

Universität Stuttgart

Masterarbeit am ifp

Lena Joachim

Wolkenerkennung in  
Satellitenaufnahmen des  
sichtbaren  
Spektralbereichs bei Nacht



Betreuer: Dr. Tobias Storch (DLR)

Prüfer: Prof. Dr.-Ing. Norbert Haala



## Zusammenfassung

Wolkenerkennung für Satellitenaufnahmen des sichtbaren Spektralbereichs und nahen Infrarots (VNIR) erfolgt bei Nacht üblicherweise anhand von zusätzlichen Beobachtungen im Thermalen Infrarot (TIR). In dieser Arbeit wird untersucht, inwiefern die Wolkenerkennung auch ausschließlich anhand der VNIR-Beobachtungen durchgeführt werden kann. Dies würde es ermöglichen, bei einer auf nächtliche VNIR-Beobachtungen spezialisierten Satellitenmission (N8) auf die Integration eines zusätzlichen TIR-Sensors zu verzichten.

Auf Grundlage der nächtlichen Satellitenbilder des VIIRS/Day-Night-Band wird ein Verfahren vorgeschlagen, das mit Random Forest eine semantische Segmentierung der Bilder in die beiden Klassen *klar* und *wolkig* durchführt. Hierfür werden auf Grundlage des Beobachtungsmodells mehrere Merkmale zur pixelweisen Beschreibung der durch Wolken beeinflussten Eigenschaften vorgeschlagen, dies sind vor allem die Albedo, die optische Dicke, Streuungseffekte und zeitliche Veränderungen. Der Fokus liegt hierbei auf der Anwendung über künstlich beleuchteten urbanen Gebieten.

Das in dieser Arbeit entwickelte Verfahren wurde sowohl an urbanen Bildausschnitten getestet, als auch an solchen, die Landbedeckungen repräsentieren, die das Verfahren besonders herausfordern. Die besten Ergebnisse konnten dem Fokus der Erwartungen entsprechend mit Gesamtgenauigkeiten von 85% bzw. 79% für zwei der urbanen Testgebiete erzielt werden.

## Abstract

Cloud detection for night-time images of the visible and near-infrared spectrum (VNIR) is usually performed based on additional observations of the thermal infrared (TIR). This thesis examines, if it is possible to perform the cloud detection solely based on the VNIR observations. This would make it possible to design a new night-time VNIR satellite mission (N8) without adding an additional TIR sensor.

Based on night-time satellite images of the VIIRS/Day-Night-Band a cloud detection method, which uses Random Forest to perform a semantic segmentation of the images into the classes *clear* and *cloudy*, is developed. To do so, several model-based features are used to describe the effects of clouds on the observations. Amongst the considered effects and cloud properties are the albedo, cloud optical thickness, scattering and temporal changes. The method is developed with focus on the application over urban areas with artificial lighting.

The proposed method is tested both over urban areas and over areas, which represent special types of landcover where difficulties are expected. According to the expectations, the best results could be achieved for two of the urban test areas, where the overall accuracies are 85% and 79%.



## **Erklärung der Urheberschaft**

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Das elektronische Exemplar stimmt mit den gedruckten Exemplaren überein.

Stuttgart, 14.11.2019

gez. Lena Joachim



# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>IX</b>
<b>Tabellenverzeichnis</b>	<b>X</b>
<b>Abkürzungsverzeichnis</b>	<b>XI</b>
<b>1. Einleitung</b>	<b>1</b>
<b>2. Grundlagen</b>	<b>5</b>
2.1. Datengrundlage . . . . .	5
2.1.1. Day-Night-Band (DNB) . . . . .	5
2.1.2. Wolkenmaske (ECM) . . . . .	8
2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen . . . . .	10
2.2.1. Allgemeines Beobachtungsmodell . . . . .	10
2.2.2. Modell für die Beleuchtung durch den Mond . . . . .	11
2.2.3. Lichtemissionen der Erdoberfläche und weitere Beleuchtungsquellen . . . . .	14
2.2.4. Einfluss von Wolken auf VNIR-Satellitenbeobachtungen bei Nacht . . . . .	15
2.3. Algorithmus zur semantischen Segmentierung – Random Forests . . . . .	20
2.3.1. CART – Classification And Regression Trees . . . . .	20
2.3.2. Bagging – Bootstrap Aggregating . . . . .	21
2.3.3. Training und Klassifikation mit Random Forests . . . . .	22
<b>3. Methodik</b>	<b>23</b>
3.1. Vorverarbeitung . . . . .	23
3.1.1. Testgebiete . . . . .	23
3.1.2. Vorverarbeitungskette . . . . .	30
3.1.3. Herausfiltern von nicht geeigneten Bildern . . . . .	34
3.2. Definition von Merkmalen . . . . .	35
3.2.1. Einzelbildmerkmale . . . . .	35
3.2.2. Bildserienmerkmale . . . . .	40
3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken . . . . .	43
3.3.1. Kriterium zur Optimierung des Verfahrens . . . . .	44
3.3.2. Klassifikationsverfahren . . . . .	45
3.3.3. Definition des Merkmalsvektors – Bestimmung der aussagekräftigsten Merkmalskombination . . . . .	46
3.3.4. Fenstergröße . . . . .	50
3.3.5. Ground Truth . . . . .	50
<b>4. Analyse der Ergebnisse</b>	<b>53</b>
4.1. Definition der betrachteten Genauigkeitsmaße . . . . .	53
4.2. Ergebnisse für verschiedene Testgebiete . . . . .	55
4.2.1. Ergebnisse pro Gebiet . . . . .	55

4.2.2. Ergebnisse pro Bild . . . . .	59
4.2.3. Einfluss von Bildserienmerkmalen auf die Ergebnisse . . . . .	68
4.3. Variation der Trainingsdaten . . . . .	69
4.4. Variation der Sensitivität des Sensors . . . . .	74
<b>5. Fazit und Ausblick</b>	<b>77</b>
5.1. Ausblick . . . . .	78
<b>Literatur</b>	<b>XIII</b>
<b>A. Weitere Ergebnisbilder</b>	<b>XIX</b>
<b>B. Informationen über die verwendeten Bilder</b>	<b>XXIX</b>

## Abbildungsverzeichnis

1.1.	Black Marble . . . . .	1
2.1.	Mit DNB bei Nacht beobachtbare Lichtquellen . . . . .	6
2.2.	Beispiel für Artefakte in ECM . . . . .	9
2.3.	Allgemeines Beobachtungsmodell für optische Nachtaufnahmen . . . . .	10
2.4.	Beispiel für lunare Strahldichten. . . . .	14
2.5.	Blitz und Polarlicht in DNB-Bildern . . . . .	15
2.6.	Beispiele für das Erscheinungsbild von Flächen mit verschiedener Albedo in DNB-Bildern . . . . .	17
2.7.	Beispiel für Streuungseffekt durch Wolken . . . . .	18
2.8.	Durch Wolken verursachter Georeferenzierungsfehler . . . . .	19
3.1.	Beispiel eines Bildausschnittes des Testgebietes <i>muenchen</i> . . . . .	24
3.2.	Beispiel eines Bildausschnittes des Testgebietes <i>stuttgart</i> . . . . .	24
3.3.	Beispiel eines Bildausschnittes des Testgebietes <i>belgien</i> . . . . .	25
3.4.	Beispiel eines Bildausschnittes des Testgebietes <i>mailand</i> . . . . .	26
3.5.	Beispiel eines Bildausschnittes des Testgebietes <i>neworleans</i> . . . . .	26
3.6.	Beispiel eines Bildausschnittes des Testgebietes <i>indien</i> . . . . .	27
3.7.	Beispiel eines Bildausschnittes des Testgebietes <i>indozean</i> . . . . .	28
3.8.	Beispiel eines Bildausschnittes des Testgebietes <i>petronius</i> . . . . .	28
3.9.	Beispiel eines Bildausschnittes des Testgebietes <i>nil</i> . . . . .	29
3.10.	Beispiele eines Bildausschnittes des Testgebietes <i>alpen</i> . . . . .	29
3.11.	Ablaufdiagramm der Vorverarbeitungskette . . . . .	31
3.12.	Ablaufdiagramm des Unterprogramms der Vorverarbeitungskette . . . . .	32
3.13.	Vergleich der Ergebnisse für verschiedene Versionen des Merkmalsvektors . . . . .	49
4.1.	Ergebnisse pro Testgebiet für Training mit <i>muenchen</i> . . . . .	56
4.2.	Beste Ergebnisse für <i>stuttgart</i> . . . . .	61
4.3.	Schlechteste Ergebnisse für <i>stuttgart</i> . . . . .	62
4.4.	Abhängigkeit der Ergebnisse von der Mondbeleuchtung für alle Testgebiete . . . . .	63
4.5.	Abhängigkeit der Ergebnisse von der Mondbeleuchtung für <i>petronius</i> . . . . .	64
4.6.	Abhängigkeit der Ergebnisse von der Mondbeleuchtung für <i>nil</i> . . . . .	65
4.7.	Beste Ergebnisse für <i>nil</i> . . . . .	66
4.8.	Schlechteste Ergebnisse für <i>nil</i> . . . . .	67
4.9.	Ergebnisse für verschiedene Trainingsdatenversionen pro Testgebiet . . . . .	71
4.10.	Ergebnisse für verschiedene Trainingsdatenversionen für die urbanen Test- gebiete . . . . .	71
4.11.	Ergebnisse für verschiedene Trainingsdatenversionen für die nicht urbanen Testgebiete . . . . .	72
4.12.	Ergebnisse für verschiedene Trainingsdatenversionen pro Testgebiet mit Fo- kus auf Bildanzahl . . . . .	73
4.13.	Ergebnisse für verschiedene radiometrische Sensitivitäten . . . . .	75
A.1.	Beste Ergebnisse für <i>belgien</i> . . . . .	XX
A.2.	Schlechteste Ergebnisse für <i>belgien</i> . . . . .	XXI

A.3. Beste Ergebnisse für <i>neworleans</i> . . . . .	XXII
A.4. Schlechteste Ergebnisse für <i>neworleans</i> . . . . .	XXIII
A.5. Beste Ergebnisse für <i>petronius</i> . . . . .	XXIV
A.6. Schlechteste Ergebnisse für <i>petronius</i> . . . . .	XXV
A.7. Weitere Ergebnisse für <i>petronius</i> . . . . .	XXVI
A.8. Beste Ergebnisse für <i>alpen</i> . . . . .	XXVII
A.9. Schlechteste Ergebnisse für <i>alpen</i> . . . . .	XXVIII

## Tabellenverzeichnis

4.1. Klassenverteilung des Testdatensatzes . . . . .	53
4.2. Definition der Konfusionsmatrix . . . . .	54
4.3. Konfusionsmatrix für die semantische Segmentierung des gesamten Testdatensatzes . . . . .	55
4.4. Konfusionsmatrix für die semantische Segmentierung von <i>stuttgart</i> . . . . .	57
4.5. Konfusionsmatrix für die semantische Segmentierung von <i>belgien</i> . . . . .	57
4.6. Konfusionsmatrix für die semantische Segmentierung von <i>neworleans</i> . . . . .	58
4.7. Konfusionsmatrix für die semantische Segmentierung von <i>petronius</i> . . . . .	58
4.8. Konfusionsmatrix für die semantische Segmentierung von <i>nil</i> . . . . .	58
4.9. Konfusionsmatrix für die semantische Segmentierung von <i>alpen</i> . . . . .	58
4.10. Vergleich von Versionen ohne und mit zeitlichen Merkmalen . . . . .	68
4.11. Trainingsdatenversionen . . . . .	70
B.1. Bildinformationen für <i>stuttgart</i> . . . . .	XXX
B.2. Bildinformationen für <i>belgien</i> . . . . .	XXXI
B.3. Bildinformationen für <i>neworleans</i> . . . . .	XXXII
B.4. Bildinformationen für <i>petronius</i> . . . . .	XXXIII
B.5. Bildinformationen für <i>nil</i> . . . . .	XXXIV
B.6. Bildinformationen für <i>alpen</i> . . . . .	XXXV

## Abkürzungsverzeichnis

BA	Balanced Accuracy
CART	Classification And Regression Trees
DLR	Deutsches Zentrum für Luft- und Raumfahrt e.V.
DNB	Day-Night-Band des Sensors VIIRS
DMSP/OLS	U.S. Air Force Defense Meteorological Satellite Program, Operational Lincan System
ECM	NOAA Enterprise Cloud Mask
GLCM	Grey Level Co-Occurrence Matrix
GT	Ground Truth
MVCM	MODIS-VIIRS Cloud Mask
N8	Nacht/Night, Name einer in der Planung befindlichen VNIR-Satellitenmission des DLR
OA	Overall Accuracy
S-NPP	Suomi National Polar-Orbiting Partnership, amerikanischer Wetter- und Umweltsatellit
TIR	Thermales Infrarot
TOA	Top-Of-Atmosphere
VCM	VIIRS Cloud Mask
VIIRS	Visible Infrared Imaging Radiometer Suite
VNIR	Visible and Near Infrared



# 1. Einleitung

Der Fußabdruck menschlicher Aktivitäten auf der Erde ist besonders deutlich bei Nacht zu erkennen, wie das Mosaik nächtlicher Satellitenbilder in Abbildung 1.1 zeigt. Dies liegt daran, dass menschliches Leben und Arbeiten bei Nacht meist mit künstlicher Beleuchtung einhergeht. Da diese in Satellitenaufnahmen des sichtbaren und nahen Infrarot-Bereiches (kurz: VNIR) sehr deutlich sichtbar ist, enthalten solche Bilder eine Vielzahl an Informationen, die so explizit aus keinem anderen Fernerkundungsprodukt abgeleitet werden können [29], [38], [54].

So können auf ihrer Basis nicht nur besiedelte Flächen kartiert und charakterisiert und die Entwicklung der Urbanisierung beobachtet werden, sondern die Beleuchtung einer Region lässt auch Rückschlüsse auf Eigenschaften der dort lebenden Bevölkerung zu. Zum Beispiel beeinflussen der Wohlstand und weitere sozioökonomische Faktoren die Art und das Ausmaß der Beleuchtung und verschiedene zeitliche Beleuchtungsmuster lassen auf kulturelle Unterschiede schließen [42]. Aus Analysen nächtlicher Beleuchtung können außerdem Aussagen und Prognosen bezüglich ihrer Auswirkungen auf die Umwelt, z.B. in Form von Lichtverschmutzung, Energieverbrauch oder Treibhausgasemissionen, abgeleitet werden.

Bisher gibt es jedoch nur wenige satellitengestützte Sensoren, die Daten für solche Analysen liefern. Die ersten nächtlichen VNIR-Satellitenbilder stammen vom amerikanischen DMSP/OLS-Sensor [32], der seit den Siebzigerjahren panchromatische Aufnahmen im VNIR-Bereich mit täglicher globaler Abdeckung und 2.7 km räumlicher Auflösung erzeug-

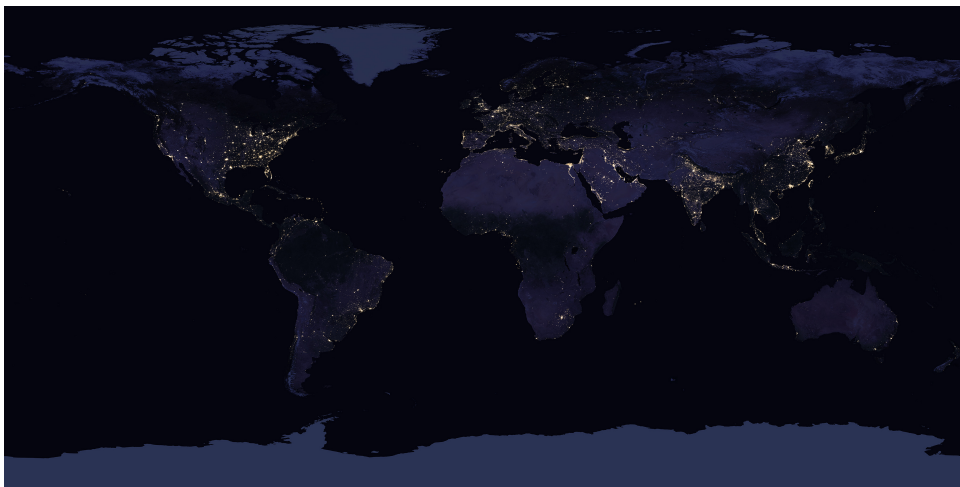


Abbildung 1.1: Black Marble: Wolkenfreies Mosaik nächtlicher Satellitenbilder des sichtbaren und nahen Infrarot-Bereichs (panchromatisch, 500 – 900 nm) aus dem Jahr 2016 [43]. Die zur Erzeugung verwendeten Bilder wurden vom VIIRS/Day-Night-Band an Bord des US-amerikanischen Erdbeobachtungssatelliten S-NPP aufgenommen (vgl. Kapitel 2.1.1). Die Farbinformation dient lediglich Visualisierungszwecken.

## 1. Einleitung

te, die allerdings nicht radiometrisch kalibriert waren. Sein in vielen Dingen ähnlicher, aber verbesserter, Nachfolger ist seit 2011 das VIIRS/Day-Night-Band (vgl. Kapitel 2.1.1) [28], das nicht nur eine deutlich bessere räumliche Auflösung von 750 m, sondern auch die ersten vollständig kalibrierten nächtlichen Satellitenbilder liefert. Seit 2018 gibt es mit LJ1-01 [30] eine chinesische CubeSat-Mission, die ebenfalls global panchromatische nächtliche Bilder aufnimmt, jedoch nicht mit täglicher Abdeckung. Die räumliche Auflösung wurde hier auf 130 m verbessert.

Neben diesen drei global operierenden Missionen, deren Daten frei verfügbar sind, gibt es einige teilweise kommerzielle Satellitenmissionen, die sich nur auf bestimmte Interessensgebiete konzentrieren, aber dafür eine bessere räumliche Auflösung und meist auch spektrale Informationen besitzen. Beispiele hierfür sind unter anderem AeroCubes [39], JL1-3B [55] oder auch die Fotografien der Astronauten auf der ISS [27].

Da frei verfügbare Daten mit globaler Abdeckung derzeit also nur in einer groben räumlichen Auflösung und als panchromatische Beobachtungen vorliegen, können kleinräumige Strukturen urbaner Gebiete nicht dargestellt und auch keine spektralen Informationen abgeleitet werden. Deshalb wurde am Deutschen Zentrum für Luft- und Raumfahrt (DLR) die Idee entwickelt, eine Satellitenmission (N8) zu entwerfen, die ausschließlich auf die Beobachtung der Erde in VNIR bei Nacht optimiert ist und es so ermöglicht, genauere Analysen der nächtlichen künstlichen Beleuchtung durchzuführen. Ein ähnlicher Missionsvorschlag wird in [13] vorgestellt. Solche sowohl räumlich als auch spektral höher aufgelösten Daten würden eine Vielzahl völlig neuer Analysemöglichkeiten eröffnen, sowie bestehende Anwendungen verbessern. Eine mögliche Anwendung wäre zum Beispiel die Charakterisierung der städtischen Straßenbeleuchtung hinsichtlich der Lage der Lampen und des Typs des Leuchtmittels [35]. Eine solche Analyse kann im globalen Kontext deutlich detailliertere Rückschlüsse auf Parameter wie den Energieverbrauch oder die Lichtverschmutzung liefern, als es mit derzeit verfügbaren Daten möglich ist.

Die Voraussetzung von auf Basis optischer Daten durchgeführten Analysen von Eigenschaften der Erdoberfläche, wie z.B. der künstlichen Beleuchtung, ist eine zuverlässige Wolkenerkennung. Diese wird benötigt, um all diejenigen Pixel aus der Analyse auszuschließen, bei denen die Erdoberfläche durch Wolken verdeckt wird. Gerade bei Nacht ist dies essenziell, da Bewölkung aufgrund der fehlenden Beleuchtung durch die Sonne im VNIR-Bereich ohne Mondbeleuchtung nicht direkt sichtbar ist. In diesem Fall kann zum Beispiel keine Aussage darüber getroffen werden, ob eine sonst künstlich beleuchtete Fläche nicht mehr beleuchtet ist, oder lediglich von einer Wolke verdeckt wird.

Die operationellen Wolkenerkennungsalgorithmen, mit denen die Wolkenmasken für die existierenden nächtlichen Satellitenbilder generiert werden, basieren deshalb in der Regel nicht auf den VNIR-Beobachtungen, sondern vor allem auf Beobachtungen im Bereich des Thermalen Infrarots (TIR). Dort lassen sich Wolken sehr gut erkennen, da sie in der Regel deutlich kälter sind als die Erdoberfläche. Da es sich hierbei um von der Erdoberfläche oder den Wolken emittierte Strahlung handelt, kann diese auch nachts unabhängig von der Beleuchtung beobachtet werden [23].

Typischerweise werden zur nächtlichen Wolkenerkennung hauptsächlich Beobachtungen im Bereich von  $11\ \mu\text{m}$  (TIR) und zusätzlich auch der kurzwelligere Infrarotbereich von ca.  $3.9\ \mu\text{m}$ , sowie deren Differenz verwendet [1], [22], [23]. So erfolgt die Wolkenerkennung für die beiden Wolkenprodukte MVCM [14] und VCM [15] für die Daten des VIIRS/Day-Night-Band (kurz: DNB) bei Nacht anhand mehrerer Schwellwerttests in diesen beiden Wellenlängenbereichen. Die jeweils konstanten Schwellwerte zur pixelweisen Unterscheidung von Bewölkung und klarem Himmel werden hierbei unter Berücksichtigung der Landbedeckung, Beleuchtung und weiteren Faktoren modellbasiert bestimmt.

