2013

Im Kapitel 4.2.3 wurden diese Tage daher ausgenommen.

³ Analyse von Hagelschäden
im Zusammenhang mit Hagelabwehrflügen
im Raum Rosenheim [P.Zentgraf]

4.2.2 Auswertung aller Schadenstage

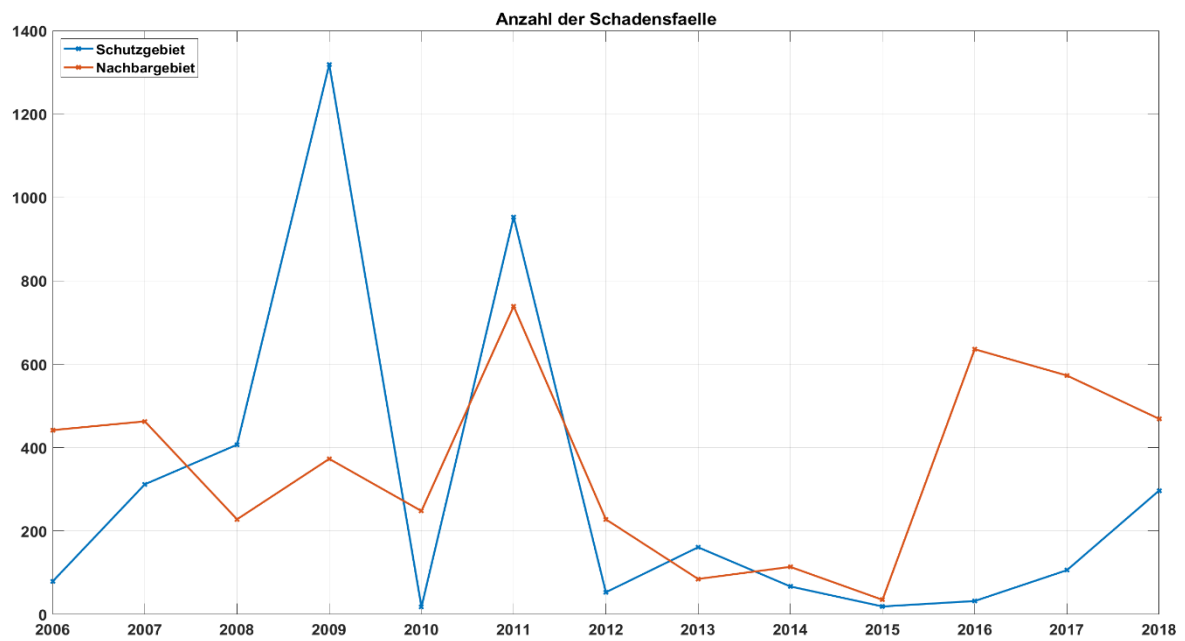


Abbildung 58: Anzahl der Schadensfaelle 2006-2018

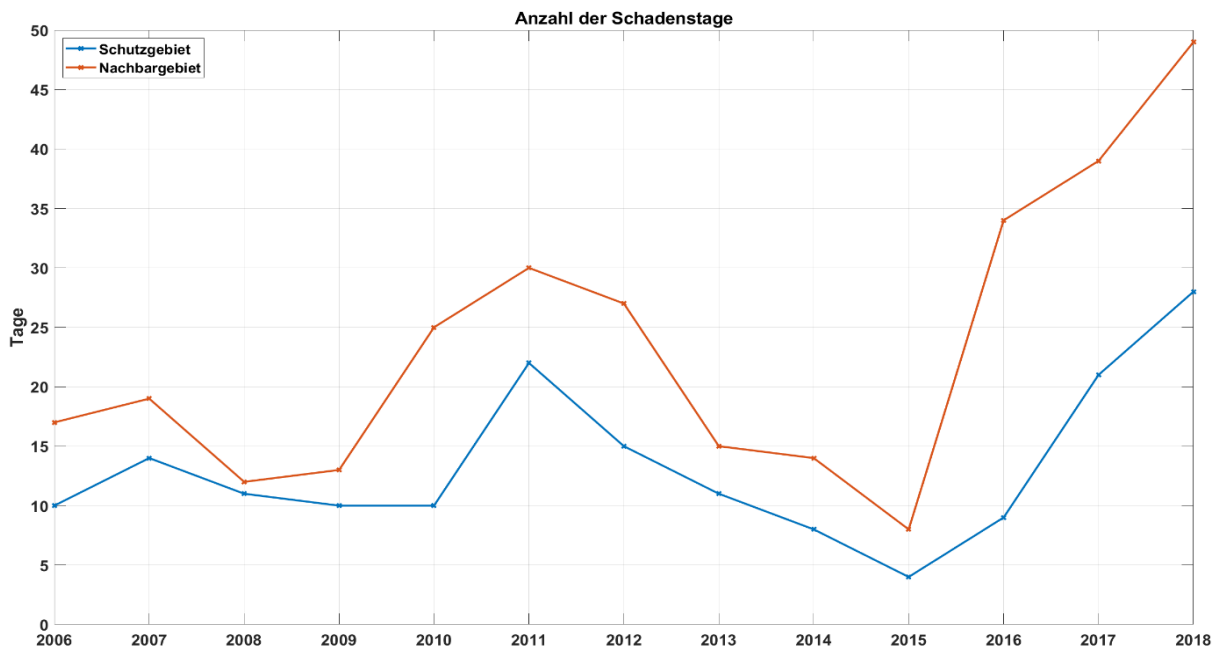


Abbildung 59: Anzahl der Schadenstage

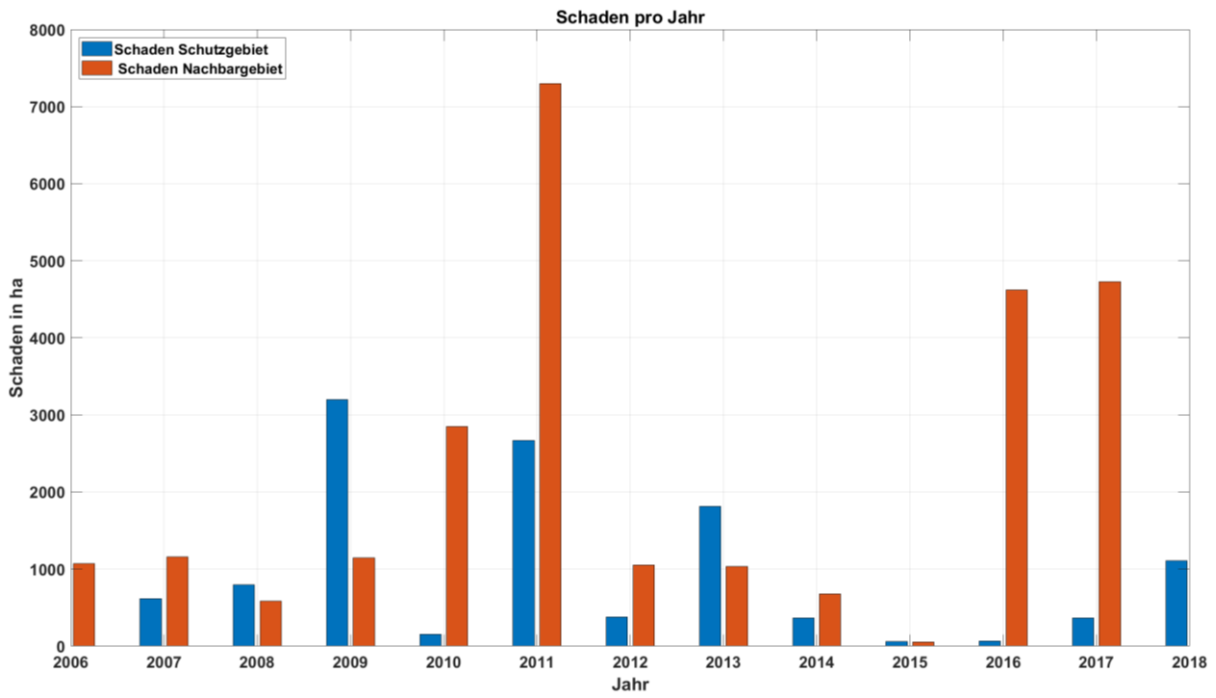


Abbildung 60: Schaden pro Jahr 2006-2018

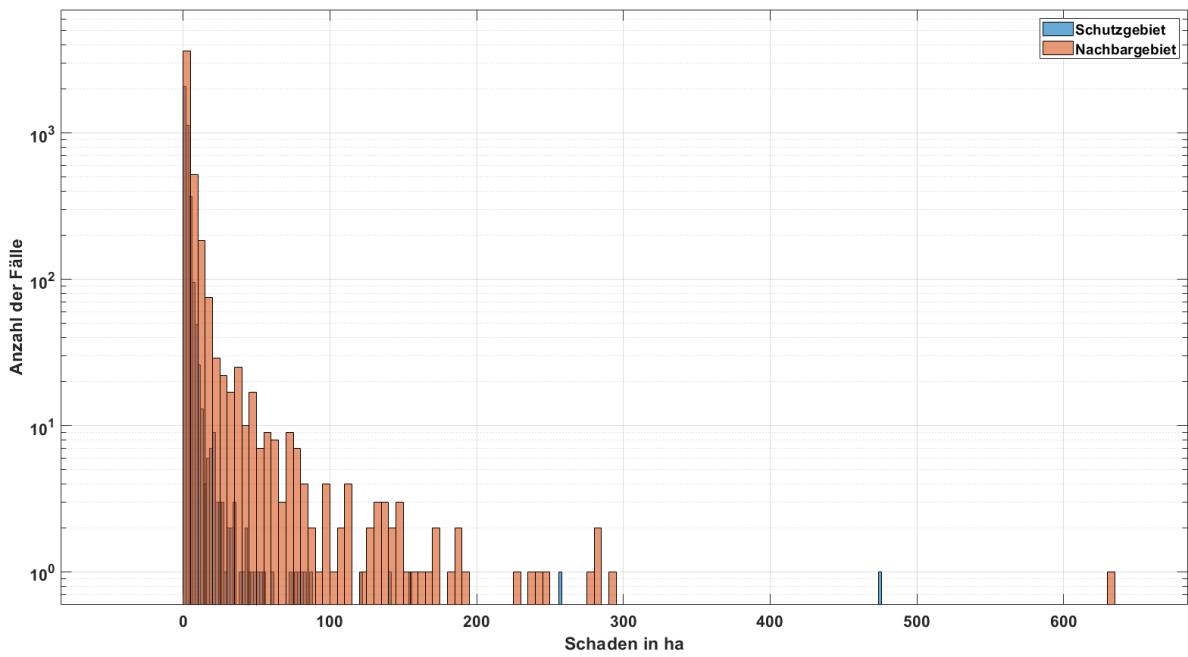


Abbildung 61: Häufigkeit der Schadengröße

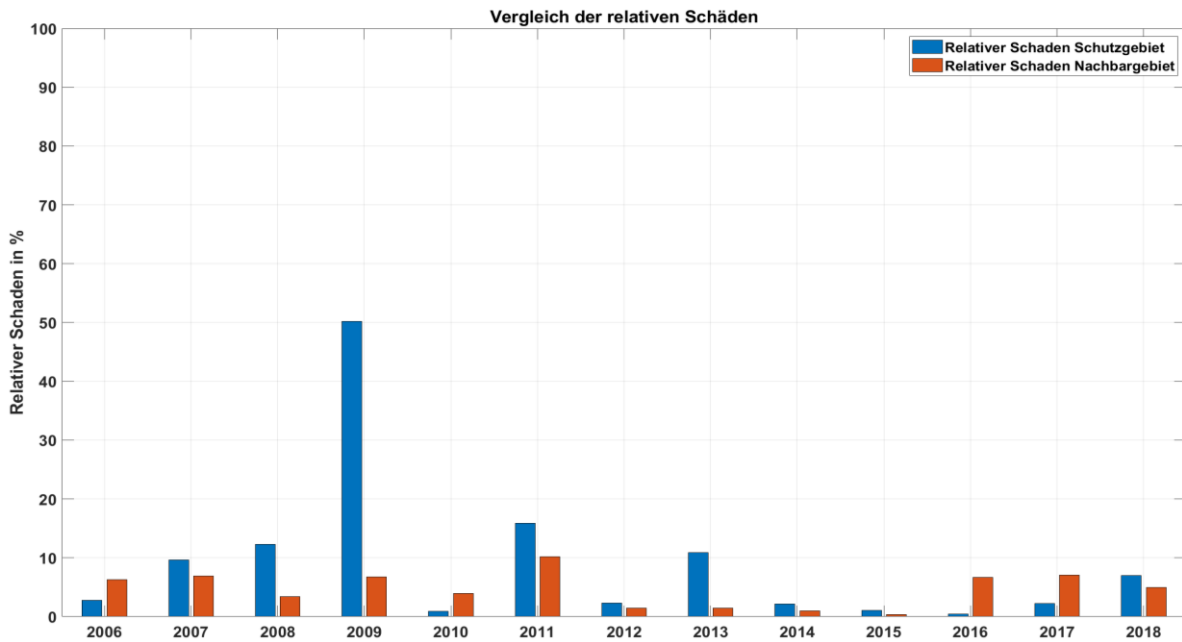


Abbildung 62: Relativ geschädigte Fläche 2006-2018

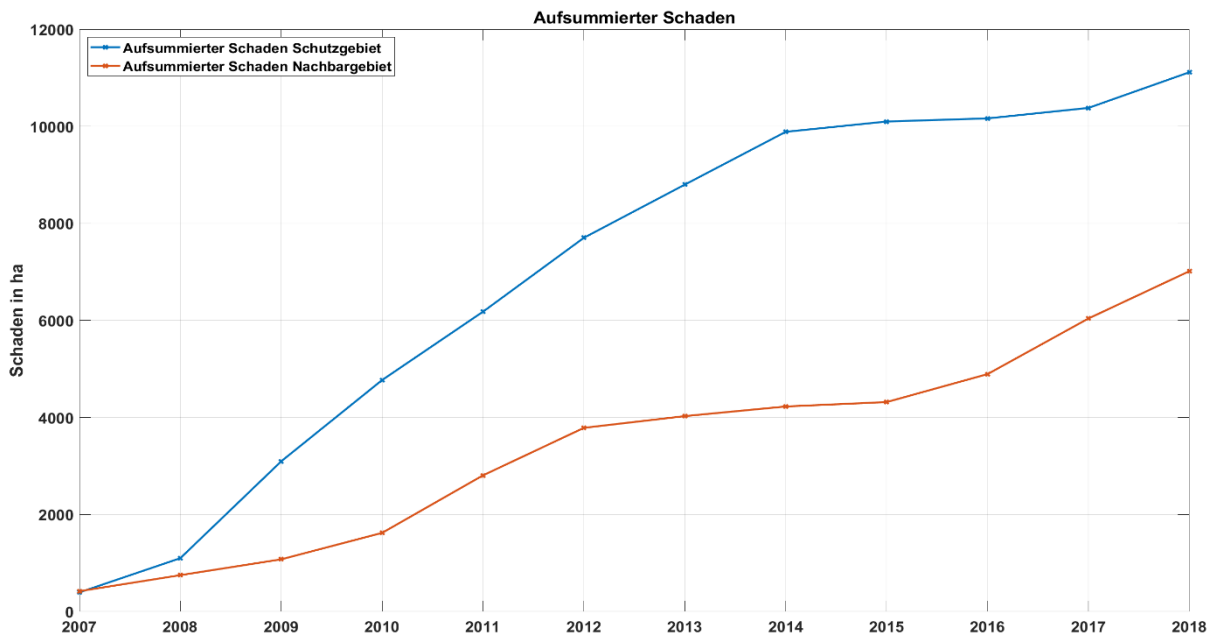


Abbildung 63: Aufsummierter Schaden

Der aufsummierte Schaden aus Abbildung 63 wurde für das Nachbargebiet um den Faktor F bereinigt. Und errechnet sich folgendermaßen:

$$\text{Aufsummierter Schaden Nachbargebiet} = \int \frac{\text{Schaden pro Jahr}}{\text{Faktor } F}$$

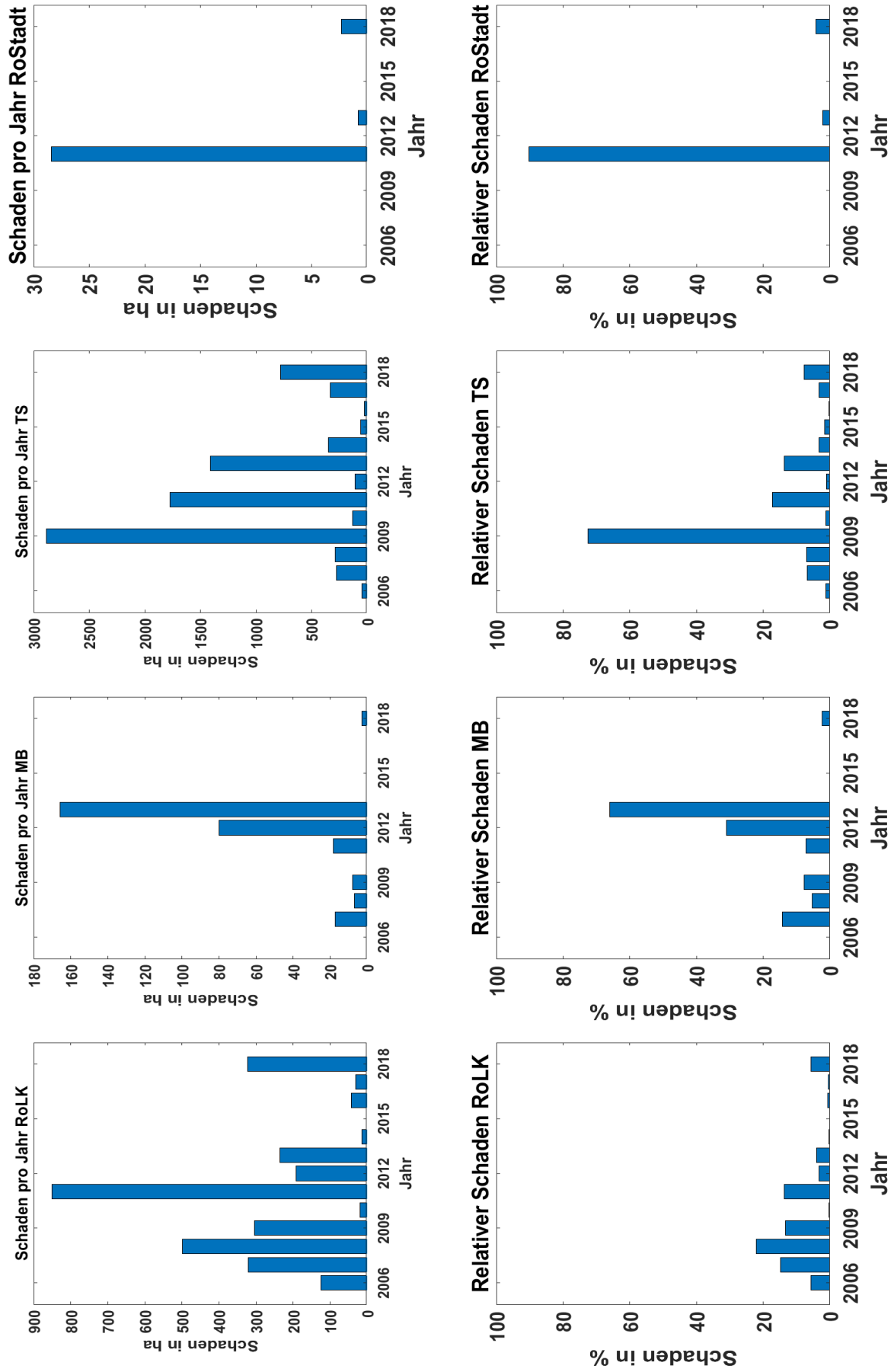


Abbildung 64: Detailplot für das Schutzgebiet

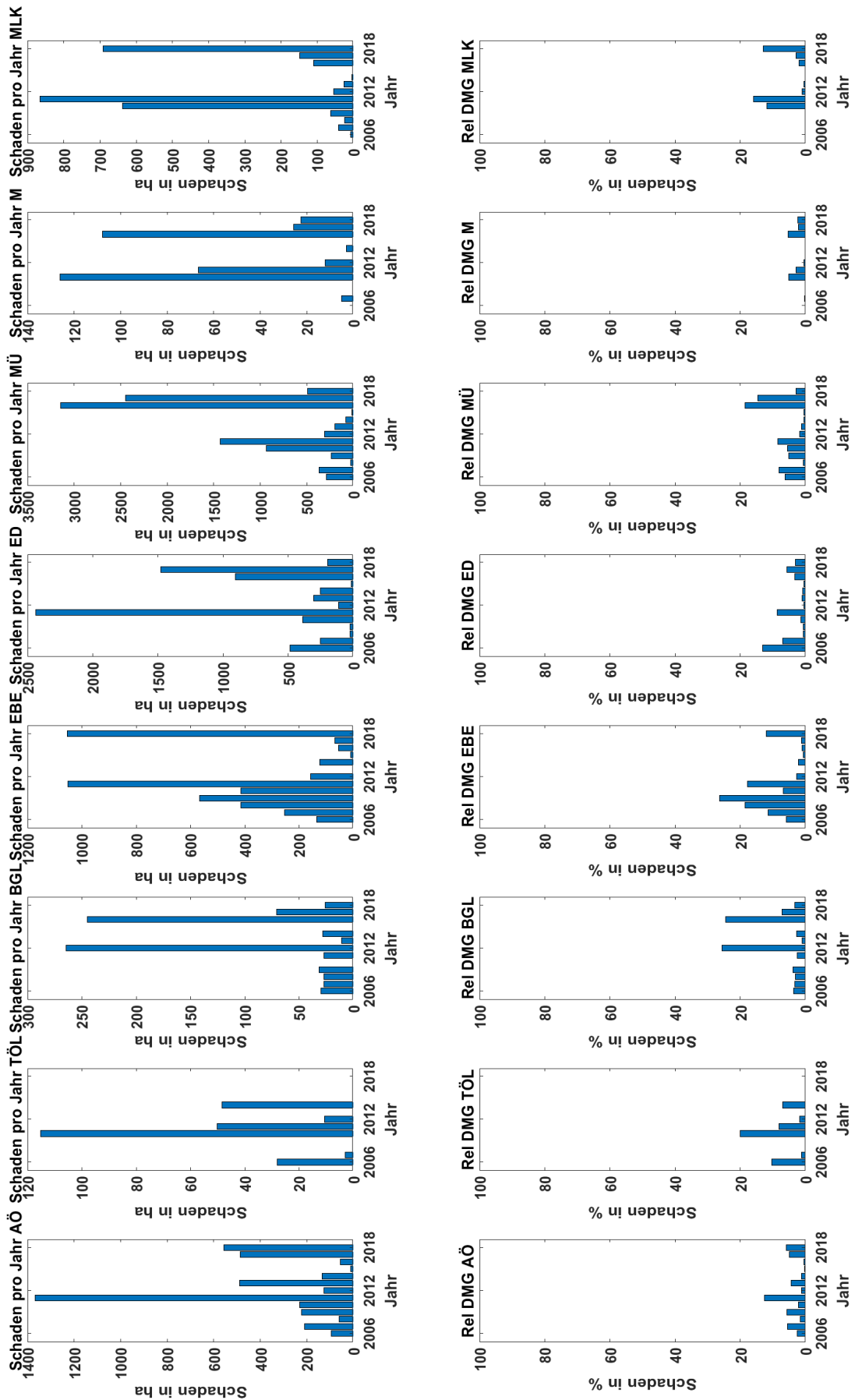


Abbildung 65: Detailplot des Nachbargebietes

4.2.3 Auswertung aller guten Flugtage ohne Extremereignisse

Um bessere Ergebnisse zu erzielen wurde die Datenbasis so sortiert, dass nur gute Flugtage, ohne die oben genannten Extremereignisse, verglichen wurden. Bei „guten“ Flugtagen handelt es sich um Aussagen der Piloten, welche, nach einem geflogenen Hagelabwehrflug, den Flug als erfolgreich oder nicht erfolgreich eingetragen haben. Ein Grund für einen nicht erfolgreichen Flug kann das zu späte starten des Hagelabwehrfluges oder der Ausfall eines AgJ-Generators sein.