Der ECM-Algorithmus [19], der ebenfalls zur Wolkenerkennung für DNB entwickelt wurde, weicht von diesen ausschließlich modellbasierten Ansätzen ab und implementiert ein Lernen der Schwellwerte anhand von zuverlässigeren Wolken­daten (vgl. Kapitel 2.1.2). Eine weitere Besonderheit von ECM ist, dass die tagsüber standardmäßig durchgeführten Tests im sichtbaren Spektralbereich (dort haben Wolken eine sehr starke Reflektivität) bei ausreichender Mondbeleuchtung auch nachts integriert werden, also nicht mehr ausschließlich Beobachtungen im TIR-Bereich berücksichtigt werden.

Vereinzelt gibt es auch experimentelle Ansätze, die zum Großteil auf VNIR basieren, diese beziehen sich jedoch nur auf die Detektion spezieller Wolkentypen, Dunst oder Aerosolen. So schlagen [25] ein Verfahren zur Detektion von Aerosolen anhand der Betrachtung des Kontrasts zwischen künstlich beleuchteten und unbeleuchteten Pixeln vor, der durch den Einfluss von Aerosolen vermindert wird. [21] und [24] identifizieren Nebel und niedrig gelegene Stratus-Wolken unter anderem anhand ihrer Reflektivität im VNIR, wodurch das Verfahren nur bei ausreichend starker Beleuchtung durch den Mond angewandt werden kann.

Die operationell und global eingesetzten Wolkenerkennungs­algorithmen basieren also vor allem auf der Verwendung von Beobachtungen im Bereich des TIR. Ein ausschließlich auf nächtlichen VNIR-Beobachtungen basierendes Verfahren existiert bisher nicht.

Einen TIR-Sensor zu verwenden, bedeutet jedoch immer, dass der Satellit ein zusätzliches Instrument tragen muss, was sich auf die Konstruktion (u.a. Gewicht, Größe, Energie), die Infrastruktur (u.a. Datenvolumen) und damit auch finanziell auf die Satellitenmission auswirkt. Bei den Satellitenmissionen, für deren Daten die zuvor vorgestellten Wolkenerkennungsverfahren entwickelt wurden, ist ein solcher Sensor zur Beantwortung anderer wissenschaftlicher Fragen jedoch ohnehin Teil des Instrumentariums. Da es sich bei N8 um einen kleinen, ausschließlich auf VNIR spezialisierten Satelliten handeln soll, stellt sich nun die Frage, ob auf einen zusätzlichen TIR-Sensor zur Wolkenerkennung auch verzichtet werden kann. Da eine präzise und robuste Wolkenerkennungsmethode für die Verwendung der generierten Satellitenbilder zu Analyse­zwecken unumgänglich ist, ist eine Voraussetzung hierfür, dass die Wolkenerkennung auch ohne einen TIR-Sensor funktioniert. In dieser Arbeit wird deshalb untersucht, inwiefern auch allein aus den Nachtaufnahmen des VNIR-Bereichs Wolkenmasken generiert werden können.

## 1. Einleitung

**Fragestellung** In dieser Arbeit wird untersucht, wie die semantische Segmentierung von Wolken unter der ausschließlichen Verwendung von nächtlichen VNIR-Beobachtungen realisiert werden kann und welche Qualität für eine so erzeugte Wolkenmaske erreichbar ist. Hierbei liegt der Fokus auf der Wolkenerkennung in urbanen, künstlich beleuchteten Gebieten, da auf diesen das Hauptaugenmerk der für N8 vorstellbaren Anwendungen liegt. Zur Entwicklung des Wolkenerkennungsverfahrens werden die DNB-Daten verwendet, da diese derzeit die einzigen als Zeitreihen global verfügbaren und kalibrierten nächtlichen VNIR-Daten sind (vgl. Kapitel 2.1).

Zur Beantwortung dieser Fragestellung werden im Folgenden in Kapitel 2 zunächst die verwendeten Daten genauer betrachtet, sowie ein allgemeines Beobachtungsmodell für nächtliche VNIR-Aufnahmen eingeführt. Anschließend wird in Kapitel 3 die Methodik zur Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken vorgestellt. In Kapitel 4 erfolgt die Analyse der Qualität der Wolkenmasken, die mit dem in dieser Arbeit entwickelten Verfahren erzeugt wurden. Abschließend gibt Kapitel 5 einen Überblick darüber, welche weitergehenden Untersuchungen auf dem Weg zu einem nächtlichen, VNIR-basierten Wolkenerkennungsverfahren durchgeführt werden können.

## 2. Grundlagen

Bevor in Kapitel 3 die Methodik zur Entwicklung des Wolkenerkennungsverfahrens erläutert wird, werden im Folgenden die Grundlagen der dort durchgeführten Betrachtungen beschrieben. Hierzu gehören neben den verwendeten Daten das Beobachtungsmodell für nächtliche Satellitenaufnahmen sowie das für den Wolkenerkennungsalgorithmus verwendete Klassifikationsverfahren.

### 2.1. Datengrundlage

Zur Entwicklung des Wolkenerkennungsverfahrens für nächtliche VNIR-Satellitenaufnahmen werden zum einen entsprechende Bilddaten benötigt und zum anderen eine Referenz darüber, wo sich in diesen Daten Wolken befinden. Hierfür werden die DNB-Daten in Kombination mit einer der für diese Daten verfügbaren Wolkenmasken (ECM) verwendet. Beide Datensätze werden im Folgenden genauer beschrieben.

#### 2.1.1. Day-Night-Band (DNB)

Das Day-Night-Band [28], kurz DNB, liefert vollständig kalibrierte nächtlichen Aufnahmen der Erde im sichtbaren Spektralbereich und im nahen Infrarot. Es ist einer der 22 Kanäle des Radiometers VIIRS (Visible Infrared Imaging Radiometer Suite), das sich an Bord der amerikanischen Erdbeobachtungssatelliten S-NPP und NOAA20 befindet [10], [9]. Der Satellit S-NPP, dessen Daten in dieser Arbeit verwendet werden, befindet sich auf einem polaren Orbit in einer Höhe von circa 830 km. VIIRS scannt die Erdoberfläche nach dem Whiskbroom-Prinzip mit einer Streifenbreite von 3060 km und erreicht damit eine zweimalige Abdeckung der Erde pro Tag (einmal mittags und einmal nachts), wodurch die Daten besonders gut für Zeitreihenanalysen geeignet sind.

Für DNB wird durch ein spezielles Verarbeitungsverfahren eine über die gesamte Streifenbreite konstante Bodenpixelgröße von circa 750 m erzielt [10]. Der panchromatische DNB-Kanal erfasst den spektralen Bereich von 500 bis 900 nm, er ist also gegenüber dem sichtbaren Spektralbereich leicht ins nahe Infrarot verschoben, wodurch blaues Licht nicht erfasst wird. Die Besonderheit von DNB ist der große Dynamikbereich von fast sieben Größenordnungen – mittels verschiedener Verstärkungsstufen können Strahldichten von  $3 \times 10^{-9}$  bis  $0.02 \text{ Wcm}^{-2}\text{sr}^{-1}$  erfasst werden. Dies ermöglicht es, dass sowohl bei Tag als auch bei Nacht Aufnahmen gemacht werden können, ohne in den Sättigungsbereich zu geraten und die trotz sehr schwacher Beleuchtung bei Nacht viele Details enthalten. In der Praxis ist die Sensitivität noch eine Größenordnung besser, als spezifiziert [31], wodurch selbst die Einflüsse sehr schwacher Lichtquellen wie das Licht einer einzelnen Straßenlampe [38] oder das Nachthimmellicht [37] detektiert werden. Abbildung 2.1 gibt eine Übersicht darüber, welche nächtlichen Lichtquellen in Abhängigkeit von der Mondbeleuchtung mit DNB erfasst werden können.

## 2. Grundlagen



Abbildung 2.1: Mit DNB bei Nacht beobachtbare Lichtquellen bei Vollmond (a) und bei Neumond (b). Grafik entnommen aus [38]. Urheber: Steve Deyo, UCAR COMET Program.

Aus den DNB-Aufnahmen werden verschiedene Produkte abgeleitet, in dieser Arbeit wurde das kalibrierte und georeferenzierte SDR (Sensor Data Record, [10]) Produkt verwendet<sup>1</sup>. Es enthält die DNB-Messungen in Form von kalibrierten Strahldichten in der Einheit  $\text{Wcm}^{-2}\text{sr}^{-1}$  (siehe hierzu S. 12). Durch die Kalibrierung können negative Strahldichten auftreten, die in dieser Arbeit auf Null gesetzt wurden, da die beobachteten Strahldichten nicht negativ sein können und diese Werte somit ungültig sind.

Zur Georeferenzierung liegen zwei Versionen von pixelweisen Lageinformationen vor, eine Version ohne und eine mit Berücksichtigung des Geländemodells [2]. Letztere wird in dieser Arbeit verwendet. Sie erreicht im Nadir eine Genauigkeit von circa einem halben Pixel und zum Streifenrand hin wächst die Unsicherheit auf über einen Kilometer an. Durch Bewölkung kann die Georeferenzierung noch ungenauer werden, siehe hierzu auch Kapitel 2.2.4 [38], [52].

**Visualisierung von DNB-Daten** Aufgrund des großen Dynamikbereiches der DNB-Bilder ist die Visualisierung der Daten ohne eine entsprechende Skalierung oft nur wenig aussagekräftig. Dunklere Details, wie auch die Reflexion des Mondlichtes durch Wolken, sind ohne deren Hervorhebung durch eine Skalierung nicht erkennbar. Zur Darstellung der DNB-Bilder wurde deshalb der von [46] vorgeschlagene Skalierungsalgorithmus in vereinfachter Form angewandt. Der Algorithmus skaliert die beobachteten DNB-Strahldichten  $L$  anhand der erwartbaren minimalen und maximalen Strahldichten  $L_{min}$  und  $L_{max}$  entsprechend Gleichung 2.1.

$$L_{skal.} = 0 \leq 255 \cdot \sqrt{\frac{L - L_{min}}{L_{max} - L_{min}}} \leq 255 \quad (2.1)$$

Die erwartbaren Strahldichten werden in [46] dynamisch anhand der Beleuchtung durch die Sonne bzw. den Mond berechnet. Da in diesem Fall ausschließlich nächtliche Aufnahmen skaliert werden sollen, wurden  $L_{min}$  und  $L_{max}$  basierend auf [46] so gewählt, dass  $L_{min} = 10^{-10} \text{Wcm}^{-2}\text{sr}^{-1}$  (entspricht ungefähr der minimal detektierbaren Strahldichte) und  $L_{max} = L_{MT2009}$  (wenn Mondbeleuchtung  $L_{MT2009} > 10^{-8} \text{Wcm}^{-2}\text{sr}^{-1}$ ) bzw.  $L_{max} = 10^{-8} \text{Wcm}^{-2}\text{sr}^{-1}$  (wenn Mondbeleuchtung  $L_{MT2009} < 10^{-8} \text{Wcm}^{-2}\text{sr}^{-1}$ ). Zur Modellierung der Mondbeleuchtung  $L_{MT2009}$  siehe Kapitel 2.2.2.

Durch diese Skalierung werden künstliche Beleuchtungsquellen wie städtische Lichter, die deutlich heller sind als die Mondbeleuchtung, gesättigt, und Bewölkung wird hervorgehoben. Bei Betrachtung der so skalierten Bilder ist jedoch zu beachten, dass diese je nach Mondbeleuchtung unterschiedlich skaliert sind.

---

<sup>1</sup>Download möglich via CLASS: [https://www.avl.class.noaa.gov/saa/products/search?sub\\_id=0&datatype\\_family=VIIRS\\_SDR](https://www.avl.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=VIIRS_SDR)

## 2. Grundlagen

### 2.1.2. Wolkenmaske (ECM)

Die ECM (NOAA Enterprise Cloud Mask) [19] ist eine für verschiedene Sensoren (neben VIIRS unter anderem auch MODIS, AVHRR etc.) entwickelte Wolkenmaske, die anhand eines Naïve Bayes Ansatzes auf Basis von Beobachtungen in verschiedensten Spektralbereichen erzeugt wird. Die Ausgabe des Algorithmus sind pixelweise Wolkenwahrscheinlichkeiten, aus denen unter anderem eine binäre Wolkenmaske (*klar/wolkig*) abgeleitet werden kann. Die ECM-Wolkenmaske wurde primär als Teil einer Prozessierungskette zur Ableitung weiterer Produkte aus den Sensordaten entwickelt, für die bewölkte Pixel eliminiert werden müssen. Die Maske selbst wird dennoch auch als eigenständiges Produkt angeboten<sup>2</sup>.

Das Prinzip des ECM-Algorithmus basiert auf der Bestimmung von Annahmen über die spektralen und räumlichen Eigenschaften der Beobachtungen bei klarem Himmel anhand von Zusatzdaten und das anschließende Testen jedes Pixels auf Abweichungen von diesen Eigenschaften. Als Zusatzdaten kommen hierfür unter anderem Informationen aus NWP (numerical weather prediction), ein Strahlungstransfermodell und Landbedeckungskarten zum Einsatz, der Algorithmus basiert also nicht nur auf den VIIRS-Daten selbst.

Die Eigenschaften der Pixel werden anhand von verschiedenen Tests bzw. Merkmalen beschrieben, von denen manche jedoch nur tagsüber anwendbar sind. Die Menge der für nächtliche Aufnahmen durchgeführten Tests umfasst vor allem spektrale Tests im Bereich des thermalen Infrarot. Die Besonderheit von ECM gegenüber den weiteren für VIIRS verfügbaren Wolkenerkennungsalgorithmen ist jedoch, dass er als einziger die nächtlichen VNIR-Beobachtungen (DNB) berücksichtigt, was sonst immer nur bei Anwendung bei Tag der Fall ist. Diese bei vorhandener Beleuchtung durch den Mond auch nachts angewandten Tests beschreiben die Differenz der beobachteten und für wolkenfreie Bedingungen modellierten Reflektanzen, die Differenz der beobachteten Reflektanz zum minimalen Wert in der Nachbarschaft sowie die Standardabweichung der Reflektanz [19].

Jedem der ECM-Tests wird in Abhängigkeit von den Zusatzdaten eine Wahrscheinlichkeitsverteilung bzw. ein Schwellwert zugeordnet, die anhand eines Trainings mit Wolkendaten von CALIPSO (Lidar Erdbeobachtungssatellit mit Fokus auf Wolken, [50]) bestimmt wurden. Neben der Wolkenwahrscheinlichkeit, die anhand der Kombination aller Tests durch den Naïve Bayes Ansatz ermittelt wird, sind auch die anhand jedes einzelnen Tests bestimmten Wolkenwahrscheinlichkeiten Teil des ECM-Produktes.

Da die ECM-Wolkenmaske durch die Verwendung der DNB-Daten (zumindest bei vorliegender Mondbeleuchtung) dem in dieser Arbeit angestrebten Verfahren am ehesten ähnelt und sich somit am besten zum Vergleich eignet, wurde sie aus den drei für DNB verfügbaren Wolkenmasken zur Verwendung als Ground Truth ausgewählt.

---

<sup>2</sup>ECM für VIIRS ist downloadbar via CLASS: [https://www.av1.class.noaa.gov/saa/products/search?sub\\_id=0&datatype\\_family=JPSS\\_GRAN](https://www.av1.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=JPSS_GRAN)

**Qualität** Mit dem ECM-Algorithmus wird eine konservative Schätzung der Bewölkung erzielt, das heißt die Precision der Klasse *wolkig* wird optimiert. Laut Spezifikation wird bei Nacht global eine Gesamtgenauigkeit von 85% über Ozean und 88% über Land für Wolken von einer optischen Dicke (vgl. Kapitel 2.2.4) von mindestens 1 erzielt. Tagsüber werden im Vergleich dazu 94% und 90% spezifiziert, bei Nacht liefert der Algorithmus also insgesamt schlechtere Ergebnisse als bei Tag [19].

Bei der Betrachtung der in dieser Arbeit verwendeten ECM-Daten fiel jedoch auf, dass manche von ihnen rechteckige Artefakte enthalten, d.h. mehrere rechteckige Bereiche haben eine deutlich andere Wolkenwahrscheinlichkeit als die Umgebung, siehe Abbildung 2.2. Dies kann in keinem Fall realen Gegebenheiten entsprechen. Diese Fehler sind nicht dokumentiert und sind anhand der Qualitätsinformationen der Daten nicht erkennbar. Eine Analyse der Fehler im Rahmen dieser Arbeit legt nahe, dass die Artefakte durch die zur Charakterisierung der Landbedeckungsarten verwendeten Zusatzdaten verursacht worden zu sein scheinen. Diese scheinen in diesen Fällen in einer unzureichenden Auflösung vorzuliegen oder fehlerhaft zu sein. Da die Ableitung der Schwellwerte, auf denen die Klassifikation mit dem ECM-Algorithmus basiert, von der Landbedeckung abhängig ist, übertragen sich diese Fehler in die Wolkenmasken.

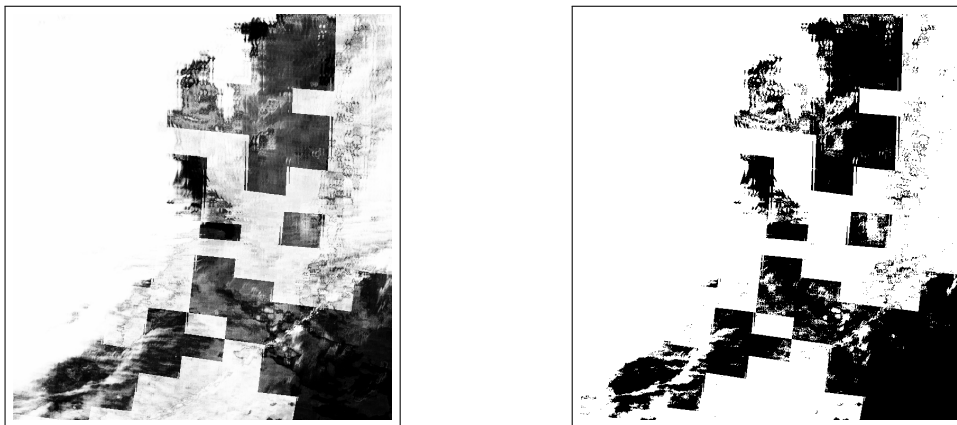


Abbildung 2.2: Beispiel für Artefakte in der ECM-Wolkenmaske für einen ca. 540 km breiten Bildausschnitt. Links: Wolkenwahrscheinlichkeit, Rechts: Binäre Wolkenmaske (weiß=*wolkig*, schwarz=*klar*).

## 2. Grundlagen

### 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

Als Grundlage für die Interpretation von VNIR-Nachtaufnahmen wird im Folgenden ein allgemeines Beobachtungsmodell eingeführt. Dieses erklärt das Erscheinungsbild von Wolken und verschiedenen Landbedeckungen in den Aufnahmen und dessen Beeinflussung durch verschiedene Faktoren. Auf Basis dieser Betrachtungen wird in dieser Arbeit das Wolkenerkennungsverfahren entwickelt.

Im Folgenden wird zunächst das gesamte Beobachtungsmodell vorgestellt, anschließend wird auf die dort identifizierten Haupteinflussfaktoren näher eingegangen.

#### 2.2.1. Allgemeines Beobachtungsmodell

Die Beobachtung des Sensors ist die sogenannte Top-Of-Atmosphere (TOA) Strahldichte, also die Strahldichte oberhalb der Atmosphäre. Diese ist die Summe aller Strahlung, die von verschiedenen Lichtquellen aus über verschiedene Wege den Sensor erreicht. Ein allgemeines, vereinfachtes Modell für diese verschiedenen Bestandteile des vom Sensor empfangenen Lichts (radiative transfer model) wird in Abbildung 2.3 dargestellt.

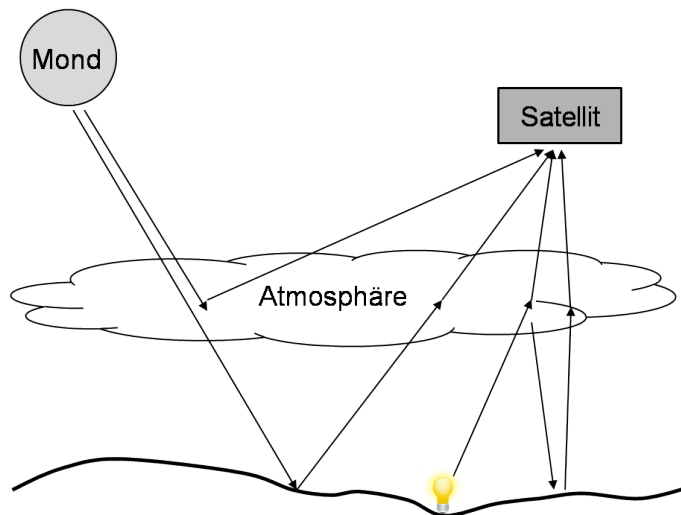


Abbildung 2.3: Vereinfachtes Beobachtungsmodell für Nachtaufnahmen im VNIR-Bereich. Die Summe der einzelnen Bestandteile am Ort des Sensors ist die beobachtete TOA-Strahldichte.

Die TOA-Strahldichte wird also vor allem durch drei Faktoren beeinflusst:

1. Beleuchtung durch den Mond:

Der Mond ist die einzige, in Abhängigkeit der Mondphase, global vorhandene nächtliche Beleuchtungsquelle und hat einen großen Einfluss auf das Erscheinungsbild von Erdoberfläche und Wolken, siehe Abschnitt 2.2.2.

## 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

2. Erdoberfläche inkl. ihrer Lichtemissionen:  
Neben den Reflexionseigenschaften der Erdoberfläche selbst, wirken sich vor allem von ihr ausgehende Lichtemissionen auf die Beobachtungen aus, die neben dem Mond weitere Lichtquellen darstellen, siehe Abschnitt 2.2.3.
3. Streuung und Reflexion in der Atmosphäre (insbesondere durch Bewölkung):  
Wolken beeinflussen den Strahlungstransfer in der Atmosphäre. Je nach Art der Bewölkung sind die in Abbildung 2.3 dargestellten Streuungs- und Reflexionseffekte und die Transmission und Absorption in der Atmosphäre unterschiedlich stark ausgeprägt, siehe Abschnitt 2.2.4.

Neben diesen Haupteinflussfaktoren gibt es weitere Einflüsse, die in diesem Modell vernachlässigt werden. Bei Nacht kann zu gewissen Zeitpunkten zum Beispiel vor allem Streulicht von der Sonne einen großen Einfluss auf die beobachteten Strahldichten haben.

### 2.2.2. Modell für die Beleuchtung durch den Mond

Das Erscheinungsbild der Erdoberfläche und der Atmosphäre hängt bei VNIR-Nachtaufnahmen von der vorhandenen Beleuchtung ab, weshalb zu ihrer Interpretation die Kenntnis der zum Aufnahmezeitpunkt vorliegenden Beleuchtungssituation notwendig ist. Bei Nacht ist die hauptsächliche Beleuchtungsquelle der Mond, weitere Quellen sind z.B. künstliches Licht, Feuer oder auch Nordlichter (vgl. Kapitel 2.2.3), diese sind jedoch räumlich beschränkt und zeitlich veränderlich, wodurch keine zur ihrer Modellierung geeigneten Daten vorliegen. Deshalb wird in dieser Arbeit für die Ableitung der Information über die vorliegende Beleuchtungssituation nur die als einzige sicher bekannte Beleuchtung durch den Mond herangezogen. Diese wird anhand eines speziell für die Auswertung von DNB-Daten entwickelten Modells [36] beschrieben, welches im Folgenden vorgestellt wird.

Die Modellierung der Beleuchtung durch den Mond ist nicht nur aufgrund des im speziellen Fall der Wolkenerkennung relevanten Einflusses auf das Erscheinungsbild von Wolken notwendig, sondern auch ganz allgemein, um die mit DNB beobachteten Strahldichten vergleichbar zu machen. Nur wenn die externe Beleuchtung durch den Mond bekannt ist, kann eine Aussage darüber getroffen werden, ob eine Fläche Licht emittiert und ob sie im Vergleich zu einer anderen hell oder dunkel ist bzw. stärker oder schwächer reflektiert. Letztere Eigenschaft wird durch die radiometrische Größe der Reflektanz beschrieben. Diese und weitere radiometrische Größen, die bei der Vorstellung des Beleuchtungsmodells relevant sind, werden im Folgenden kurz vorgestellt.

## 2. Grundlagen

**Grundlegende radiometrische Größen** Die soeben angesprochene Größe der Reflektanz  $R$  beschreibt das Verhältnis zwischen der von einer Fläche reflektierten ( $L_{refl}$ ) und der auf sie einfallenden Strahldichte  $L_{einf}$ .

$$R = \frac{L_{refl}}{L_{einf}} \quad (2.2)$$

Die Strahldichte gibt die Energie von Strahlung pro Flächeneinheit und Raumwinkel an (Einheit  $\text{Wm}^{-2}\text{sr}^{-1}$ ). Integriert man die Strahldichte über einen bestimmten Raumwinkel, erhält man die Bestrahlungsstärke  $E$ , die die Energie der Strahlung pro Flächeneinheit angibt. Im Falle eines Lambertischen Strahlers und der Betrachtung der gesamten Hemisphäre vereinfacht sich dieser Zusammenhang zu

$$E = \pi L. \quad (2.3)$$

Der Faktor  $\pi$  ist hierbei der projizierte Raumwinkel des Halbraums (Hemisphäre) [34]. Im Falle der ausschließlichen Berücksichtigung der Mondbeleuchtung und der Vernachlässigung der Atmosphäre ist die reflektierte Strahldichte einer Fläche die von DNB gemessene Größe und die von der Fläche empfangene Strahldichte ist die eintreffende Beleuchtung durch den Mond. Um die auf die Mondbeleuchtung bezogenen Reflektanzen  $R_m$  der beobachteten Flächen zu bestimmen, wird also ein Modell der einfallenden Strahldichte des Mondes  $L_{einf}$  benötigt, welches aus einem Modell für die Bestrahlungsstärke des Mondes berechnet werden kann. Ein solches Modell ist das von Miller und Turner [36].

**Modell für die lunare Bestrahlungsstärke von Miller und Turner** Als Modell für die Beleuchtung durch den Mond wird das von Miller und Turner [36] entwickelte Modell für die TOA spektrale lunare Bestrahlungsstärke (*top-of-atmosphere spectral lunar irradiance*) verwendet. Dieses Modell wurde eigens für die Anwendung auf DNB-Daten entwickelt. Ziel der Autoren war es, zu ermöglichen, dass tagsüber gängige, auf der Reflektanz im sichtbaren Bereich beruhende Verfahren auch nachts angewandt werden können.

Die Modellierung der Mondbeleuchtung ist komplexer als die Modellierung der Beleuchtung durch die Sonne, da deutlich mehr Faktoren einen Einfluss auf sie haben. So ist die Bestrahlungsstärke des Mondes abhängig vom veränderlichen Abstand des Mondes zu Erde und Sonne (elliptische Orbits), der Mondphase, Variationen in der Beleuchtung durch die Sonne selbst und die aufgrund der inhomogenen Mondoberfläche veränderliche Reflektanz des Mondes. Das Modell von Miller und Turner berücksichtigt die meisten dieser Einflüsse, jedoch manche nur in vereinfachter Form. Sehr kleine Einflüsse, wie zum Beispiel der Effekt der Libration, werden vollständig vernachlässigt. Außerdem wird der Mond als Lambertischer Strahler betrachtet. Zu den Details sei hierzu auf [36] verwiesen. Insgesamt haben die aus dem Modell abgeleiteten Bestrahlungsstärken laut Analysen der Autoren eine Unsicherheit von 7-12 %.

Die Autoren stellen das Modell in Form einer Look-Up Tabelle für die spektrale Bestrahlungsstärke des Mondes mit 1 nm Auflösung im Wellenlängenbereich von 0.3 bis 12  $\mu\text{m}$  für alle Mond-Phasenwinkel in 1°-Schritten und bezogen auf eine Standardgeometrie zwischen

## 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

Sonne, Erde und Mond bereit<sup>3</sup>. Für die spezielle Anwendung auf DNB-Daten wurde ein weiterer Datensatz erzeugt, bei dem die spektralen Bestrahlungsstärken bereits mit der Spectral Response Function von DNB gefaltet wurden. Dieser Datensatz wird in dieser Arbeit verwendet.

Zur Anwendung des Modells stellen die Autoren neben der Look-Up Tabelle einen zusätzlichen Datensatz mit stündlichen Werten für den Phasenwinkel, den Abstand Sonne-Erde und den Abstand Mond-Erde für die Jahre 2010-2030 sowie Informationen, wie aus diesen Daten die spektrale Bestrahlungsstärke des Mondes für einen gegebenen Zeitpunkt berechnet werden kann, zur Verfügung. Die Berechnung des Modells wurde auf Basis dieser bereitgestellten Informationen realisiert. Es sind folgende Schritte nötig:

1. Lineare Interpolation des Phasenwinkels des Mondes und der Distanzen Sonne-Erde und Mond-Erde für den gesuchten Zeitpunkt (in Schritt 3 zum Übergang von Standardgeometrie auf tatsächliche Situation benötigt).
2. Interpolation der zum Phasenwinkel gehörigen Bestrahlungsstärke für die Standardgeometrie aus der DNB-spezifischen look-up Tabelle.
3. Skalierung der Standard-Bestrahlungsstärke mit aktueller Sonne-Erde-Mond Geometrie und dem Phasenwinkel.

Aus der so berechneten lunaren Bestrahlungsstärke des Mondes  $E_m$  kann über folgende Beziehung die lunare Strahldichte  $L_{MT2009}$  abgeleitet werden (vgl. auch [53]).

$$L_{MT2009} = E_m \cdot \frac{\cos(\theta_m)}{\pi} \quad \text{mit} \quad \theta_m = \text{Zenitwinkel des Mondes} \quad (2.4)$$

Der im Vergleich zur Gleichung (2.3) zusätzliche Faktor  $\cos(\theta_m)$  bezieht  $E_m$  auf die zur Beleuchtungsrichtung orthogonale Projektion der beleuchteten Fläche. Dies ist die Anwendung des Lambertischen Kosinusetzes für die Beleuchtung geneigter Flächen [34]. Die Neigung der Fläche wird in dieser Arbeit durch den mittleren Zenitwinkel des Mondes  $\theta_m$  im Bildausschnitt beschrieben, welcher aus den Metadaten des DNB-Bildes berechnet werden kann.

Bei einem Zenitwinkel des Mondes von  $\theta_m > 90^\circ$  wird die Strahldichte  $L_{MT2009}$  gleich Null gesetzt, da sich der Mond dann unterhalb des Horizontes befindet und die betrachtete Fläche nicht beleuchtet. Dies ist während circa der Hälfte eines Mondzyklus der Fall. Abbildung 2.4 zeigt einen typischen Verlauf der anhand dieses Modells berechneten lunaren Strahldichten.

---

<sup>3</sup><https://ieeexplore.ieee.org/document/4895690/media#media>, siehe [36]

## 2. Grundlagen

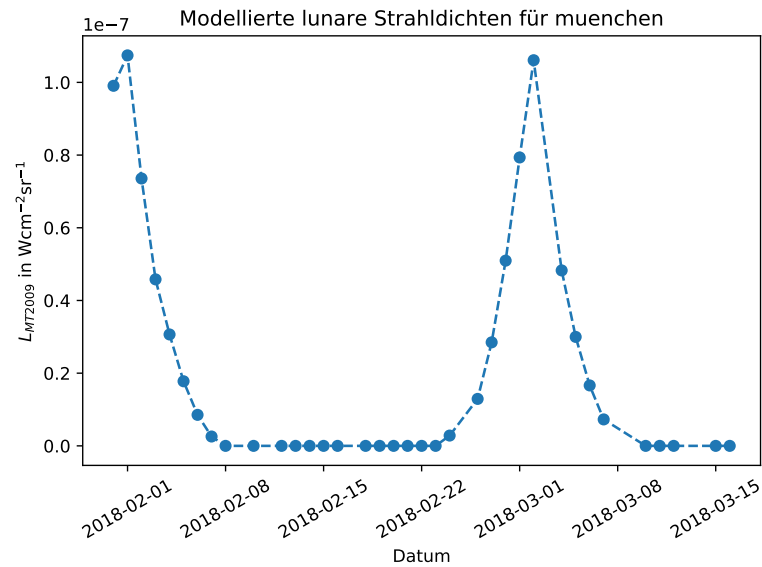


Abbildung 2.4: Beispiel für modellierte lunare Strahldichten  $L_{MT2009}$  innerhalb von einhalb Mondzyklen für einen Bildausschnitt von München.

Unter der Annahme, dass der Mond die einzige Beleuchtungsquelle ist und unter Vernachlässigung der Atmosphäre, lässt sich die Reflektanz  $R_m$  einer mit DNB beobachteten Fläche schließlich aus dem Verhältnis zwischen beobachteter ( $L_{DNB}$ ) und vom Mond aus einfallender Strahldichte  $L_{MT2009}$  berechnen:

$$R_m = \frac{L_{DNB}}{L_{MT2009}} \quad (2.5)$$

Ist der Mond die einzige Lichtquelle in der Szene, dann liegt  $R_m$  im Wertebereich  $[0, 1]$ . Höhere Werte treten auf, wenn die Annahme, dass ausschließlich der Mond als Beleuchtungsquelle vorhanden ist, nicht zutrifft [45].

### 2.2.3. Lichtemissionen der Erdoberfläche und weitere Beleuchtungsquellen

Neben der Beleuchtung durch den Mond werden die von einem VNIR-Sensor bei Nacht beobachteten TOA-Strahldichten durch weitere Lichtquellen beeinflusst, die in Abbildung 2.1 auf Seite 6 zusammengefasst werden. Es kann hierbei zwischen natürlichen und menschengemachten Lichtquellen unterschieden werden.

Die in nächtlichen VNIR-Aufnahmen auffälligsten Lichtemissionen stammen von künstlicher Beleuchtung von urbanen Strukturen und Verkehrswegen. Ein Beispielbild hierfür befindet sich in Abbildung 2.5 links. Weitere beobachtbare menschengemachte Lichtquellen sind zum Beispiel beleuchtete Fischerboote oder Gasfackeln, letztere verursachen die hellsten Lichtemissionen der Erdoberfläche [12].

## 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

Zu den Lichtquellen natürlichen Ursprungs gehören unter anderem Feuer, Blitze, Polarlichter und das Nachthimmelleuchten (engl. nightglow) [38]. Letzteres kann in DNB-Daten bei Neumond trotz der geringen Intensität nicht nur direkt beobachtet werden, sondern auch der durch Wolken reflektierte Anteil ist detektierbar. Dadurch können Wolken unter diesen Bedingungen auch ohne Mondbeleuchtung sichtbar werden [37].

Blitze erscheinen in DNB-Bildern aufgrund des scannenden Aufnahmeprinzips als helle Streifen mit der Breite einer Scanlinie [12], siehe Abbildung 2.5. Ein Beispiel für das Erscheinungsbild von Polarlichtern in DNB-Daten findet sich ebenfalls in Abbildung 2.5.

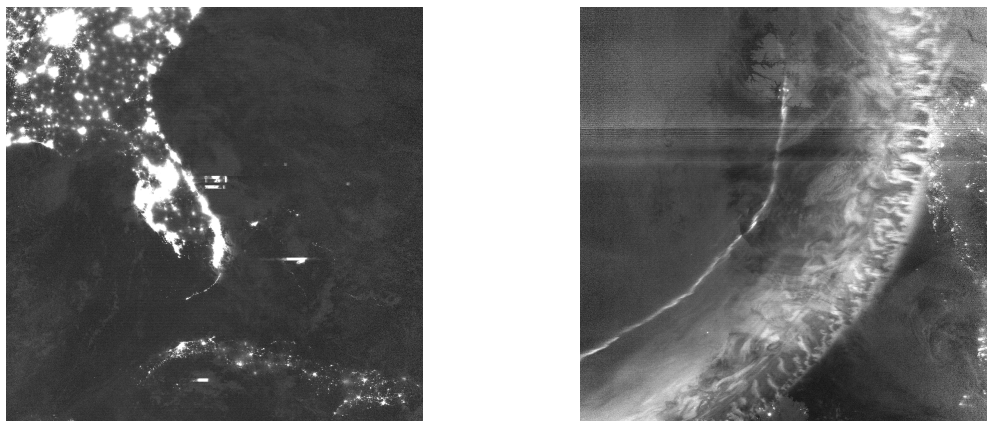


Abbildung 2.5: Links: Mehrere Blitze vor der Ostküste Floridas, die anhand der Streifenform mit Ausdehnung in Richtung der Scanlinie deutlich erkennbar sind. Deutlich sichtbar ist außerdem die künstliche Beleuchtung der urbanen Strukturen. Rechts: Polarlichter nordwestlich von Skandinavien. Am rechten Bildrand befindet sich Norwegen.

### 2.2.4. Einfluss von Wolken auf VNIR-Satellitenbeobachtungen bei Nacht

Der Einfluss von Wolken auf nächtliche VNIR-Satellitenaufnahmen kann durch einige besonders charakteristische optische Eigenschaften von Wolken beschrieben werden, die im Folgenden kurz erläutert werden. Wie sich diese theoretischen Eigenschaften tatsächlich in den Nachtaufnahmen äußern, wird jeweils am Beispiel von DNB-Bildern veranschaulicht.

**Reflektanz/Albedo** Wolken zeichnen sich gegenüber der Erdoberfläche im Allgemeinen durch eine höhere Reflektanz aus, sie reflektieren also deutlich mehr Mondlicht, als die Erdoberfläche [1]. Ganz allgemein kann der prozentuale Anteil des von einer Fläche reflektierten Lichts am auf sie einfallenden Licht durch die Größe der Albedo ausgedrückt werden. Hierbei werden zwar etwaige Richtungsabhängigkeiten vernachlässigt, die Albedo-Werte verschiedener Oberflächen für den sichtbaren Spektralbereich dienen aber bei vorhandener Mondbeleuchtung trotzdem als gute Richtwerte für ihr Erscheinungsbild in Nachtaufnahmen.

## 2. Grundlagen

Die Albedo von verschiedenen Wolkentypen liegt typischerweise im Bereich von ca. 0.5 (Cirrostratus) bis zu 0.9 (Cumuliform), lediglich für sehr dünne Cirrus-Wolken, Dunst etc. ist sie geringer. Die gängigen Landbedeckungsarten der Erdoberfläche haben im Vergleich zu Wolken eine geringe Albedo, zum Beispiel liegen Grasflächen bei ca. 0.2, Asphalt und Wasserflächen nur bei ca. 0.1. Lediglich Schnee (bis zu 0.95) und Sand (bis zu 0.45) haben eine annähernd so hohe Albedo wie Wolken und sind deshalb auch in optischen Nachtaufnahmen bei starker Mondbeleuchtung deutlich sichtbar [16]. Abbildung 2.6 zeigt einige Beispiele für typische DNB-Nachtaufnahmen bei starker Beleuchtung durch den Mond, in denen die unterschiedlich starken Reflexionsvermögen von Wolken, Gras, Asphalt, Wasserflächen und Schnee gut erkennbar sind.

**Optische Dicke/COT (Cloud Optical Thickness)** Die optische Dicke einer Wolke beschreibt die gesamte Abschwächung (durch Absorption und Streuung) von Licht auf dessen Weg durch die Wolke [20]. Je nach Wolkentyp haben Wolken eine unterschiedliche optische Dicke, auch abhängig von ihrer geometrischen Dicke. Sie haben jedoch immer eine eher hohe optische Dicke, sind also eher lichtundurchlässig [41]. Der von der Erdoberfläche ausgehende Teil der beobachteten Strahldichte verringert sich also auf dem Weg zum Sensor durch eine Wolke. Dadurch erscheint unter Vernachlässigung einer eventuell vorhandenen Beleuchtung durch den Mond ein DNB-Bild von der Erdoberfläche mit Wolken dunkler als ohne oder die Wolke ist gar vollkommen lichtundurchlässig und die Erdoberfläche ist nicht sichtbar. Ist eine Beleuchtung durch den Mond vorhanden, kommt zur Beobachtung das von den Wolken reflektierte Mondlicht hinzu, wodurch ein bewölktetes Pixel wiederum heller erscheinen kann als im unbewölkten Fall.

## 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

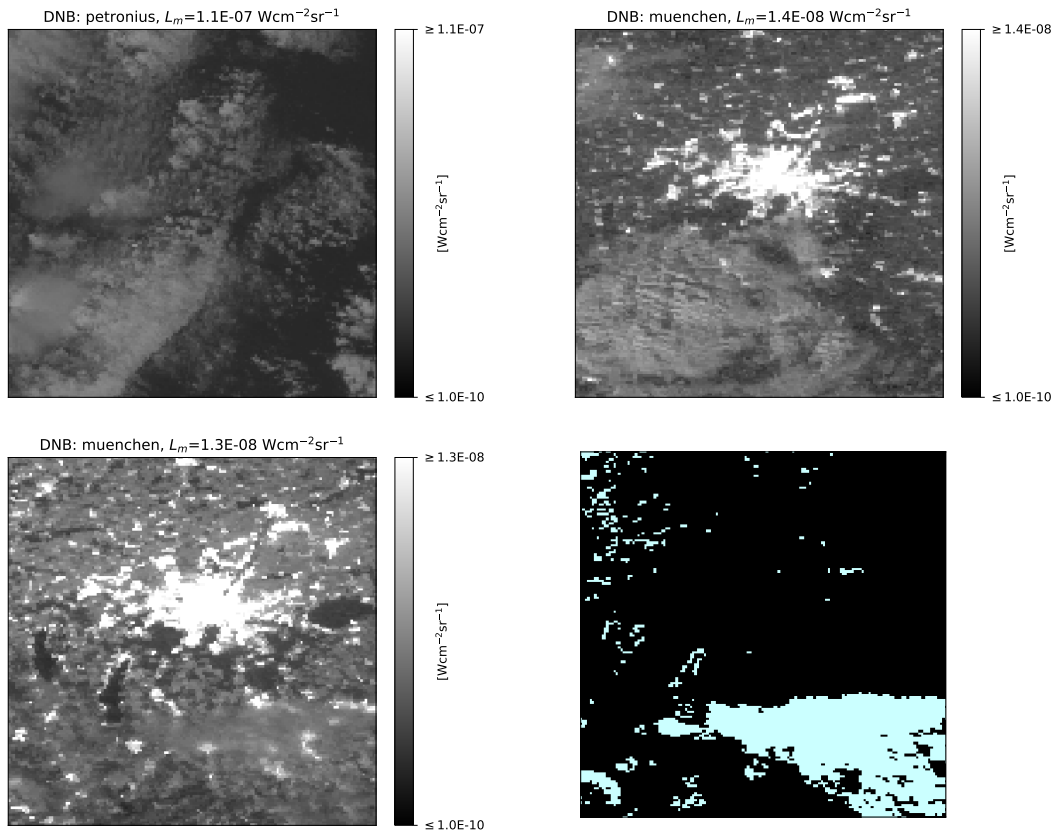


Abbildung 2.6: Beispiele für das Erscheinungsbild von Flächen mit verschiedener Albedo in DNB-Bildern bei Beleuchtung durch den Mond ( $L_{MT2009}$  wird in den Abbildungstiteln mit  $L_m$  abgekürzt).