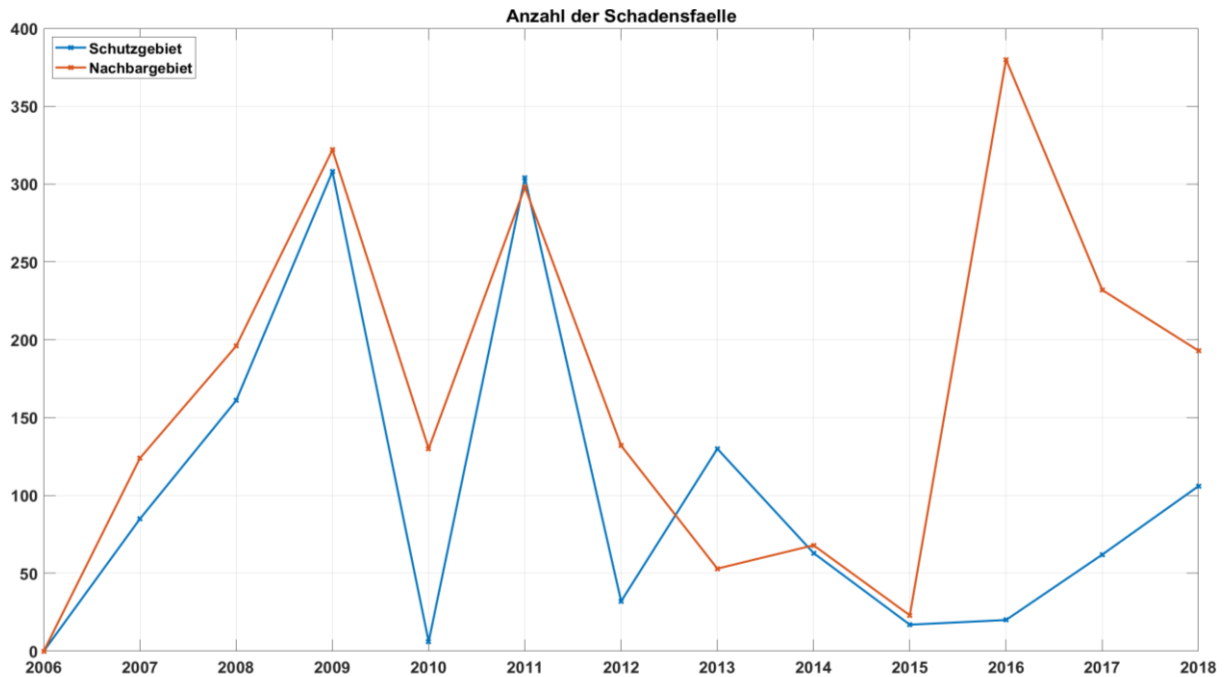


Abbildung 66: Anzahl der Schadensfälle 2006 bis 2018. Gute Flugtage ohne Extremereignisse

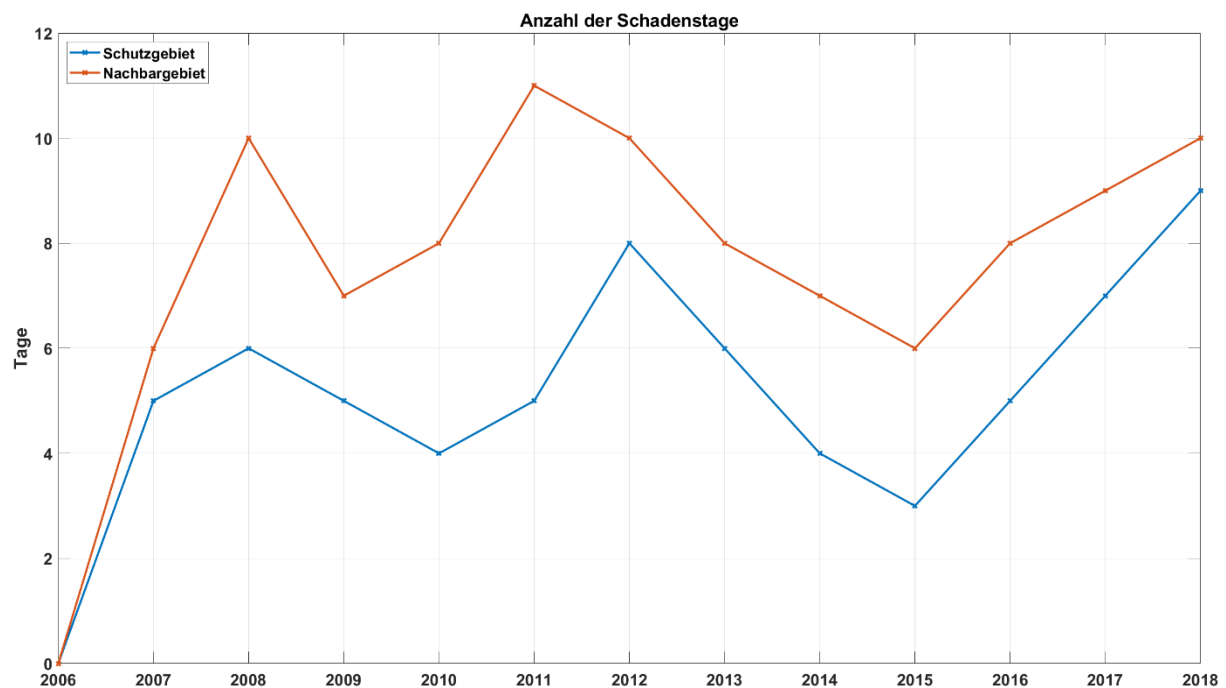


Abbildung 67: Anzahl der Schadenstage 2006 bis 2018. Gute Flugtage ohne Extremereignisse

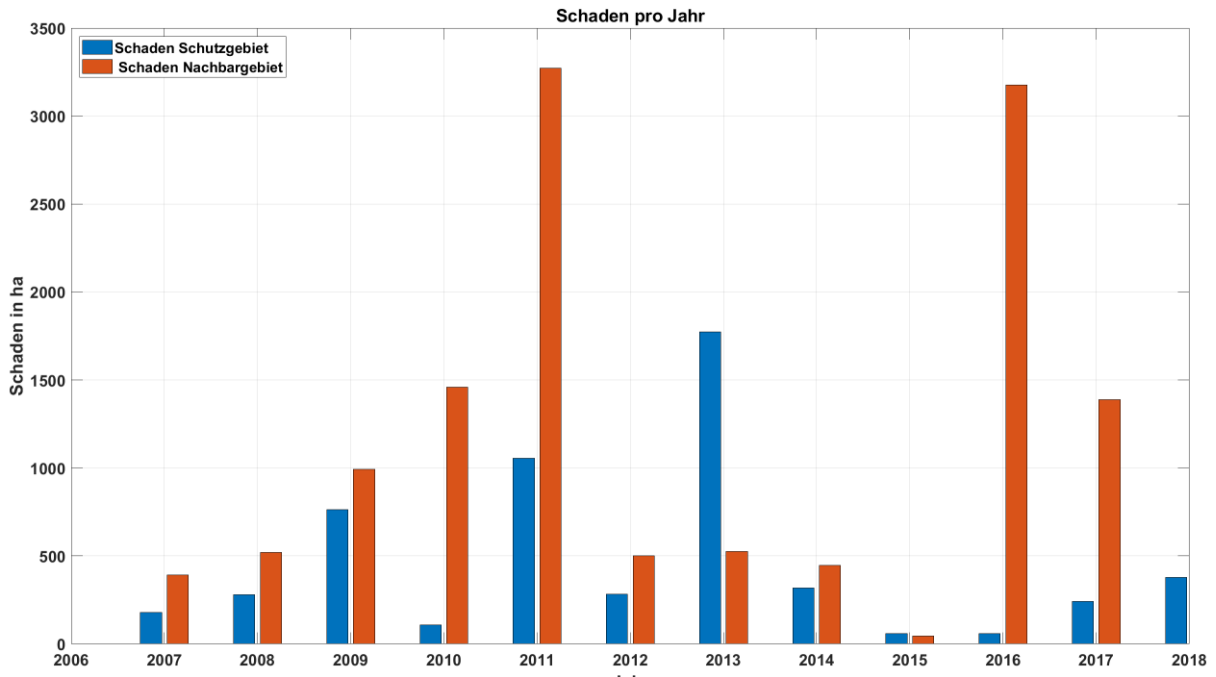


Abbildung 68: Schaden pro Jahr. Gute Flugtage ohne Extremereignisse

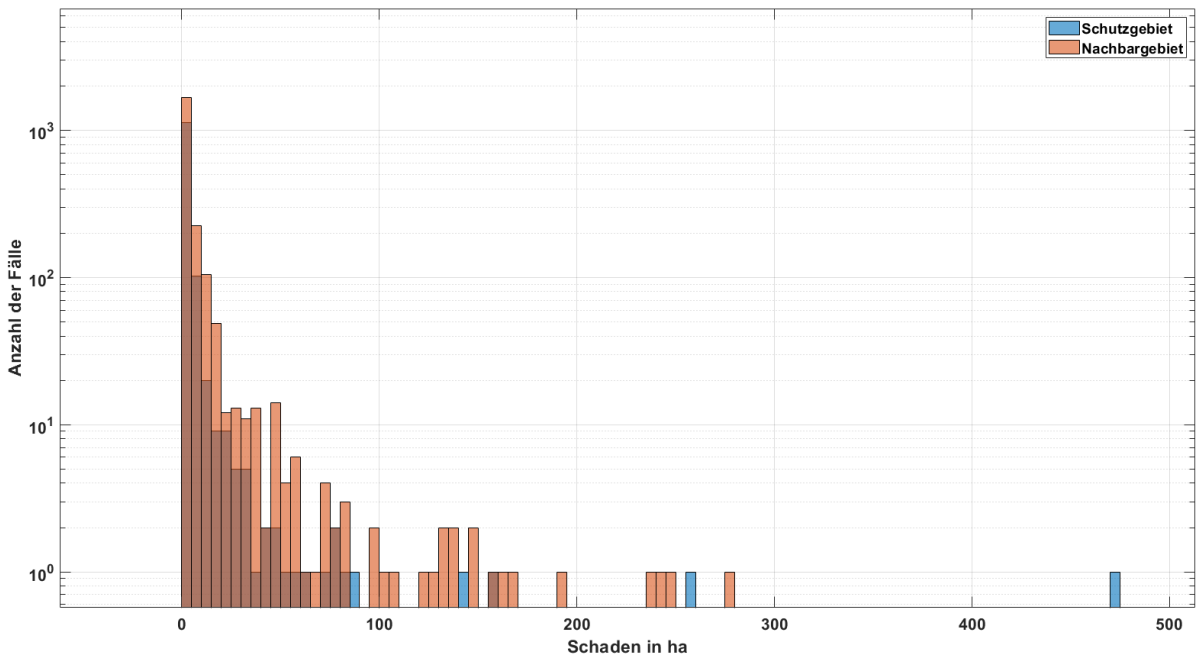


Abbildung 69: Häufigkeit der Schadengröße

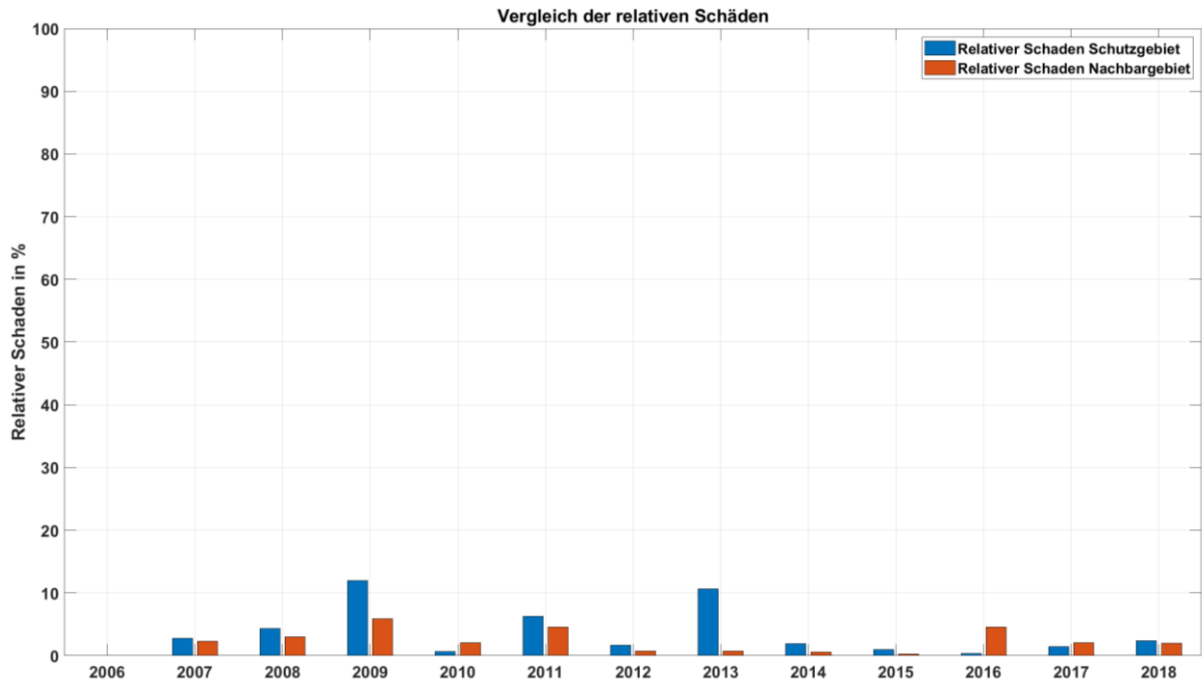


Abbildung 70: Vergleich der relativen Schäden. Gute Flugtage ohne Extremereignisse

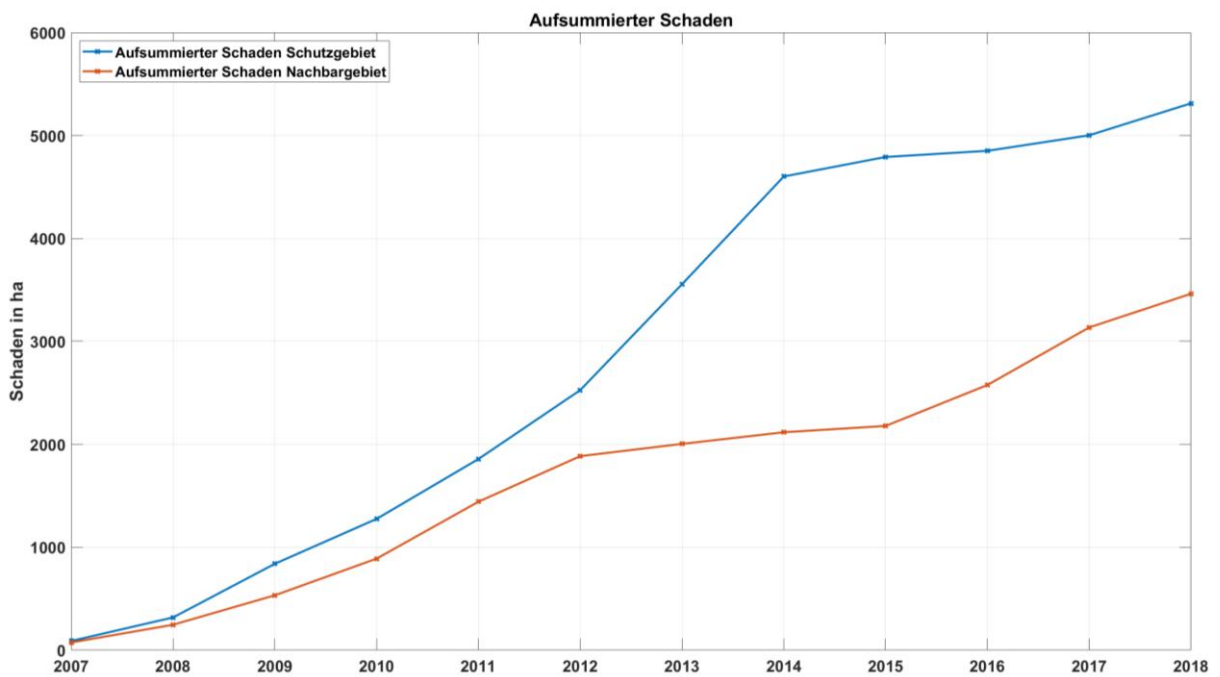


Abbildung 71: Aufsummierter Schaden. Gute Flugtage ohne Extremereignisse

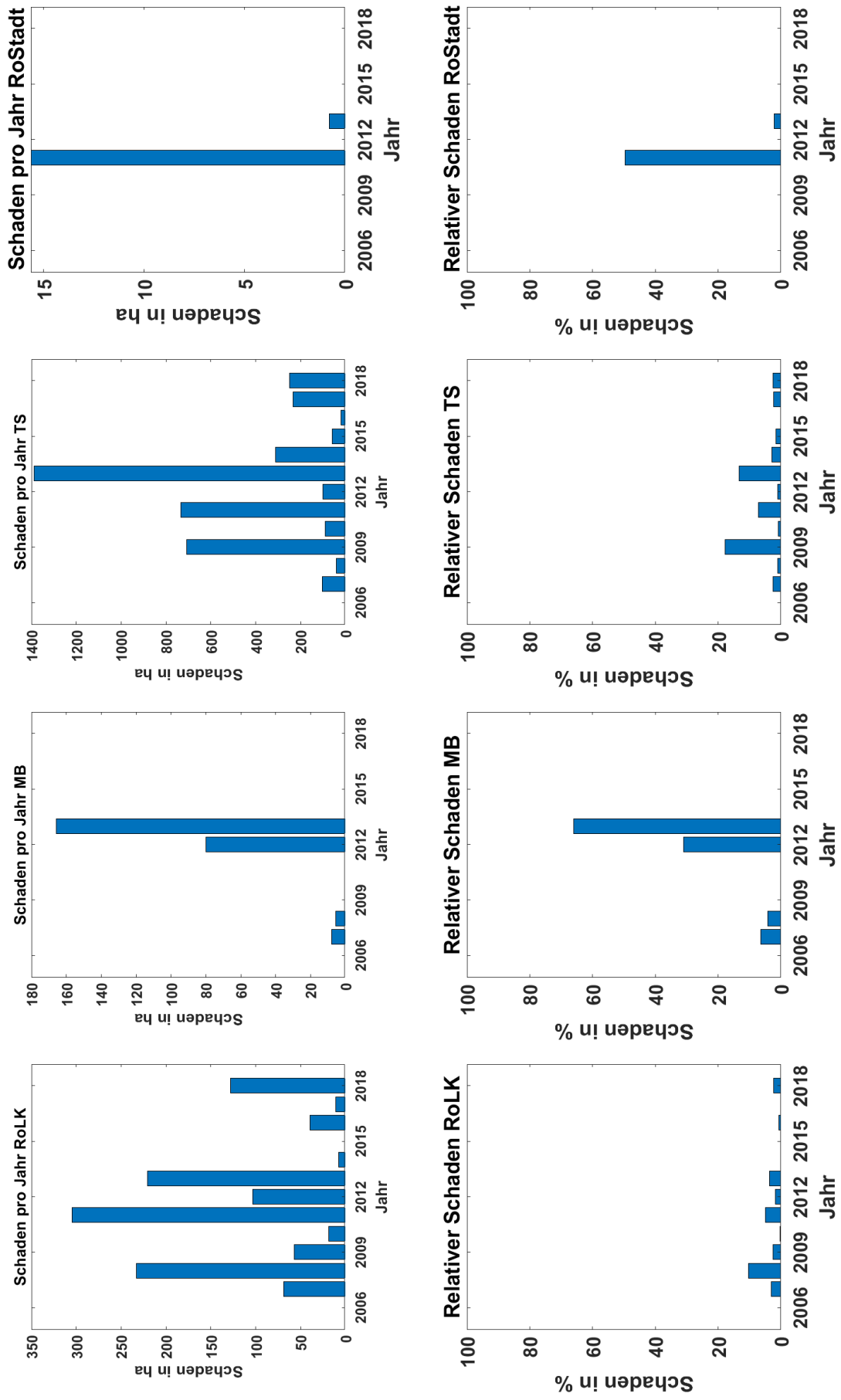


Abbildung 72: Detailplot Schutzgebiet. Gute Flugtage ohne Extremereignisse

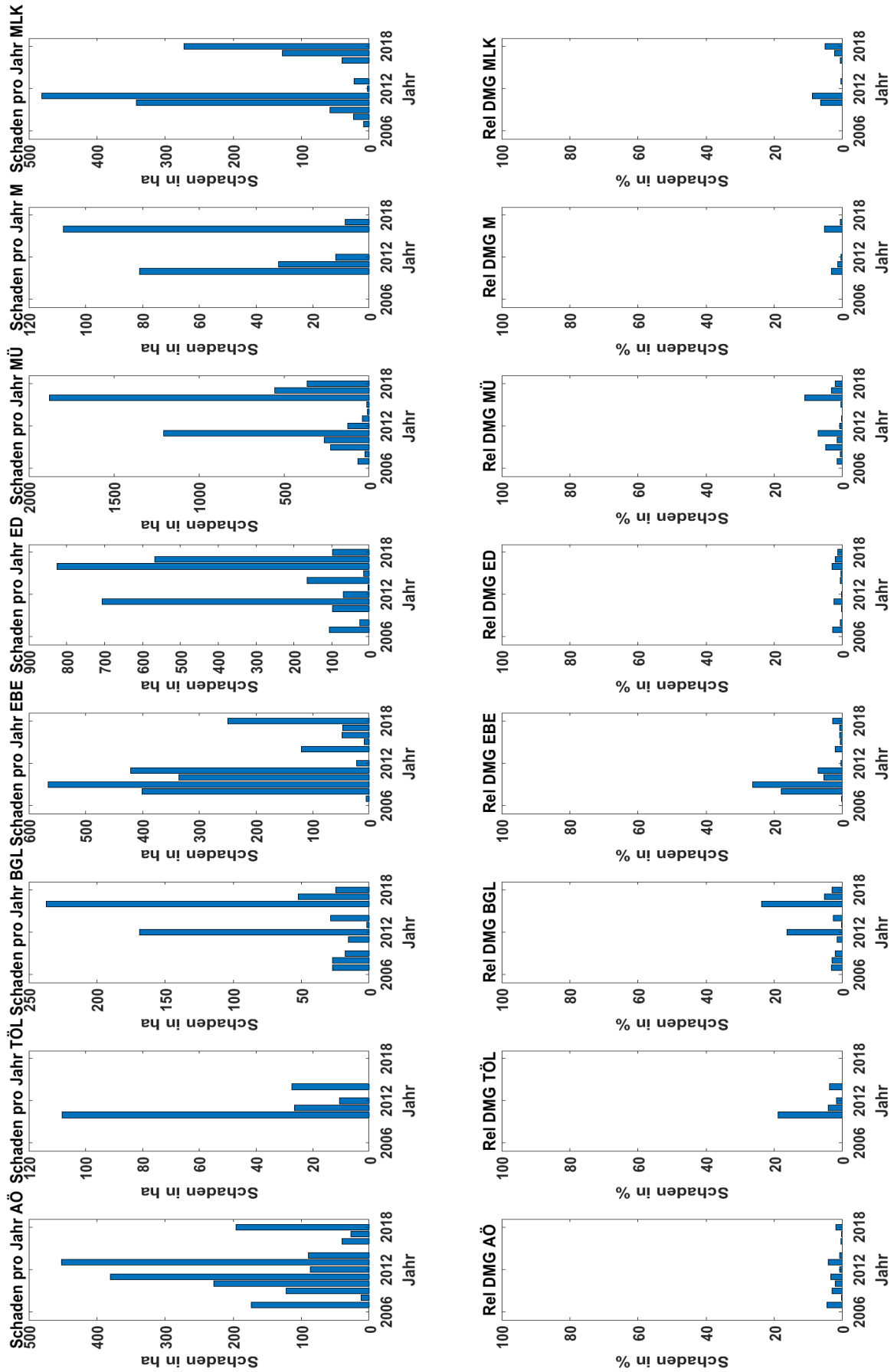


Abbildung 73: Detailplot Nachbargebiet

4.3 Multiple lineare Regressionsanalyse der Versicherungsdaten

Um eine Reduzierung des Schadens durch die Hagelabwehr nachzuweisen, wird eine multiple lineare Regressionsanalyse durchgeführt.

Die versicherte Fläche für jeden Landkreis ist bekannt, daher kann eine relative Schädigung gebildet werden. Die relative Schädigung ist die abhängige Variable (Responsevariable), für welche Prädiktoren (Regressoren) bestimmt werden sollen. Es wird eine multiple lineare Regressionsanalyse für alle Flugtage durchgeführt. Daher werden nur Tage betrachtet, an denen ein Schaden entstanden ist und mindestens ein Flug stattgefunden hat.

Als Prädiktoren (Regressoren) für die relative Schädigung (Responsevariable) wurden folgende Variablen gewählt:

- Landkreisnummer (als Faktor codiert)
- Monat der Schädigung (als Faktor codiert)
- Schutzgebiet (als Faktor codiert, boolesche Variable)
- Flugdauer in Minuten
- Ausgebrachtes AgJ in Liter
- Anzahl der Flüge
- Erfolg des Fluges (als Faktor codiert, entweder 0,1 oder gemischt⁴)
- Alle Fruchtarten (als Faktor codiert, entweder 0 oder 1)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
date	LK	RelDmg	Month	SGebiet	FlugdauerMin	AgJinL	FlugAnzahl	FlugErfolgreich	Futterpflanzen	Ganzpflanzenslage	Getreide	Kartoffeln	Mais	Raps	Spargel	Obst	SonstigeFruchtarten	Olfruchte	Hulsenfruchte	ErbsenBohnen	Ruben	Gemüse
12.07.2006	9187	1,006466877	7	1	58	9	1	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0
14.07.2006	9187	0,041009464	7	1	164	25	2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
14.07.2006	9189	0,051735016	7	1	164	25	2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
05.05.2007	9187	0,021283255	5	1	236	37	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
26.05.2007	9189	0,468075117	5	1	200	30	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15.06.2007	9182	0,145696401	6	1	183	36	2	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0
15.06.2007	9187	3,145226917	6	1	183	36	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20.06.2007	9189	0,773082942	6	1	168	31	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
25.06.2007	9187	0,752895149	6	1	368	66	4	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
25.06.2007	9189	0,352269171	6	1	368	66	4	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0
20.07.2007	9182	0,122222222	7	1	193	50	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
20.07.2007	9187	0,298591549	7	1	193	50	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
02.06.2008	9187	0,00942813	6	1	247	40	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
25.06.2008	9189	0,305873261	6	1	259	38	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
03.07.2008	9189	0,03446677	7	1	276	35	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0

Abbildung 74: Datensatz für die Multiple lineare Regression, der Datensatz enthält 334 Einträge

Mithilfe der multiplen linearen Regression lässt sich der Einfluss von den verschiedenen Einflussvariablen (Prädiktoren) auf eine Responsevariable nachweisen.

Die multiple lineare Regression wird durch folgende Gleichung beschrieben:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + e$$

Die Parameter $\beta_0, \beta_1, \beta_2 \dots$ bezeichnet unbekannte Parameter, welche den Einflusswert auf die Ergebnisvariable Y festlegen. Bei der einfachen linearen Regression kann β_0 auch als y-Achsenabschnitt einer Geraden und β_1 als die Steigung dieser Geraden verstanden werden, welche eine Punktwolke nähert. Der Parameter e bezeichnet einen Fehlerterm, welcher nicht kontrolliert werden kann. Um an die Thematik der multiplen linearen Regression hinzuführen, soll zuerst eine einfache lineare Regression gezeigt werden.

Für die multiple lineare Regression soll die ausgebrachte AgJ-Menge als Prädiktor verwendet werden. Bei einigen Flügen ist die ausgebrachte AgJ-Menge jedoch unbekannt, da nach einem absolvierten Hagelabwehrflug vergessen wurde, diesen in das Flugprotokoll einzutragen. Deshalb wird eine einfache

⁴ Wenn es mehr als einen Flug gab und einer der beiden Flüge als nicht erfolgreich gewertet wurde, dann ist der Erfolg des Gesamteinsatzes als „gemischt“ zu werten.

lineare Regression durchgeführt um diese fehlenden AgJ-Mengen zu bestimmen. Es soll der Zusammenhang zwischen der ausgebrachten AgJ-Menge und der Flugzeit untersucht werden. Dafür wird angenommen, dass ein linearer Zusammenhang zwischen Flugzeit und ausgebrachter AgJ-Menge besteht.

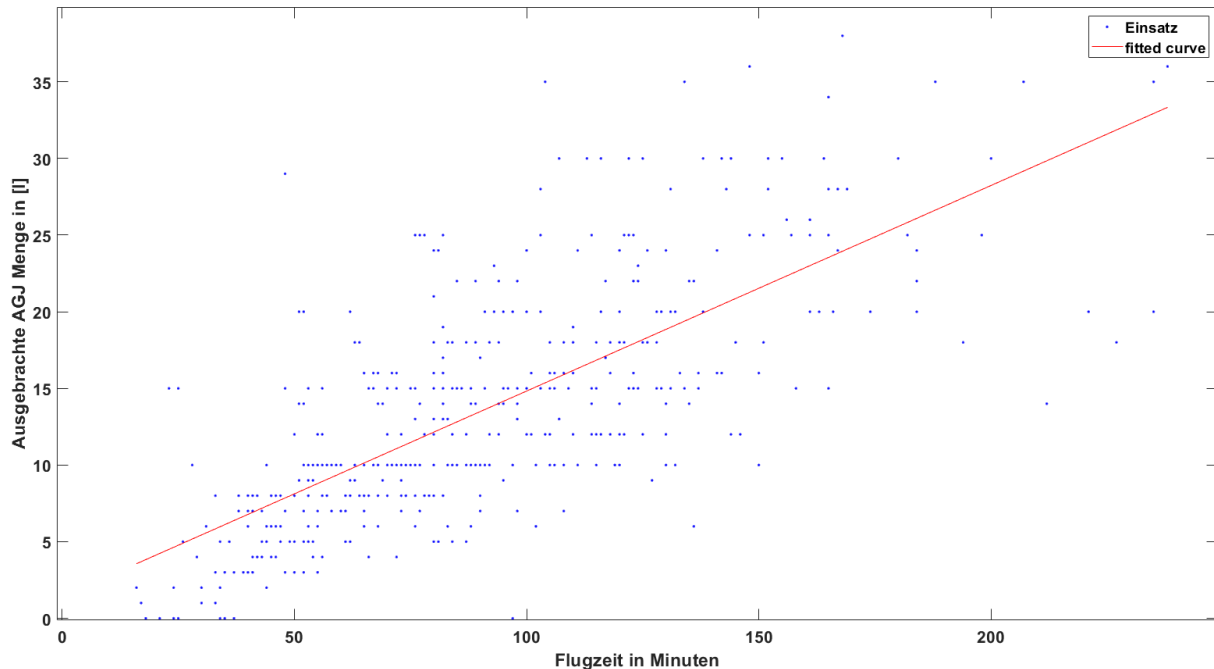


Abbildung 75: Einfache lineare Regression für die ausgebrachte AgJ-Menge in Abhängigkeit der Flugzeit

Die Koeffizienten des einfachen linearen Regressionsmodell, für die ausgebrachte AgJ-Menge, ergeben die in Abbildung 75 zu sehende Gerade. Fehlende Werte für die ausgebrachte AgJ-Menge werden nach der daraus folgenden Geradengleichung bestimmt:

$$Y = \beta_0 + \beta_1 x_1$$
$$AgJ = 1.417 + 0.1341 * Flugzeit$$

Das Prinzip für die multiple lineare Regression funktioniert gewissermaßen ähnlich, nur kann hier bei einem Modell, mit mehr als zwei Prädiktoren, keine Visualisierung mehr erfolgen.

Im Folgenden ist die Auswertung der multiplen linearen Regressionsanalyse zu sehen. Um deren Bedeutung verstehen zu können, sollen alle notwendigen Koeffizienten in den folgenden Kapiteln erklärt werden. Auf nachfolgender Seite ist das Ergebnis der multiplen linearen Regressionsanalyse zu sehen.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.435964	0.961634	0.453	0.6506
LK9163	-0.577292	1.300410	-0.444	0.6574
LK9171	0.081039	0.742466	0.109	0.9132
LK9172	0.012776	0.813124	0.016	0.9875
LK9173	-0.062710	0.961848	-0.065	0.9481
LK9175	0.129131	0.745309	0.173	0.8626
LK9177	0.078742	0.757478	0.104	0.9173
LK9182	-0.335531	1.026003	-0.327	0.7439
LK9183	0.113298	0.737787	0.154	0.8781
LK9184	0.053371	0.758793	0.070	0.9440
LK9187	0.242158	0.894216	0.271	0.7867
LK9189	0.949099	0.868688	1.093	0.2755
Month6	-0.011116	0.307099	-0.036	0.9711
Month7	0.052770	0.295891	0.178	0.8586
Month8	0.010973	0.341454	0.032	0.9744
Month9	-0.072024	1.096142	-0.066	0.9477
FlugdauerMin	-0.003760	0.004106	-0.916	0.3605
AgJinL	0.049153	0.021730	2.262	0.0244 *
FlugAnzahl	-0.247958	0.285552	-0.868	0.3859
Futterpflanzen	0.708511	3.171038	0.223	0.8234
Ganzpflanzensilage	-0.246247	1.075978	-0.229	0.8191
Getreide	0.140051	0.366451	0.382	0.7026
Kartoffeln	-0.245061	0.564485	-0.434	0.6645
Mais	-0.619073	0.484771	-1.277	0.2026
Raps	-0.017122	0.588588	-0.029	0.9768
Spargel	0.019659	1.993078	0.010	0.9921
Obst	0.008197	1.303559	0.006	0.9950
SonstigeFruchtarten	-0.061056	0.331924	-0.184	0.8542
Olfruchte	0.063326	0.342180	0.185	0.8533
Hulsenfruchte	-0.185600	0.491313	-0.378	0.7059
ErbsenBohnen	0.171988	0.574904	0.299	0.7650
Ruben	-0.200177	1.439872	-0.139	0.8895
Gemuse	0.167476	1.203872	0.139	0.8895

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

Residual standard error: 1.764 on 301 degrees of freedom
 Multiple R-squared: 0.1003, Adjusted R-squared: 0.004688
 F-statistic: 1.049 on 32 and 301 DF, p-value: 0.3999

4.3.1.1 Estimate:

Die Spalte für den *Estimate* $\hat{\beta}$ enthält *Anzahl der Prädiktoren+1* Zeilen. Der Koeffizient *Intercept* ist hierbei der Schnittpunkt einer Geraden mit der y-Achse (bei einer einfachen linearen Regression). Der Schnittpunkt ist der erwartete Wert der relativen Schädigung, wenn die Mittelwerte aller Prädiktoren berücksichtigt werden. Die durchschnittliche relative Schädigung beträgt hier 0,435964. Ein Prädiktor mit einem positiven Vorzeichen im *Estimate* lässt einen positiven Einfluss auf die Ergebnisvariable zu. Zum Beispiel liefert der Prädiktor *FlugdauerMin*, für jede theoretisch länger geflogene Minute, eine Reduzierung der relativen Schädigung um 0,3760 %.

Die Berechnung des *Estimate* $\hat{\beta}$ erfolgt üblicherweise mittels Matrizenberechnung.

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & X_{1,1} & X_{1,2} & \dots & X_{1,p} \\ 1 & x_{2,1} & X_{2,2} & \dots & X_{2,p} \\ 1 & X_{3,1} & X_{3,2} & \dots & X_{3,p} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{n,1} & X_{n,2} & \dots & X_{n,p} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, e = \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_{1,1} + \beta_2 X_{1,2} + \beta_p X_{1,p} + e_0 \\ \beta_0 + \beta_1 X_{2,1} + \beta_2 X_{2,2} + \beta_p X_{2,p} + e_1 \\ \beta_0 + \beta_1 X_{3,1} + \beta_2 X_{3,2} + \beta_p X_{3,p} + e_2 \\ \vdots \\ \beta_0 + \beta_1 X_{n,1} + \beta_2 X_{n,2} + \beta_p X_{n,p} + e_n \end{bmatrix}$$

Die Responsevariable ist für die multiple lineare Regressionsanalyse eine $n \times 1$ Matrix. Die Einflussgrößenmatrix (enthält alle Prädiktoren Werte) muss zu einer $n \times (p + 1)$ Matrix zusammengefasst werden, n beschreibt hierbei die Anzahl der Beobachtungen und p die Anzahl der Prädiktoren. Die erste Spalte in der Einflussgrößenmatrix wird durchgehend mit 1 belegt, mit dieser Spalte wird unter anderem der *Intercept* β_0 errechnet.

Ein besonderes Augenmerk liegt auf der $n \times 1$ Matrix für die unabhängigen Zufallsfehlervariablen e , da dem linearen Modell der Regressionsanalyse die Vorstellung zugrunde liegt, dass jede Beobachtung Y der Responsevariable um den von den Regressoren erzeugten Wert $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$ schwankt. Die Differenz zwischen geschätztem Wert und tatsächlichem Wert wird auch Residuen genannt. Dieser, von den Regressoren erzeugte, Wert kann auch in der Fehlerquadratsumme $SQE(\beta)$ ausgedrückt werden. Ziel ist es nun, die einzelnen Parameter $\beta_0, \beta_1 \dots \beta_n$ so zu bestimmen, dass die Fehlerquadratsumme möglichst klein wird (Pruscha, Statistisches Methodenbuch, 2006).

$$SQE(\beta) = \sum_{i=1}^n (Y_i - (\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n))^2$$

$$SQE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

\hat{Y}_i kennzeichnet hierbei das Ergebnis der Schätzung für die bestimmten Koeffizienten von $\hat{\beta}$.

$$\beta = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{\widehat{Cov}(X, Y)}{\widehat{Var}(X)}$$

4.3.1.2 Std. Error σ_e :

Der Koeffizient *Standardfehler* misst den Durchschnittswert, den der Estimatewert vom tatsächlichen Durchschnittswert, der Reaktionsvariablen, abweicht. Es wird idealerweise eine niedrigere Zahl, im Verhältnis zu den Estimatekoeffizienten, erwartet.

4.3.1.3 *t-Value*:

Der Koeffizient *t-Value* ist ein Maß dafür, wie viele Standardabweichungen die Koeffizientenschätzung von 0 entfernt ist. Es wird erwartet, dass die Schätzung weit von Null entfernt ist, da dies bedeuten würde, dass ein Zusammenhang zwischen Prädiktor und relativer Schädigung besteht. Bei dieser Analyse sind die *t*-statistischen Werte relativ weit von 0 entfernt und klein, im Vergleich zum Standardfehler, was darauf hindeuten kann, dass keine Beziehung zwischen Prädiktor und Response besteht. Im Allgemeinen werden *t-Value* auch zur Berechnung von *p-Werten* verwendet. Der *t-Value* errechnet sich folgendermaßen:

$$t_n = \frac{\beta_n}{\sigma_{e_n}}$$

4.3.1.4 *Pr(>t)*:

In der Modellausgabe gibt der Koeffizient *Pr(>t)* die Wahrscheinlichkeit an, mit der ein Wert gleich oder größer als *t* zu beobachten ist. Ein kleiner *p*-Wert zeigt an, dass es unwahrscheinlich ist, dass eine zufällige Beziehung zwischen den Prädiktorvariablen und der Responsevariable, der relativen Schädigung beobachtet werden kann. Typischerweise ist ein *p*-Wert von kleiner gleich 5 % ein gutes Signifikanzniveau. In dieser Auswertung liegen die *p*-Werte sehr weit von 0 entfernt, was zur Aussage führt, dass fast keiner der verwendeten Prädiktorvariablen für diesen Datensatz einen Einfluss zulässt. Die Signifikanz Codes *Pr(>t)*, die jeder Schätzung zugeordnet sind, markieren die Prädiktoren, welche den größten Einfluss auf die relative Schädigung haben. Drei „*“ neben einem Prädiktor stehen für einen hoch signifikanten *p*-Wert.

4.3.1.5 *Residual Standard Error*:

Prinzipiell besitzt jedes Modell einen unkontrollierbaren Fehlerterm, durch welchen logischerweise keine perfekte Vorhersage der Responsevariable, mit den zur Verfügung stehenden Prädiktoren, möglich ist. Der *Residual Standard Error* gibt die mittlere Abweichung für alle Prädiktoren vom tatsächlichen Modell an, welches ein Maß für die Qualität des erzeugten statistischen Modells ist. Für die Vorhersage der relativen Schädigung mit diesem Datensatz würde die Abweichung 1,764 % betragen.

$$RSE = \sqrt{\frac{\sum(Y - \hat{Y})^2}{n - 2}}$$

4.3.1.6 Multiple R-squared, Adjusted R-squared

Der R^2 -Koeffizient (Bestimmtheitsmaß) ist ein Maß dafür, wie gut das Modell zu den tatsächlichen Daten passt. R^2 ist ein Maß für die Beziehung zwischen den Prädiktorvariablen und der Responsevariable. R^2 liegt immer zwischen 0 und 1 (d.h. eine Zahl nahe 0 stellt eine Regression dar, die die Varianz in der Responsevariable nicht gut erklärt und eine Zahl nahe 1 erklärt die beobachtete Varianz in der Responsevariable). In dieser Auswertung beträgt der R^2 -Koeffizient 8,932 %. Das heißt, dass 8,932% der Datenvarianz durch die Prädiktoren erklärt werden kann. Es kann auch vereinfacht gesagt werden, dass der R^2 -Koeffizient ein Maß für den Beitrag der Prädiktoren an der Variation der Responsevariable ist. Der $R_{Adjusted}^2$ -Koeffizient ist eine verbesserte Version des einfachen R^2 -Koeffizienten. Der $R_{Adjusted}^2$ besteht aus dem einfachen R^2 , nur wird dieser für jede zusätzliche hinzugefügte Variable (Prädiktor) beeinflusst. Der $R_{Adjusted}^2$ kann so verstanden werden, dass dieser nur bei hinzufügen von Prädiktoren anwächst, welche das Modell, bezüglich seiner Güte, deutlich verbessern. Bei Hinzufügen von wirkungslosen Prädiktoren kommt es zu einer Verschlechterung vom $R_{Adjusted}^2$. Folgende Formel stellt die Berechnung vom $R_{Adjusted}^2$ dar, wobei n die Anzahl der Beobachtungen und p die Anzahl der unabhängigen Variablen (Prädiktoren) kennzeichnet:

$$R_{Adjusted}^2 = 1 - (1 - R^2) * \frac{n - 1}{n - p - 1}$$

4.3.1.7 F-Statistics und p-Value

Der F -Statistic Wert kann als Verhältnis zwischen der Anzahl an bestimmaren Varianzen und unbestimmaren Varianzen im Modell interpretiert werden. Je kleiner F -Statistic, desto größer sind die Varianzen im Modell, welche nicht durch die vorhandenen Prädiktoren bestimmt werden können. Der p -Value ist ein Gesamtmaß für die Signifikanz des erstellten Modells. Bei der hier vorliegenden Auswertung wird ein p -Value von 39,99 % erzielt, das bedeutet, dass 39,99 % des Modells zufällig sind und nicht durch die Prädiktoren kontrolliert werden können. Daher je kleiner der p -Value, desto besser das Modell.

$$F = \frac{R^2}{1 - R^2}$$

4.4 Erweiterte Analyse einer Multiplen linearen Regression (Residuenanalyse)

Das in Kapitel 4.3 gezeigte Ergebnis der multiplen linearen Regressionsanalyse, stellt nur einen Teil des Ergebnisses der statistischen Modellierung dar. Um einen tieferen Einblick in ein statistisches Modell zu erhalten ist es notwendig, den Residuenvektor e genauer zu untersuchen.

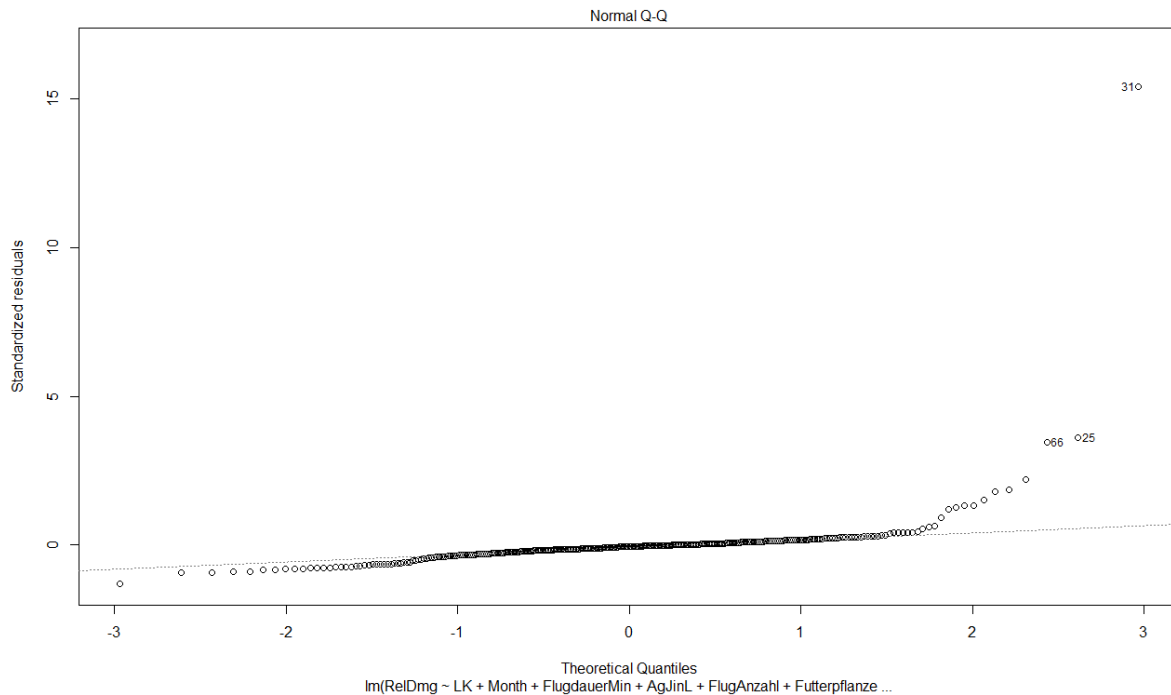


Abbildung 76: QQ-Plot der multiplen linearen Regression der Versicherungsdaten

Der *QQ-Plot* oder *Quantil-quantil-Diagramm*, ist ein grafisches Werkzeug, um zu beurteilen, ob ein Datensatz plausibel aus einer theoretischen Verteilung, wie einer normalen oder exponentiellen Verteilung, stammt. Wenn beispielsweise eine statistische Analyse erstellt wird, die davon ausgeht, dass die Response Variable normalverteilt ist, kann ein *Normal Q-Q-Diagramm* verwendet werden, um diese Annahme zu überprüfen. Es kann hierbei nur eine visuelle Kontrolle erfolgen und ist daher subjektiv zu betrachten. Ein *Q-Q-Diagramm* erlaubt eine grafische Überprüfung, ob eine Annahme plausibel ist, und wenn nicht, wie die Annahme verletzt wird und welche Datenpunkte zur Verletzung beitragen. Der in Abbildung 76 gezeigte *QQ-Plot* bestätigt, dass ein Großteil der Datenpunkte sich auf einer Geraden befinden. Dies lässt darauf schließen, dass die Daten aus einer Normalverteilung stammen. Es muss hier aber dennoch beachtet werden, dass der *QQ-Plot* ein rein graphisches Mittel ist um schnell beurteilen zu können, ob Daten aus einer Normalverteilung stammen. Um Daten auf eine bestimmte Verteilung zu überprüfen kann auch ein *Shapiro-Wilk-Test* oder ein *Kolmogorov-Smirnov-Test* durchgeführt werden. An den Enden driften die Datenpunkte in Abbildung 76 stark auseinander, was zur Aussage führt, dass der Datensatz durch Extremereignisse bzw. Ausreißer verfälscht wird.

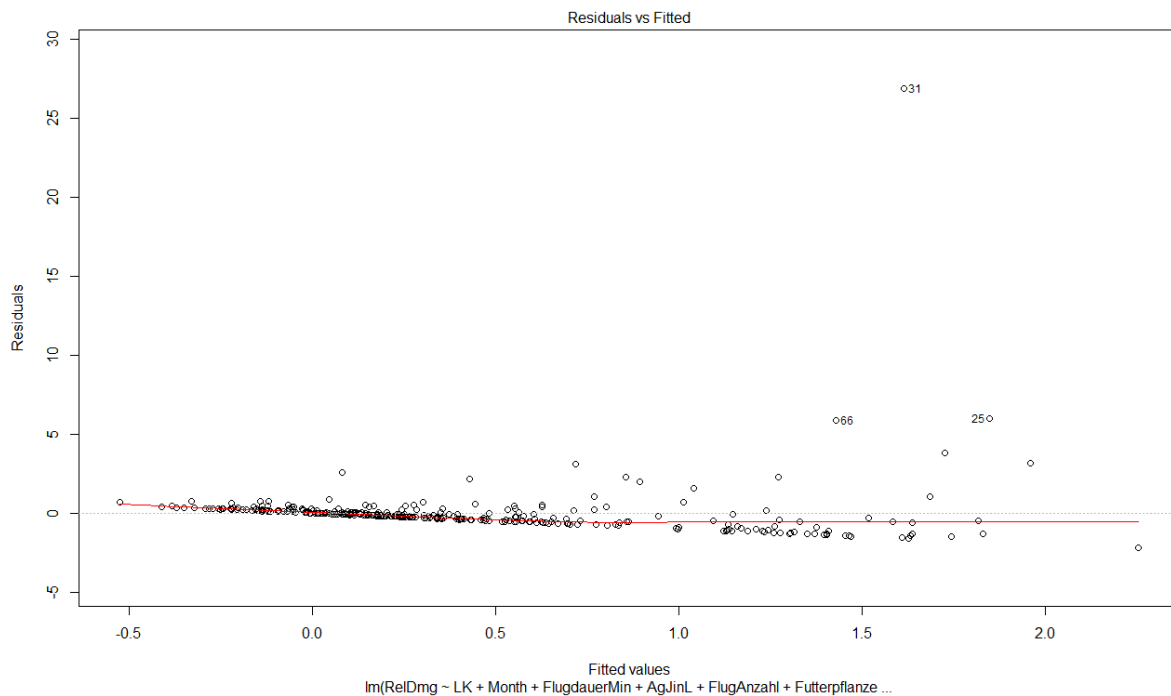


Abbildung 77: Residuenplot zeigt die Prädiktorwerte und die dazugehörigen Residuen

Der *Residuals vs. Fitted Plot* (oder einfach nur Residuenplot) aus Abbildung 77 zeigt die Abweichungen der Beobachtungswerte Y_i von dem generierten statistischen Modell. Ein Residuenplot ist ein graphisches Werkzeug um zu prüfen, ob in den Daten ein nichtlinearer Einfluss vorhanden ist, welcher noch nicht durch einen Prädiktor erklärt wird. Für diesen Fall streuen die einzelnen Punkte gleichmäßig um eine Gerade, was auf eine gleichmäßige Verteilung schließen lässt.

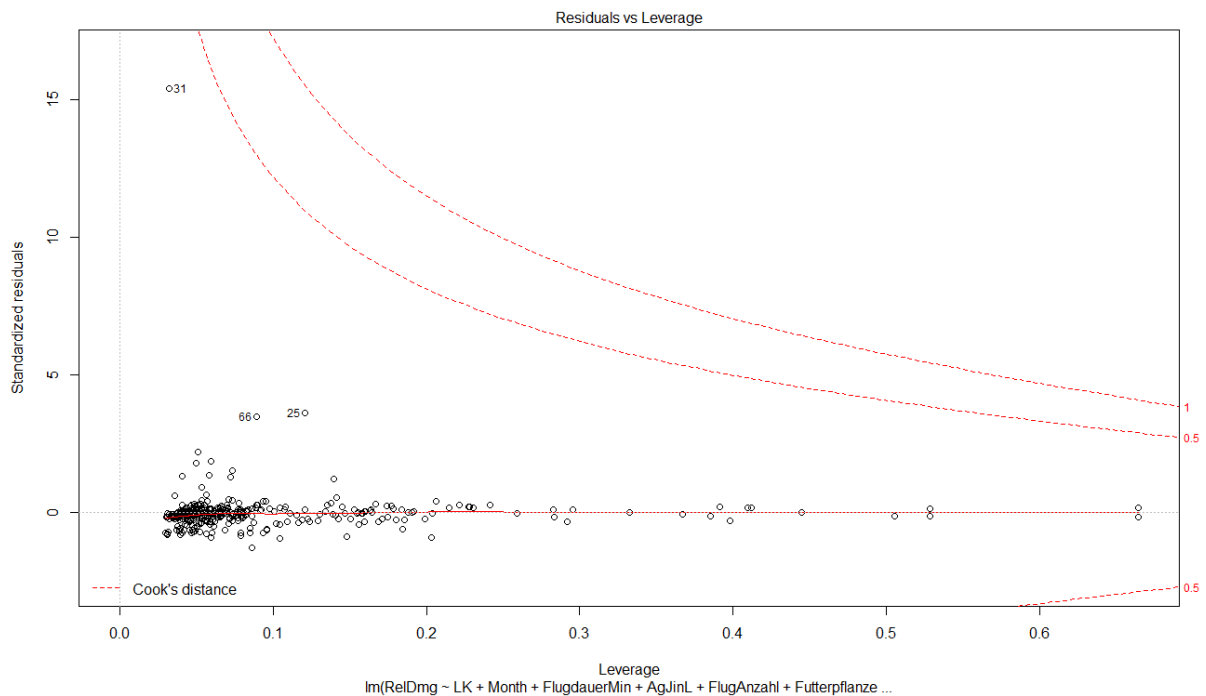


Abbildung 78: Residuals vs. Leverage zeigt die Einflussstärke einzelner Residuen auf das statistische Modell

Der, in Abbildung 78 zu sehende, *Residuals vs. Leverage Plot* zeigt den sogenannten *Leverage-Effekt*, gegenüber den standardisierten Residuen. Der *Leverage-Effekt* (zu Deutsch Hebeleffekt) beschreibt den Einfluss, den eine Beobachtung auf die Responsevariable besitzt. Mithilfe des *Residuals vs. Leverage Plot* kann der Einfluss einer Beobachtung auf das statistische Gesamtmodell bestimmt werden. In Abbildung 78 ist zu sehen, dass der Großteil der Variablen einen geringen Einfluss auf das Modell besitzt. Die rot eingezeichnete gestrichelte Linie, wird als *Cook-Abstand* bezeichnet. Wenn sich Beobachtungen außerhalb des *Cookschenabstands* befinden ist dies ein Indikator dafür, dass diese Beobachtung einen signifikanten Einfluss auf das Modell besitzen.