Links oben: Wolken über Ozean.

Rechts oben: Wolken über bebautem Gebiet und Ackerland in der Gegend um München.

Unten: Wolken über dem schneebedeckten Umland von München. Sehr gut heben sich die beiden Seen (Ammersee und Starnberger See) am linken Bildrand von der schneebedeckten Umgebung ab. Die Wolke in der rechten unteren Bildecke hebt sich kaum von Untergrund ab, da sie ähnlich stark reflektiert wie der Schnee (vgl. Wolkenmaske rechts unten, blau=*wolkig*, schwarz=*klar*).

## 2. Grundlagen

**Streuung** Die Streuung des einfallenden Lichtes (VNIR) an den Wassertröpfchen oder Eiskristallen, aus denen Wolken bestehen, ist der Haupteffekt bei der Betrachtung der Interaktion von Licht mit Wolken. Die beiden zuvor beschriebenen Eigenschaften von Wolken basieren auf dieser Streuung, die auch als Mie-Streuung bekannt ist (Streuung von Licht an Partikeln mit Durchmessern in der Größenordnung der Wellenlänge des Lichts) [41].

Befindet sich eine Wolke mit relativ geringer optischer Dicke über einer künstlichen oder natürlichen Lichtquelle (also Feuer, Straßenbeleuchtung etc.) tritt durch die Streuung ein Effekt zum Vorschein, der für nächtliche VNIR-Satellitenaufnahmen von urbanen Gebieten besonders charakteristisch ist. Durch die Streuung des Lichts beim Durchdringen der Wolke erscheinen darunter liegende beleuchtete Strukturen nicht nur dunkler, sondern auch verschwommen und weiter ausgedehnt, als bei Beobachtung mit wolkenfreier Sicht. Die Wolke wirkt also wie ein Tiefpassfilter [26] und vermindert den Kontrast des Bildes ([23], [25]), was auf Abbildung 2.7 gut zu erkennen ist.

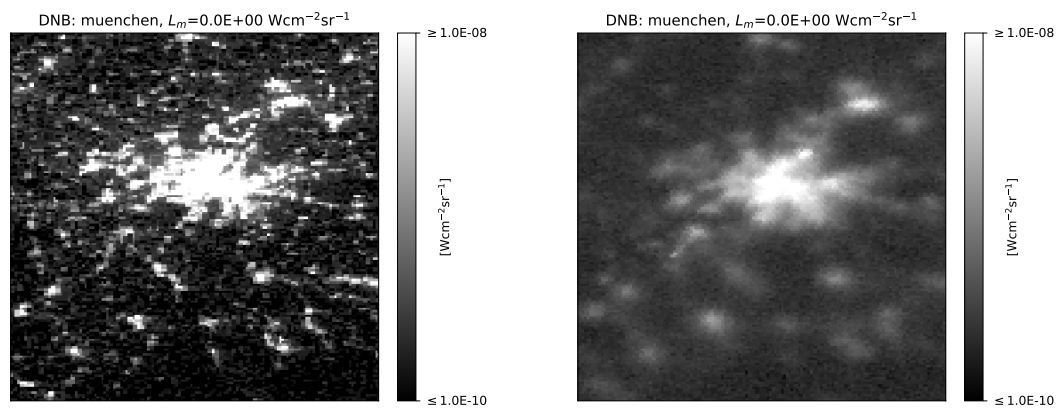


Abbildung 2.7: Beispiel für den Streuungseffekt durch Wolken. Dargestellt ist zweimal der gleiche Bildausschnitt von München ohne Beleuchtung durch den Mond. Links bei klarem Himmel, rechts bei komplett bewölktem Himmel. Die Bewölkung hat eine relativ geringe optische Dicke, wodurch die Lichter der Stadt zwar noch erkennbar sind, aber stark gestreut werden.

Der Effekt der Streuung hat außerdem einen Einfluss auf die Genauigkeit der Georeferenzierung von optischen Nachtaufnahmen, da er bei Off-Nadir Beobachtungen zu einem geometrischen Versatz der sich unter einer Wolke befindenden beleuchteten Szene führt. Die schematische Darstellung in Abbildung 2.8 zeigt die Ursache für diesen Versatz. Der Sensor auf dem Satelliten erfasst an Stelle der beleuchteten Szene auf der Erdoberfläche die durch die Szene beleuchtete Wolke. Diese befindet sich jedoch in einer deutlich größeren Höhe, als die Szene selbst, und wird dadurch bei der Georeferenzierung durch Schnitt der Blickrichtung des Pixels mit dem Geländemodell an einer falschen Stelle der Erdoberfläche abgebildet, wie es bei Off-Nadir Beobachtungen von Wolken immer der Fall ist. Dadurch, dass nur die leuchtende Wolke und nicht die Stadt selbst vom Satelliten aus sichtbar ist,

## 2.2. Beobachtungsmodell für VNIR-Nachtaufnahmen

tritt dieser geometrische Versatz bei Nacht auch für die Stadt auf. Es handelt sich hierbei um einen typischen Höhenfehler, wie er zum Beispiel auch beim Orthorektifizieren von Bildern bei Gebäuden, die nicht Teil des verwendeten Oberflächenmodells sind, auftritt.

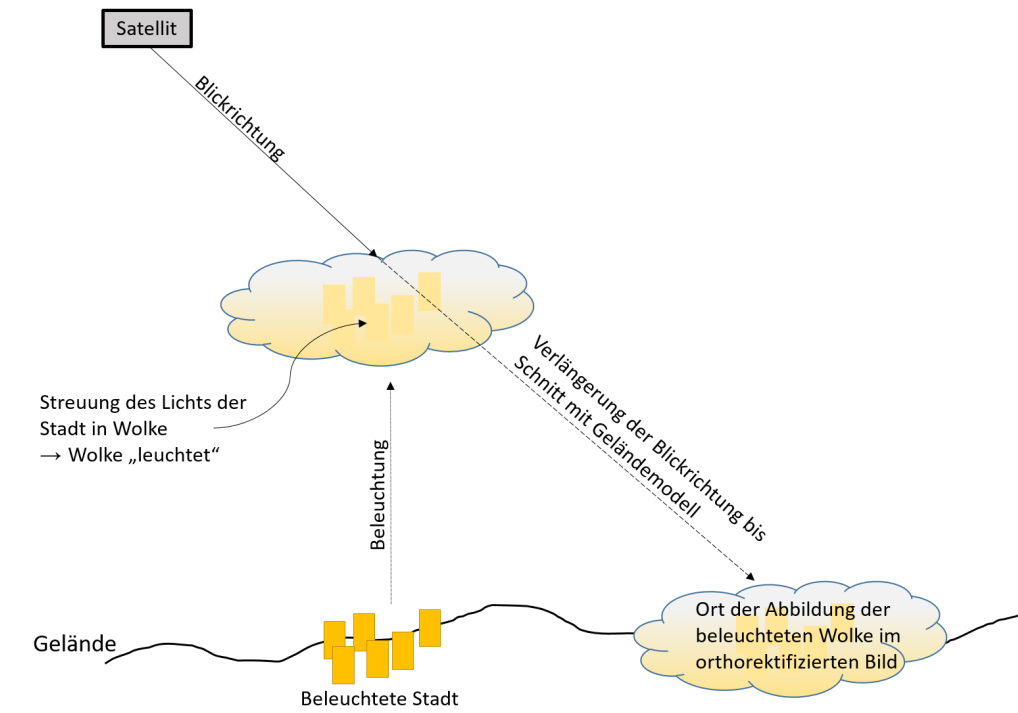


Abbildung 2.8: Schematische Darstellung des durch Wolken verursachten Fehlers in der Georeferenzierung (nicht maßstäblich). Durch die Streuung des von der Stadt emittierten Lichts in der Wolke wird die Stadt bei der Georeferenzierung genau wie die nicht im Geländemodell enthaltene Wolke auf einen falschen Ort auf der Erdoberfläche projiziert.

### 2.3. Algorithmus zur semantischen Segmentierung – Random Forests

Für die Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken wurde mit Random Forests ein Standardverfahren des Maschinellen Lernens als Klassifikator gewählt (zur Begründung der Auswahl siehe Kapitel 3.3.2). Im Folgenden wird das Prinzip von Random Forests kurz dargestellt, um die in Kapitel 3 verwendeten Begrifflichkeiten und Parameter zu definieren.

Random Forests basieren auf dem Prinzip des Ensemble-Lernens und wurden von Breiman [8] als Anwendung von Bagging auf CART-Entscheidungsbäume (beides ebenfalls Konzepte von Breiman et al., [7] und [6]) vorgestellt.

Das Ensemble-Lernen ist ein allgemeines Konzept, bei dem mehrere, nicht ideale Modelle (die sogenannten Basislerner – bei Klassifikationsaufgaben sind dies einzelne Klassifikatoren) miteinander kombiniert werden, anstatt ein einziges von ihnen auszuwählen und nur dieses zum Beispiel zur Klassifikation zu verwenden. Durch diese Kombination der Ergebnisse der Basislerner kann meist ein besseres und robusteres Ergebnis erzielt werden, als mit dem besten der einzelnen Klassifikatoren. Voraussetzung hierfür ist, dass die einzelnen Basislerner eine möglichst große Varianz aufweisen. Das heißt, dass die Basislerner alle möglichst verschiedene Ergebnisse liefern sollen [17]. Breiman [8] leitet in seiner Arbeit zu Random Forests her, dass durch eine große Varianz (bzw. kleine Korrelation) und gleichzeitig große Stärke bzw. Genauigkeit der Basislerner der Generalisierungsfehler des Modells minimiert werden kann. Es können aber auch mit Basislernern, die nur etwas besser als der Zufall sind, gute Ergebnisse erzielt werden, solange ihre Varianz groß genug ist.

Um ein solches Modell mit möglichst unkorrelierten Basislernern zu erzeugen, wird bei Random Forests an verschiedenen Stellen eine gewisse Zufälligkeit in das Training eingebracht. Dies geschieht zum einen durch die Verwendung von CART als Basislerner und zum anderen durch deren Kombination mittels Bagging. Beide Konzepte werden im Folgenden kurz vorgestellt, bevor die Funktionsweise von Random Forests genauer erläutert wird.

#### 2.3.1. CART – Classification And Regression Trees

CART (Classification And Regression Trees) [6] ist eine Methode zum Lernen von binären Entscheidungsbäumen. Sie integriert sowohl die Verwendung von Zufälligkeit, als auch die Berücksichtigung der Klassenverteilungen beim Lernen und ist so in der Lage, relativ starke Entscheidungsbäume zu generieren.

Bei CART wird beim Lernen des Baumes der Trainingsdatensatz rekursiv an jedem Knoten anhand eines Schwellwert-Tests beziehungsweise einer Trennfläche in zwei Teilmengen aufgeteilt. Hierfür wird an jedem Knoten eine zufällig gewählte Teilmenge der Merkmale betrachtet, wobei die Anzahl der verwendeten Merkmale vom Nutzer gewählt werden muss. Für diese Merkmale werden einige zufällige Trennebenen beziehungsweise Schwellwert-Tests generiert. Anschließend wird derjenige Test ausgewählt, der den meisten Informationsgewinn bringt. Das bedeutet, dass dieser Test die Daten so trennt, dass die Klassenverteilungen in den beiden Kind-Knoten möglichst rein werden. Als Maß für diese Reinheit

### 2.3. Algorithmus zur semantischen Segmentierung – Random Forests

bzw. Unreinheit der Klassenverteilung in einem Knoten wird zum Beispiel die Entropie oder der Gini-Index (beim originalen CART-Algorithmus) verwendet.

$$\text{Entropie} = - \sum_{c=1}^C P(c|n_t) \cdot \log(P(c|n_t)) \quad (2.6)$$

$$\text{Gini-Index} = 1 - \sum_{c=1}^C P(c|n_t)^2 \quad (2.7)$$

Hierbei ist  $C$  die Anzahl der Klassen und  $P(c|n_t)$  die Wahrscheinlichkeit für Klasse  $c$  im Knoten  $n_t$  des Baumes  $t$ .

Der Informationsgewinn berechnet sich schließlich aus der Differenz der Unreinheit des aktuellen Knotens und der Unreinheit der beiden neu entstehenden Kindknoten.

Auf diese Weise wird der Trainingsdatensatz immer weiter in Teilmengen aufgeteilt, bis eines der Abbruchkriterien erreicht wird. Als Abbruchkriterien werden unter anderem die maximale Baumtiefe, eine Mindestanzahl für Datenpunkte pro Knoten, ein minimaler Informationszuwachs oder die Reinheit eines Knotens (das heißt alle Daten gehören schon zu einer Klasse) verwendet. Die Werte für diese Kriterien sind Hyperparameter des Verfahrens und werden vor dem Training festgelegt. Wenn eine der Abbruchbedingungen erreicht wird, wird der Knoten zu einem Blatt. Für dieses Blatt werden dann die Eigenschaften der sich dort befindenden Trainingsdaten gespeichert (z.B. die Verteilung der Klassen).

Mit diesem Prinzip können aufgrund der Berücksichtigung der Klassenverteilungen beim Lernen der Schwellwert-Tests relativ starke Entscheidungsbäume generiert werden. Da der beste der Tests immer aus einer zufällig erzeugten Menge von Tests und Merkmalen ausgewählt wird, kann anhand der Häufigkeit und der Position im Baum, bei der ein Merkmal für einen solchen Test verwendet wird, außerdem bestimmt werden, wie wichtig die einzelnen Merkmale für die Klassifikation sind. Diese Merkmals-Wichtigkeiten können genutzt werden, um das gesamte Klassifikationsverfahren zu optimieren, indem zum Beispiel unwichtige Merkmale weggelassen werden [17].

#### 2.3.2. Bagging – Bootstrap Aggregating

Das Ensemble-Lernen mit CART-Bäumen wird bei Random Forests durch Bagging realisiert. Bagging (Bootstrap aggregating) bezeichnet das Training von mehreren Basislernern anhand von Bootstrap-Datensätzen und die Kombination ihrer Ergebnisse durch Mittelbildung oder Abstimmung [7].

Ein Bootstrap-Datensatz wird aus einer Menge an Trainingsdaten erzeugt, indem aus ihr für jeden Basislerner eine zufällige Teilmenge gezogen wird (nach dem Prinzip „Ziehen mit Zurücklegen“). Das heißt, jeder Basislerner wird nur auf einem Teil der Trainingsdaten trainiert. Da besonders beim Lernen von Entscheidungsbäumen eine starke Abhängigkeit des gelernten Modells von den Trainingsdaten vorliegt, erhöht dies die Varianz der Basislerner und steigert damit die Genauigkeit des Modells [8].

## 2. Grundlagen

### 2.3.3. Training und Klassifikation mit Random Forests

Beim Training eines Random Forest werden also mehrere CART-Entscheidungsbäume anhand von Bootstrap-Datensätzen trainiert. Die Anzahl der Bäume  $T$  ist hierbei ein Hyperparameter des Verfahrens und muss vor dem Training festgelegt werden. Breiman [8] zeigt, dass der Generalisierungsfehler eines Random Forest mit steigender Anzahl von Bäumen zunächst ansteigt und schließlich einen Grenzwert erreicht. Es besteht in dieser Hinsicht also keine Gefahr des Overfittings (Überanpassung des Klassifikators an die Trainingsdaten). Dennoch sollte die Anzahl der Bäume eingeschränkt werden, da sie den Trainingsaufwand sehr stark erhöhen kann.

Die Klassifikation eines Datenpunktes erfolgt schließlich, indem dessen Merkmalsvektor durch alle CART-Bäume des Waldes geschleust wird. Die Kombination der Ergebnisse der einzelnen Bäume ergibt dann das Ergebnis der Klassifikation.

Für die Kombination der Einzelergebnisse gibt es verschiedene Ansätze, zum Beispiel kann jeder Baum eine Stimme für die laut ihm wahrscheinlichste Klasse abgeben oder er berücksichtigt bei seiner Stimme die Wahrscheinlichkeiten für alle Klassen. In der in dieser Arbeit verwendeten Implementierung im Python-Modul *Scikit-learn* [40] wird letztere Variante verwendet, da sie die Wahrscheinlichkeitsverteilung für die einzelnen Klassen genauer abbildet. Die Kombination der Einzelergebnisse erfolgt hier durch Mittellung der prädierten Klassenwahrscheinlichkeiten jedes Baumes für jede Klasse und Auswählen derjenigen Klasse, für die die so berechnete Wahrscheinlichkeit  $P(c|X)$  maximal wird.

$$P(c|X) = \frac{1}{T} \sum_{t=1}^T P(c|n_t) \quad (2.8)$$

Hierbei ist  $T$  die Anzahl der Bäume und  $P(c|n_t)$  die Wahrscheinlichkeit, dass ein Sample, das im Blatt  $n_t$  des Baumes  $t$  ankommt, zur Klasse  $c$  gehört. Diese Wahrscheinlichkeit wird jeweils aus der Klassenverteilung der Trainingsdaten im Blatt abgeleitet [17]. Die Klassenwahrscheinlichkeiten  $P(c|X)$  können auch direkt als Ergebnis des Random Forests verwendet werden und sind ein gutes Indiz dafür, wie sicher der Datenpunkt  $X$  einer Klasse zugeordnet werden kann.

Durch das Prinzip der Kombination von CART und Bagging hat ein Random Forest einige vorteilhafte Eigenschaften. So ist er im Vergleich zu anderen Verfahren, wie zum Beispiel der Support Vector Machine [5], kaum anfällig für Overfitting. Außerdem ist er sehr robust gegenüber Ausreißern und Rauschen, wodurch auch mit fehlerbehafteten Trainingsdaten noch gute Ergebnisse erzielt werden können. Bisher nicht erwähnt wurde die Möglichkeit, einen Random Forest anstatt zur Klassifikation zur Regression kontinuierlicher Werte einzusetzen. Dies kann im Falle der Anwendung zur semantischen Segmentierung genutzt werden, falls als Trainingsdaten für eine binäre Klassifikation kontinuierliche Wahrscheinlichkeiten für das Auftreten der gesuchten Klasse vorhanden sind. Hinsichtlich der Implementierung ist ein Random Forest durch die Verwendung von vielen voneinander unabhängigen CART-Entscheidungsbäumen zudem einfach parallelisierbar und kann dadurch auch mit großen Datensätzen effizient genutzt werden [17].

## 3. Methodik

Der Prozess zur Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken in nächtlichen Satellitenaufnahmen kann in drei Schritte unterteilt werden. Der erste Schritt beinhaltet die Vorbereitung der verwendeten Daten. Im zweiten Schritt werden Merkmale definiert, mit denen die Eigenschaften der Daten charakterisiert werden können. Auf dieser Grundlage wird schließlich im dritten Schritt das eigentliche Wolkenerkennungsverfahren definiert und optimiert.

### 3.1. Vorverarbeitung

Als Grundlage für die Entwicklung des Wolkenerkennungsverfahrens wurde ein Workflow zur Vorverarbeitung von DNB- und ECM-Daten entwickelt, um diese als Trainingsdaten bzw. Ground Truth Daten nutzbar zu machen. Dies schafft außerdem die Basis für weitere Analysen auf diesen Daten.

In einem ersten Schritt wird zunächst eine Auswahl von Gebieten definiert, die in den Untersuchungen betrachtet werden sollen. Im zweiten Schritt wird die Vorgehensweise zur Vorverarbeitung der entsprechenden Daten erläutert.