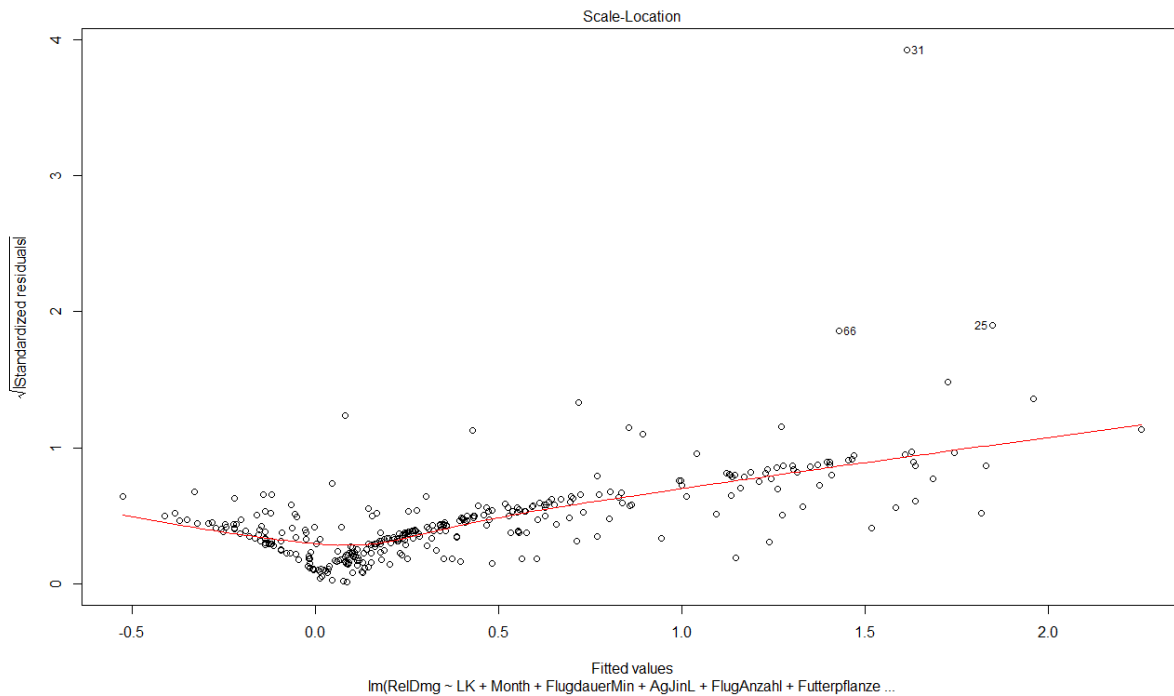


Abbildung 79: Scale-Location Plot

Der, in Abbildung 79 zu sehende, *Scale-Location Plot*, ermöglicht eine Analyse der Standardabweichung der Residuen, was eine Aussage über die Verteilung der einzelnen Varianzen zulässt. Hierbei kann insbesondere geprüft werden, ob die Varianzen einer *Varianzheterogenität (Heteroskedastizität)* oder einer *Varianzhomogenität (Homoskedastizität)* unterliegen. Bei einer *Varianzheterogenität* streuen die Residuen konstant um eine Gerade, bei einer *Varianzhomogenität* hingegen, ist die Streuung nicht konstant. Abbildung 80 veranschaulicht diese beiden Begrifflichkeiten.

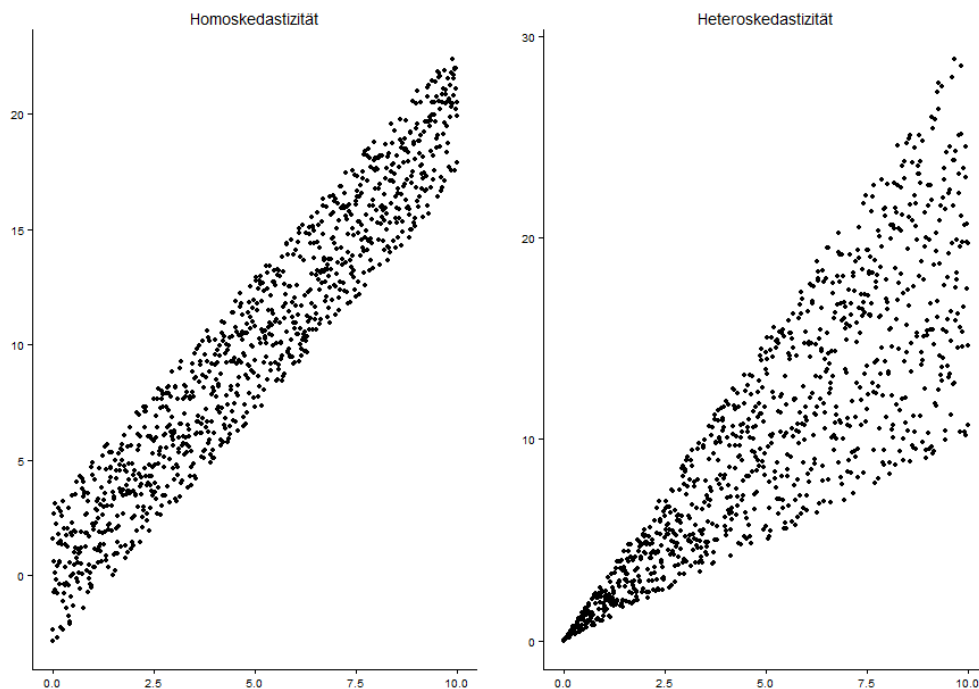


Abbildung 80: Varianzhomogenität, die Streuung der Punkte ist konstant (links). Varianzheterogenität, die Streuung der Punkte ist nicht konstant (rechts) (Schlegel, 2016)

Für die Auswertung der Versicherungsdaten ist Abbildung 79 so zu interpretieren, dass die Residuen ungleichmäßig um die rote Gerade streuen.

4.4.1.1 Zusammenfassung der Auswertung der multiplen linearen Regressionsanalyse

Das, mit dieser Auswertung generierte, statistische Modell ist für einen Nachweis der Wirksamkeit der Hagelabwehr nicht geeignet, stellt aber auch keinen Gegenbeweis dar. Das Modell weist einen F -Wert von 1,049 auf, was bedeutet, dass ein Großteil der Varianzen nicht kontrolliert werden können. Der p -Value von 39,99 % ist ein Indiz dafür, dass ein Großteil der Vorhersage mit diesem Modell zufällig passiert. Das Modell weist einen R^2 -Wert von lediglich 10,03 % auf. Dies bedeutet, dass die vorliegenden Prädiktoren, die Daten unzureichend fitten. Daher kann die Aussage getroffen werden, dass die Prädiktoren (oder deren Daten) die Responsevariable, die relative Schädigung, nicht ausreichend beeinflussen um eine statistische Signifikanz zu besitzen. Das niedrigste Signifikanzniveau im Modell weist die ausgebrachte AgJ-Menge mit 2,44 % auf. Der Estimate-Koeffizient für die ausgebrachte AgJ-Menge ist wiederum positiv. Dies bedeutet, dass das Ausbringen von AgJ eine Erhöhung des Responses, der relativen Schädigung, bewirkt. Dieser Effekt kann so erklärt werden, dass die Hagelabwehripiloten auch für Extremereignisse aufsteigen, an denen die Hagelabwehr nur wenig bis gar nichts ausrichten kann. Speziell an solchen Tagen wird eine unwahrscheinlich große Menge an AgJ ausgebracht, was zu einer Korrelation zwischen Schäden und ausgebrachter AgJ Menge führt. Das Skript zur multiplen linearen Regressionsanalyse kann im Anhang A gefunden werden.

4.5 Analyse der Versicherungsdaten mithilfe eines generalisierten linearen Modells

Eine Erweiterung des klassischen Regressionsansatzes im linearen Modell findet sich im generalisierten oder verallgemeinerten linearen Modell (GLM) wieder. Eine gute Einführung zu generalisierten linearen Modellen bietet (Lindsey P., 2007).

Lineare Regressionsmodelle unterliegen gewissen Annahmen bezüglich Normalität, Homoskedastizität und Linearität. Diese Annahmen werden durch generalisierte Modelle abgeschwächt.

Klassischerweise wird das normale Regressionsmodell der multiplen linearen Regression in folgender Form geschrieben:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + e$$

Wobei auch folgende Form zulässig ist:

$$\mu_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

μ_i kennzeichnet hierbei den Erwartungswert für einen Response, mit der konstanten Varianz σ^2 .

Im linearen Modell wird die Annahme getroffen, dass die Fehler normalverteilt sind. Im *GLM* hingegen kann die Annahme getroffen werden, dass die Fehler einer gewissen Verteilung entsprechen (z.B. Normal-, Binomial-, Poisson-, Gamma- und inverse Gaußverteilung). Die Zufälligkeiten in einem statistischen Modell lassen sich für Versicherungskontexte mit einem *GLM* oftmals besser beschreiben, als mit einem einfachen linearen Modell.

Das bei einem generalisierten linearen Modell erzeugte Modell besteht aus einer *zufälligen Komponente*, einer *systematischen Komponente* und einer zugehörigen *Linkfunktion* g (Lindsey J. K., Applying Generalized Linear Models, 1997, S. 18).

4.5.1.1 Linearer Prädiktor, systematische Komponente

Die linearen Prädiktoren $\eta = (\eta_1, \dots, \eta_i)$ bezeichnen die systematischen Komponenten im Modell. Diese sind definiert durch:

$$\eta := x\beta \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$x\beta$ definiert die Lage (μ_i) des Responses in Abhängigkeit der restlichen Prädiktoren. Der Parameter β muss, ähnlich zur multiplen linearen Regressionsanalyse, geschätzt werden. Im Gegensatz zur multiplen linearen Regressionsanalyse kann der Wert für β nicht mehr explizit gelöst werden (Lindsey J. K., Applying Generalized Linear Models, 1997, S. 18) (siehe auch Kapitel 4.5.1.4).

4.5.1.2 Response Verteilung, zufällige Komponente

Y_i kennzeichnet eine unabhängige Zufallsvariable, welche den Erwartungswert μ_i besitzt. Der Zufallsvektor Y_i wird als Vektor $Y = (Y_1 \dots Y_n)^T$ beschrieben und stellt die zu beobachtende Variable dar (Lindsey, 2019 S.18).

4.5.1.3 Link Funktion

Die Verbindung von zufälliger und systematischer Komponente wird durch eine Transformation des Erwartungswerts geschaffen. Die Verbindung zwischen diesen beiden Komponenten wird mit einer sogenannten Linkfunktion g vollzogen. Dazu wird der lineare Prädiktor η als Funktion des Erwartungswerts μ_i beschrieben.

$$g(\mu_i) = \eta_i = x_i^T \beta$$

Die Linkfunktion ist die Inverse der Responsefunktion, damit gilt $\eta_i = g(\mu_i)$. Für jede Verteilungsannahme muss eine dazugehörige Linkfunktion gewählt werden, z.B. *Gamma-Verteilung*, *Poisson-Verteilung* etc.

4.5.1.4 Schätzung der Modellparameter mithilfe der ML-Methode

Die Modellparameter β von generalisierten linearen Modellen werden im Zusammenhang mit einem numerischen Verfahren der *Maximum-Likelihood* Methode gelöst. Bevor dieses Verfahren vorgestellt wird, soll die einfache *Maximum-Likelihood* Methode kurz erläutert werden.

4.5.2 Maximum-Likelihood Methode

Die *Maximum-Likelihood* Methode stellt ein Schätzverfahren dar, um eine bestimmte Stichprobe mit den Werten x_1, x_2, \dots, x_n mit einer angenommenen Verteilung möglichst plausibel abzubilden. Ziel ist es, für eine Stichprobe eine gemeinsame Dichtefunktion f_v (*Likelihood*-Funktion) anzunehmen und für diese einen unbekannt Parameter ϑ zu bestimmen (Pantle, 2004). Der bestimmende Parameter ϑ ist der Parameter, der die Annahme der Dichtefunktion f_v mit den Werten aus der Stichprobe x_1, x_2, \dots, x_n am wahrscheinlichsten (engl. likeli) abbildet. Um den Parameter ϑ zu finden ist der Gradient der *Likelihood*-Funktion zu bestimmen. Dafür muss die erste Ableitung (*Score*-Funktion) der *Likelihood*-Funktion maximiert werden. Die *Likelihood*-Funktion L ist bei unabhängiger Stichprobenvariable folgendermaßen definiert:

$$L(\vartheta) = \prod_{i=1}^n f_v(x_i)$$

Die *Likelihood*-Funktion ist algebraisch identisch zur gemeinsamen Dichtefunktion $f_v(x_i; \vartheta)$. Die Bestimmung einer Maximum *Likelihood*-Funktion soll anhand folgenden Beispiels gezeigt werden:

„Die Anzahl der Tore zweier Fußballmannschaften in einem Spiel kann mit einer *Poisson-Verteilung* modelliert werden. Die Mannschaft A erzielt in einem Spiel 3 Tore und die Mannschaft B erzielt 5 Tore“.

Die Dichte der *Poisson-Verteilung* ist folgendermaßen definiert:

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

Der für diese Stichprobenverteilung zu bestimmendem Parameter ist der Koeffizient λ , welcher den Erwartungswert und gleichzeitig die Varianz der Verteilung beschreibt.

Die *Likelihood*-Funktion bildet sich folgendermaßen:

$$L(\lambda) = f(\{x_1 = 3\}) \cap f(\{x_1 = 5\}) = f(x_1 = 3) * f(x_2 = 5)$$

$$L(\lambda) = \frac{\lambda^3}{3!} e^{-\lambda} * \frac{\lambda^5}{5!} e^{-\lambda}$$

$$L(\lambda) = \frac{\lambda^8}{3! 5!} e^{-2\lambda}$$

Bildung der 1. Ableitung nach λ (*Score*-Funktion)

$$\frac{d}{d\lambda} L(\lambda) = \frac{1}{3! 5!} * (8\lambda^7 * e^{-2\lambda} - 2\lambda^8 * e^{-2\lambda}) = 0$$

$$\frac{2\lambda^7 e^{-2\lambda}}{3! 5!} * (4 - \lambda) = 0$$

Die Nullstellen für λ ergeben sich zu $\lambda=0$ und $\lambda=4$. Für $\lambda=4$ besitzt die *Likelihood*-Funktion ihr Maximum, was den *Maximum-Likelihood* Schätzwert für diese Stichprobe bestimmt. Das Torverhalten der beiden Mannschaften lässt sich dementsprechend mit folgender Verteilung beschreiben:

$$f(x) = \frac{4^x}{x!} e^{-4}$$

4.5.3 Bestimmung der *Score*-Funktion für die Parameterschätzer bei GLMs

Für die Lösung von *GLMs* muss die Ableitung der *Likelihood*-Funktion S (*Score*-Funktion) nach β gelöst werden:

$$S(\beta) = \frac{d}{d\beta} L(\beta) = 0$$

Um das Maximum zu finden wird die *Score*-Funktion gleich Null gesetzt. Der Schätzer für den Regressionsparameter β wird also durch das Lösen mehrerer *Score*-Gleichungen gefunden.

Um die hierbei nicht mehr explizit lösbare Gleichung berechnen zu können, wird der Iterative *Weighted-Least-Squares-Algorithmus* (Lindsey, 2019 S.200) verwendet, welcher eng mit dem *Newton-Raphson-Verfahren* verwandt ist.

4.6 Ergebnis des GLMs für die Versicherungsdaten

Für die Auswertung der Versicherungsdaten, mithilfe eines generalisierten linearen Modells, wurde der gleiche Datensatz wie für die multiple lineare Regressionsanalyse herangezogen. Als Response wurde wieder der relative Schaden betrachtet. Als Linkfunktion wurde eine logarithmische Verteilung angenommen, da diese für die Betrachtung von Versicherungsschäden eine übliche Verteilungsfunktion vorsieht.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-3.721612	0.942753	-3.948	9.81e-05	***
LK9163	-0.643435	1.309348	-0.491	0.62349	
LK9171	1.353287	0.747829	1.810	0.07134	.
LK9172	0.563876	0.819930	0.688	0.49216	
LK9173	0.136595	0.966362	0.141	0.88769	
LK9175	1.677046	0.749330	2.238	0.02594	*
LK9177	1.371486	0.757961	1.809	0.07137	.
LK9182	1.347954	1.030261	1.308	0.19174	
LK9183	1.546030	0.741767	2.084	0.03797	*
LK9184	1.130952	0.763721	1.481	0.13968	
LK9187	2.119920	0.899357	2.357	0.01905	*
LK9189	2.755710	0.874983	3.149	0.00180	**
FlugdauerMin	-0.001202	0.004109	-0.292	0.77017	
AgJinL	0.055634	0.021767	2.556	0.01108	*
FlugAnzahl	-0.413887	0.287828	-1.438	0.15147	
Futterpflanzen	2.606315	3.176144	0.821	0.41252	
Ganzpflanzensilage	-2.050444	1.080616	-1.897	0.05871	.
Getreide	1.022048	0.363536	2.811	0.00525	**
Kartoffeln	0.074778	0.560918	0.133	0.89403	
Mais	-0.912087	0.487274	-1.872	0.06219	.
Raps	0.900562	0.587038	1.534	0.12605	
Spargel	-0.451970	2.009263	-0.225	0.82217	
Obst	-0.866013	1.295621	-0.668	0.50437	
SonstigeFruchtarten	0.823707	0.333184	2.472	0.01397	*
Olfruchte	-0.263057	0.341393	-0.771	0.44158	
Hulsenfruchte	-0.902889	0.492400	-1.834	0.06768	.
ErbsenBohnen	-0.052979	0.575839	-0.092	0.92676	
Ruben	0.249795	1.447288	0.173	0.86308	
Gemuse	0.266479	1.213847	0.220	0.82638	

(Dispersion parameter for Gamma family taken to be 3.169022)

Null deviance: 1229.07 on 333 degrees of freedom

Residual deviance: 773.64 on 305 degrees of freedom

AIC: -427.07

Number of Fisher Scoring iterations: 23

Das generalisierte lineare Modell lässt, im Gegensatz zur multiplen linearen Regressionsanalyse, die Fruchtart *Getreide* mehr in den Fokus rücken. Der positive Estimate lässt darauf schließen, dass Getreide eher anfälliger ist gegen Hagel, als z.B. Mais. In der multiplen linearen Regressionsanalyse ist der Prädiktor für den Landkreis Traunstein (9189) mit einem geringen *p-Value* und einem positiven Estimate aufgetreten. Für das *GLM* kommen die beiden Landkreise Berchtesgaden (9175) und Ebersberg (9177) hinzu, welche beide im Nachbargebiet liegen. Die Anzahl der Hagelabwehrflüge und die Flugdauer sind für den *p-Value* nicht signifikant genug, um einen Einfluss auf das *GLM* zu besitzen. Die ausgebrachte AgJ-Menge hingegen ist signifikant und tritt mit einem positiven Estimate hervor. Ein positiver Estimate für die ausgebrachte AgJ-Menge hat eine höhere, zu erwartende, relative Schädigung zur Folge. Im Intercept ist der Landkreis München (Stadt) kodiert, welcher eine hohe Signifikanz für ein geringes Schädigungspotential aufweist.

Das Skript für das Generalisierte lineare Modell kann im Anhang A gefunden werden.

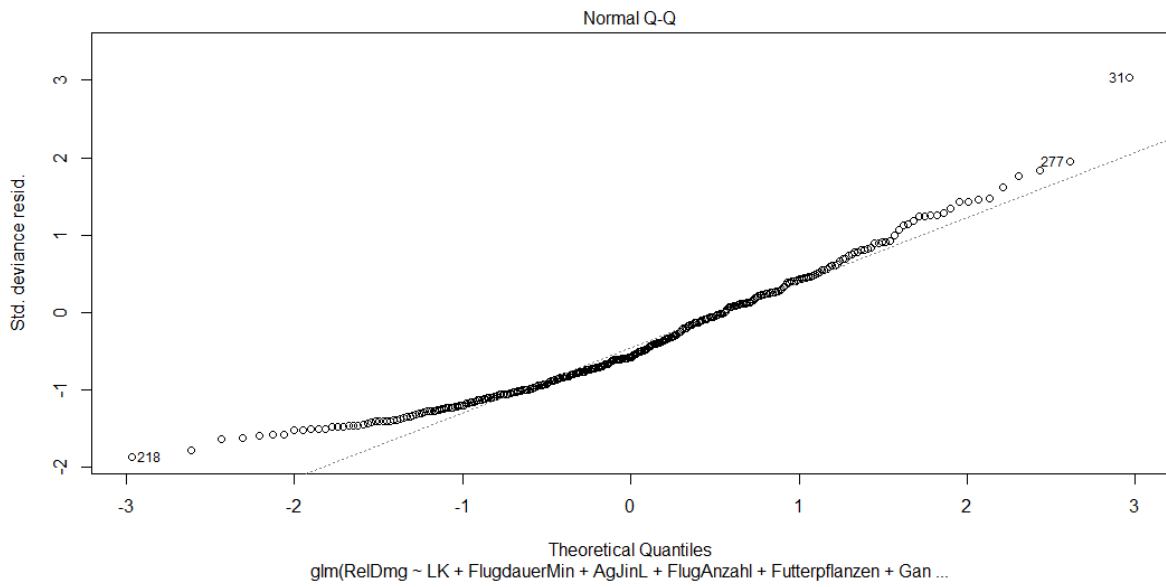


Abbildung 81: QQ-Plot des GLMs

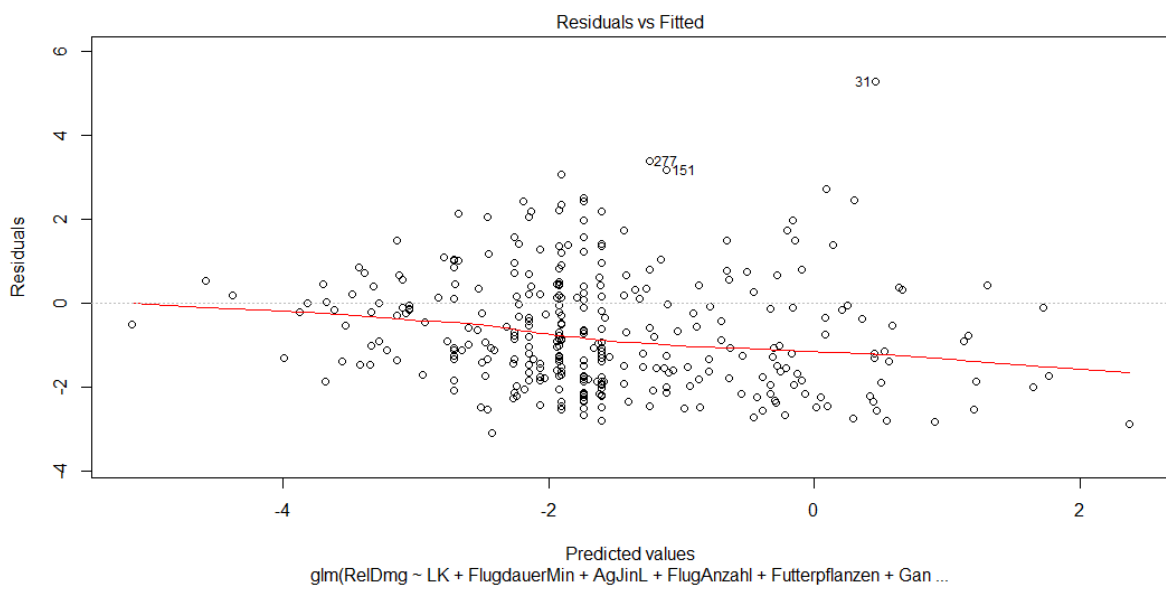


Abbildung 82: Residuals vs. Fitted GLM

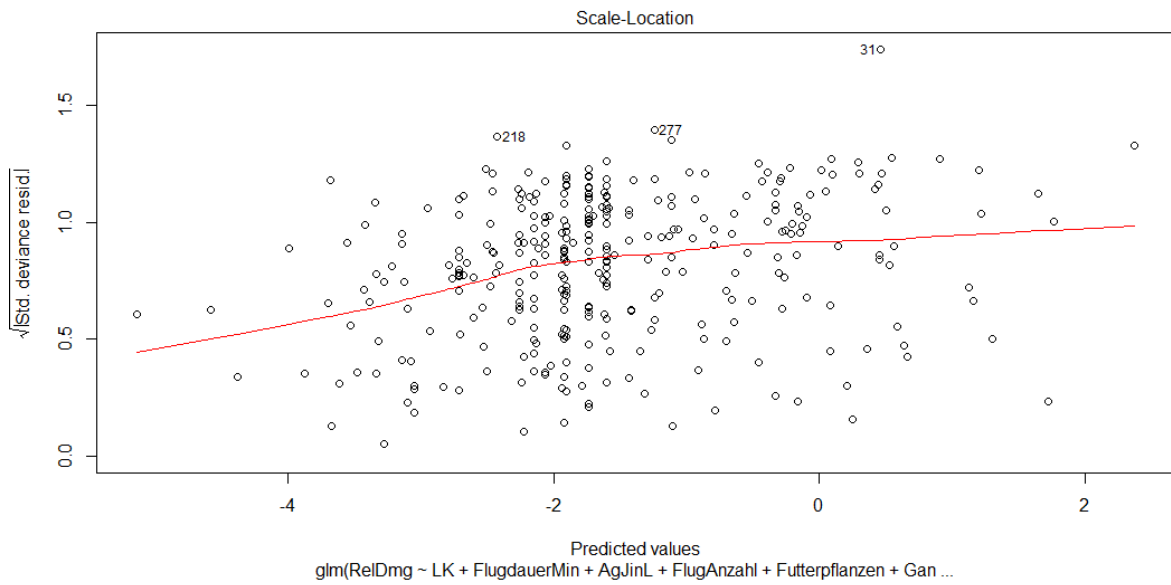


Abbildung 83: Scale-Location GLM

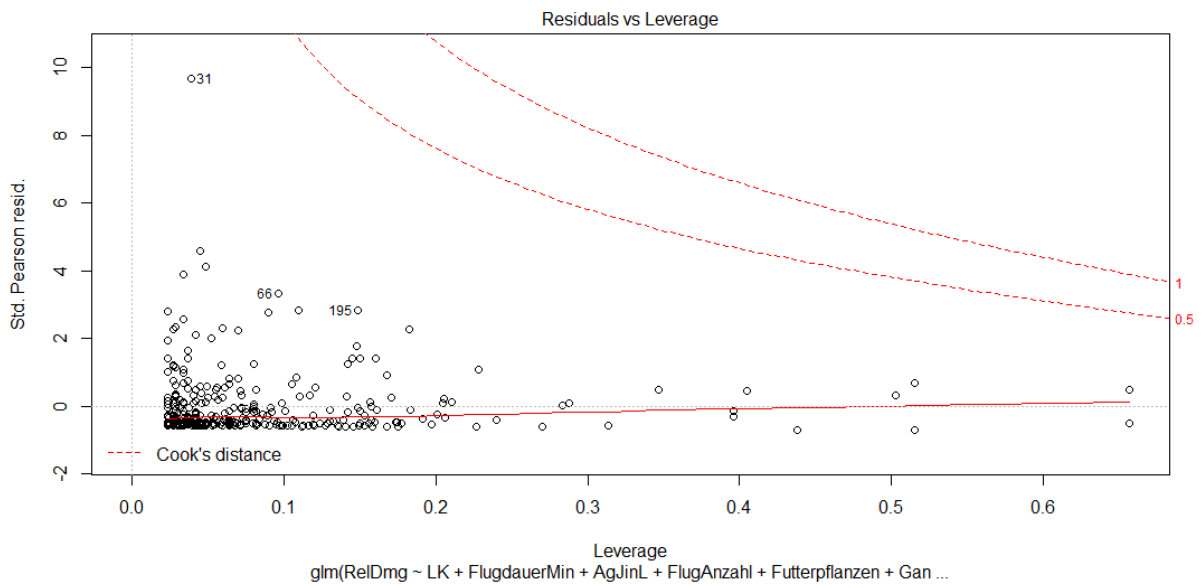


Abbildung 84: Residuals vs. Leverage

4.7 Analyse der Versicherungsdaten mithilfe eines bereinigten GLM

Im vorigen Kapitel 4.6 wurde ein generalisiertes lineares Modell vorgestellt, im Bezug zu den Versicherungsdaten. Die Daten aus Kapitel 4.6 waren jedoch nicht von Ausreißern bereinigt. Im hier vorliegenden Kontext handelt es sich jedoch um Versicherungsdaten, hierbei ist eine Bereinigung von Ausreißern eine übliche Vorgehensweise. Die Bestimmung eines Ausreißers geschieht über Schwellenwerte. Bei Elementarereignissen, wie Hagel auch eines ist, ist die Definition von Extremereignissen von Versicherungen selbst bestimmt. Da hier keine Größenordnungen in Erfahrung gebracht werden konnte, wurde sich auf eine allgemeine Definition von Ausreißern festgelegt (Wellisch, 2018):

„Alle Beobachtungswerte einer Stichprobe, die größer als das obere Quartil plus das 1,5-fache des Quartil Abstands sind und alle Beobachtungswerte einer Stichprobe, die kleiner als das untere Quartil minus das 1,5-fache des Quartil Abstands sind.“

Für Versicherungsschäden erfolgt nur eine Bereinigung für das obere *Quartil*. Abbildung 85 zeigt mithilfe eines *Box-Whisker-Plot* die Verteilung der unbereinigten relativen Schädigungen für Schutzgebiet und Nachbargebiet. Diese Beobachtungen wurden nach beschriebener Methodik bereinigt und sind in Abbildung 86 zu sehen.