#### 3.1.1. Testgebiete

Die im Rahmen dieser Arbeit beschriebenen Trainings und Tests von Wolkenklassifikatoren werden exemplarisch anhand einiger besonders charakteristischer Bildausschnitte ( $150 \times 150$  Pixel) durchgeführt, die als Beispiele für verschiedene Typen von Landbedeckungen dienen sollen. Da der Fokus auf der Wolkenerkennung für urbane Gebiete liegt, werden als Testgebiete vor allem Bildausschnitte verwendet, die solche Landbedeckungen in verschiedener Ausprägung zeigen. Als Kontrast dazu kommen zwei Testgebiete hinzu, die eine Referenz ohne künstliche Beleuchtung darstellen. Zusätzlich werden Beispielgebiete ausgewählt, die bestimmte, bei der nächtlichen Wolkenerkennung besonders herausfordernde Situationen repräsentieren, um die Performance des entwickelten Verfahrens unter solchen Bedingungen evaluieren zu können. Im Folgenden werden diese verwendeten Testgebiete genauer vorgestellt.

#### Urbane Testgebiete

1. München (im Folgenden bezeichnet als *muenchen*)
2. Stuttgart (*stuttgart*)

München und Stuttgart wurden als Beispiele für typische europäische urbane Gebiete mittlerer Größe ausgewählt, die sowohl sehr dicht besiedelte und dadurch stark beleuchtete Gebiete als auch weniger dicht besiedeltes Umland enthalten. Die beiden Städte repräsentieren somit eine gute Schnittmenge verschiedener Ausprägungen von

### 3. Methodik

Besiedelung und Beleuchtung. Abbildung 3.1 zeigt ein wolkenfreies Beispiel des gewählten Bildausschnittes für München und Abbildung 3.2 ein vergleichbares Beispiel für Stuttgart.

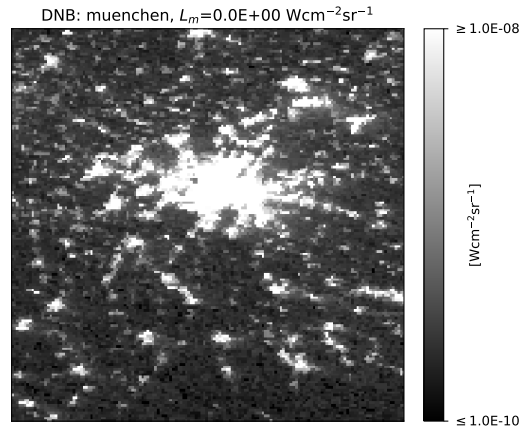


Abbildung 3.1: Beispiel eines Bildausschnittes des Testgebietes *muenchen*. Aufgenommen am 21.03.2018 um ca. 02:03 Uhr (UTC) bei wolkenfreien Bedingungen und ohne Mondbeleuchtung.

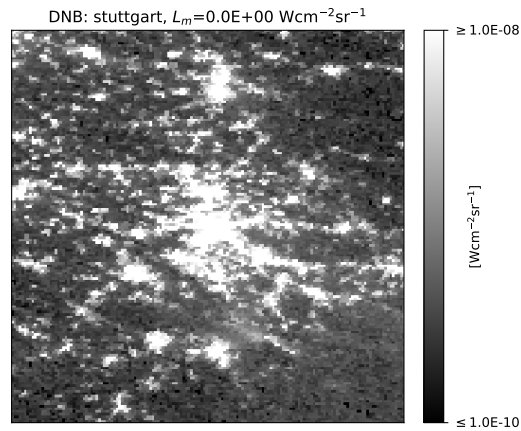


Abbildung 3.2: Beispiel eines Bildausschnittes des Testgebietes *stuttgart*. Aufgenommen am 15.01.2018 um ca. 02:43 Uhr (UTC) bei wolkenfreien Bedingungen und ohne Mondbeleuchtung.

3. Belgien (*belgien*)

Wie in Abbildung 3.3 gut zu erkennen, ist Belgien im Vergleich zu den anderen beiden urbanen Gebieten München und Stuttgart sehr stark beleuchtet. Dies liegt unter anderem an der dort üblichen Beleuchtung der Autobahnen. Belgien wurde deshalb als Beispiel für sehr stark beleuchtete Regionen ausgewählt.

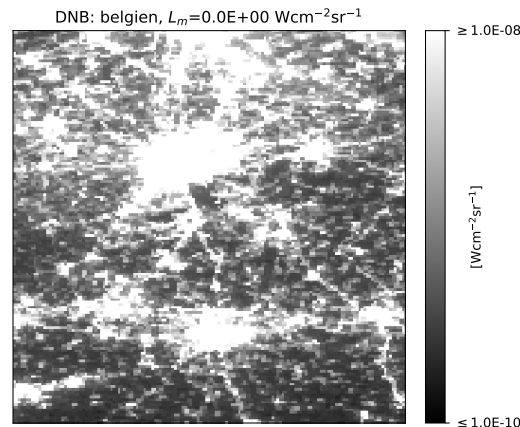


Abbildung 3.3: Beispiel eines Bildausschnittes des Testgebietes *belgien*. Aufgenommen am 14.01.2018 um ca. 01:24 Uhr (UTC) bei wolkenfreien Bedingungen und ohne Mondbeleuchtung.

4. Mailand (*mailand*)

Das Gebiet *mailand* ist ein weiteres Beispiel für stark beleuchtete urbane Regionen, siehe Abbildung 3.4.

5. New Orleans (*neworleans*)

Das Testgebiet *neworleans* besteht nicht wie die zuvor vorgestellten Gebiete aus einem urbanen Ballungsraum mit dünner besiedeltem Umland, sondern ist von starken Kontrasten zwischen unterschiedlichen Landbedeckungen geprägt. Im Norden grenzt die Stadt an einen See, wodurch ein direkter Übergang von einem dicht besiedelten Gebiet auf eine Wasserfläche gegeben ist. Abbildung 3.5 zeigt außerdem, wie gut der Verlauf des Mississippi (z.B. in der rechten unteren Bildecke) durch die angrenzende beleuchtete Besiedelung erkennbar ist.

6. Indien (*indien*)

Mit den Bundesstaaten Kerala und Tamil Nadu wurde ein Gebiet ausgewählt, das unterschiedlich dichte Besiedelungen aufweist und zusätzlich an den Ozean grenzt, wodurch sehr unterschiedliche Beleuchtungsverhältnisse im Testgebiet enthalten sind, wie in Abbildung 3.6 gut erkennbar ist. Speziell Indien wurde hierfür gewählt, um verschiedene Klimazonen und somit verschiedene Bewölkungsverhältnisse zu repräsentieren. Außerdem stellt das Gebiet im Gegensatz zu den anderen urbanen Gebieten ein Beispiel für die urbanen Strukturen in einem Schwellenland dar.

### 3. Methodik

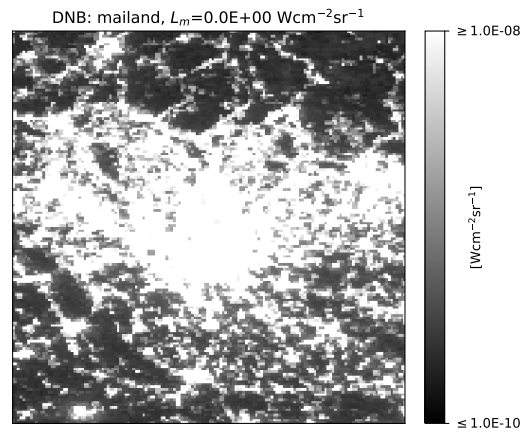


Abbildung 3.4: Beispiel eines Bildausschnittes des Testgebietes *mailand*. Aufgenommen am 14.01.2018 um ca. 01:24 Uhr (UTC) bei größtenteils wolkenfreien Bedingungen und ohne Mondbeleuchtung.

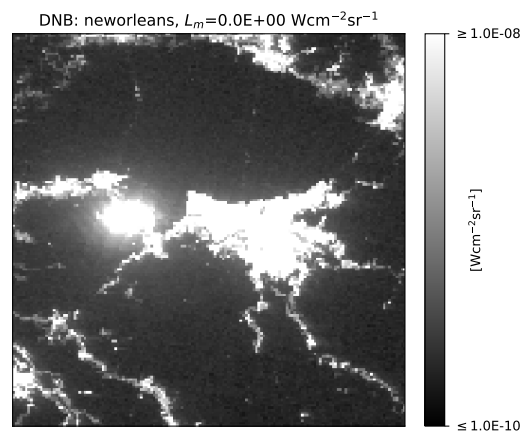


Abbildung 3.5: Beispiel eines Bildausschnittes des Testgebietes *neworleans*. Aufgenommen am 21.01.2018 um ca. 07:42 Uhr (UTC) bei größtenteils wolkenfreien Bedingungen und ohne Mondbeleuchtung.

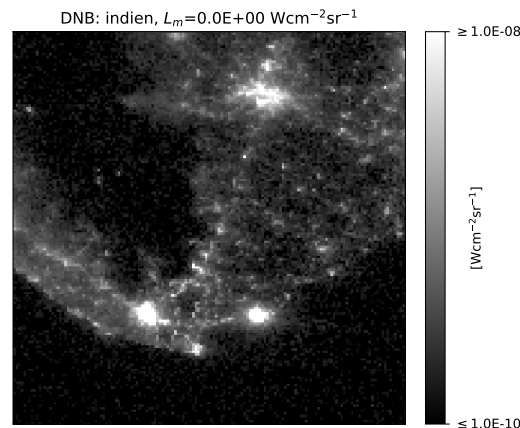


Abbildung 3.6: Beispiel eines Bildausschnittes des Testgebietes *indien*. Aufgenommen am 17.01.2018 um ca. 20:55 Uhr (UTC) bei größtenteils wolkenfreien Bedingungen und ohne Mondbeleuchtung.

**Testgebiete mit besonderen Landbedeckungen** Um besondere Herausforderungen der nächtlichen Wolkenerkennung im Spektralbereich des VNIR hervorzuheben, werden einige Gebiete betrachtet, die sich durch ihre Landbedeckung von den bisher vorgestellten Gebieten stark unterscheiden. Dies sind zum einen zwei Testgebiete, die den Ozean darstellen und somit als Referenzen ohne künstliche Beleuchtung dienen. Zum anderen wird jeweils ein Gebiet der beiden Landbedeckungsarten mit der höchsten Albedo, nämlich Wüste und Schnee, betrachtet. Diese Landbedeckungen können visuell oft leicht mit Bewölkung verwechselt werden, da sie im Bild sehr ähnlich aussehen. Dadurch stellen sie auch für den entwickelten Algorithmus potenzielle Schwierigkeiten dar.

1. Indischer Ozean (*indozean*)

Als erstes Testgebiet, das ausschließlich den Ozean enthält, wurde ein Gebiet im indischen Ozean, südlich vom Testgebiet *indien* ausgewählt. Eine wolkenfreie Aufnahme dieses Gebietes wird in Abbildung 3.7 dargestellt.

2. Ölplattform Petronius (*petronius*)

Als zweites Testgebiet, das einen Ozean enthält, wurde eine Region im Golf von Mexiko, südöstlich von New Orleans, verwendet. Das Besondere hieran ist, dass sich in diesem Gebiet einige Ölplattformen befinden, darunter eine der weltweit größten mit dem Namen Petronius. Solche Ölplattformen sind konstant hell beleuchtet und erscheinen in DNB-Aufnahmen als punktförmige Lichtquellen, weshalb sie sich zum Beispiel besonders gut für Analysen der Streuungseffekte von Wolken eignen. Abbildung 3.8 zeigt das Testgebiet bei klarem Himmel.

### 3. Methodik

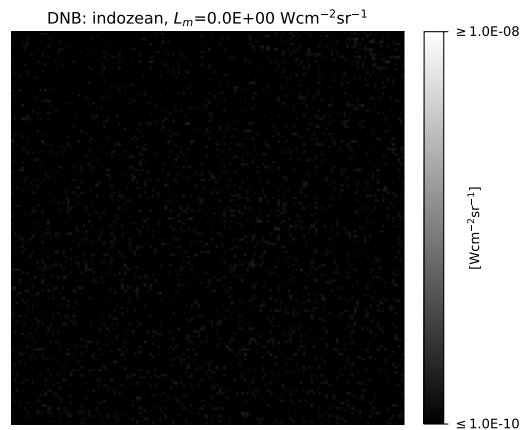


Abbildung 3.7: Beispiel eines Bildausschnittes des Testgebietes *indozean*. Aufgenommen am 19.01.2018 um ca. 20:20 Uhr (UTC) bei grostenteils wolkenfreien Bedingungen und ohne Mondbeleuchtung.

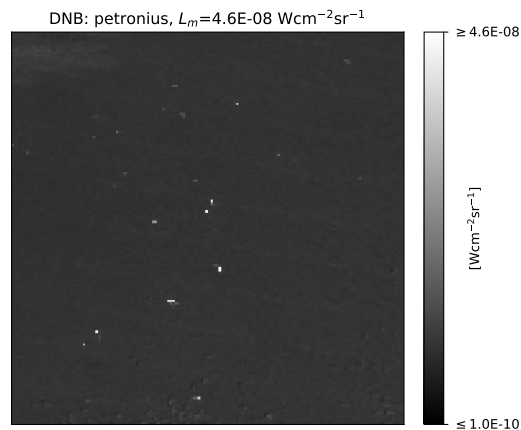


Abbildung 3.8: Beispiel eines Bildausschnittes des Testgebietes *petronius*. Aufgenommen am 05.01.2018 um ca. 07:43 Uhr (UTC) bei grostenteils wolkenfreien Bedingungen und mit Mondbeleuchtung.

#### 3. Nil (*nil*)

Als Beispiel fur ein Wustengebiet wurde ein Bildausschnitt gewahlt, der den Nil und die angrenzende Wustenregion zeigt. Wie in Kapitel 2.2.4 erlautert, zeichnet sich eine solche Landbedeckung gegenuber den in den anderen Bildausschnitten vorherrschenden Ackerflachen, bebauten Gebieten oder Wasserflachen durch eine hohere Albedo aus, sie reflektiert also deutlich mehr Mondlicht. Dies ist in Abbildung 3.9 gut zu erkennen.

### 3.1. Vorverarbeitung

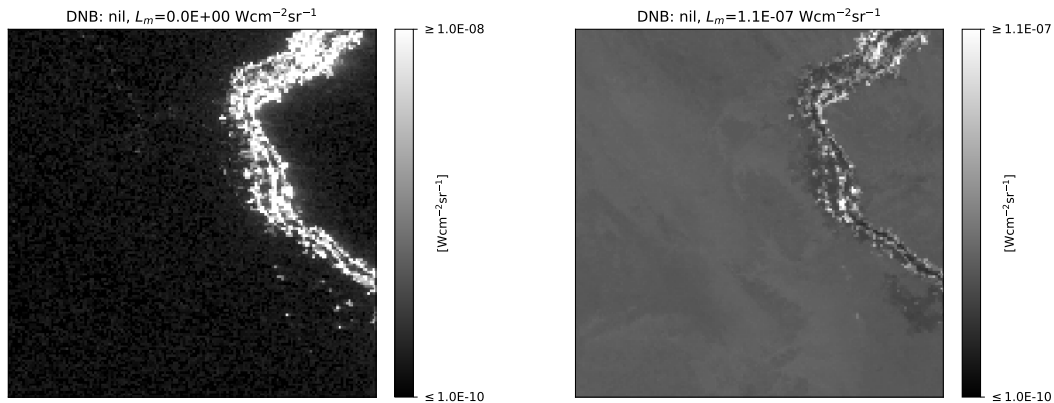


Abbildung 3.9: Beispiel eines Bildausschnittes des Testgebietes *nil*. Beide Bilder wurden bei wolkenfreien Bedingungen aufgenommen. Linkes Bild: Aufgenommen am 15.01.2018 um ca. 23:12 Uhr (UTC) ohne Mondbeleuchtung. Rechtes Bild: Aufgenommen am 02.01.2018 um ca. 23:52 Uhr (UTC) mit Mondbeleuchtung. Zu beachten ist die unterschiedliche Skalierung, siehe 2.1.1.

#### 4. Alpen (*alpen*)

Ein Ausschnitt der Alpenregion im Winter wurde als Beispiel für Gebiete mit Schneebedeckung gewählt. Die beiden Beispielbilder in Abbildung 3.10 zeigen deutlich den Unterschied des Erscheinungsbildes von Schnee ohne und mit Beleuchtung durch den Mond, der durch die hohe Albedo von Schnee verursacht wird.

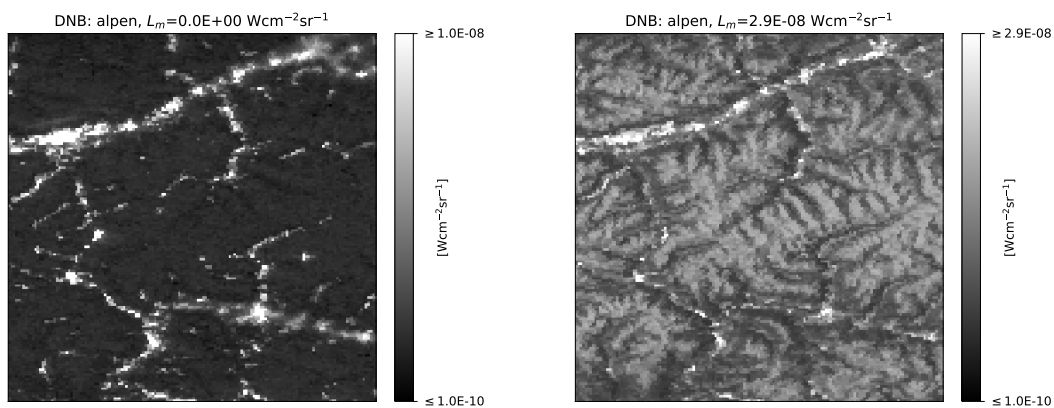


Abbildung 3.10: Beispiele eines Bildausschnittes des Testgebietes *alpen*, beide mit größtenteils wolkenfreien Bedingungen. Linkes Bild: Aufgenommen am 14.01.2018 um ca. 01:24 Uhr (UTC) ohne Mondbeleuchtung. Rechtes Bild: Aufgenommen am 29.01.2018 um ca. 01:40 Uhr (UTC) mit Mondbeleuchtung. Beim durch den Mond beleuchteten Bild ist die Schneebedeckung eindeutig erkennbar. Zu beachten ist die unterschiedliche Skalierung, siehe 2.1.1.

### 3. Methodik

#### 3.1.2. Vorverarbeitungskette

Die verwendeten DNB- und ECM-Daten werden in unterschiedlich großen Kacheln und verschiedenen Koordinatensystemen bereitgestellt, sie können also nicht ohne weitere Verarbeitung zueinander in Bezug gesetzt werden. Dies ist jedoch die Voraussetzung für ihre Nutzung als Trainings- und Ground Truth Daten. Da in dieser Arbeit nur kleine Ausschnitte aus den verfügbaren Daten verwendet werden sollen, müssen diese zudem einheitlich aus den bereitgestellten Daten ausgeschnitten werden. Um eine zeitliche Analyse der Daten zu ermöglichen, muss dabei zudem sichergestellt werden, dass sich alle Daten der Zeitreihe auf das gleiche Koordinatenraster beziehen, das heißt ein Pixel einer bestimmten Position muss sich immer auf ein und dieselbe Koordinate auf der Erdoberfläche beziehen.

Im Rahmen dieser Arbeit wurde deshalb eine Vorverarbeitungskette entwickelt, die all diese Schritte durchführt und flexibel für unterschiedliche Bildausschnitte angewandt werden kann. Dies ermöglicht es, auch für sich anschließende Untersuchungen unkompliziert Daten zur Verfügung zu stellen. Realisiert wurden die verschiedenen Vorverarbeitungsschritte mit SNAP [47] (Desktoanwendung, Python-Modul *snappy* sowie SNAP Graph Processing Tool) und Matlab.

In den Abbildungen 3.11 und 3.12 befindet sich der Ablaufplan der Vorverarbeitungskette für ein Testgebiet. Die wichtigsten Schritte werden im Folgenden näher erläutert. Ausgangspunkt der Vorverarbeitungsschritte sind alle DNB- und ECM-Daten des gewünschten Zeitraums, die das Testgebiet enthalten und innerhalb der möglichen Zeitspanne des nächtlichen Überflugs des S-NPP-Satelliten aufgenommen wurden.

1. Austausch Georeferenzierung:

Im Datensatz der DNB-Daten sind zwei Versionen der Informationen zur pixelweisen Georeferenzierung vorhanden. Eine, die sich nur auf das Referenzellipsoid bezieht und eine, die ein Geländemodell mit berücksichtigt (sog. *terrain corrected*). Standardmäßig wird die ungenauere Variante ohne die Berücksichtigung des Geländes verwendet, weshalb vor der Projektion der Daten in ein anderes Koordinatensystem der verwendete Georeferenzierungs-Kanal ausgetauscht werden muss, um Fehler in der Abbildung der Bilder zu verringern [52].

2. Masterbild:

Da alle Bildausschnitte eines Gebietes das selbe Pixelraster verwenden sollen, wird ein Bild benötigt, das als Referenz für alle anderen dient. Als ein solches Masterbild wird ein DNB-Bild verwendet, bei dem das zu extrahierende Gebiet möglichst im Nadir des Scans liegt. Der Grund hierfür sind die zum Rand des Scanstreifens hin anwachsenden Ungenauigkeiten in der Georeferenzierung.