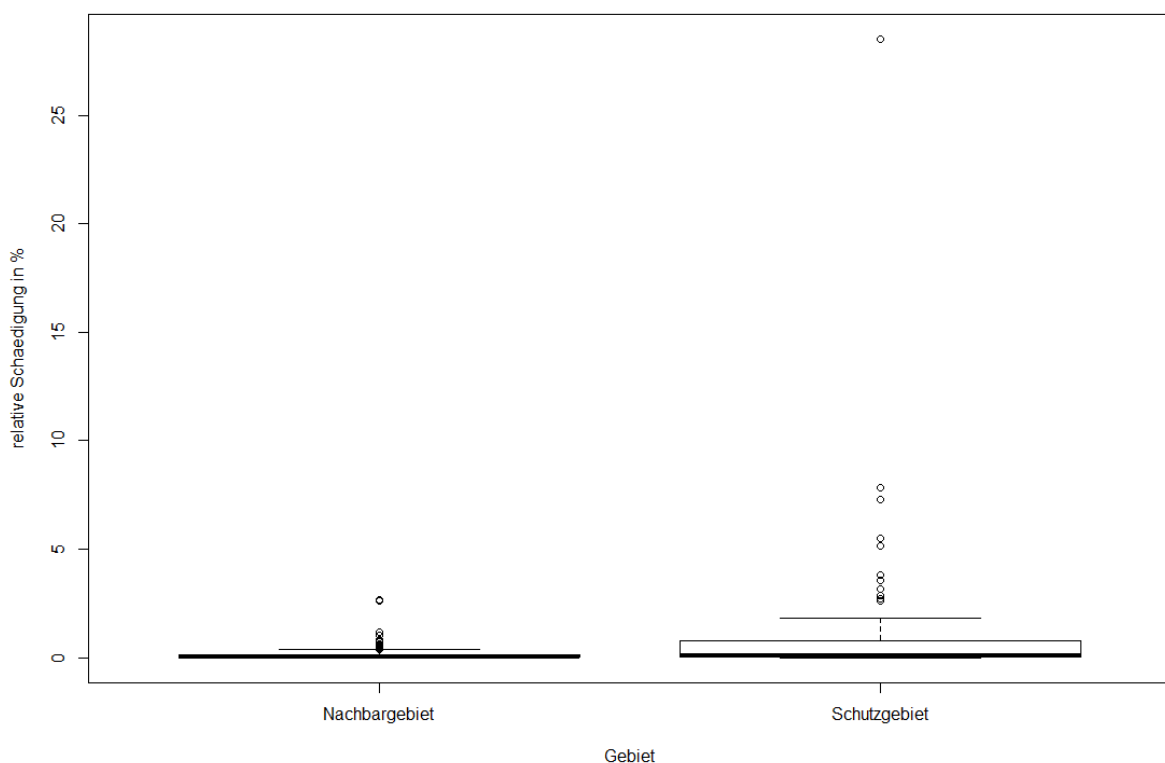


Abbildung 85: Vergleich der relativen Schädigung. Für das beide Gebiete sind ganz klar Ausreißer zu erkennen

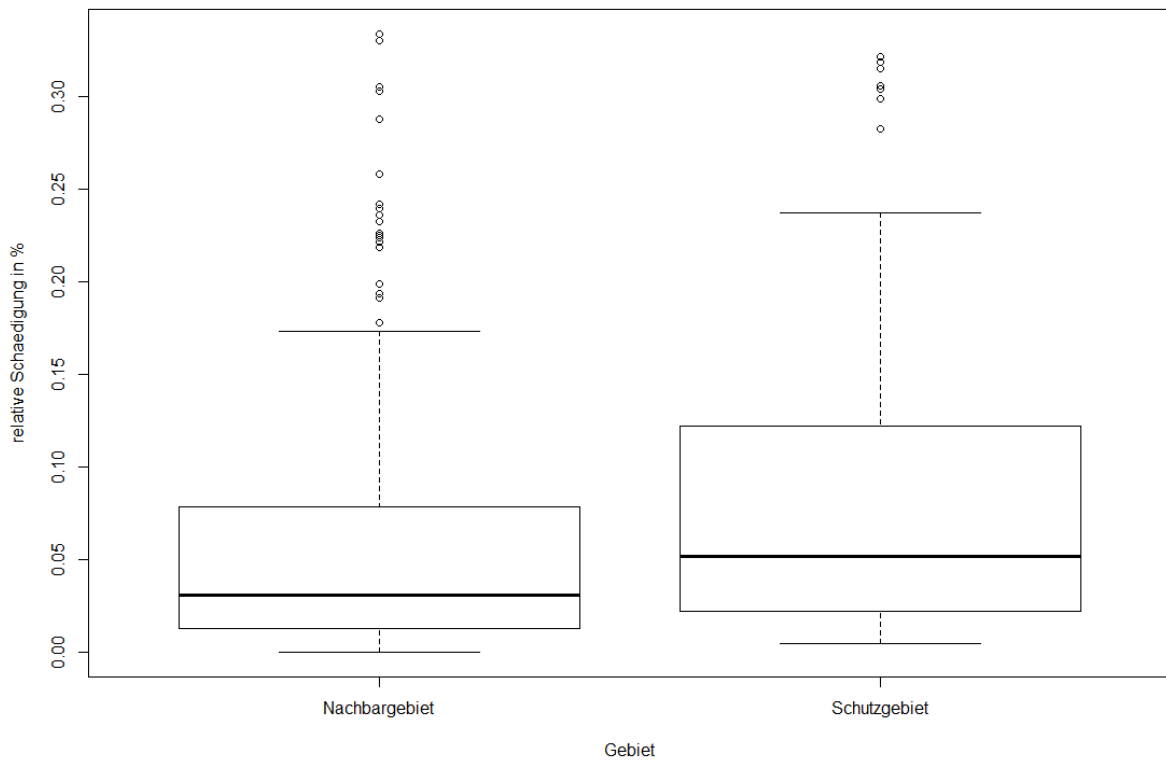


Abbildung 86: Bereinigte relative Schädigung

Für die so bereinigten Daten wurde erneut ein *GLM* gebildet. Abbildung 87 zeigt eine Verteilung der relativen Schäden der Jahre 2011 bis 2018 für die einzelnen Landkreise, mit dem bereinigten Datensatz. Auffällig hier ist, dass keiner der Mediane, welcher eine relative Schädigung im Schutzgebiet aufzeigt, geringer ist, als ein Median aus dem Nachbargebiet.

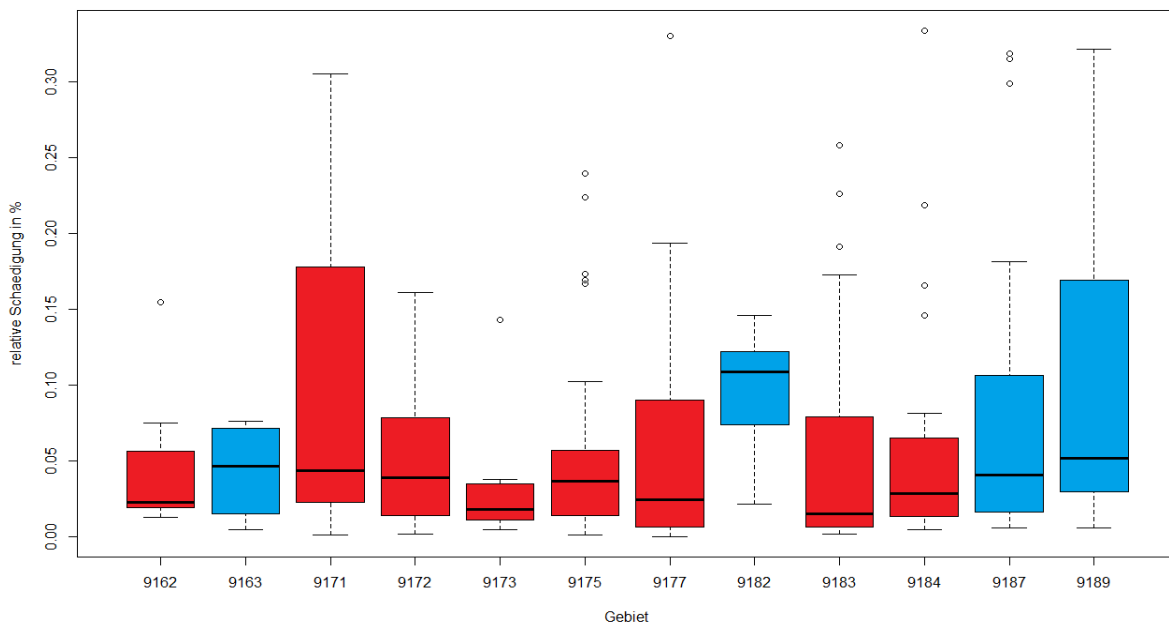


Abbildung 87: Um Ausreißer bereinigte Schadensverteilung. Schutzgebiet in blau, Nachbargebiet in rot

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.905396	0.648234	-4.482	1.14e-05	***
LK9163	-0.746788	0.958798	-0.779	0.43682	
LK9171	0.517629	0.488618	1.059	0.29049	
LK9172	0.021112	0.540413	0.039	0.96887	
LK9173	-0.463784	0.626845	-0.740	0.46010	
LK9175	-0.004007	0.497377	-0.008	0.99358	
LK9177	0.080572	0.498462	0.162	0.87172	
LK9182	0.194482	0.796523	0.244	0.80731	
LK9183	-0.189277	0.490845	-0.386	0.70012	
LK9184	0.068038	0.497719	0.137	0.89138	
LK9187	0.078481	0.675210	0.116	0.90757	
LK9189	0.272471	0.639900	0.426	0.67063	
FlugdauerMin	0.002504	0.003309	0.757	0.44982	
AgJinL	0.004206	0.018330	0.229	0.81870	
FlugAnzahl	-0.185834	0.271862	-0.684	0.49491	
Futterpflanzen	2.783614	2.133783	1.305	0.19329	
Ganzpflanzensilage	-0.849388	0.703199	-1.208	0.22827	
Getreide	0.039495	0.252891	0.156	0.87603	
Kartoffeln	0.346543	0.386280	0.897	0.37055	
Mais	-0.014385	0.349972	-0.041	0.96725	
Raps	-0.362149	0.418693	-0.865	0.38793	
Spargel	0.213332	1.327647	0.161	0.87248	
Obst	-0.743120	0.840511	-0.884	0.37751	
SonstigeFruchtarten	0.174258	0.239143	0.729	0.46691	
Olfruchte	-0.189424	0.247017	-0.767	0.44392	
Hulsenfruchte	0.084068	0.347494	0.242	0.80904	
ErbsenBohnen	-0.704760	0.405316	-1.739	0.08335	.
Ruben	2.299215	1.002445	2.294	0.02267	*
Gemüse	-2.986267	1.061843	-2.812	0.00532	**

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

(Dispersion parameter for Gamma family taken to be 1.32187)

Null deviance: 391.31 on 269 degrees of freedom
Residual deviance: 347.31 on 241 degrees of freedom
AIC: -881.65

Number of Fisher Scoring iterations: 16

Das bereinigte *GLM* weist gegenüber dem unbereinigten *GLM*, speziell für die Prädiktoren, welche die Hagelabwehr betreffen, keinerlei Signifikanz mehr auf. Einzig die Prädiktoren, welche die Fruchtgattungen Gemüse und Rüben betreffen müssen noch als die einzigen signifikanten Prädiktoren betrachtet werden. Im Intercept ist erneut der Prädiktor für die Stadt München codiert, welcher eine hohe Signifikanz aufweist und ein geringes Schadensrisiko voraussagt.

Verbesserung des statistischen Modells mithilfe des Akaike Informationskriterium

Statistische Modelle versuchen immer eine Korrelation zwischen kategoriellen oder metrischen Variablen zu bilden. Da der Zusammenhang oftmals unbekannt ist, kann theoretisch eine unbegrenzte Anzahl von Prädiktoren hinzugezogen werden, von welchen einer garantiert in der Lage sein wird, ein Modell gut zu fitten. Hierbei sollte vor allem auf die Gefahr des Over-Fittens eines Modells, in Anbetracht seiner Anpassungsgüte (R^2 bei der Multiplen linearen Regressionsanalyse), geachtet werden.

Für die Beurteilung der Auswirkung auf die relative Schädigung von Hagel, in den einzelnen Landkreisen, könnte beispielsweise auch als Prädiktor die Anzahl der roten Autos, in den einzelnen Landkreisen, herangezogen werden. Bei der hieraus möglichen Korrelation (Scheinkorrelation) könnte womöglich der Response, die relative Schädigung, besser erklärt werden, als die momentan verwendeten Prädiktoren. Die Modellgüte könnte so scheinbar noch weiter erhöht werden, indem eine Vielzahl weiterer Prädiktoren hinzugefügt werden, welche einen scheinbaren Einfluss auf die Hagelabwehr besitzen.

Um ein Over-Fitten in dieser Hinsicht zu vermeiden, kann das *Akaike-Informationskriterium* betrachtet werden (Abk.: AIC). Das AIC berechnet sich in R folgendermaßen (The R Foundation, 2019):

$$AIC = -2 * \ln(\hat{L}) + k * n$$

\hat{L} repräsentiert hierbei den maximalen Wert der *Likelihood*-Funktion für das statistische Modell. Der Faktor k kennzeichnet die Anzahl der Prädiktoren und n die Anzahl der Beobachtungen. Aus der Formel ist ersichtlich, dass ein Hinzufügen einer Vielzahl von Prädiktoren k den AIC erhöht und damit verschlechtert. Wichtig ist, dass der AIC nicht in der Lage ist, die absolute Güte eines Modells zu beurteilen, sondern lediglich einen relativen Vergleich zwischen verschiedenen statistischen Modellen bereitstellt. Daher gilt: Je kleiner ein AIC für ein Modell ist, desto besser erklärt es einen gegebenen Zusammenhang.

Um das Modell zu finden, welches den besten AIC aufweist wurde in „R“ mit dem, in der *Mass-Bibliothek* enthaltenen, Algorithmus *stepAIC* gearbeitet. Hierbei werden bei einem bestehenden Modell solange Prädiktoren entfernt und hinzugefügt, bis dessen AIC sich nicht mehr verbessert.

4.7.1 Verbesserung des GLM mithilfe des stepAIC

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-4.08843	0.87792	-4.657	4.74e-06	***
LK9163	-0.92072	1.27757	-0.721	0.47164	
LK9171	1.35850	0.74708	1.818	0.06995	.
LK9172	0.60212	0.80555	0.747	0.45534	
LK9173	0.19144	0.96761	0.198	0.84329	
LK9175	1.69294	0.75027	2.256	0.02473	*
LK9177	1.41432	0.75796	1.866	0.06298	.
LK9182	1.31574	1.01691	1.294	0.19666	
LK9183	1.57861	0.74013	2.133	0.03371	*
LK9184	1.14679	0.76364	1.502	0.13417	
LK9187	2.04482	0.88722	2.305	0.02183	*
LK9189	2.69053	0.86658	3.105	0.00208	**
AgJinL	0.05398	0.01357	3.977	8.68e-05	***
FlugAnzahl	-0.46167	0.28024	-1.647	0.10047	
Ganzpflanzensilage	-0.81043	0.56450	-1.436	0.15209	
Getreide	1.06392	0.35215	3.021	0.00272	**
Mais	-0.73462	0.46013	-1.597	0.11137	
Raps	1.09318	0.48277	2.264	0.02423	*
SonstigeFruchtarten	0.97474	0.29737	3.278	0.00116	**
Hulsenfruchte	-0.80902	0.42856	-1.888	0.05998	.

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

(Dispersion parameter for Gamma family taken to be 3.202045)

Null deviance: 1229.07 on 333 degrees of freedom
 Residual deviance: 780.67 on 314 degrees of freedom
 AIC: -441.14

Number of Fisher Scoring iterations: 16

Das verbesserte *GLM* weist nach dem *stepAIC* Algorithmus nur noch Prädiktoren auf, welche zu einer Verbesserung des *AIC* führen. Es wurde eine Vielzahl der Prädiktoren, welche für eine bestimmte Fruchtart stehen, aussortiert. Das *AIC* des verbesserten Modells beträgt nun -441.14, das nicht verbesserte Modell weist einen *AIC* von -427.07 auf. Der *p-Value* für die ausgebrachte Silberjodidmenge stellt nun einen Prädiktor dar, welcher einen noch größeren Einfluss auf das Modell besitzt. Die Anzahl der Flüge besitzt, im Vergleich zu den anderen Modellen, den besten *p-Value* mit 10 % und weist einen negativen Estimate auf, was ein deutlicher Hinweis auf einen positiven Einfluss der Hagelabwehr ist. Das Skript, für das mit dem *stepAIC* arbeitenden Algorithmus, kann im Anhang A gefunden werden.

4.7.1.1 Rechenbeispiel für das erzeugte Modell

Mit dem, in Kapitel Verbesserung des GLM mithilfe des stepAIC4.7.1 vorgestellten, Modell, soll eine beispielhafte Vorhersage, für eine relative Schädigung, erfolgen. Der geschätzte Wert, der relativen Schädigung, wird mit dem Erwartungswert $\hat{\mu}$ gekennzeichnet. Die Modellgleichung aus Kapitel 4.7.1, ist mit den numerischen Werten für die ausgebrachte Silberjodidmenge x_1 und die Anzahl der Flüge x_2 folgendermaßen zu beschreiben:

$$\log(\hat{\mu}) = \beta_{Intercept} + \beta_{AgJ} * x_1 + \beta_{Fluganzahl} * x_2 + \beta_{Fruchtart} * 1 + \beta_{Landkreis} * 1$$

$$\hat{\mu} = e^{\beta_{Intercept} + \beta_{AgJ} * x_1 + \beta_{Fluganzahl} * x_2 + \beta_{Fruchtart} * 1 + \beta_{Landkreis} * 1}$$

Die, von der Hagelabwehr beeinflussbaren, Parameter sind die ausgebrachte Silberjodidmenge und die Anzahl der Flüge. Abbildung 88 und Abbildung 89 zeigen den zu erwartenden relativen Schaden für den Landkreis Miesbach und Rosenheim, mit und ohne einen Hagelabwehrflug, bei welchem eine Silberjodidmenge von 6 Litern freigesetzt wird.

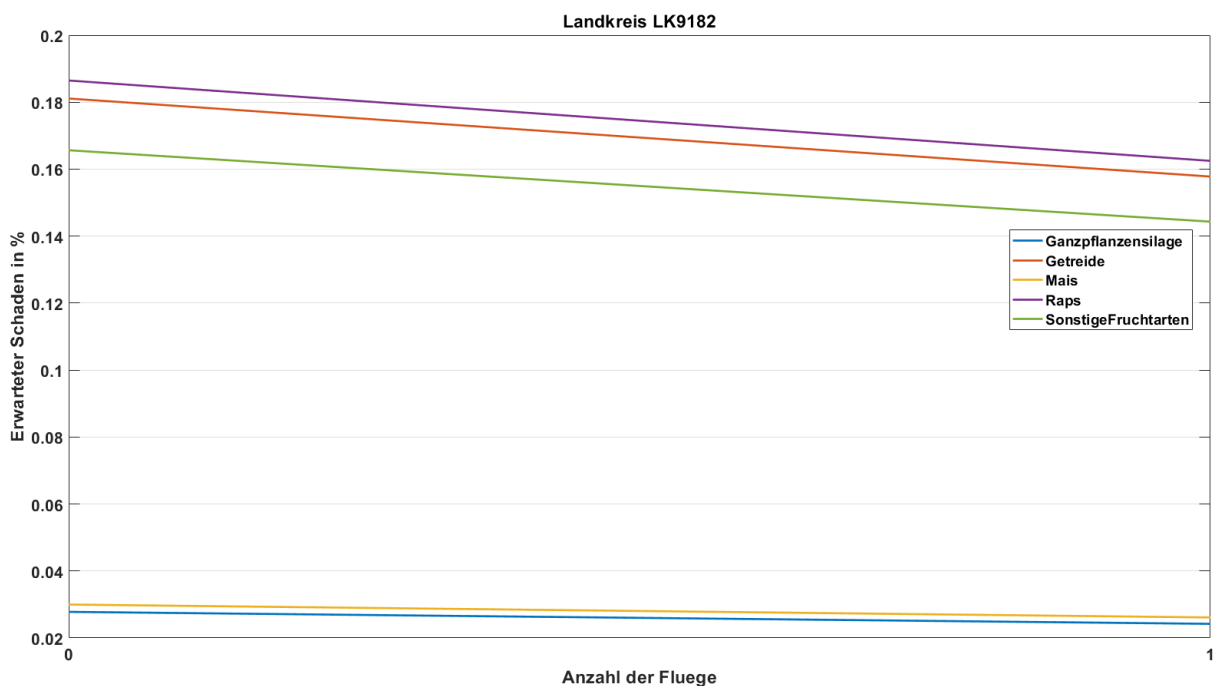


Abbildung 88 Es ist eine deutliche Abnahme des zu erwartenden relativen Schadens, durch einen Hagelabwehreinsetz für verschieden Fruchtarten, zu sehen.

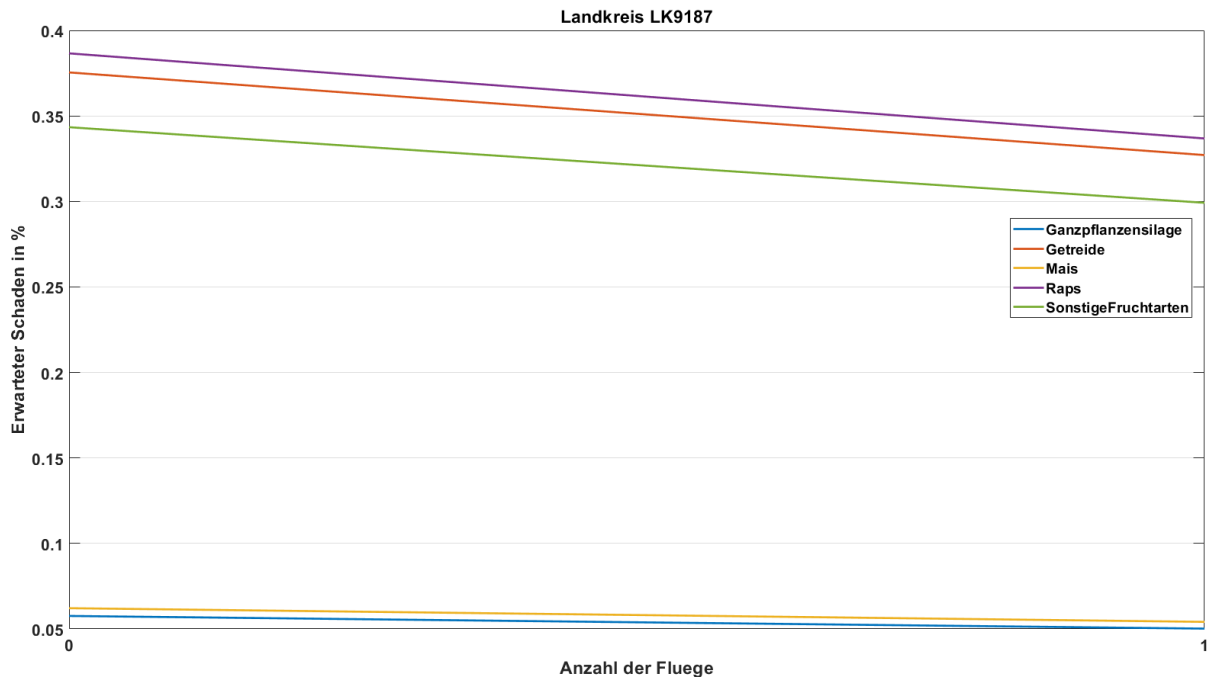


Abbildung 89: Abnahme des zu erwartenden Schadens für den Landkreis Rosenheim.

Für alle anderen Landkreise innerhalb des Schutzgebietes konnten ähnliche Schadensreduktionen nachgewiesen werden. Das, mit dem *stepAIC*, generierte Modell ist ein Nachweis für die Wirksamkeit der Hagelabwehr. Abbildung 90 zeigt die Differenz zwischen den einzelnen Gebieten, einmal mit und einmal ohne Hagelabwehr. Gebiete, welche eigentlich dem Nachbargebiet zuzuordnen sind, wurden so behandelt, als hätten sie eine Hagelabwehr. Diese Annahme wird unter anderem durch die Tatsache gestützt, dass die Piloten nicht zwischen Schutzgebiet und Nachbargebiet unterscheiden (siehe Kapitel 4.9).