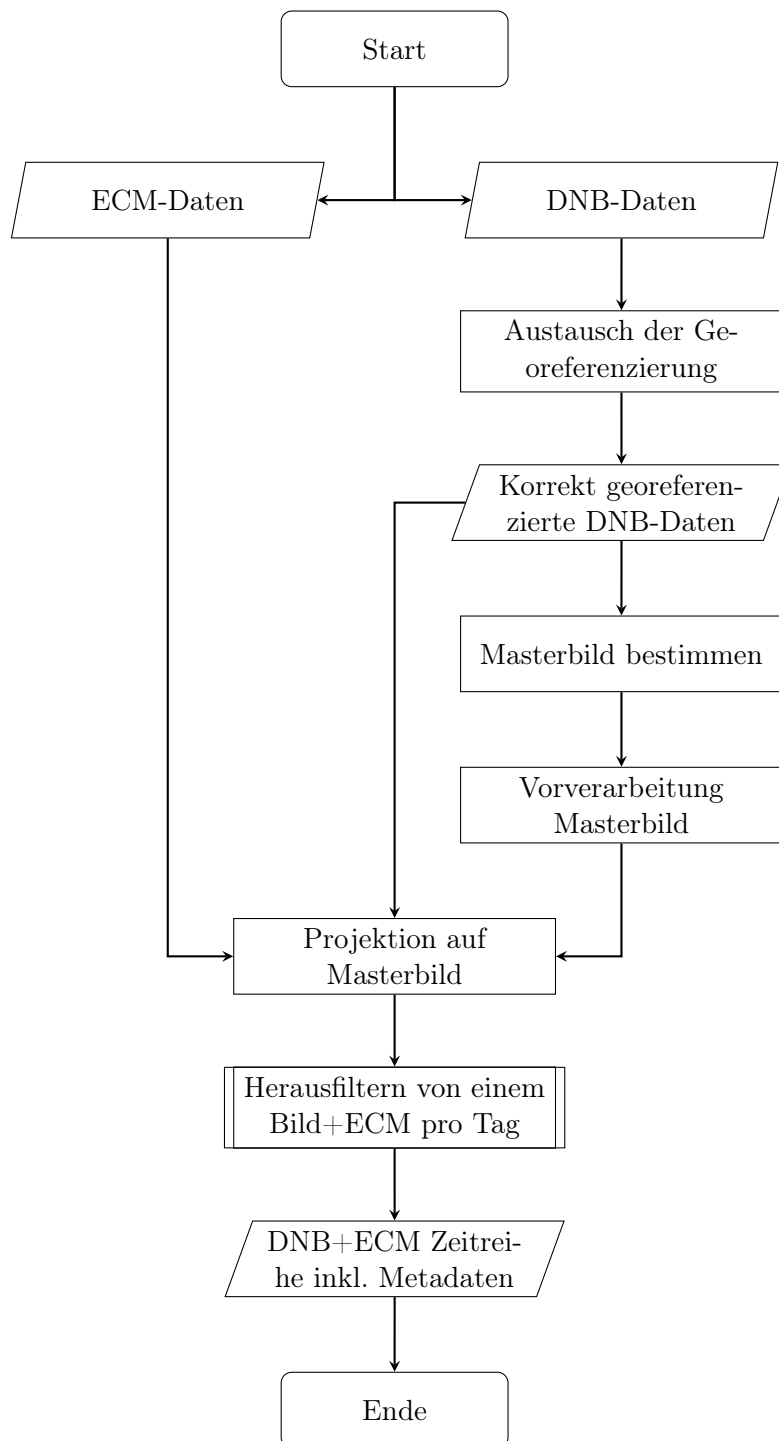


Abbildung 3.11: Ablaufdiagramm der Vorverarbeitungskette für ein Gebiet.

### 3. Methodik

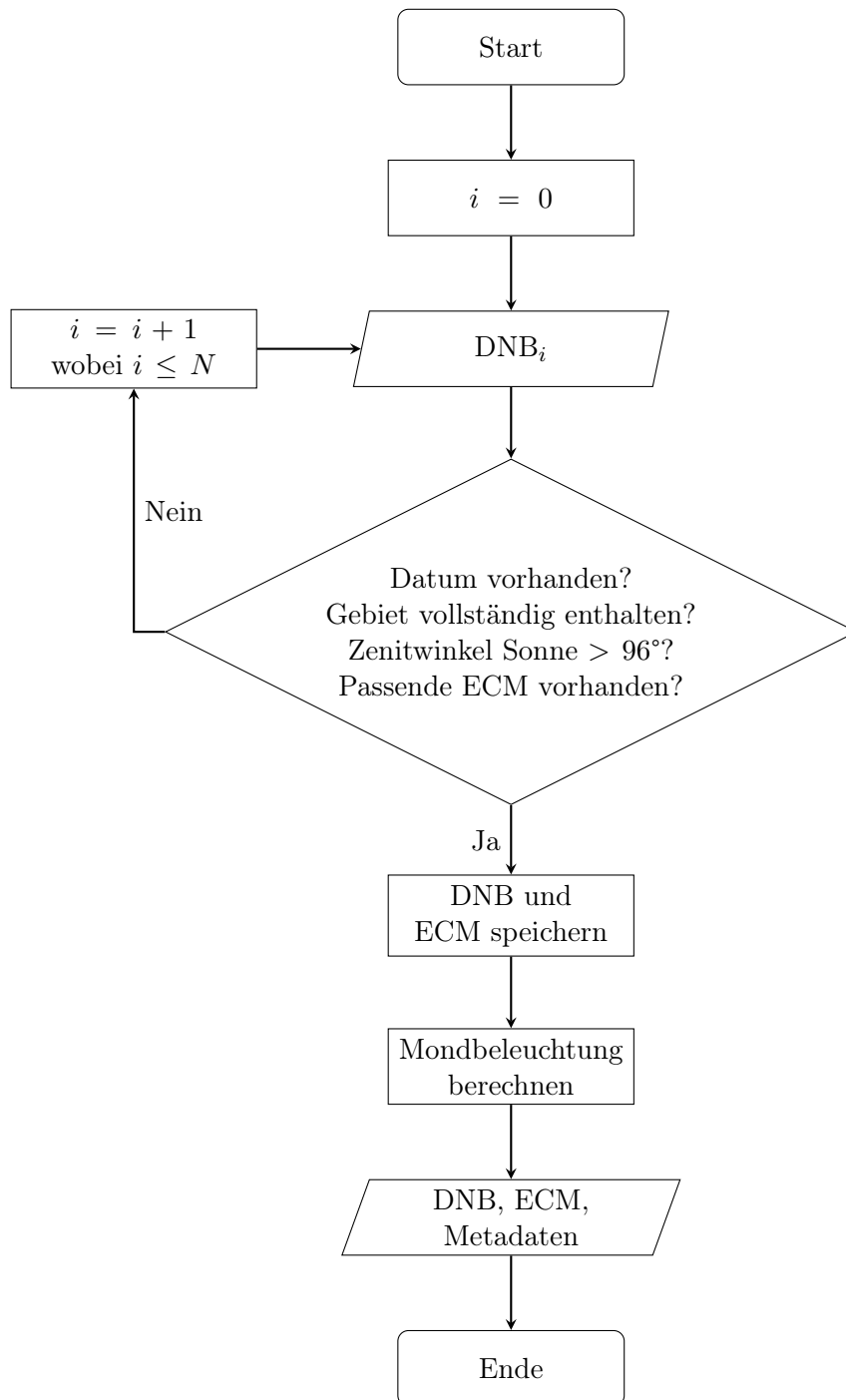


Abbildung 3.12: Ablaufdiagramm des Unterprogramms *Herausfiltern von einem Bild+ECM pro Tag* der Vorverarbeitungskette, siehe Abbildung 3.11.

#### 3. Vorverarbeitung Masterbild:

Das Masterbild wird zunächst in das WGS84 Koordinatenraster projiziert. Anschließend wird der durch ein rechteckiges Polygon spezifizierte Bildausschnitt aus der DNB-Kachel ausgeschnitten. Dieser Bildausschnitt dient nun als Referenz für alle weiteren Bilder des Gebietes. Die für diese Arbeit gewählte Größe der Bildausschnitte ist  $150 \times 150$  Pixel. Dies entspricht einer Seitenlänge von ungefähr 110 km und damit in etwa der Streifenbreite des N8-Satelliten.

#### 4. Projektion auf Masterbild:

Alle weiteren DNB- und ECM-Daten werden in gleicher Weise wie das Masterbild projiziert und zugeschnitten. Bei der Projektion werden die Daten mittels Interpolation auf das Pixelraster des Masterbildes abgebildet. Hierfür wird die Nearest-Neighbor Interpolation verwendet, damit die diskreten Werte der Wolkenmaske erhalten bleiben. Im Rahmen dieses Schrittes werden zusätzlich alle für diese Anwendung nicht benötigten Kanäle der DNB- und ECM-Daten verworfen, um die Datenmenge zu reduzieren.

#### 5. Herausfiltern von einem Bildausschnitt pro Tag:

Für die Analysen soll pro Tag nur ein Bildausschnitt mit zugehöriger Wolkenmaske verwendet werden, die heruntergeladenen Daten enthalten aufgrund des Verfahrens zur Datenauswahl jedoch meist mehrere Bilder pro Tag. Deshalb werden von der Gesamtmenge der  $N$  vorhandenen DNB-Bildausschnitte nur diejenigen beibehalten, die alle folgenden Herausforderungen erfüllen:

- Es ist noch kein Bild des gleichen Datums vorhanden.
- Das Testgebiet ist vollständig im Bildausschnitt enthalten (Test auf ungültige Werte, die entstehen, wenn das Testgebiet am Rand einer Bildkachel liegt und dadurch nicht vollständig auf das Masterbild projiziert werden kann).
- Der Sonnenstand ist passend (siehe hierzu Abschnitt 3.1.3).
- Es ist eine zum DNB-Bildausschnitt passende Wolkenmaske vorhanden (Suche nach einer ECM mit passendem Aufnahmezeitpunkt und Überprüfung des ECM-Ausschnitts auf ungültige Werte).

#### 6. Berechnung der lunaren Bestrahlungsstärke:

Pro Bild wird ein Wert für  $L_{MT2009}$  anhand des in Kapitel 2.2.2 beschriebenen Modells berechnet.

#### 7. Output:

Als Ergebnis der Vorverarbeitung liegt ein Bildpaar (DNB und ECM) pro Tag (insofern für diesen Tag passende Daten verfügbar sind) inklusive der zugehörigen Metadaten (Modell der Mondbeleuchtung, Aufnahmezeitpunkt etc.) vor.

### 3. Methodik

#### 3.1.3. Herausfiltern von nicht geeigneten Bildern

Die verwendeten DNB- und ECM-Daten sind in bestimmten Fällen nicht für die Verwendung zur Wolkenerkennung geeignet. Der erste Fall betrifft eine zum Aufnahmezeitpunkt unter Umständen vorliegende Beleuchtung durch die Sonne, was einige DNB-Daten für die Entwicklung von auf Nachtaufnahmen bezogenen Methoden ungeeignet macht. Der zweite Fall betrifft die Wolkendaten, die teilweise fehlerhaft sind.

**Sonneneinstrahlung** Da je nach Jahreszeit und Breitengrad sehr unterschiedliche Beleuchtungsverhältnisse durch die Sonne vorliegen, werden aus den Produkten diejenigen herausgefiltert, die bei Tageslicht aufgenommen wurden. Dies schließt sowohl diejenigen mit ein, bei denen die Sonne noch über dem Horizont steht, als auch diejenigen, die in der Phase der bürgerlichen Dämmerung aufgenommen wurden.

Die bürgerliche Dämmerung ist definiert als der Zeitraum, in dem die Sonne einen Zenitwinkel von  $90^\circ - 96^\circ$  hat. Durch die Refraktion der Atmosphäre kommt in diesem Zeitraum immer noch relativ viel Sonnenlicht auf der Erde an, weshalb er bei der Aufteilung der Beleuchtung in Tag/Nacht meist noch zum Tag gezählt wird (vgl. [3]). Deshalb wird auch in dieser Arbeit die  $96^\circ$  Grenze als Tag/Nacht-Grenze verwendet. Obwohl auch die beiden darauf folgenden Dämmerungsphasen (nautische Dämmerung ( $96^\circ < \text{Zenitwinkel} < 102^\circ$ ) und astronomische Dämmerung ( $102^\circ < \text{Zenitwinkel} < 108^\circ$ )) immer noch nicht komplett frei von Sonnenlicht sind, werden die in diesen Phasen aufgenommenen Bilder verwendet, da sie dunkel genug sind, damit die typischen nächtlichen Effekte beobachtbar sind. Außerdem würde mit einer Einschränkung des Verfahrens auf absolute Nacht (ab einem Zenitwinkel von  $> 108^\circ$ ) auch eine große Beschränkung der verfügbaren Daten einhergehen, da dieser Zustand für nördliche Breiten ( $> 49^\circ$ ) im Sommer überhaupt nicht eintritt.

Der Zenitwinkel, anhand dessen die Bilder herausgefiltert werden, ist das Mittel der als Teil der Georeferenzierungsdaten von DNB pixelweise bereitgestellten Sonnen-Zenitwinkel innerhalb des Bildausschnitts. Die Verwendung eines solchen Mittelwertes approximiert die tatsächlichen Verhältnisse unter Vernachlässigung der Ausdehnung des Gebietes und des Geländes. Für das Herausfiltern von zu hellen Szenen ist dies aber völlig ausreichend, da die durchschnittliche maximale Abweichung der Zenitwinkel innerhalb eines Bildausschnittes nur ungefähr  $1.12^\circ$  beträgt.

**Fehlerhafte Wolkendaten** Bei Verwendung der ECM-Wolkenmasken ist zu beachten, dass diese in wenigen Fällen Artefakte beinhalten können, die die Maske für Analysen und Training unbrauchbar machen (vgl. Kapitel 2.1.2). Da diese Fehler nicht in den Qualitätsinformationen der Daten vermerkt sind, wurden die fehlerhaften Wolkenmasken auf Basis visueller Beurteilung aussortiert. In diesem Zuge wurden ebenfalls solche Bilder aussortiert, bei denen die ECM-Wolkenmaske offensichtlich falsch ist, zum Beispiel wenn eine im DNB-Bild klar erkennbare Wolke nicht in der Wolkenmaske enthalten ist.

### 3.2. Definition von Merkmalen

Da mit Random Forests ein überwachtes, merkmalsbasiertes Verfahren zur semantischen Segmentierung gewählt wurde (siehe hierzu Abschnitt 3.3.2), müssen zunächst Merkmale definiert werden, anhand derer das Training und die Klassifikation mit dem Random Forest erfolgen soll. Im folgenden Abschnitt werden deshalb zunächst einige für die Charakterisierung von Wolken geeignete Merkmale vorgeschlagen.

Die Merkmale sollen die in Kapitel 2.2.4 beschriebenen Einflüsse von Wolken auf optische Nachtaufnahmen möglichst charakteristisch beschreiben. Es werden also Merkmale zur Charakterisierung der Albedo und der optischen Dicke und zur Beschreibung der durch die Streuung verursachten Effekte benötigt. Da als Datengrundlage ausschließlich Satellitenbeobachtungen im VNIR-Bereich sowie das Modell der Mondbeleuchtung verwendet werden sollen, müssen alle Merkmale aus dieser begrenzten Menge an Information (d.h. jeweils ein Wert für die beobachtete und die einfallende Strahldichte pro Pixel) ableitbar sein.

Im Folgenden werden auf Basis dieser Voraussetzungen zunächst Merkmale vorgestellt, die für jedes Bild bzw. jedes Pixel separat berechnet werden können und anschließend solche, die die zeitliche Dimension der Einflüsse von Wolken berücksichtigen.

#### 3.2.1. Einzelbildmerkmale

**Optische Dicke und Albedo** Die optische Dicke und die Albedo einer Wolke beeinflussen generell die Helligkeit eines Pixels, also die mit dem Sensor beobachtete Strahldichte. Das erste mögliche Merkmal ist somit der direkt gemessene Grauwert des Pixels. Da Wolken meist weiträumig ausgedehnt sind, liegt alternativ die Verwendung der mittleren Strahldichten  $\mu_{L_{DNB}}$  in einem Fenster um das betrachtete Pixel herum nahe (zur Wahl der Fenstergröße siehe Kapitel 3.3.4).

$$\mu_{L_{DNB}} = \frac{1}{N_{pix}} \sum_{i=1}^{N_{pix}} L_{DNB,i} \quad \text{mit} \quad N_{pix} = \text{Anzahl Pixel im Fenster} \quad (3.1)$$

Wie in Kapitel 2.2.2 beschrieben, werden diese Beobachtungen jedoch sehr stark von der vorliegenden Beleuchtung durch den Mond beeinflusst. Die absolute beobachtete Strahldichte eines Pixels, das eine Wolke abbildet, ist also bei Neumond und Vollmond sehr verschieden. Deshalb bietet sich die ebenfalls in Kapitel 2.2.2 vorgestellte Größe der Reflektanz als Merkmal an, sie setzt schließlich die Beobachtung des Sensors ins Verhältnis zur Beleuchtung durch den Mond und kann somit als relatives Maß für die Helligkeit des Pixels verwendet werden. Da die Reflektanz nur für  $L_{MT2009} > 0$  definiert ist (vgl. Gleichung (2.5)), ist sie jedoch für etwa die Hälfte der Zeit nicht bestimmbar, da sich dann der Mond zum Beobachtungszeitpunkt unterhalb des Horizontes befindet oder die Mondphase dem Neumond sehr nahe ist und somit keine Beleuchtung durch den Mond vorliegt.

Um ein für alle Fälle definiertes Maß für die Reflektanz bzw. Albedo zu erhalten, wurde

### 3. Methodik

deshalb eine Alternative zum Maß der Reflektanz entwickelt. Dieses Maß ( $\Delta L_{norm}$ ) ist die normierte Differenz der Strahldichten des Mondes und der Beobachtung.

$$\Delta L_{norm} = \frac{L_{MT2009} - L_{DNB}}{L_{MT2009} + L_{DNB}} \quad (3.2)$$

Das Merkmal nimmt im Fall von  $L_{MT2009} = 0$  den Wert  $-1$  an. In allen anderen Fällen beschreibt ein kleinerer Wert für  $\Delta L_{norm}$ , ausgehend vom Maximalwert  $\Delta L_{norm} = 1$  (beobachtete Strahldichte ist gleich 0), eine größere anteilige Reflexion des Mondlichtes, also immer größeren Reflektanzen. Bei  $\Delta L_{norm} = 0$  sind die vom Mond aus einfallende und die empfangene Strahldichte gleich groß. Dies entspräche theoretisch der Beobachtung einer Fläche mit einer Reflektanz von 1. Werte zwischen 0 und  $-1$  beschreiben im Verhältnis zum Mond immer größere empfangene Strahldichten, also Reflektanzen größer als Eins. Dies ist bei Lichtemissionen von der Erdoberfläche oft der Fall. Bei der Interpretation von  $\Delta L_{norm}$  wie auch der Reflektanz sollte generell berücksichtigt werden, dass solche weiteren Beleuchtungsquellen bei der Berechnung vernachlässigt wurden.

Da die beschriebene Normierung der Differenz immer zur Folge hat, dass  $\Delta L_{norm}$  in Fällen ohne Mondbeleuchtung immer  $-1$  ist, und damit für diese Fälle keine Aussage trifft, wurde die nicht normierte Differenz  $\Delta L$  ebenfalls als Merkmal getestet.

$$\Delta L = L_{MT2009} - L_{DNB} \quad (3.3)$$

Hierbei besteht jedoch der Nachteil, dass die Differenz ohne Normierung von den Absolutwerten von  $L_{MT2009}$  und  $L_{DNB}$  beeinflusst wird. Das heißt zum Beispiel, dass unter der Annahme vom Mond als einzige Beleuchtungsquelle und einer Wolke mit 50% Reflektanz die Differenz  $\Delta L$  bei doppeltem  $L_{MT2009}$  auch doppelt so groß wird, da sich  $L_{DNB}$  ebenfalls verdoppelt. Die reine Betrachtung der Differenz trifft also keine eindeutige Aussage über den Anteil der empfangenen Strahldichte an der Mondbeleuchtung, ist aber in der Lage, grobe Tendenzen anzugeben.

Zur Beschreibung der Reflexion von Mondlicht und der Dämpfung des von der Erdoberfläche emittierten Lichts werden somit folgende Merkmale vorgeschlagen:

- $L_{DNB}$ : Beobachtete Strahldichte (Grauwert des Pixels)
- $\mu_{L_{DNB}}$ : Mittelwert der beobachteten Strahldichten in Pixelnachbarschaft
- $\Delta L$ : Differenz der Strahldichte des Mondes und der Beobachtung
- $\Delta L_{norm}$ : Normierte Differenz der Strahldichte des Mondes und der Beobachtung

**Streuung** Der Effekt der Streuung durch eine Wolke beeinflusst, wie in Kapitel 2.2.4 bereits erläutert, den Kontrast und die Textur des Bildes, welche durch verschiedene nachbarschaftsbasierte Merkmale beschrieben werden können [25].

Das erste Merkmal für den Kontrast beziehungsweise die Homogenität eines Bildausschnittes ist die Varianz der Grauwerte  $\sigma_{L_{DNB}}^2$ , die die quadratische Abweichung der beobachteten Strahldichten von deren Mittelwert in einem Fenster um das betrachtete Pixel herum beschreibt (zur Wahl der Fenstergröße siehe Kapitel 3.3.4).

$$\sigma_{L_{DNB}}^2 = \frac{1}{N_{pix}} \sum_{i=1}^{N_{pix}} (L_{DNB,i} - \mu_{L_{DNB}})^2 \quad \text{mit} \quad N_{pix} = \text{Anzahl Pixel im Fenster} \quad (3.4)$$

Umso kleiner die Varianz, umso homogener sind die vorkommenden Grauwerte. Eine sehr kleine Varianz tritt bei VNIR-Nachtaufnahmen mit wolkenfreiem Himmel zum Beispiel bei Wasserflächen auf, eine sehr hohe in künstlich beleuchteten urbanen Gebieten. Genauso gegensätzlich, wie die Varianzen verschiedener Landbedeckungen in den Bildern sind, ist auch der jeweilige Einfluss von Wolken auf diese. Bei homogenen Flächen wie z.B. Wasser und vorhandener Mondbeleuchtung führt eine Wolke zu einer Erhöhung der Varianz, da die Wolke das Mondlicht deutlich inhomogener streut als die Wasserfläche. Liegt eine sehr inhomogene Fläche (z.B. eine Stadt) vor, glättet eine Wolke deren Erscheinungsbild und vermindert somit die Varianz.

Um diesen Effekt von Wolken auf künstlich beleuchtete Szenen, die bei klarem Himmel von sehr starken Kanten geprägt sind, zu beschreiben, kann als weiteres Merkmal ein Kantenfilter verwendet werden. Dort, wo sehr starke Kanten auftreten, befindet sich sehr wahrscheinlich keine Wolke, da diese eine kontrastarme Textur besitzen und meist auch räumlich eher unscharf abgegrenzt sind. Es wurde ein abgewandelter Laplace-Filter der Fenstergröße  $3 \times 3$ , der Kanten in Form der zweiten Ableitung der Grauwerte an der Stelle des Pixels beschreibt, ausgewählt. Gleichung (3.5) stellt die verwendete Filtermaske dar.

$$Laplace = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.5)$$

Der hier verwendete Filter berücksichtigt im Gegensatz zum Standard-Laplace-Operator neben der horizontalen und vertikalen zweiten Ableitung auch die diagonalen Richtungen. Der Laplace-Operator wurde anstatt des ebenfalls oft üblichen Sobel-Filters (1. Ableitung) ausgewählt, da er auch für die Detektion von Blobs, also einzelne hellere oder dunklere Pixel, geeignet ist. Diese treten in Nachtaufnahmen der betrachteten Auflösung in Form von einzelnen stark beleuchteten Objekten wie Booten oder Ölplattformen auf. Als alternatives Merkmal wurde zudem der Betrag des Ergebnisses der Laplace-Filterung getestet.

Die Varianz im Bildausschnitt beschreibt zwar, wie stark die Unterschiede der Grauwerte sind, jedoch nicht deren Anordnung innerhalb des Bildausschnittes. Diese wird durch die Textur beschrieben. Die Textur von Wolken ist zwar relativ kontrastarm, aber trotzdem

### 3. Methodik

unruhig und erinnert in ihrem Erscheinungsbild an Rauschen. Die Textur der Erdoberfläche kann sehr homogen sein oder aber auch sehr kontrastreich und strukturiert, wie zum Beispiel bei beleuchteten Straßen (vgl. Beispiele in Kapitel 2.2.4). Wolken verändern in beiden Fällen die Textur, wie bereits bei der Vorstellung der Varianz als Merkmal erläutert wurde. Als weitere Merkmale kommen also Merkmale in Frage, die die Textur in der Nachbarschaft des Pixels beschreiben.

Sehr gängige Texturmerkmale sind die sogenannten Haralick-Merkmale [18]. Sie basieren auf der Annahme, dass die Textur durch die gegenseitigen räumlichen Beziehungen der verschiedenen Grauwerte im Bild oder Bildausschnitt beschrieben werden kann. Diese räumlichen Beziehungen werden durch so genannte *grey-tone spatial dependence matrices* (oft auch *grey level co-occurrence matrices*, kurz GLCM) beschrieben, aus denen anschließend verschiedene Texturmerkmale abgeleitet werden.

Eine GLCM-Matrix besteht aus einer Spalte und einer Zeile pro möglichem Grauwert, bei einem 8bit-Bild handelt es sich also um eine  $256 \times 256$  große Matrix. Abhängig von der betrachteten Richtung und Distanz beschreibt ein Matrix-Eintrag  $(i, j)$ , wie oft eine bestimmte Grauwertkombination  $i, j$  im Bildausschnitt vorkommt. Die folgende Darstellung verdeutlicht dieses Prinzip am Beispiel eines Bildes mit 4 möglichen Grauwerten und einer GLCM-Matrix mit Distanz 1 und Winkel  $90^\circ$ .

$$\text{Bild: } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 1 \\ 2 & 2 & 2 & 1 \end{bmatrix} \Rightarrow \text{GLCM}(\alpha = 90^\circ, d = 1) : \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Der erste Eintrag (Stelle (1,1)) der GLCM-Matrix gibt an, wie oft auf den Grauwert 1 der Grauwert 1 folgt, wenn jeweils von oben nach unten (Richtung= $90^\circ$ ) immer der nächste Grauwert (Distanz=1) betrachtet wird. Dies ist in diesem Bild drei Mal der Fall. Der zweite Eintrag (Stelle (1,2)) gibt an, wie oft auf den Grauwert 1 der Grauwert 2 folgt, was hier nur einmal der Fall ist. Nach diesem Prinzip kann die GLCM-Matrix für verschiedene Richtungen und Distanzen berechnet werden.

Für die Berechnung einer GLCM-Matrix müssen also, neben der Größe des betrachteten Bildausschnitts, die Distanz und die Richtung festgelegt werden. Je nach Textur ist es sinnvoll, mehrere Distanzen und mehrere Richtungen pro Distanz zu betrachten, wobei die Distanz an die Größe der zu beschreibenden Textur angepasst wird. Zur Betrachtung der Wolkentextur wird in dieser Arbeit lediglich die Distanz 1 verwendet, da die Verschiedenheit jedes einzelnen Pixels gerade die Wolkentextur ausmacht. Bezüglich der Wahl der Richtungen schlägt Haralick vor, pro Distanz den Mittelwert und die Varianz des aus dem für verschiedene Richtungen bestimmten Merkmals zu verwenden, um das Texturmerkmal rotationsunabhängig zu machen. Dies ist im Fall der Wolkentextur nicht nötig, da es sich ohnehin um sehr unstrukturierte, richtungsunabhängige Texturen handelt. Um aber die in urbanen Gebieten vorkommenden linearen Strukturen (z.B. beleuchtete Straßen) auf jeden Fall zu erfassen, wird trotzdem der Mittelwert aus den Richtungen  $\alpha_1 = 0^\circ$  und  $\alpha_2 = 90^\circ$  (d.h. horizontal und vertikal) genutzt, um die Merkmale zusammenzufassen und damit kompakter zu machen.

Der in dieser Arbeit zur Berechnung der GLCM-Matrix verwendete Bildausschnitt ist ein  $5 \times 5$  großes Fenster mit dem betrachteten Pixel im Zentrum. Diese Fenstergröße wurde gewählt, da sie schon ausreichend Bildinformation zur Beschreibung der Textur enthält, und trotzdem noch klein genug ist, um nicht zu stark von angrenzenden anderen Texturen beeinflusst zu werden.

Im Falle der DNB-Bilder kann die GLCM-Matrix nicht direkt aus den Bildern berechnet werden. Die Voraussetzung für die Berechnung der GLCM-Matrix ist schließlich, dass die Grauwerte des Bildes in quantisierter Form vorliegen, was für die verwendeten DNB-Bilder nicht der Fall ist. Deshalb wird vor der Berechnung der Haralick-Features die bereits in Kapitel 2.1.1 beschriebene Quantisierung der Bilder in ein 8-bit Format durchgeführt. Die Anzahl der betrachteten Grauwerte (levels) ist also 256. Zu beachten ist, dass hierbei auch Texturinformation verloren geht.

In den GLCM-Matrizen sind nach Haralick alle Texturinformationen des Bildes enthalten, weshalb aus ihnen verschiedene Texturmerkmale abgeleitet werden können. Für diese Arbeit wurden die Texturmerkmale *contrast*, *dissimilarity* und *energy* ausgewählt, da sie am besten die Veränderung der Textur durch Wolken beschreiben können. In anfänglichen Tests wurden auch die drei weiteren gängigen, teilweise sehr ähnlichen Merkmale *homogeneity*, *entropy* und *correlation* getestet, diese erwiesen sich aber im Vergleich zu den ausgewählten Merkmalen als weniger geeignet und werden deshalb nicht weiter erläutert. Die folgende Darstellung der Berechnung der Merkmale folgt der verwendeten Implementierung im Python-Modul *scikit-image* [51].

1. *contrast*: Das Merkmal *contrast* ist ein Maß für den Kontrast der Textur beziehungsweise für die Stärke der Grauwertvariation. Es entspricht der Summe der einzelnen GLCM-Einträge, jeweils gewichtet mit der Differenz der Grauwerte. Dadurch fallen Einträge für große Grauwertdifferenzen besonders stark ins Gewicht. Der Wert für den Kontrast wird also hoch, wenn oft sehr unterschiedliche Grauwerte aufeinander folgen.

$$contrast = \sum_{i,j=0}^{levels-1} GLCM_{\alpha,d}(i,j)(i-j)^2 \quad \text{mit } (i,j) = \text{Grauwerte} \quad (3.6)$$

2. *dissimilarity*: Die *dissimilarity* ist ebenfalls ein Maß für den Kontrast, bzw. übersetzt die „Unähnlichkeit“ der vorhandenen Grauwerte. Sie unterscheidet sich vom *contrast* lediglich durch die lineare statt quadratische Gewichtung bei der Aufsummierung.

$$dissimilarity = \sum_{i,j=0}^{levels-1} GLCM_{\alpha,d}(i,j)|i-j| \quad (3.7)$$

3. *energy*: Das Merkmal *energy* ist ein Maß für die Homogenität der Textur. Ist die Energie klein, gibt es im Bildausschnitt viele verschiedene Grauwertkombinationen. Die GLCM-Matrix hat in diesem Fall viele kleine Einträge, das Bild ist also inhomogen.

### 3. Methodik

Bei einer großen Energie kommen im Bild nur sehr wenige Grauwertkombinationen vor, die dafür aber oft auftreten, das Bild ist also homogen.

$$energy = \sum_{i,j=0}^{levels-1} (GLCM_{\alpha,d}(i,j))^2 \quad (3.8)$$

Zur Beschreibung der Effekte der Streuung durch Wolken werden insgesamt folgende Merkmale vorgeschlagen:

- $\sigma_{L_{DNB}}^2$ : Varianz in Pixelnachbarschaft
- Laplace-Filter
- Betrag des Ergebnisses der Laplace-Filterung
- *contrast*: Haralick-Merkmal für den Kontrast der Textur
- *dissimilarity*: Haralick-Merkmal für die Ähnlichkeit der Grauwerte der Textur
- *energy*: Haralick-Merkmal für die Homogenität der Textur

Die soeben vorgestellten Merkmale sind zwar Maße für die durch Wolken beeinflussten Eigenschaften von VNIR-Nachtaufnahmen, sie sind aber nicht in der Lage, die teilweise gegensätzlichen Einflüsse auf diese Eigenschaften, also die Veränderung selbst, zu beschreiben. Ein Beispiel ist das bereits betrachtete Verhalten der Varianz unter Einfluss von Wolken. Diese wird im Vergleich zu wolkenfreier Sicht je nach Mondbeleuchtung und von der Erdoberfläche emittierter Beleuchtung entweder verringert oder vergrößert. Um diese Veränderung ebenfalls in Form von Merkmalen ausdrücken zu können, werden im folgenden Abschnitt Merkmale vorgestellt, die die zeitlichen Unterschiede der Beobachtungen beschreiben.

#### 3.2.2. Bildserienmerkmale

Ein einzeln betrachtetes panchromatisches Bild liefert nur begrenzt viele Informationen, die durch die zuvor vorgestellten Merkmale bereits recht gut ausgeschöpft werden. Deshalb liegt die Idee nahe, die zeitliche Komponente als weitere Informationsquelle zu nutzen und die durch Wolken verursachte Veränderungen zu betrachten.

Das Ziel ist also, Maße für die zeitlichen Veränderungen des Bildes oder Bildausschnittes beziehungsweise der daraus abgeleiteten Einzelbildmerkmale zu bestimmen. Die Herausforderung hierbei ist die Wahl der für den Vergleich verwendeten Referenz. Da die Wolkenbedeckung der vorangegangenen Bilder ebenfalls als unbekannt angenommen wird, steht kein wolkenfreies Referenzbild zur Verfügung, anhand dessen die Veränderungen der Beobachtungen bestimmt werden könnten. Eine solche Referenz ist jedoch nötig, um die Beobachtungen eines einzelnen Bildes eindeutig interpretieren zu können, da verschiedene Situationen auf der Erdoberfläche verschieden und oft gegensätzlich durch Bewölkung

beeinflusst werden. Es kann zum Beispiel keine klare Aussage getroffen werden, ob ein beobachteter Grauwert von einer schwach beleuchteten Stadt bei wolkenfreier Sicht, einer von einer Wolke überdeckten sehr hell beleuchteten Stadt oder von einer stark reflektierenden Wolke über Gewässer verursacht wird. Liegt eine wolkenfreie Referenz vor, ist die Ausgangssituation bekannt und es kann anhand der Differenz zu dieser klar bestimmt werden, ob eine Wolke vorhanden ist.

In der Literatur finden sich zwei Ansätze, wie eine solche Referenz bestimmt werden kann. Jedlovec [23] setzt wolkenfreie künstliche Bilder aus den Pixeln vorheriger Bilder zusammen, die am wahrscheinlichsten wolkenfrei sind. Dies sind jene Pixel, die innerhalb der betrachteten Zeitreihe den Extremwert eines Merkmals besitzen, für das eine klare Aussage getroffen werden kann, dass es bei klarem Himmel maximal wird. Hierfür verwendet Jedlovec die Beobachtung des Thermalkanals, da die Temperatur von Wolken meist geringer ist als die der Erdoberfläche. Für den wärmsten Wert der Zeitreihe für ein Pixel liegen also sehr wahrscheinlich wolkenfreie Bedingungen vor, weshalb dieser als Referenzwert für dieses Pixel verwendet wird.

Lyapustin et. al [33] basieren ihre Methode zur Bestimmung der Referenz auf der Annahme, dass sich das Erscheinungsbild der Erdoberfläche bei wolkenfreien Bedingungen nur langsam verändert, wohingegen Wolken dieses Erscheinungsbild in sehr kurzen Zeiträumen stark verändern. Dieses Prinzip der durch Bewölkung erhöhten temporalen Varianz von Bildern wird auch in weiteren Arbeiten zur Wolkenerkennung verwendet ([11],[44]). Zur Quantifizierung der Veränderung wird die Korrelation zweier zeitlich versetzter Bildausschnitte berechnet. Sind zwei Bildausschnitte stark korreliert, sind sehr wahrscheinlich beide wolkenfrei, da eine Bewölkung in einem oder beiden Bildausschnitten das Bild immer verändern und somit die Korrelation verringern würde. Auf Basis dieses Prinzips bestimmen Lyapustin et al. dynamisch die beiden am stärksten korrelierten Bildausschnitte in der nahen Vergangenheit des aktuellen Bildes und verwenden diese als Referenz.

In dieser Arbeit werden diese beiden Ansätze auf die Anwendung auf panchromatische nächtliche Satellitenbilder übertragen.

Für den Ansatz von [23] ist dies nicht direkt möglich, da es kein Merkmal gibt, welches eine klare Tendenz in Richtung eines Extremwertes für wolkenfreie Bedingungen besitzt. Dies wird durch die bereits beschriebene Problematik der gegensätzlichen Einflüsse von Wolken in Abhängigkeit der Landbedeckung und Mondbeleuchtung verursacht. Auf Basis der Überlegungen von [23] wurde deshalb die Idee entwickelt, als Referenzwert den Mittelwert eines Merkmals in der nahen Vergangenheit (maximal 30 Bilder, also mindestens ein kompletter Mondzyklus) zu verwenden. Hierzu wird die Annahme getroffen, dass in diesem Zeitraum jedes betrachtete Pixel Teil beider Klassen war. Ist dies der Fall, dann stellt die Differenz des Merkmals und des entsprechenden Mittelwertes in den vergangenen Bildern zwar immer noch keine eindeutige Aussage über die wahrscheinlichste Klassenzugehörigkeit dar, in Kombination mit dem Absolutwert des Merkmals kann aber eine eindeutigere Aussage getroffen werden.

Beim Beispiel der Varianz wäre im Fall einer Wolke über einer homogenen nicht selbst leuchtenden Erdoberfläche und vorliegender Mondbeleuchtung der Absolutwert deutlich

### 3. Methodik

größer als der Mittelwert, da die Wolke die Varianz erhöht (wenn man davon ausgeht, dass die Fläche auch in wolkenfreiem Zustand beobachtet wurde). Eine solche große Differenz kann aber auch auftreten, wenn ein wolkenfreier Bildausschnitt mit durch künstliche Beleuchtung verursachter großer Varianz mit dem Mittelwert der Vergangenheit verglichen wird, da hier die in der Vergangenheit vorhandene Bewölkung die Varianz verringert. Als alleinstehendes Merkmal ist die Differenz zum Mittelwert also nur bedingt geeignet, in Kombination mit dem Absolutwert kann sie aber durchaus einen Informationsgewinn bieten.

Berechnet wird die Differenz zum Mittelwert der vergangenen Merkmale für alle zuvor vorgestellten Einzelbildmerkmale, wodurch die Möglichkeit besteht, jedes Merkmal durch eine zeitliche Komponente zu stützen.

Der zweite von [33] entwickelte Ansatz auf Basis der Korrelation ist direkt auf die vorliegende Problemstellung übertragbar und liefert somit im Gegensatz zum zuvor vorgestellten Ansatz auch im Falle von VNIR-Nachtaufnahmen wahrscheinlich wolkenfreie Referenzen. Die Differenz des Merkmals für das betrachtete Pixel zum Merkmal für die beiden Referenzbilder trifft also auch ohne Einbezug des Absolutwertes eine Aussage zur Wahrscheinlichkeit von Bewölkung. Umso größer die Differenz zur Referenz, umso wahrscheinlicher liegt eine Bewölkung vor.

Die Bestimmung der beiden Referenzbilder erfolgt in dieser Arbeit durch die Berechnung des Pearson-Korrelationskoeffizienten für maximal 30 der Vorgängerbilder des betrachteten Bildes. Die Korrelation wird anhand derjenigen Fenstergröße berechnet, die auch zur Berechnung des Merkmals verwendet wurde. Die beiden am stärksten korrelierten Bildausschnitte werden als Referenz für den aktuellen Bildausschnitt verwendet. Wie beim vorherigen Ansatz können dann die Differenzen aller Einzelbildmerkmale zu diesen Referenzbildern berechnet und als neue Merkmale verwendet werden. Da in diesem Fall zwei Referenzbilder vorliegen, wird der Mittelwert der Differenzen verwendet.

Die Annahme, die letzterem Ansatz zu Grunde liegt, kann auch direkt in ein zusätzliches Merkmal übersetzt werden. Wenn die Korrelation zweier Bildausschnitte maximal wird, wenn beide frei von Wolken sind, ist die maximale Korrelation des betrachteten Bildausschnittes mit seinen Vorgängern ein Maß für die Wahrscheinlichkeit, dass sowohl der aktuelle, als auch der vergangene Bildausschnitt wolkenfrei sind. Die maximale Korrelation kann also direkt als Merkmal verwendet werden. Sie trifft lediglich in dem Fall, dass der betrachtete Bildausschnitt in der betrachteten Zeitspanne nie wolkenfrei war, keine gute Aussage. Dann ist die maximale Korrelation nämlich unabhängig davon klein, ob im betrachteten Bild Bewölkung vorliegt oder nicht.

Zur Berechnung des Merkmals der maximalen Korrelation wurden die Korrelationskoeffizienten des betrachteten Fensters (zur Auswahl der Fenstergröße siehe Kapitel 3.3.4) mit den entsprechenden Fenstern der letzten Bilder (maximal 30) bestimmt und der Maximalwert als Merkmal verwendet.