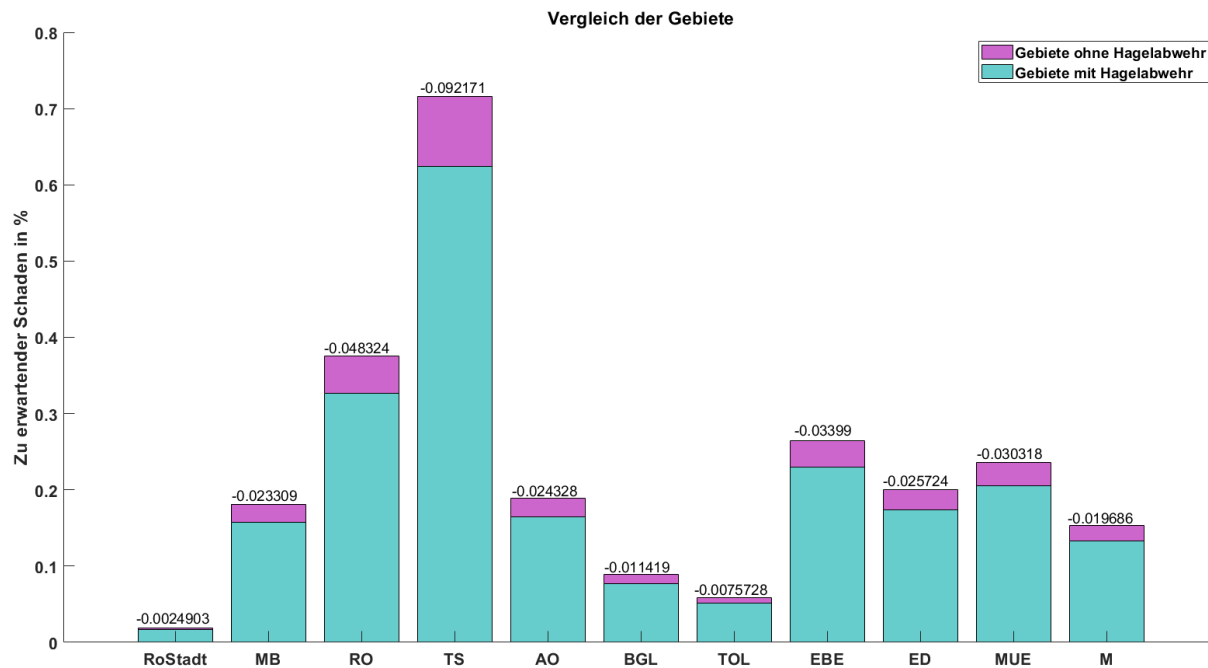


Abbildung 90: Unterschied zwischen den Gebieten

4.8 Vergleich der Versicherungsdaten mithilfe eines Scoring-Verfahrens

Eine weitere Variante, das Schutz- und Nachbargebiet miteinander zu vergleichen, ist ein sehr einfaches Scoring-Verfahren. Hierbei wurden alle Flugtage und Nichtflugtage herangezogen und die Tage gefiltert, an denen sowohl im Schutzgebiet, als auch im Nachbargebiet ein Schaden stattgefunden hat. Die spezielle Wahl der Tage stellt sicher, dass an diesen Tagen über beide Gebiete eine Hagelfront gezogen sein muss. Der Scoring-Algorithmus erfasst die, an einem Tag erfolgten, Schäden für beide Gebiete und bildet, mithilfe der versicherten Fläche, den relativen Schaden. Die beiden relativen Schäden werden im Anschluss miteinander verglichen. Der geringere relative Schaden für diesen Tag erzielt einen Score für sein Gebiet. Ein Beispiel:

Am 25.06.2006 tritt ein relativer Schaden im Schutzgebiet von 0,65 % auf. Im Nachbargebiet beträgt der Schaden an diesem Tag 1,21 %. Da der Relative Schaden im Schutzgebiet geringer, als der relative Schaden im Nachbargebiet ist, wird ein Punkt für das Schutzgebiet erzielt. Das Scoring-Verfahren wurde auf zwei Arten von Tage angewandt. Einmal nur auf Tage, welche Einsatztage sind und in beiden Gebieten ein Schaden vorliegt. Und für Tage, an denen kein Flug stattgefunden hat und es in beiden Gebieten ein Schaden stattgefunden hat.

4.8.1.1 Vergleich der Flugtage

Der Score für das Schutzgebiet ist, im Gegensatz zum Nachbargebiet, für die genannten Tage in fast jedem Jahr geringer (siehe Abbildung 91). Wenn der Score für das Schutzgebiet und das Nachbargebiet über alle Jahre aufsummiert wird, erhält man für das Schutzgebiet einen Gesamtscore von 21 und für das Nachbargebiet 39.

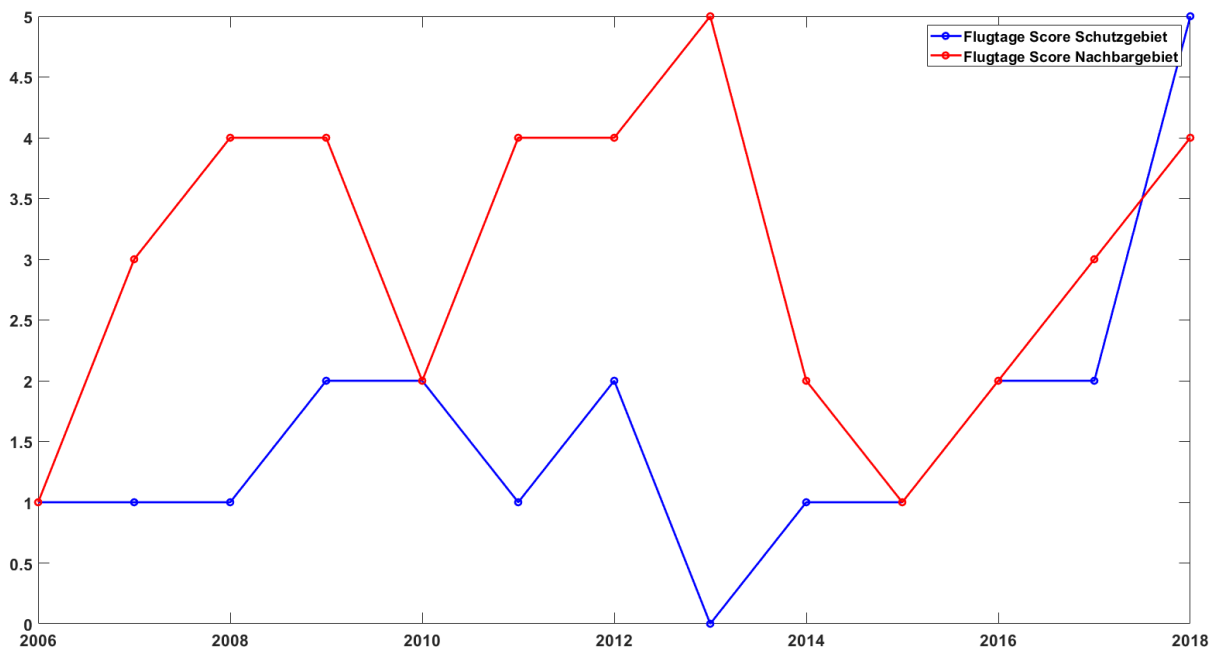


Abbildung 91: Das Schutzgebiet unterliegt dem Nachbargebiet bis zum Jahr 2015 eindeutig

4.8.1.2 Vergleich der nicht Flugtage

Wenn der Score für das Schutzgebiet und das Nachbargebiet über alle Jahre aufsummiert wird, erhält man für das Schutzgebiet einen Gesamtscore von 24 und für das Nachbargebiet von 23.

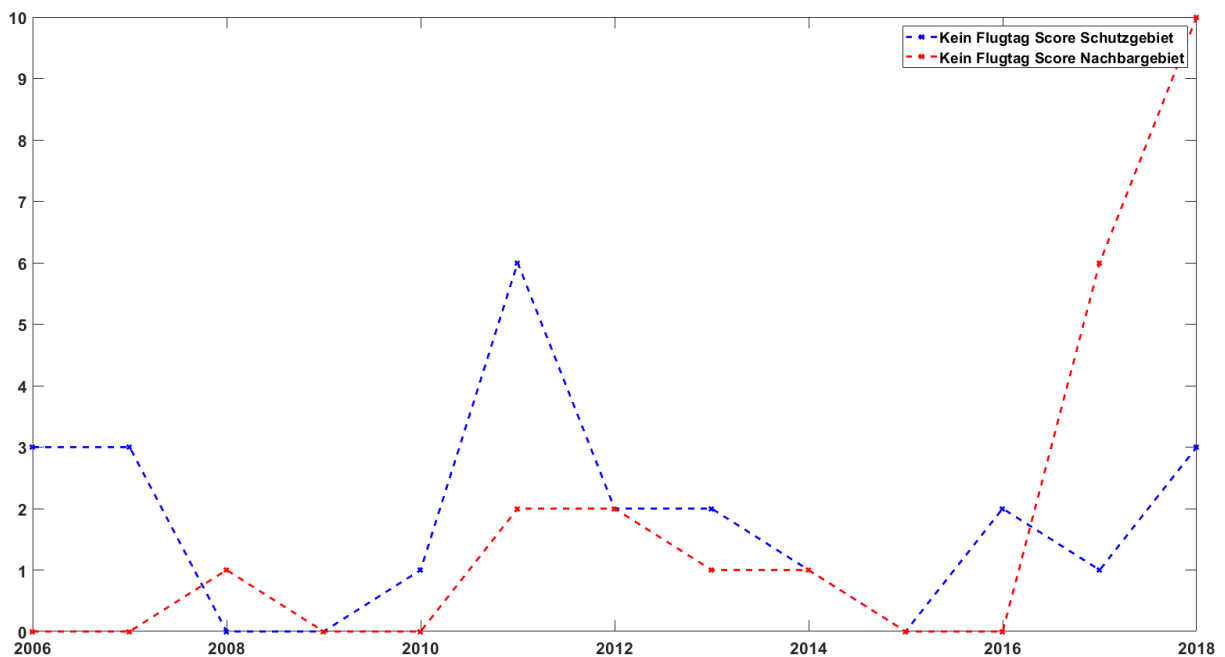


Abbildung 92: Für die Tage mit Schädigung und ohne Hagelabwehr hebt sich keiner der beiden Scores bis 2016 klar vom anderen ab

4.9 Zusammenfassung des Scoringverfahrens

Wie in den beiden Abbildungen zu sehen ist, ist der Score für die Flugtage nicht ausgeglichen zwischen Flugtage und nicht Flugtagen. An den nicht geflogenen Tagen ist der Gesamtscore fast gleich. Aufgrund der geringen Menge an Tagen, an welchen in beiden Gebieten ein Schaden aufgetreten ist, kann diese Methodik nicht als ausreichend aussagekräftig gewertet werden. Hinzukommend ist die Tatsache, dass die Piloten nicht zwischen Schutzgebiet und Nachbargebiet entscheiden. Wenn eine Gewitterwolke sich im Nachbargebiet anfängt aufzubauen und droht ins Schutzgebiet zu fliegen, dann erfolgt schon eine Impfung im Nachbargebiet. Die harten Grenzen, welche Nachbargebiet und Schutzgebiet definieren sind daher äußerst fraglich. Das Skript für das Scoringverfahren kann im Anhang A gefunden werden.

4.10 Untersuchung der Schadensverteilung der einzelnen Gebiete

Die beiden Gebiete Schutzgebiet und Nachbargebiet sollen anhand ihrer relativen Schadensverteilung untersucht werden. Es sollen folgende Schädigungen speziell betrachtet werden:

- Analyse des Schadens im Nachbargebiet. Untersuchung ob dieser aus einer anderen Verteilung kommt, als der Schaden an Flugtagen im Schutzgebiet
- Analyse des Schadens im Schutzgebiet. Untersuchung ob der Schaden von Flugtagen im Schutzgebiet aus einer anderen Verteilung kommt, als an Tagen im Schutzgebiet ohne Flug

Zur Analyse werden hierzu drei verschiedenen Signifikanztests verwendet, welche nur kurz beschrieben werden sollen.

4.10.1 T-Test - Vergleich zweier Mittelwerte

Vergleich eines theoretischen Erwartungswert aus einer Stichprobe. Bei einem Mittelwertvergleich wird geprüft, ob zwei Gruppen aus derselben Population (auch Grundgesamtheit genannt) stammen (Mittelwertvergleich). Existieren mehr als zwei Gruppen, muss statt dem T-Test ein anderes Testverfahren, wie beispielsweise eine einfache Varianzanalyse gerechnet werden. Für den T-Test sind folgende Voraussetzungen zu erfüllen (Pruscha, Statistisches Methodenbuch, 2006):

- Annähernd Gauß-verteilt (Annahme)
- Varianzhomogenität
- Unabhängig

4.10.2 F-Test - Vergleich auf Varianzheterogenität

Der *F*-Test testet zwei Stichproben anhand ihrer *Varianzen*, ob diese aus derselben Grundgesamtheit kommen.

4.10.3 Wilcoxon Vorzeichen-Rangsummentest (Mann-Whitney Test)

Der Wilcoxon Vorzeichen-Rangsummentest ist ein weiterer Signifikanztest. Dieser Test untersucht anhand der Bildung von Rangzahlen, ob zwei Stichproben sich voneinander unterscheiden. Der Rangsummentest ordnet jedem Wert aus der ersten Stichprobe eine Rangzahl *R* zu. Der kleinste Wert erhält die Zahl 1, der zweitkleinste die Rangzahl 2 usw. Aus der Summe dieser Rangzahlen, der ersten Stichprobe wird die Testgröße *T* gebildet.

$$T = \sum_{i=1} R_i$$

Für die zweite Stichprobe wird identisch vorgegangen und es wird die Testgröße *t* gebildet. Wenn die Unterschiede zwischen den Rangzahlensummen der Testgrößen *T* und *t* zu groß ist, dann kann folglich die Hypothese, dass die beiden Stichproben gleichwertig sind, abgelehnt werden. Der Unterschied zwischen den beiden Testgrößen muss hierbei größer oder kleiner sein, als dass dieser durch Zufallszahlen erklärt werden könnte (Sheldon M, 2018).

4.11 Ergebnis der Verteilungsanalyse zwischen den Schäden im Nachbargebiet und im Schutzgebiet, an Flugtagen

Die Signifikanztests in „R“ wurden folgendermaßen durchgeführt:

```
myT=t.test(SGVert$RelDmg,NGVert$RelDmg, alternative = "two.sided",var.equal = FALSE,paired = FALSE,conf.level = 0.95)
myF=var.test(SGVert$RelDmg,NGVert$RelDmg,alternative = "two.sided",paired = FALSE,conf.level = 0.95)
myW=wilcox.test(SGVert$RelDmg,NGVert$RelDmg,alternative = "two.sided",paired = FALSE,conf.level = 0.95)
```

Der hier zu betrachtende Parameter ist der p-Value, welcher mit dem Konfidenzintervall $\alpha = 0.05$ verglichen wird. Wenn der p-Value kleiner ist, als das Konfidenzintervall α , dann kann die Hypothese (H_0), dass die beiden Gebiete gleichwertig sind, verworfen werden.

Für die einzelnen Tests wurden die in Tabelle 7 gezeigten p-Values erzielt:

Tabelle 7: Ergebnis der einzelnen statistischen Tests

	p-Value	Ergebnis	Interpretation
T-Test	0,005867833	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich
F-Test	0	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich
Wilcox-Test	2,399381e-09	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich

Die Signifikanztests haben für die gebietsweisen Schädigungen ergeben, dass die Werte aus zwei unterschiedlichen Grundgesamtheiten stammen.

4.12 Ergebnis der Verteilungsanalyse zwischen den Schäden im Schutzgebiet an Flug und Nichtflugtagen

Die Untersuchung der Schäden, welche sich nur im Schutzgebiet zugetragen haben, sollen untersucht werden. Vor allem ist hier interessant, ob sich die relativen Schäden, welche sich an Flugtagen ereignet haben, von denen unterscheiden, welche an Tagen passiert sind, an denen kein Hagelflug stattgefunden hat.

Es wurde dieselbe Methodik verwendet, wie im Kapitel zuvor. Die Signifikanztests haben für die gebietsweisen Schädigungen auch ergeben, dass die Werte aus zwei unterschiedlichen Grundgesamtheiten stammen. Für die einzelnen Tests wurden die, in Tabelle 8 gezeigten, p-Values erzielt:

Tabelle 8: Ergebnis der einzelnen statistischen Tests

	p-Value	Ergebnis	Interpretation
T-Test	0,005364716	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich
F-Test	0	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich
Wilcox-Test	1,603632e-05	$p < \alpha \rightarrow H_1$ annehmen	Gebiete sind ungleich

4.13 Ergebnis der Verteilungsanalysen

Die Verteilungsanalyse hat die Schäden im Nachbargebiet und Schutzgebiet an Flugtagen untersucht. Ebenso wurden die Schäden im Schutzgebiet an Flugtagen und an nicht Flugtagen betrachtet. Für beide Betrachtungsfälle sind die Gebiete für alle Signifikanztests ungleich ausgefallen. Dies ermöglicht zwar keinen Aufschluss über die Wirksamkeit der Hagelabwehr, stellt jedoch eine interessante Tatsache fest. Nämlich, dass die Schäden an Flugtagen und nicht Flugtagen im Schutzgebiet sich unterscheiden, was ein Indiz für einen Effekt durch die Hagelabwehr ist. An Tagen wo nicht geflogen wurde und es zu Schäden kommt würde es sich dann eher um Unwetterfronten handeln, welche zu schnell aufgezogen sind, um sie effektiv bekämpfen zu können. Die Ungleichheit zwischen Schutzgebiet und Nachbargebiet kann vermutlich mit dem topographischen Unterschied, zwischen Nachbargebiet und Schutzgebiet, erklärt werden. Auch hier muss wieder erwähnt werden, dass die Piloten nicht zwischen den beiden Gebieten unterscheiden (siehe Kapitel 4.9). Das Skript zur Verteilungsanalyse in R kann im Anhang A gefunden werden.

4.14 Möglichkeit der Generierung eines Faktors zur Klassifikation, der Unwetterstärke

Da Gewitter unterschiedliche Intensitäten aufweisen, soll ein Faktor gebildet werden, welcher versucht die Stärke eines Unwetters zu berücksichtigen. Dieser Faktor könnte dann in die relative Schädigung miteinfließen und diese unterschiedlich stark gewichten. Zum Beispiel kann der Großteil eines Unwetters nördlich von München sein komplettes Schädigungspotenzial entfalten und Richtung Rosenheim ziehen, wo es nur noch abgeschwächten Schaden verursacht. Der Schaden, welcher durch das abgeschwächte Gewitter passiert müsste anders gewichtet werden, als der Schaden für das Nachbargebiet. Für verschiedene Naturphänomene existieren schon Bewertungsskalen, wie z.B. die Beaufort-Skala, welche die Windgeschwindigkeit bewertet oder die Saffir-Simpson-Hurrikan-Skala (Thomas Sävert, 2011). Problematisch bei diesen bestehenden Skalen ist, dass keine historischen Daten zu vergangenen Unwettern vorhanden sind oder, dass eine Skala ein Naturphänomen nicht metrisch erfasst, wie z.B. die Fujita Tornado-Skala (Thomas Sävert, 2011). Auch ist es fraglich, inwiefern für die Bewertung von Hagelunwetter, z.B. die Windgeschwindigkeit, herangezogen werden kann. Die Hagelabwehr Rosenheim besitzt jedoch seit einigen Jahren ein Messsystem, welches die atmosphärische Impulsstrahlung misst. Für die Hagelabwehr selbst wird dieses Messsystem als Indikator genutzt, ob sich ein größeres Unwetter aufbaut.

4.14.1 Atmosphärische Impulsstrahlung

Bei Atmosphärischer Impulsstrahlung (kurz AIS) handelt es sich um elektromagnetische Entladungen, welche von Blitzen verursacht werden. Solche auftretenden Entladungen können über das in Europa verteilte LINET System, welches von der Firma nowCast GmbH entwickelt wurde, gemessen und lokalisiert werden. Das LINET System wird unter anderem von Versicherern, Energieversorgern und Wetterdiensten genutzt, um Gewitter besser vorhersagen zu können. Im Falle von Versicherern und Energieversorgern wird LINET verwendet, um einen Überspannungsschaden, welcher durch einen Blitz verursacht wurde, nachzuweisen. Die Firma nowCast hat sich freundlicherweise bereit erklärt historische Daten zur Verfügung zu stellen.

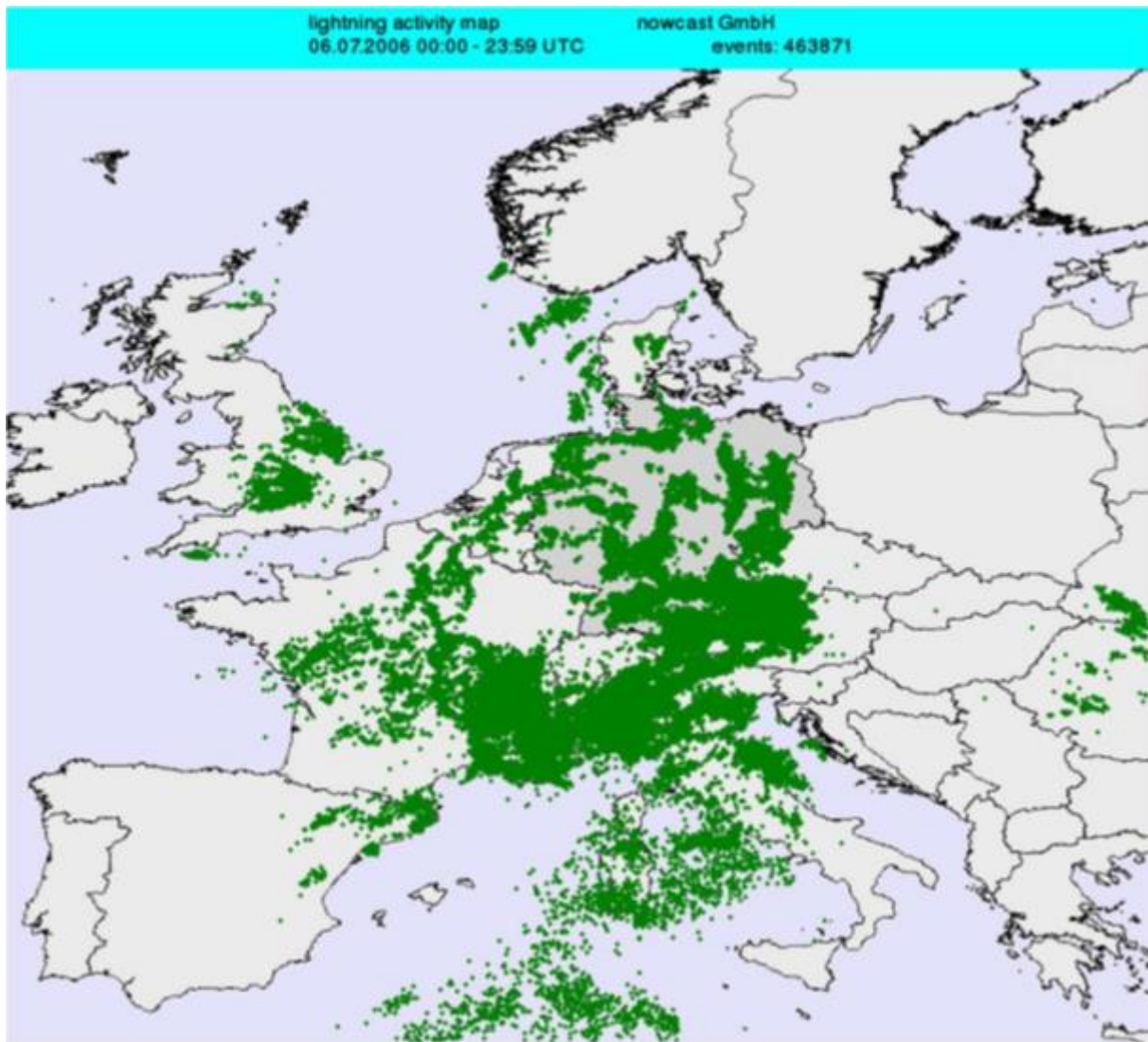


Abbildung 93: Karte welches, die vom LINET gemessene Blitzaktivität vom 06.07.2006 zeigt. (Betz, 2009)

Die zur Verfügung stehenden Daten decken rasterweise einen Großteil von Südbayern, und damit auch das Schutzgebiet und Nachbarggebiet, ab. Die Daten stehen für alle Flugtage zur Verfügung und beinhalten die gemessenen atmosphärischen Entladungen je Raster (siehe Abbildung 94), für einen Zeitraum von einer Stunde vor dem stattfinden des Hagelabwehrfluges, bis eine Stunde nach Abschluss des Hagelabwehrfluges. Die, an allen Tagen erfassten, Zeiträume betragen im Mittelwert 3 Stunden und 52 Minuten, mit einer Standardabweichung von einer Stunde. Die so erfassten Zeiträume werden daher als lang genug angenommen, um das an dem jeweiligen Tag stattgefundenene Hagelunwetter zu erfassen.

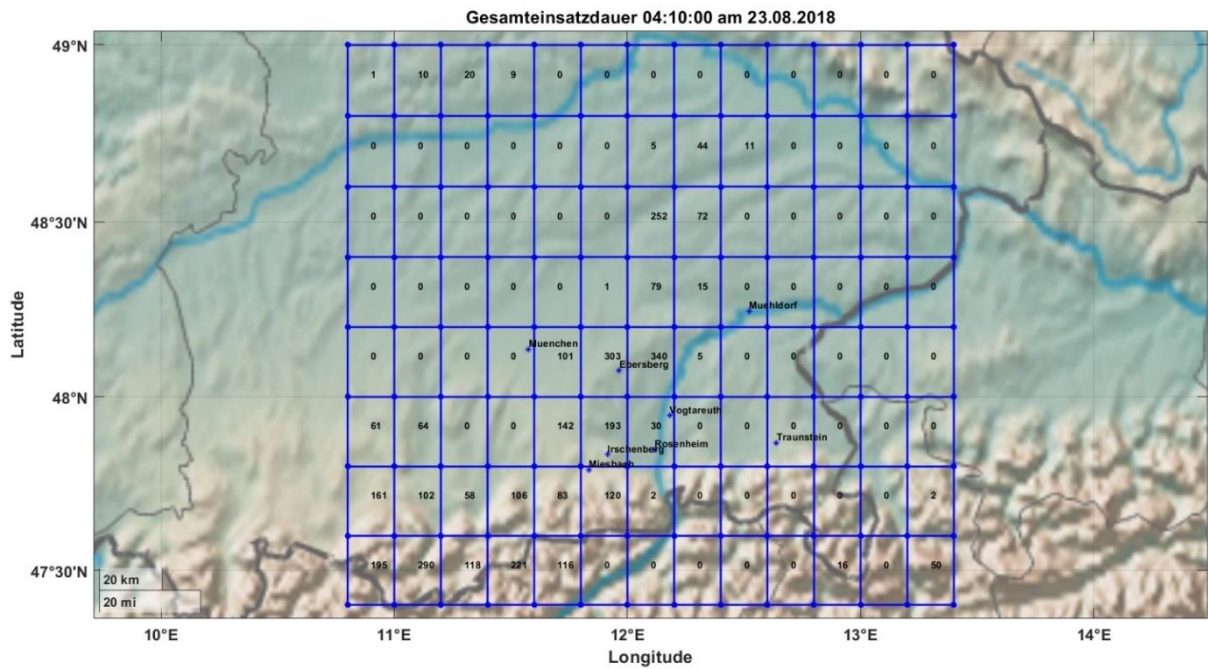


Abbildung 94: Erfasstes Gebiet. Die Zahl pro Raster zeigt die Anzahl der detektierten Blitze

Es wurde ein Skript in MATLAB geschrieben (siehe Anhang A), welches in der Lage ist, für jeden erfassten Hagelabwehrflug die Anzahl der detektierten Blitze für jedes Raster darzustellen. Die so, pro Hagelabwehreinsatz, erfassten Blitze wurden aufsummiert und es wurde untersucht, ob diese in Korrelation mit den erfolgten absoluten Schäden, in Schutzgebiet und Nachbargebiet, stehen. Hierfür wird ein linearer Zusammenhang angenommen. Es wird versucht zwischen folgenden Daten ein Vorhersagemodell zu bestimmen:

- AIS Aktivität – Absoluter Schaden Schutzgebiet und Nachbargebiet (summiert)
- AIS Aktivität – Absoluter Schaden Schutzgebiet
- AIS Aktivität – Absoluter Schaden Nachbargebiet

Aufgrund der unterschiedlichen Schadensgrößen werden die Ergebnisse logarithmisch skaliert. Die Schäden wurden hierbei nicht um Ausreißer bereinigt, sondern es handelt sich um alle absoluten Schäden an Flugtagen, zu denen AIS Aktivitätsdaten vorliegen. Das Fitting wurde in MATLAB programmiert und kann im Anhang A gefunden werden.

4.14.2 AIS-Aktivität – Absoluter Schaden Schutzgebiet und Nachbargebiet

Für die Bewertung des Modells wird der Goodness-of-Fit Parameter R^2 erneut betrachtet. Der Parameter R^2 misst, wie auch bei der multiplen linearen Regression, wie gut der Fit die Variation der Daten erklärt. Der R^2 nimmt Werte zwischen 0 und 1 an, wobei ein Wert von 1 die Daten zu 100% erklären würde. Das in Abbildung 95 gezeigte Modell erzielt lediglich einen Wert von 0,12. Das bedeutet, dass nur 12% der Daten durch dieses Modell erklärt werden können. Ein so geringer Wert kann als ungenügend betrachtet werden.

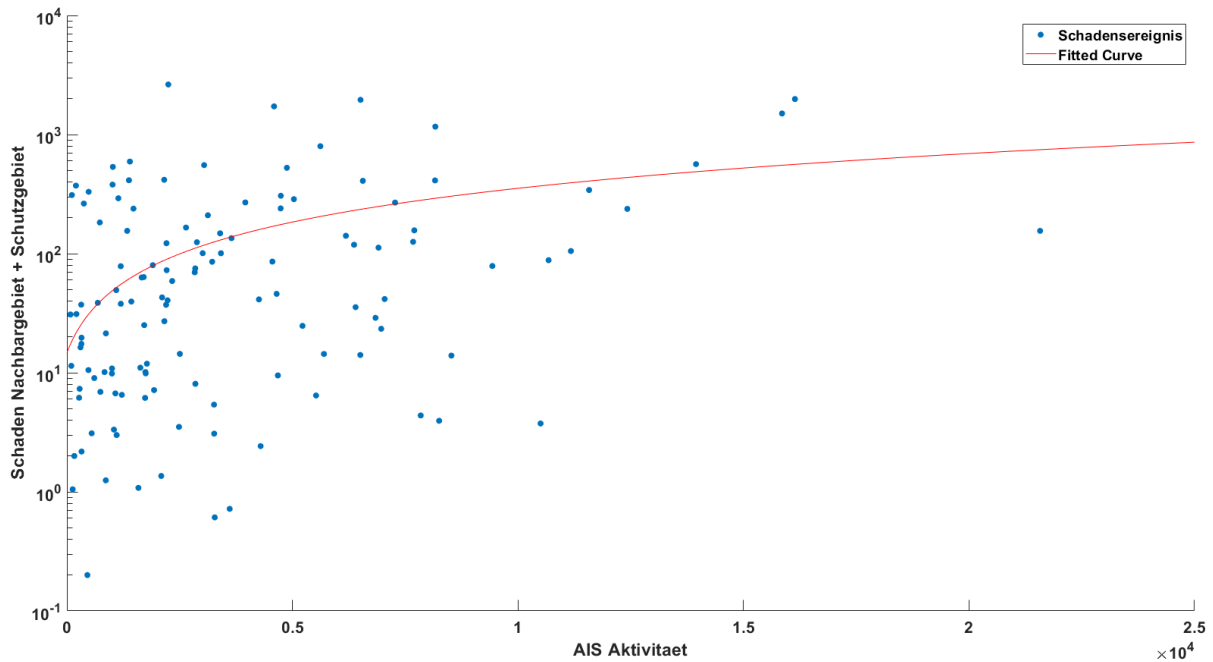


Abbildung 95: Fit für die absoluten Schäden Schutzgebiet und Nachbargebiet

4.14.3 AIS-Aktivität – Absoluter Schaden Schutzgebiet

Für die Schäden im Schutzgebiet konnten ein R^2 von 0,17 erzielt werden. Dies stellt zwar eine Verbesserung dar, ist aber immer noch zu gering.

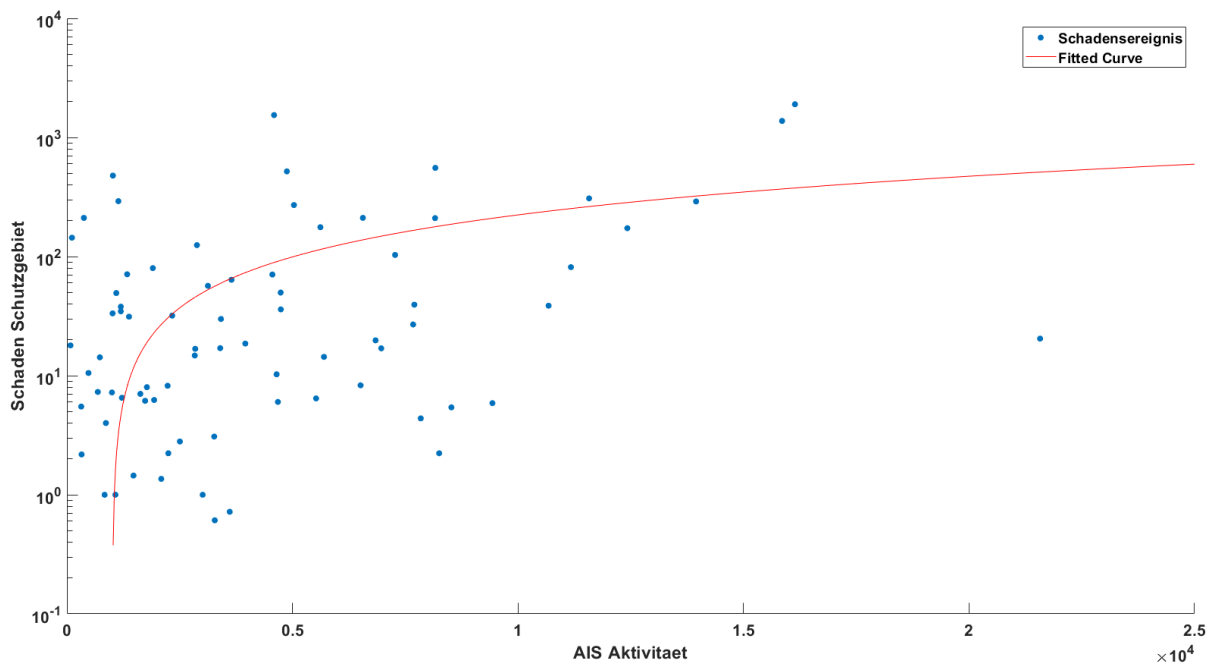


Abbildung 96: Fit für die absoluten Schäden Schutzgebiet

4.14.4 AIS-Aktivität – Absoluter Schaden Nachbargebiet

Für die absoluten Schäden im Nachbargebiet konnte ein R^2 von 0,0156 erzielt werden, was 1,56 % entspricht und ebenfalls als zu gering einzustufen ist.

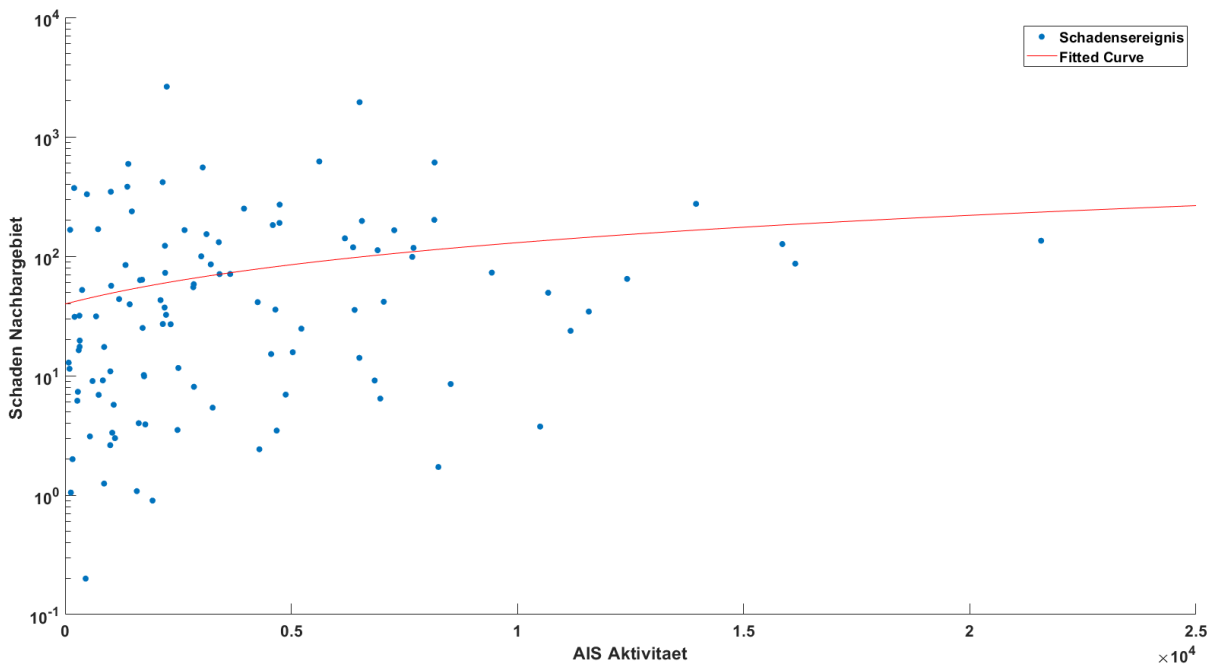


Abbildung 97: Fit für die absoluten Schäden Nachbargebiet

4.14.5 AIS-Aktivität – Absoluter Schaden Schutzgebiet und Nachbargebiet bereinigt

Da die AIS-Aktivität mit den unbereinigten Daten keinen Zusammenhang erkennen lässt, wurden die Daten um Ausreißer bereinigt. Als Ausreißer wurden hierbei alle Tage gekennzeichnet, welche größer als das 1,5-fache des oberen Quartils plus das 1,5-fache des Quartilabstands ist (Wellisch, 2018).

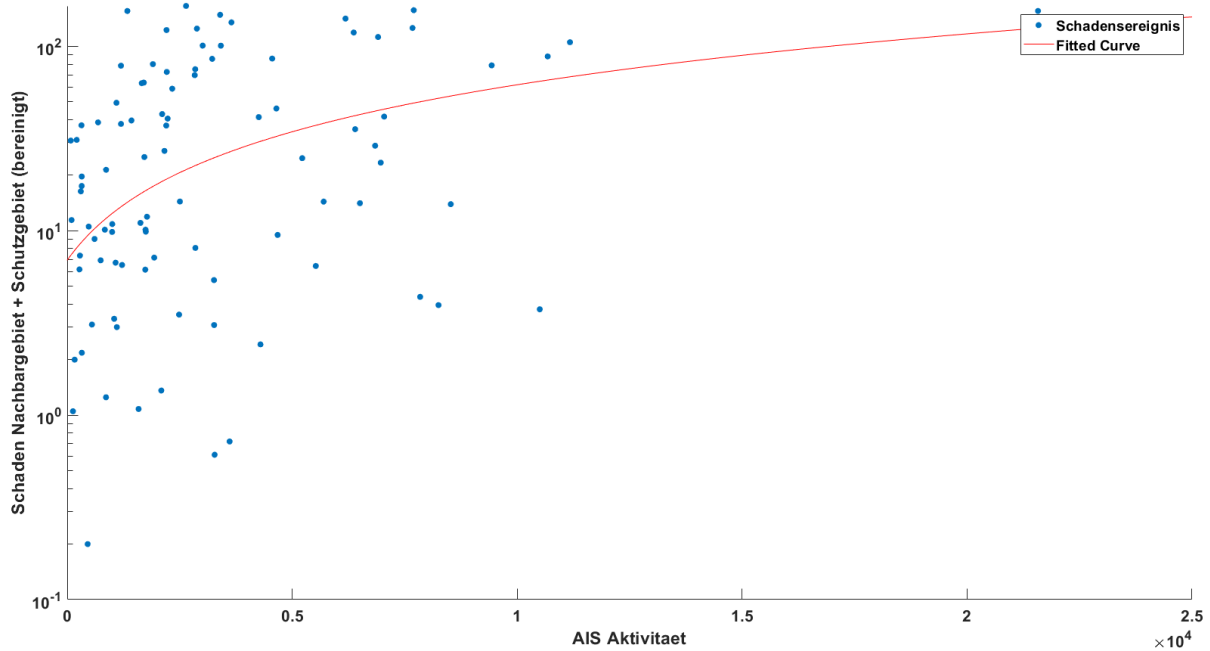


Abbildung 98: Fit für die absoluten bereinigten Schäden Schutzgebiet und Nachbargebiet. Das erzeugte Modell besitzt einen R^2 von 0,1695

4.14.6 AIS-Aktivität – Absoluter Schaden Schutzgebiet um Ausreißer bereinigt

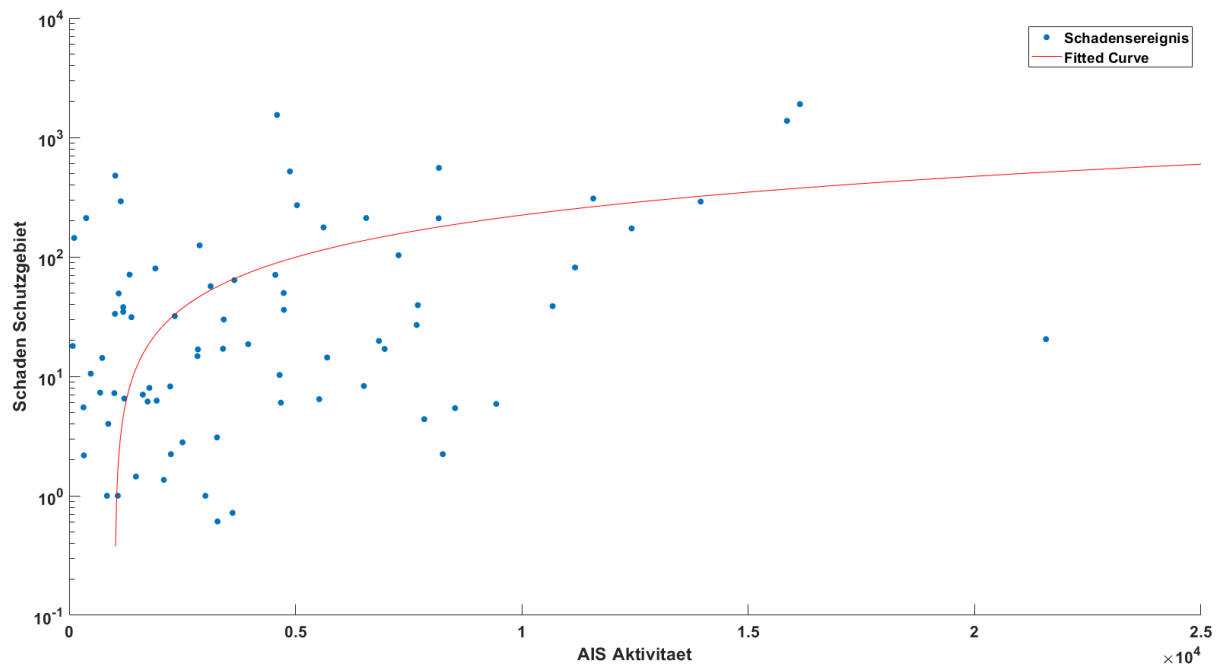


Abbildung 99: Fit für die absoluten bereinigten Schäden im Schutzgebiet. Das erzeugte Modell besitzt einen R^2 von 0,0267

4.14.7 AIS-Aktivität – Absoluter Schaden Nachbargebiet um Ausreißer bereinigt

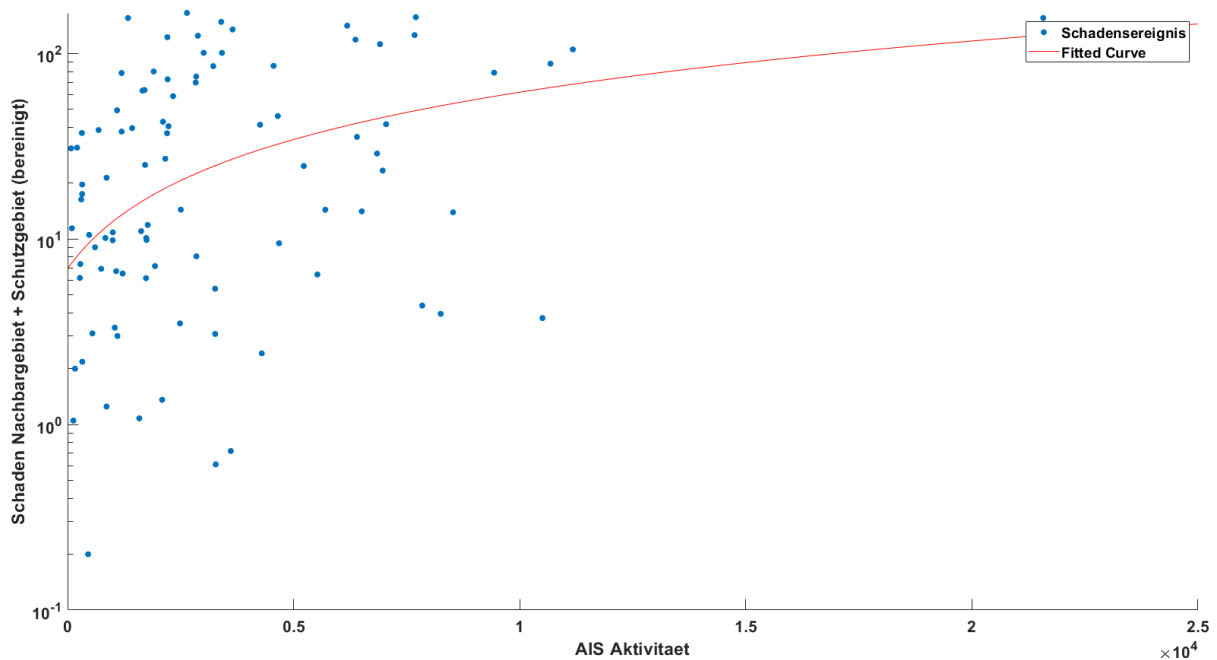


Abbildung 100: Fit für die absoluten bereinigten Schäden im Schutzgebiet. Das erzeugte Modell besitzt einen R^2 von 0,1034

4.14.8 Fazit zur AIS-Aktivität im Zusammenhang mit den zu erwartenden Schäden

Abbildung 95 bis Abbildung 100 zeigen alle statistischen Modelle, welche einen ungenügenden R^2 aufweisen. Die AIS-Aktivität ist zwar geeignet zur Vorhersage von Unwettern, stellt aber eine schlechte Möglichkeit dar, die Schwere eines Hagelunwetters zu erfassen. Aus diesen Daten kann daher kein Faktor gebildet werden, welcher zur Beurteilung der Schwere von Hagelunwettern dient.

4.15 Fazit zur Analyse der Versicherungsdaten

Die Analyse der Versicherungsschadensdaten stellte einen möglichen Ansatz dar, die Wirksamkeit der Hagelabwehr nachzuweisen. Hierbei konnte jedoch kein wirksamer Effekt durch die Hagelabwehr nachgewiesen werden. Die erzeugten statistischen Modelle aus Kapitel 4.3, 4.6 und 4.7 besitzen entweder eine zu geringe statistische Signifikanz, für die Hagelabwehr relevanten Prädiktoren, oder die signifikanten Prädiktoren, wie die ausgebrachte Silberjodid Menge aus Kapitel 4.7.1 ist hoch signifikant und besitzt einen positiven Estimate. Dies kann so interpretiert werden, dass je mehr Silberjodid ausgebracht wird, desto mehr relativer Schaden ist theoretisch zu erwarten. Der ausgebrachten Silberjodidmenge steht jedoch gegenüber, dass die Fluganzahl ein signifikanter Prädiktor ist und einen negativen Estimate besitzt, was für eine Schadensreduktion spricht. Das ansteigend des Schadens durch die Silberjodidmenge kann daher eher auf extrem Ereignisse zurückgeführt werden, an denen die Hagelabwehr sich einer Superzelle gegenüber sieht, welche auch schon im Nachbargebiet entstehen. An solchen Tagen wird besonders viel Silberjodid ausgebracht, obwohl die Gewitterzelle sich im Nachbargebiet schon voll ausgebildet hat.

5 Zusätzliche Ergänzungen zum ROBERTA-Projekt

Da nicht beide Flugzeuge der Hagelabwehr Rosenheim ihre GPS-Position während eines Einsatzes senden können, war es notwendig, eine Alternative zu finden, um dennoch diese Flüge auswerten zu können. Im Folgenden sollen zwei Möglichkeiten gezeigt werden, wie vergangene Flüge erfasst werden können.

5.1 Datenaquise von *FlightAware*

Das Unternehmen *FlightAware* betreibt eine Website, die eine kostenlose Flugverfolgung, von Privat- und Verkehrsflugzeugen, rund um den Globus ermöglichen. Sie ist derzeit die weltweit größte Flugverfolgungsplattform, mit einem Netzwerk von über 23.000 ADS-B-Bodenstationen in 197 Ländern (FlightAware, 2019). Da die Internetseite www.flightaware.com auch vergangene Hagelabwehrflüge zeigt, wurde ein Programm in MATLAB geschrieben, welches in der Lage ist, vergangene Flüge (siehe Abbildung 101) von *FlightAware* zu extrahieren und im Anschluss zu visualisieren.

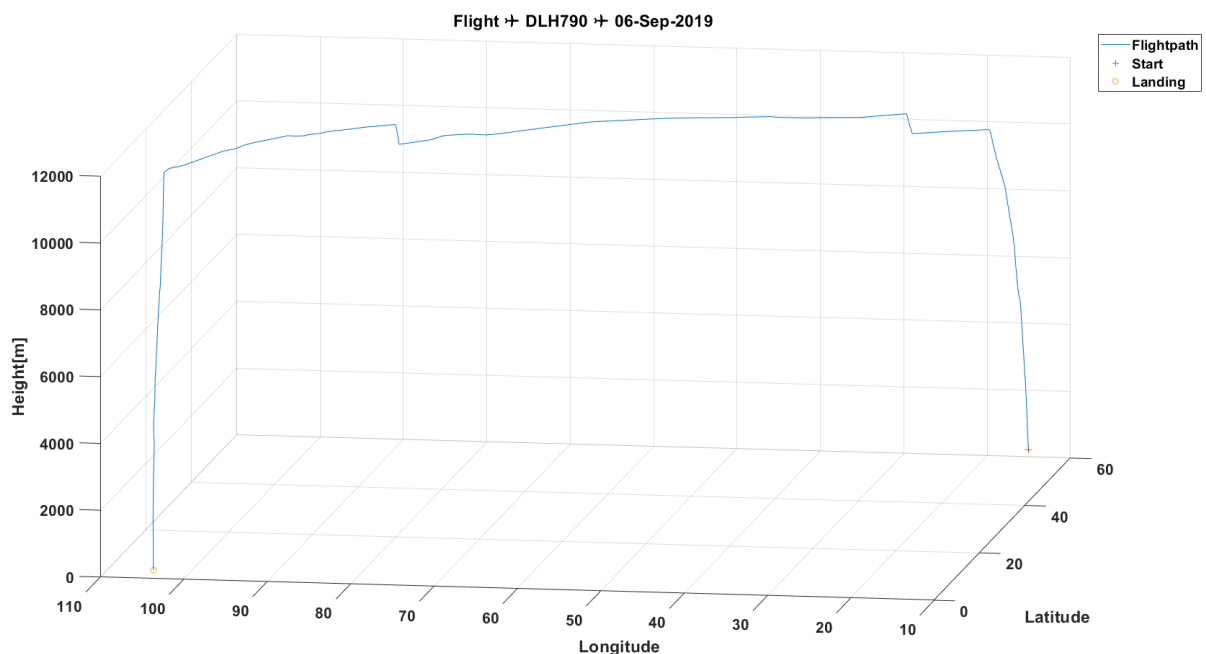


Abbildung 101: Ein vergangener Lufthansa Flug von München nach Singapur

5.2 Raspberry ADS-B Empfänger Datenbeschaffung

Flightradar24.com ist ein Online-Service zur Echtzeit-Positionsanzeige von Flugzeugen und arbeitet ähnlich zu dem vorgestellten Dienst aus Kapitel 5.1. Das Flightradar24-Netzwerk arbeitet mithilfe von ADS-B Empfängern, welche von Privatpersonen betrieben werden und die Positionsdaten von zivilen Flugmaschinen erfassen. Diese Positionsdaten werden auf www.flightradar24.com zu Verfügung gestellt. Um Daten im CSV Format zu erhalten ist es notwendig, für den Service von Flightradar24 zu bezahlen oder einen eigenen ADS-B Empfänger zur Verfügung zu stellen und diesen in das Netzwerk von Flightradar24 zu integrieren. Hierzu wurde ein Raspberry Pi 3 mit einer ADS-B Antenne gekoppelt, wie Abbildung 102 zeigt. Eine detaillierte Anleitung zum Aufsetzen eines Raspberry Pis zu einer ADS-B Empfänger kann unter (Flightradar24, 2019) gefunden werden. Im Folgenden soll der Export von Flightradar24 Daten erklärt werden.



Abbildung 102: Raspberry-Pi mit ADS-B Antenne

5.2.1 Flightradar24 Export

Flightradar24 kann über www.Flightradar24.com erreicht werden. Hier zeigt sich die Standardkartenansicht, welche alle Flüge für jeden beliebigen Nutzer (auch ohne Premium Account) darstellt (siehe Abbildung 103). Über den Reiter rechts oben kann auf *Fleetview* geklickt werden. *Fleetview* lässt nach bestimmten Airlines sortieren, z.B. Lufthansa oder ANA. Unter Custom Fleets sind eigene Flugzeuge hinterlegt. Momentan sind die beiden Flieger, der Hagelabwehr Rosenheim, mit dem Callsign *D-GITY* und *D-GOGO* hinterlegt. Die Hagelabwehr der WGV mit dem Callsign *D-EWGV* wurde ebenfalls angelegt (WGV, 2019). Durch klicken auf eines der Flieger im *Fleetview* können vergangene Flüge betrachtet werden (siehe Abbildung 104).

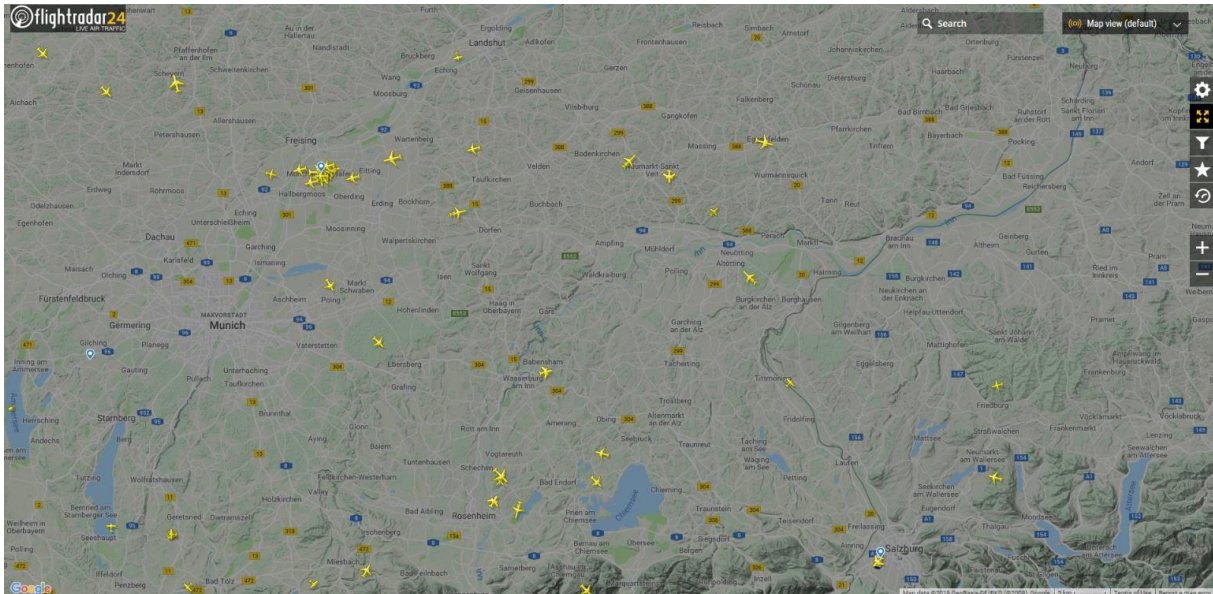


Abbildung 103: Standardansicht der Karte von flightradar24

Bei Auswahl eines bestimmten Fliegers, können dessen Daten als KML oder als CSV exportiert werden. Durch klicken auf den Play-Button kann der Flug jederzeit in Flightradar24 betrachtet werden (siehe Abbildung 105).

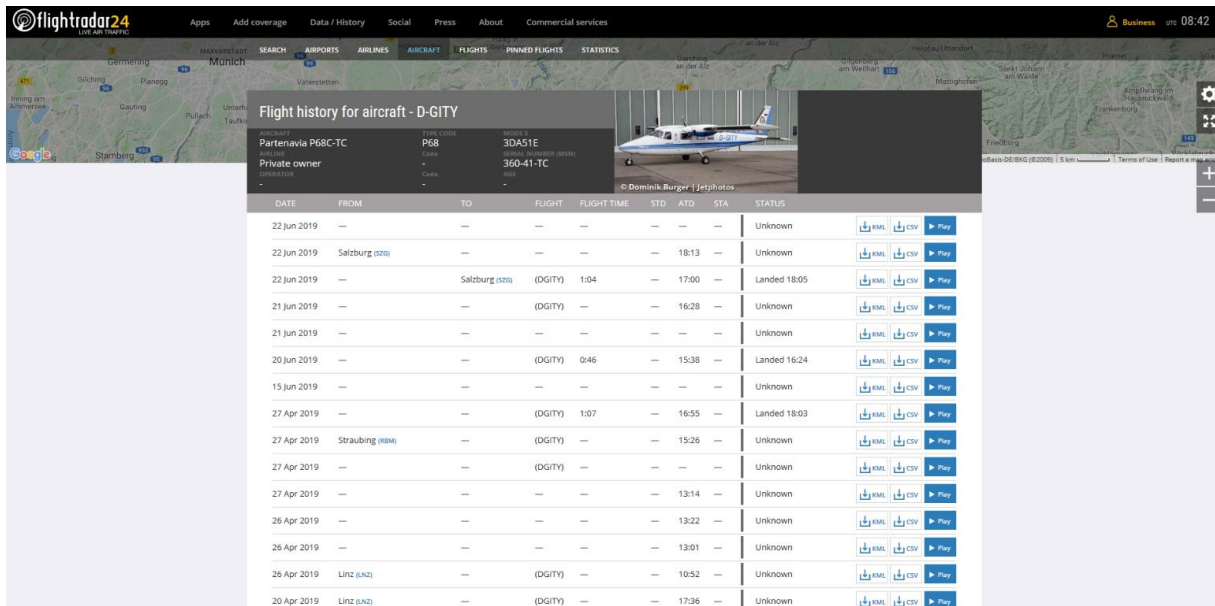


Abbildung 104: Ansicht aller vergangenen Hagelabwehrflüge der Maschine D-GITY

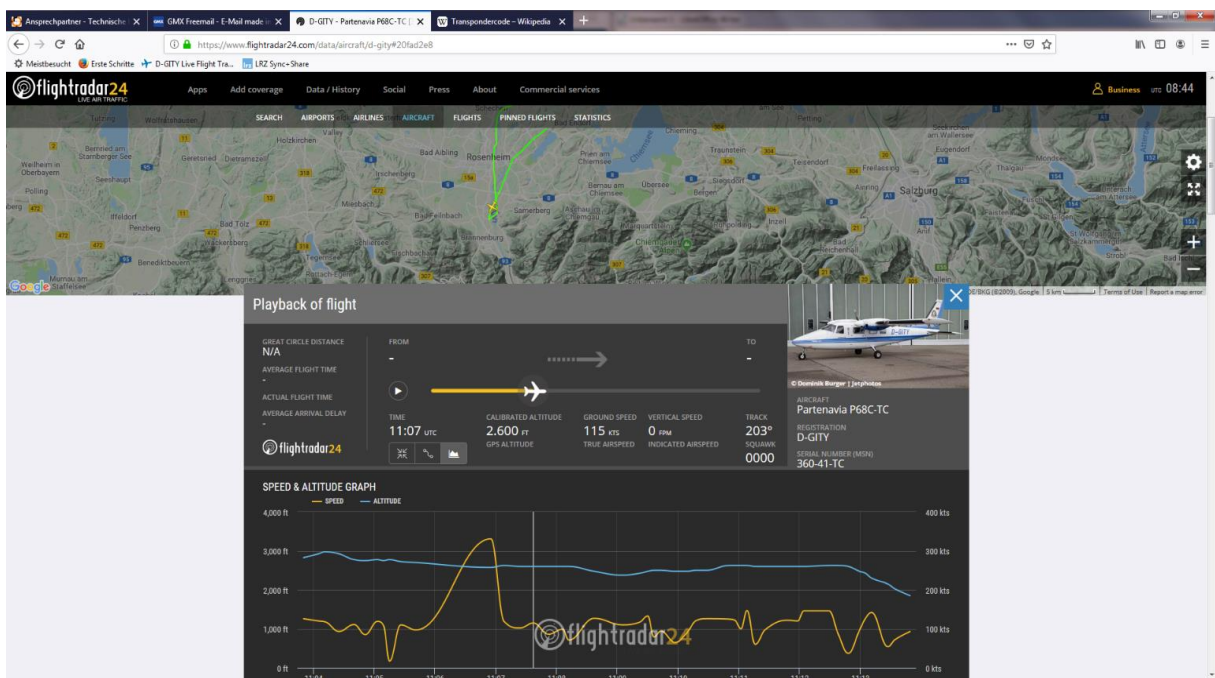


Abbildung 105: Detailsicht eines vergangenen Hagelabwehrfluges für die Maschine D-GITY

Tabelle 9 zeigt einen Überblick über alle Daten, welche Flightradar24 zur Verfügung stellt.

Tabelle 9: Von Flightradar zur Verfügung gestellte Daten

Flugzeugdaten	Kartendaten
Calibrated Altitude	Vulkanische Eruptionen (Aschewolke)
Vertical Speed	Wetter (Sonne, Regen, Temperatur)
GPS Altitude	Wolkenfelder
Track	Niederschlagsmenge
GroundSpeed	Windgeschwindigkeiten für verschiedene Höhen
True Airspeed	Blitze
Indicated Airspeed	Airmet/Sigmet ⁵
Mach	Vorausgesagte Vereisungen für bestimmte Höhen
Windspeed & Direction	Wolkenturbulenzen
Temperature	Luftturbulenzen
Transpondercode	

⁵AIRMET (englisch *AIRman's METeorological Information*) ist ein Wetterberatungssystem für den Flugverkehr. Im Vergleich zu SIGMETs decken AIRMETs weniger gefährliche Wettererscheinungen ab und betreffen sichere Flugdurchführungen unter FL245. AIRMETs werden nur bei Bedarf herausgegeben

6 Fazit

In dieser Masterarbeit wurden zwei verschiedene Datenarten mit unterschiedlichen Methoden untersucht, um die Wirksamkeit der Hagelabwehr untersuchen zu können. Der erste Ansatz betrachtet die Auswirkung von Hagelabwehrflügen auf die Reflexivitätsstufen von Gewitterwolken, welche geimpft werden. Die Methodik aus Kapitel 2.3.3 lieferte hier keine vielversprechenden Ergebnisse (siehe Kapitel 3.1.2). Dies ist darauf zurückzuführen, dass die Impf- bzw. Beobachtungsgeometrie als zu statisch in ihrer Größe angenommen wird und sich fest um das Flugzeug befindet. Die daraufhin entwickelte Methodik aus Kapitel 2.7.5 stellt eine bessere Methodik zur Untersuchung von Gewitterzellen dar, da hier die Geometrie auf die Gewitterzelle konzentriert wird und nicht auf das Hagelabwehrflugzeug. Die hier erzielten Ergebnisse aus Kapitel 3.1.3, für die untersuchten Flüge, weisen negative Driftkennzahlen auf. Dies spricht für eine Wirksamkeit von Hagelabwehrflügen. Dennoch sollte die hier vorgestellte Methodik weiterentwickelt werden. Hierzu sollen die, in Kapitel 2.3.1 und 2.7.6 gezeigten, Algorithmen zur Bildung eines konvexen Polygons genutzt werden. Dadurch könnten Gewitterzellen als ein sich dynamisch veränderndes Objekt erfasst werden und eine maschinelle Auswertung, mittels künstlicher neuronaler Netze, kann erfolgen. Die dadurch zu Gewitterclustern zusammengeführten Objekte sollten aufgrund ihrer enthaltenen Echostufen noch einen Schweregrad zugeordnet bekommen.

Der zweite Ansatz beruht auf Versicherungsdaten, welche von zwei großen Versicherern zur Verfügung gestellt wurden. Es wurden hierbei mehrere statistische Modelle gebildet. Das hier erzeugte Modell, für die multiple lineare Regressionsanalyse, weist eine zu geringe statistische Gesamtsignifikanz auf, wie Kapitel 4.3 zeigt. Die generalisierten linearen Modelle, aus Kapitel 4.6 und 4.8, weisen den Prädiktor, die ausgebrachte AgJ-Menge, mit einer erhöhten Signifikanz aus. Das, mithilfe des *AIC* verfeinerte, generalisierte lineare Modell, zeigt eine Wirkung der Hagelabwehr (siehe Kapitel 4.7.1.1). Die ausgebrachte AgJ-Menge besitzt hierbei jedoch einen positiven Estimate, welcher somit nicht für ein geringeres Schädigungspotenzial spricht. Dem Prädiktor ausgebrachte Silberjodidmenge steht jedoch gegenüber, dass die Fluganzahl ebenfalls ein signifikanter Prädiktor ist und einen negativen Estimate besitzt, was für eine Schadensreduktion spricht. Das Ansteigen des Schadens durch die Silberjodidmenge kann daher auf Ereignisse zurückgeführt werden, an denen die Hagelabwehr sich voll ausgebildeten Gewitterzellen gegenübersteht, welche schon im Nachbargebiet entstehen. An solchen Tagen wird besonders viel Silberjodid ausgebracht, obwohl sich die Gewitterzelle im Nachbargebiet bereits vollständig ausgebildet hat.

Das Ziel der Masterarbeit war es, einen Einfluss der Hagelabwehr auf Hagelschäden zu untersuchen. Es konnten statistische Nachweise für ein positives Wirken der Hagelabwehr auf mehrere Arten erbracht werden (siehe Zusammenfassungstabelle). Speziell die, in dieser Masterarbeit erdachten, Methoden für den Radaransatz müssen unbedingt weiterentwickelt werden. Da auf diesen aufbauend, eine tiefere Analyse der Thematik erfolgen kann.

Tabelle 10 zeigt eine Zusammenfassung der durchgeführten Methoden und ihrer Ergebnisse. Die grau markierten Methoden, stehen für Methoden, bei denen keine Aussage über die Wirksamkeit der Hagelabwehr getroffen werden kann. Die grünen Felder zeigen eine Wirksamkeit der Hagelabwehr auf.

Tabelle 10: Ergebnis der durchgeführten Methodiken

Kapitel	Methodik	Ergebnis
3.1.2	Auswertung des Geometrieansatzes	Kein Nachweis aufgrund von zu statischen Geometrien möglich
3.1.3	Auswertung des Cloudtargetansatzes	Anzeichen einer Wirksamkeit der Hagelabwehr
3.1.3 und 3.1.2	Test auf Zufallszahlen	Unbrauchbar aufgrund von vermutlich zu hoher Quantisierung der Radardaten.
4.3	Multiple lineare Regressionsanalyse	Zu geringe statistische Signifikanz des Modells
4.5	Generalisiertes lineares Modell	Zu geringe statistische Signifikanz der einzelnen Hagelabwehr relevanten Prädiktoren
4.7	Bereinigtes Generalisiertes lineare Modell	Zu geringe statistische Signifikanz der einzelnen Hagelabwehr relevanten Prädiktoren
4.7.1	Verbesserung des GLM mithilfe des stepAIC	Signifikanter p-Value für die Fluganzahl mit negativen Estimate. Hohe statistische Signifikanz für die ausgebrachte AgJ-Menge mit positiven Estimate
4.8.1.1	Vergleich der relativen Schäden mithilfe eines Scoring-Verfahrens für Flugtage	Nur sehr wenige vergleichbare Tage
4.8.1.2	Vergleich der relativen Schäden mithilfe eines Scoring-Verfahrens für nicht Flugtage	Nur sehr wenige vergleichbare Tage
4.11	Analyse des Schadens im Nachbargebiet, ob dieser aus einer anderen Verteilung kommt, als der Schaden an Flugtagen im Schutzgebiet	Schäden kommen aus zwei unterschiedlichen Verteilungen
4.12	Untersuchung ob der Schaden von Flugtagen im Schutzgebiet aus einer anderen Verteilung kommt, als an Tagen im Schutzgebiet ohne Flug	Schäden kommen aus zwei unterschiedlichen Verteilungen
4.14	Generierung eines Faktors zur Bewertung der Schwere von Hagelunwettern	Kein Zusammenhang zwischen Schädigung und AIS-Aktivität

7 Literaturverzeichnis

- B. Federer, A. W. (1986). *Main Results of Grossversuch IV*. Schweiz.
- Bartels, R. (März 1982). *The Rank Version of von Neumann's Ratio Test for Randomness*.
researchgate.net. Von
https://www.researchgate.net/publication/230639951_The_Rank_Version_of_von_Neumann's_Ratio_Test_for_Randomness abgerufen
- Betz, P. D.-D. (2. 02 2009). <https://www.sciencedirect.com/science/article/pii/S0169809508002305#!>
Von <https://www.sciencedirect.com/science/article/pii/S0169809508002305> abgerufen
- Bundesamt für Kartographie und Geodäsie. (07. 06 2019). www.bkg.bund.de. Von
<https://www.bkg.bund.de/DE/Produkte-und-Services/Shop-und-Downloads/Digitale-Geodaten/Quasigeoid/quasigeoid.html> abgerufen
- Burtsev, G. K. (1979). Hail process investigation and hail suppression activities in the USSR. USSR.
Von <http://journalofweathermodification.org/index.php/JWM/article/view/497/539> abgerufen
- Cairo, F. (20. Februar 2015). cran.r-project.org. Von <https://cran.r-project.org/web/packages/randtests/randtests.pdf> abgerufen
- Ellen Ferrante, A. M. (29. 06 2011). www.nsf.gov. Von
https://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=120925 abgerufen
- Eurocontrol & IfEN. (12. 02 1998). www.icao.int. Von
<https://www.icao.int/safety/pbn/documentation/eurocontrol/eurocontrol%20wgs%2084%20implementation%20manual.pdf> abgerufen
- FlightAware. (07. 09 2019). <https://de.flightaware.com/>. Von <https://de.flightaware.com/> abgerufen
- Flightradar24. (07. 09 2019). www.flightradar24.com. Von <https://www.flightradar24.com/build-your-own> abgerufen
- Frassek, D. B. (20. 06 2019). www.aa.quae.nl. Von <https://www.aa.quae.nl/en/reken/grootcirkel.html> abgerufen
- GIS Geography. (20. 06 2019). www.gisgeography.com. Von <https://gisgeography.com/cylindrical-projection/> abgerufen
- Google Inc. (12. 06 2019). www.google.com. Von <https://www.google.de/maps> abgerufen
- Günster, K. (26. 08 2019). www.karlheinz-guenster.de. Von <https://www.karlheinz-guenster.de/serendipity/archives/348-Weterradarturm-in-Schnaapping.html> abgerufen
- Hagelabwehr Rosenheim. (08. 09 2019). www.hagelabwehr-rosenheim.de. Von
<http://www.hagelabwehr-rosenheim.de/geschichtliches.php> abgerufen
- Hochschule Flensburg. (20. 08 2019). Von www.iti.fh-flensburg.de: <http://www.iti.fh-flensburg.de/lang/algorithmen/geo/index.htm> abgerufen
- Kompf, M. (22. 5 2019). www.kompf.de. Von <https://www.kompf.de/gps/distcalc.html> abgerufen

- Landratsamt Rems-Murr-Kreis. (2015/2016). *Hagelabwehr in der Region Stuttgart*. Von https://www.rems-murr-kreis.de/fileadmin/Dateien/Dateien/%C3%84mter/Hagelabwehr/RMK_Wissenschaftliche-Begleituntersuchung-2017_web.pdf abgerufen
- Liebau, F. (07. 06 2019). *kartenkunde-leichtgemacht*. Von <http://kartenkunde-leichtgemacht.de/handbuch.php?page=Kartennetzentwurf> abgerufen
- Liebau, F. (20. 06 2019). *kartenkunde-leichtgemacht.de*. Von <http://kartenkunde-leichtgemacht.de/handbuch.php?page=Koordinatensysteme> abgerufen
- Lindsey, J. K. (1997). In J. K. Lindsey, *Applying Generalized Linear Models* (S. 18). Springer.
- Lindsey, J. K. (12. 05 2019). *leg.ufpr.br*. Von <https://bit.ly/2ZmDr0c> abgerufen
- Lindsey, P. (2007). *Applying Generalized Linear Models*. Springer.
- Manfred Wasshuber. (08. 09 2019). *www.astroimages.at*. Von <http://www.astroimages.at/photography/gewitterwolke.htm> abgerufen
- Martin Hagen,DLR. (2007). *www2.meteo.uni-bonn.de*. Von https://www2.meteo.uni-bonn.de/projekte/SPPMeteo/wiki/lib/exe/fetch.php?id=cops_2a&cache=cache&media=cops-talk-hagen.ppt abgerufen
- MathWorks. (17. August 2019). *MathWorks*. Von https://de.mathworks.com/help/database/index.html?s_tid=CRUX_lftnav abgerufen
- MathWorks. (18. August 2019). *www.mathworks.com*. Von <https://de.mathworks.com/products/parallel-computing.html> abgerufen
- MathWorks. (18. 08 2019). *www.mathworks.com*. Von <https://de.mathworks.com/help/parallel-computing/gpu-support-by-release.html;jsessionid=4f9ea7e6f4fdddf23abefae9d18> abgerufen
- Matuschek. (11. 06 2019). *www.matuschek.net*. Von <http://www.matuschek.net/interpolation/> abgerufen
- Movable. (09. 06 2019). *www.movable-type.co.uk*. Von <https://www.movable-type.co.uk/scripts/latlong-vincenty.html#direct> abgerufen
- NASA. (19. 08 2019). *www.lpdaacsvc.cr.usgs.gov*. Von <https://lpdaacsvc.cr.usgs.gov/appears/> abgerufen
- Oracle. (17. August 2019). *Oracle*. Von <https://www.oracle.com/technetwork/java/javase/jdbc/index.html> abgerufen
- Pantle, U. (14. 07 2004). Von [mathematik.uni-ulm.de/stochastik/lehre/ss04/statistik1/skript/node25.html](http://www.mathematik.uni-ulm.de/stochastik/lehre/ss04/statistik1/skript/node25.html) abgerufen
- proj4. (2. 5 2019). <https://proj4.org/>. Von <https://proj4.org/> abgerufen
- Pruscha, H. (2006). In *Statistisches Methodenbuch* (S. 320). München: Springer.
- Pruscha, H. (2006). *Statistisches Methodenbuch*. Berlin Heidelberg: Springer.

- Pruscha, H. (2006). Statistisches Methodenbuch. In H. Pruscha, *Statistisches Methodenbuch* (S. 18). München: Springer.
- Schlegel, B. (05. 06 2016). *www.politikwissenschaften.ch*. Von <https://www.politikwissenschaften.ch/blog.php?id=19> abgerufen
- Sheldon M, R. (2018). Statistik für Ingenieure. In R. Sheldon M, *Statistik für Ingenieure* (S. 474). Springer.
- Smith, R. K. (08. 11 2004). *www.meteo.physik.uni-muenchen.de*. Von https://www.meteo.physik.uni-muenchen.de/lehre/roger/Einfuehrung_Teil_III/Einfuehrung_Teil_III_041108.pdf abgerufen
- Stieglmair, W. (03. 09 2019). *www.sturmwetter.de*. Von <http://www.sturmwetter.de/texte/hagelentstehung.htm> abgerufen
- Süddeutsche Zeitung. (13. 06 2019). Großer Hagelschaden. *Süddeutsche Zeitung*. Von <https://www.sueddeutsche.de/bayern/unwetter-bayern-schaden-versicherung-1.4485727> abgerufen
- The R Foundation. (29. 08 2019). *www.rdocumentation.org*. Von <https://www.rdocumentation.org/packages/stats/versions/3.6.1/topics/AIC> abgerufen
- Thomas Sävert, A. W. (Oktober 2011). *www.unwetterzentrale.de*. Von <http://www.unwetterzentrale.de/uwz/901.html> abgerufen
- TH-Rosenheim. (04. 09 2019). *www.play.google.com/store*. Von <https://play.google.com/store/apps/details?id=de.fhrosenheim.roberta.hailsimandroid&hl=de> abgerufen
- Uni-Rostock. (07. 06 2019). *Geoinformatik.uni-rostock*. Von <http://www.geoinformatik.uni-rostock.de> abgerufen
- Wagner, M. (2018). *Vorlesungsmanuskript Bildverarbeitung in der Produktion*. Rosenheim: Wagner, Michael.
- Wellisch, U. (2018). *Einführung in die Datenanalyse und Statistik mit R*. Rosenheim: Technische Hochschule Rosenheim.
- WGV. (07. 09 2019). <https://www.wgv.de>. Von <https://www.wgv.de/hagelabwehr/> abgerufen
- Wikipedia. (11. 06 2019). *www.wikipedia.de*. Von https://en.wikipedia.org/wiki/Bilinear_interpolation abgerufen
- Wikipedia. (19. 09 2019). *www.wikipedia.org*. Von <https://de.wikipedia.org/wiki/Impulsvolumen> abgerufen
- Wolff, C. (14. 2 2019). Von www.radartutorial.eu/04.history/hi04.de.html abgerufen
- Wolff, C. (14. 2 2019). *radartutorial*. Von <http://www.radartutorial.eu> abgerufen
- Zuo, W. (19. 06 2019). *www.mayr.informatik.tu-muenchen.de*. Von http://www.mayr.informatik.tu-muenchen.de/konferenzen/Jass05/courses/2/Zuo/Zuo_paper.pdf abgerufen

8 Anhang

Die einzelnen Anhänge und Programmcodes sind auf der beiliegenden CD in folgenden Ordnern zu finden:

Datenbank Konnektivitäts Beispiel	A\Matlab\SQLTests\ZentgrafDemo
Laufmängenkodierung	A\C++\RunLength
Distanzberechnung	A\Matlab\Distanzberechnung
Mercatortransformation	A\Matlab\Echoansatz\Koordinatensystem A\C++\Mercator
1D Interpolation	A\Matlab\Mathematik\Mapinterpolate\Interpolation
2D Interpolation	A\Matlab\Mathematik\Mapinterpolate\Interpolation
Graham-Scan-Algorithmus	A\Matlab\GrahamScan A\C++\GrahamScan
Particleflow-Simulation	A\Matlab\Echoansatz\ParticleFlowSimulation
Turning-point-test	A\Matlab\Mathematik\Zufallszahlen
Scoring Verfahren	A\Matlab\Versicherungsansatz\Auswertung
Hailshow	A\Matlab\Echoansatz\HailShow
VirtFlightCreator	A\Matlab\Echoansatz\VirtFlightCreator
Multiple lineare Regression	A\R\Regressionen\VersicherungsDaten\AllDmg
GLM	A\R\Regressionen\VersicherungsDaten\AllDmg
GLM mit stepAIC	A\R\Regressionen\VersicherungsDaten\AllDmg
Zufallszahlentest	A\R\Regressionen\EchoDaten
Auswertung des Echoansatzes	A\R\Regressionen\EchoDaten
Kritische Zahlen für Bartels Rank Test	B\Statistik
Verwendete statistische Methoden	B\Statistik
Theorie zu AIS	B\Blitze
Untersuchungen zur Hagelabwehreffizienz	B\Hagelabwehreffizienz
Theorie Radar	B\Radar
Theorie Kartentransformationen	B\Kartentransformationen
Wolkenmeteorologie	B\Meteorologie
Datenbankdokumentation	B\Datenbank
Flugberichte	B\Flugberichte
Anleitung DatabaseExplorer(Matlab)	B\Datenbank\DatabaseExplorerAnleitung
Raspberry-Pi Flihtadar ISO	B\Raspberry