### 3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken

Zur Beschreibung des zeitlichen Aspekts werden also drei Typen von Merkmalen vorgeschlagen:

- $\Delta_{mean}$ : Differenz des Mittelwertes eines Merkmals  $\bar{f}_t$  für die vorherigen Bilder (max. 30 Stück) zum Merkmal für den betrachteten Zeitpunkt  $f_t$ .

$$\Delta_{mean}(f_t) = f_t - \bar{f}_t \quad (3.9)$$

- $\Delta_{ref}$ : Mittlere Differenz des Merkmals  $f$  für die beiden am wahrscheinlichsten wolkenfreien Referenzbilder  $f_{ref_1}$  und  $f_{ref_2}$  zum betrachteten Merkmal  $f_t$

$$\Delta_{ref}(f_t) = \frac{f_t - f_{ref_1} + f_t - f_{ref_2}}{2} \quad (3.10)$$

- $Max_{corr}$ : Maximale Korrelation des betrachteten Bildausschnitts  $B_t$  mit den vorherigen Bildausschnitten  $B_{t-j}$  (max. 30 Stück betrachtet).

$$\rho_{max}(B_t) = \max \left( \rho(B_t, B_{t-x}) = \frac{Cov(B_t, B_{t-x})}{\sigma(B_t)\sigma(B_{t-x})} \right) \quad \text{für } x = [1, 30] \quad (3.11)$$

Voraussetzung für diese zeitlichen Betrachtungen ist generell eine ausreichend genaue Georeferenzierung und die Abbildung auf das immer gleiche Koordinatenraster, sodass sich das in der Zeitreihe betrachtete Pixel immer auf die gleiche Fläche auf der Erde bezieht. Sind diese Aspekte fehlerhaft, treffen auch die hergeleiteten Merkmale keine genauen Aussagen.

### 3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken

Nachdem alle Grundlagen des entwickelten Verfahrens erläutert wurden, kann nun der Algorithmus zur semantischen Segmentierung von Wolken in nächtlichen VNIR-Satellitenaufnahmen definiert werden.

Die in diesem Abschnitt erläuterten Betrachtungen wurden zur Beschränkung des Rechenaufwands nur an zwei der in Kapitel 3.1.1 definierten Versuchsgebiete durchgeführt, nämlich *muenchen* als Trainingsgebiet und *stuttgart* als Testgebiet (zur zusätzlichen Validierung der Ergebnisse wurden Trainings- und Testdatensatz zudem vertauscht). Es wurden urbane Gebiete gewählt, da diese im Fokus der Betrachtungen stehen (vgl. Kapitel 1). *muenchen* und *stuttgart* enthalten hierbei als global betrachtet durchschnittliche, mittelgroße urbane Gebiete sowohl sehr dicht besiedelte und stark beleuchtete Regionen als auch weniger stark beleuchtetes Umland und repräsentieren somit eine gute Schnittmenge verschiedener Ausprägungen von Besiedelung. Hinsichtlich des zeitlichen Aspekts wurden pro Gebiet 100 Bilder (ca. Januar bis März 2018) verwendet, sodass mehrere volle Mondzyklen enthalten sind und die Anzahl der Bilder trotzdem klein genug ist, um die Rechenzeit nicht zu stark zu erhöhen.

### 3. Methodik

Im Folgenden wird das auf dieser Basis entwickelte Verfahren vorgestellt. Hierfür wird zunächst definiert, in welcher Hinsicht das Ergebnis der semantischen Segmentierung optimiert werden soll. Anschließend wird die Wahl des Klassifikators erläutert und die drei im Entwicklungsprozess variierten Stellgrößen Merkmalsvektor, Größe der zur Merkmalsberechnung verwendeten Nachbarschaft und Art der Referenz betrachtet.

#### 3.3.1. Kriterium zur Optimierung des Verfahrens

Um die Merkmale und die Hyperparameter des Klassifikationsverfahrens zu wählen, muss zunächst definiert werden, was als optimales Klassifikations-Ergebnis angestrebt wird. Bei Wolkenerkennungsalgorithmen gibt es hierbei zwei Strategien.

Die erste Strategie ist der sogenannte *clear-sky conservative* Ansatz, der zum Beispiel beim Vorgänger der ECM-Wolkenmaske VCM [15] gewählt wurde (vgl. Kapitel 1). Hierbei ist das Ziel, eine möglichst reine Klasse *klar* zu generieren, beziehungsweise Wolken möglichst vollständig zu erkennen. Das heißt, ein Pixel wird als wolzig klassifiziert, sobald nur einer der durchgeführten Tests es mit einer hohen Sicherheit als wolzig klassifiziert. Diese Strategie entspricht also einer Optimierung des Recalls der Klasse *wolzig* und der Precision der Klasse *klar* (zur Definition von Recall und Precision siehe Tabelle 4.2 auf Seite 54). Dieser Ansatz bietet sich an, wenn das Ziel der Wolkenerkennung die Erzeugung einer Maske für klaren Himmel ist, die für Analysen auf den wolkenfreien Bildbereichen genutzt werden soll. Hier würde jede fälschlicherweise in der wolkenfreien Klasse verbliebene Wolke die Analysen negativ beeinflussen.

Die zweite Strategie ist der *cloud-conservative* Ansatz, der beim ECM-Algorithmus gewählt wurde (vgl. Kapitel 2.1.2). Hier ist das Ziel, eine möglichst korrekte Wolkenmaske zu generieren, das heißt Fehlzuordnungen zur Klasse *wolzig* zu vermeiden, beziehungsweise die Precision dieser Klasse zu maximieren. Dies kann unter Umständen mit einem schlechteren Recall der Klasse *wolzig*, also mit einer weniger vollständigen Erkennung der Wolken, einhergehen. In diesem Ansatz wird der Fokus also mehr auf die Wolkenmaske, als auf die wolkenfreie Maske gelegt.

Für die Entwicklung des Wolkenerkennungsalgorithmus in dieser Arbeit wurde die Strategie von ECM gewählt, da das Ergebnis mit der Maske von ECM verglichen werden soll und somit die gleiche Anforderung zu Grunde gelegt wird. Außerdem soll in dieser Arbeit zunächst allgemein das Potenzial der Wolkenerkennung evaluiert werden, sodass das Hauptaugenmerk auf die Wolken-Klasse gelegt werden kann. Sollte die resultierende Maske später zum Beispiel für Analysen auf klaren Pixeln verwendet werden, kann durch eine Anpassung der Parameter das Verhalten des Klassifikators entsprechend hin zu einem größeren Recall der Klasse *wolzig* verändert werden.

Das Optimierungskriterium für die Entwicklung des Algorithmus ist somit die Fokussierung auf die Erzeugung einer möglichst korrekten Wolkenmaske, das heißt die Maximierung der Precision der Klasse *wolzig* und generell eine höhere Gewichtung der Ergebnisse der Klasse *wolzig* gegenüber den Ergebnissen der Klasse *klar*.

#### 3.3.2. Klassifikationsverfahren

Zur semantischen Segmentierung der Bilder in die beiden Klassen *wolkig* und *klar* wird ein Klassifikator benötigt, der jedes einzelne Pixel einer dieser Klassen zuordnet. Die anfangs vorgestellten gängigen Wolkenerkennungs-Algorithmen nutzen zur Klassifikation größtenteils Schwellwert-Tests, die unter Berücksichtigung zahlreicher Fallunterscheidungen und zusätzlicher Daten miteinander verknüpft werden. Die verwendeten Schwellwerte basieren meist auf einer detaillierten Modellierung der physikalischen Sachverhalte. Der Nachteil solcher Verfahren ist, dass Fehler im Modell oder den Zusatzdaten direkt zu Fehlern in der Klassifikation führen (vgl. hierzu Kapitel 2.1.2).

Deshalb soll in dieser Arbeit ein Verfahren verwendet werden, dem das theoretische Verständnis des Beobachtungsmodells zwar zu Grunde liegt, das jedoch nicht allein von der Modellierung abhängig ist. Aus diesem Grund wird ein merkmalsbasiertes Klassifikationsverfahren gewählt. Ein solches Verfahren basiert zwar auf Merkmalen, die auf Basis des Modellwissens entwickelt wurden, die Entscheidungsgrenzen selbst müssen aber nicht modelliert werden. Da Trainingsdaten zur Verfügung stehen, kann ein überwachtes Klassifikationsverfahren verwendet werden.

Dass auf Grundlage dieser Überlegungen Random Forests [8] (vgl. Kapitel 2.3) als Klassifikator ausgewählt wurde, hat mehrere Gründe. Zum einen basiert das Verfahren auf dem gleichen Ansatz der hierarchischen Unterteilung des Merkmalsraumes, wie es auch die gängigen Verfahren mit ihren Schwellwert-Tests tun und kann somit als eine Weiterentwicklung oder Abwandlung dieser Verfahren verstanden werden. Viel entscheidender sind aber die charakteristischen Vorteile von Random Forests, die auch schon in Kapitel 2.3 erläutert wurden.

Da das Training anhand der ECM-Daten erfolgt, die selbst gewisse Ungenauigkeiten enthalten, ist es wichtig, dass das Verfahren robust gegenüber Fehlern beziehungsweise Rauschen in den Trainingsdaten ist. Dies ist bei Random Forests aufgrund des Prinzips des Bagging der Fall. Aus praktischer Sicht ist es vorteilhaft, dass mit einem Random Forest sowohl binäre als auch Mehrklassenprobleme direkt gelöst werden können und er auch zur Regression kontinuierlicher Werte eingesetzt werden kann. Dies ist in diesem Fall relevant, da die ECM-Wolkenmaske neben den Versionen mit zwei und vier Klassen auch in Form von Wolkenwahrscheinlichkeiten zur Verfügung steht und somit getestet werden kann, welche dieser Repräsentationen sich am besten zum Training eignet. Ein weiterer praktischer Vorteil ist die gute Parallelisierbarkeit des Trainings und der Klassifikation mit einem Random Forest, was auch die Verarbeitung großer Datenmengen effizient ermöglicht. Dies ist für ein potenziell global angewandtes Verfahren zwingend erforderlich.

**Wahl der Hyperparameter** Wie in Kapitel 2.3 bereits erläutert, müssen vor dem Training eines Random Forests verschiedene Hyperparameter festgelegt werden, die den Aufbau des Waldes und die Abbruchbedingungen festlegen und somit die Performance des Klassifikators beeinflussen. Zunächst wurden diese mit einer Random Search [4] (Testen mehrerer Parameterkombinationen, die zufällig aus vorgegebenen Wertebereichen bestimmt werden)

### 3. Methodik

grob bestimmt und iterativ nach Veränderungen der im Folgenden beschriebenen Aspekte mit einer Grid Search (Testen aller möglichen Kombinationen aus der Menge an Werten, die für jeden Hyperparameter vorgegeben werden) verfeinert. Genau wie zur Implementierung des Random Forest selbst wurde hierfür auf das Python-Modul *scikit-learn* [40] zurückgegriffen.

Das der Hyperparameterwahl zu Grunde gelegte Optimierungskriterium wurde im vorherigen Abschnitt 3.3.1 erläutert. Die auf dieser Basis ermittelten am besten geeigneten Parameter sind:

- Anzahl der Bäume: 70
- Maximale Tiefe: Keine
- Minimale Anzahl an Datenpunkten pro Blatt: 100
- Minimale Anzahl an für die Teilung eines Knotens erforderlichen Datenpunkten: 2
- Maximal betrachtete Anzahl an Merkmalen pro Test:  $\sqrt{\text{Anzahl Merkmale}}$
- Kriterium für den Informationsgewinn: Gini-Index

Das hauptsächliche Abbruchkriterium neben der vollständigen Reinheit eines Knotens ist also die minimale Anzahl an Datenpunkten pro Blatt. So wird die Klassenverteilung pro Blatt immer anhand von mindestens 100 Datenpunkten bestimmt, was zur Generalisierung des Verfahrens beiträgt. Eine weitere Generalisierung könnte durch die Begrenzung der Tiefe der Bäume erreicht werden, dies führt aber im betrachteten Fall zur Verschlechterung der Ergebnisse. Die genannten Werte für die Hyperparameter werden für alle weiteren Betrachtungen verwendet.

#### 3.3.3. Definition des Merkmalsvektors – Bestimmung der aussagekräftigsten Merkmalskombination

Die im Abschnitt 3.2 definierten Merkmale bilden die Ausgangsmenge, aus der diejenige Merkmalskombination ausgewählt werden soll, die unter Verwendung des soeben definierten Random Forest die besten Klassifikationsergebnisse liefert. Um den Suchraum einzuschränken, wird zunächst nur die beste Kombination der Einzelbildmerkmale bestimmt, die dann in einem zweiten Schritt durch die Bildserienmerkmale ergänzt oder ersetzt wird.

**Beste Kombination – Einzelbildmerkmale** Von den Einzelbildmerkmalen wurden zunächst diejenigen Merkmale, die einen ähnlichen Effekt beschreiben und sehr stark korrelieren, zu Gruppen zusammengefasst, was den Suchraum weiter einschränkt. Hierfür wurden, neben den Überlegungen zur Definition der Merkmale, die anhand des Trainingsdatensatzes berechneten Korrelationskoeffizienten zwischen allen Merkmalen berücksichtigt.

### 3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken

Es ergeben sich damit folgende sechs Gruppen:

1.  $L_{DNB}$ ,  $\Delta L$ ,  $\Delta L_{norm}$
2. Varianz  $\sigma_{L_{DNB}}^2$
3. Mittelwert  $\mu_{L_{DNB}}$
4. Laplace, Betrag(Laplace)
5. *contrast*, *dissimilarity*
6. *energy*

Zur Bestimmung der besten Kombination aus Einzelbildmerkmalen wurden nun in mehreren Durchgängen fast alle möglichen Kombinationen der Merkmale als Merkmalsvektoren für die Klassifikation mit dem zuvor definierten Random Forest getestet. Hierbei wurde aus jeder Gruppe immer nur maximal ein Merkmal gleichzeitig verwendet und in verschiedenen Durchgängen wurde getestet, einzelne Gruppen komplett wegzulassen. Um zu bestimmen, bei welchen Gruppen dies sinnvoll ist, wurden die beim Training des Random Forests ebenfalls berechneten Merkmals-Wichtigkeiten (vgl. Kapitel 2.3) berücksichtigt. Insgesamt gehören hierbei  $\Delta L_{norm}$  und *contrast* immer zu den wichtigsten Merkmalen.

Aufgrund dieser Untersuchungen haben sich  $\Delta L_{norm}$  und der Mittelwert  $\mu_{L_{DNB}}$  zur Beschreibung der Albedo und optischen Dicke und die Varianz  $\sigma_{L_{DNB}}^2$ , der Laplace-Filter, *contrast* und *energy* zur Beschreibung der Effekte der Streuung als die beste Kombination der Einzelbildmerkmale herausgestellt. Der Merkmalsvektor  $M_{Einzelbild}$  für die ausschließliche Betrachtung von Einzelbildmerkmalen ist somit folgendermaßen aufgebaut:

$$\Rightarrow M_{Einzelbild} = [\mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, Laplace, \Delta L_{norm}, contrast, energy] \quad (3.12)$$

**Beste Kombination – Einzelbild- und Bildserienmerkmale** Ausgehend von dieser Kombination von Einzelbildmerkmalen wurde anschließend die beste Merkmalskombination unter Hinzunahme der Bildserienmerkmale bestimmt. Getestet wurde hierfür zum einen die Ergänzung der Merkmale der Einzelbildbetrachtung  $f$  durch eines oder mehrere ihrer zeitlichen Äquivalente ( $\Delta_{mean}(f)$ ,  $\Delta_{ref}(f)$ ) und zum anderen das Ersetzen von einem oder mehreren Einzelbildmerkmalen durch ihr zeitliches Äquivalent. Hierbei wurden diejenigen Merkmale  $\Delta_{mean}(f)$ ,  $\Delta_{ref}(f)$ , die sich auf das gleiche Merkmal  $f$  beziehen, in Gruppen zusammengefasst, damit pro Kombination immer nur maximal eines der beiden verwendet wird. Das nicht direkt zu einem Einzelbildmerkmal gehörende Merkmal  $Max_{corr}$  wurde bei beiden Versionen als optionales Merkmal mit betrachtet.

Die Ergebnisse dieser Versuche identifizieren zwei Bildserienmerkmale, die eine signifikante Verbesserung der Klassifikationsergebnisse bewirken. Dies ist zum einen  $\Delta_{mean}(\sigma_{L_{DNB}}^2)$  als einziges der zeitlichen Äquivalente. Es funktioniert sowohl als Ersatz zu  $\sigma_{L_{DNB}}^2$ , als auch als dessen Ergänzung sehr gut. Dass gerade die zeitliche Betrachtung der Varianz gut funktioniert, entspricht der Annahme, die bei der Motivation des Merkmals  $\Delta_{mean}(f)$  in

### 3. Methodik

Kapitel 3.2.2 am Beispiel der Varianz erläutert wurde. Das zweite Merkmal ist die maximale Korrelation  $Max_{corr}$ .

Abgesehen von diesen beiden Merkmalen waren unter den besten der getesteten Merkmalskombinationen auch solche, die zusätzlich z.B.  $\Delta_{ref}(energy)$ ,  $\Delta_{mean}(Laplace)$  oder  $\Delta_{mean}(\bar{L}_{DNB})$  enthalten. Da diese Merkmale aber keine signifikanten Verbesserungen verursachen und ihr Weglassen das Ergebnis kaum verschlechtert, wurden lediglich die zwei anfangs genannten Bildserienmerkmale ausgewählt.

Der Merkmalsvektor  $M_{Bildserie}$ , der die Bildserienmerkmale mit einbezieht, ist somit folgendermaßen aufgebaut:

$$\Rightarrow M_{Bildserie} = [\mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, \Delta_{mean}(\sigma_{L_{DNB}}^2), Laplace, \Delta L_{norm}, contrast, energy, Max_{corr}] \quad (3.13)$$

Dieser Merkmalsvektor wird für alle in Kapitel 4 beschriebenen Analysen des entwickelten Verfahrens verwendet.

**Signifikanz der einzelnen Merkmale der besten Merkmalskombination** Um die Signifikanz der einzelnen Merkmale, aus denen die beste Merkmalskombination  $M_{Bildserie}$  besteht, zu beurteilen, wurde analysiert, wie sich die Ergebnisse verändern, wenn nur Teilmengen dieser Kombination verwendet werden. Hierfür wurden zunächst alle möglichen Teilmengen getestet, die ein, zwei oder drei Merkmale enthalten. Dabei wurden eindeutig die Merkmale *contrast* und  $\Delta L_{norm}$  als die beiden aussagekräftigsten Merkmale identifiziert, da sie auch schon alleinstehend beziehungsweise in Kombination mit wenigen weiteren Merkmalen gute Ergebnisse liefern. Darauf, dass diese beiden Merkmale eine hohe Aussagekraft haben, weisen auch die bereits erwähnten hohen Merkmals-Wichtigkeiten bei Verwendung des vollständigen Merkmalsvektors hin. Dies ist nicht überraschend, da diese beiden Merkmale die beiden Haupteffekte von Wolken am eindeutigsten beschreiben.

Inwiefern die weiteren ausgewählten Merkmale zu einer Verbesserung des Ergebnisses beitragen, zeigt Abbildung 3.13. Sie veranschaulicht die Entwicklung der Klassifikationsgenauigkeit, wenn ausgehend von den beiden wichtigsten Merkmalen sukzessiv das jeweils nächst-wichtige Merkmal zum Merkmalsvektor hinzugefügt wird. Abweichend hiervon werden die beiden Bildserienmerkmale zum Schluss hinzugefügt, um den Unterschied der beiden Varianten (Einzelbildmerkmale, Bildserienmerkmale) zu verdeutlichen. Es werden also immer größere Teilmengen des finalen Merkmalsvektors betrachtet, welche für die dargestellten Versionen 0-7 folgendermaßen aufgebaut sind:

- $M_0 = [contrast]$
- $M_1 = [contrast, \Delta L_{norm}]$
- $M_2 = [contrast, \Delta L_{norm}, \mu_{L_{DNB}}]$
- $M_3 = [contrast, \Delta L_{norm}, \mu_{L_{DNB}}, \sigma_{L_{DNB}}^2]$
- $M_4 = [contrast, \Delta L_{norm}, \mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, energy]$

### 3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken

- $M_5 = [\textit{contrast}, \Delta L_{norm}, \mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, \textit{energy}, \textit{Laplace}]$
- $M_6 = [\textit{contrast}, \Delta L_{norm}, \mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, \textit{energy}, \textit{Laplace}, \textit{Max}_{corr}]$
- $M_7 = [\textit{contrast}, \Delta L_{norm}, \mu_{L_{DNB}}, \sigma_{L_{DNB}}^2, \textit{energy}, \textit{Laplace}, \textit{Max}_{corr}, \Delta_{mean}(\sigma_{L_{DNB}}^2)]$

Betrachtet man ausschließlich die Einzelbildmerkmale (Versionen 0-5), fällt auf, dass die beiden Merkmale *energy* und *Laplace* im Vergleich zu den restlichen Merkmalen nur relativ kleine Verbesserungen bewirken. Gerade hinsichtlich des *Laplace*-Merkmals ist dies nicht verwunderlich, da es schließlich auf die Charakterisierung von beleuchteten urbanen Gebieten mit starken Kanten spezialisiert ist und dadurch vor allem an den Rändern solcher Gebiete greift, die nur einen kleinen Anteil des Bildinhalts ausmachen.

Betrachtet man die Bildserienmerkmale, zeigt sich, dass durch ihr Hinzufügen nur eine leichte Verbesserung der Ergebnisse erzielt wird und auch ohne die Berücksichtigung des zeitlichen Aspekts bereits gute Ergebnisse erreicht werden. Wenn das Klassifikationsverfahren besonders einfach und schnell berechenbar sein soll, könnte in dem getesteten Fall auf diese also ohne größere Verluste verzichtet werden.

Diese Beobachtung widerspricht der in Kapitel 3.2.2 formulierten Hypothese, dass die zeitliche Betrachtung eine große Verbesserung des Verfahrens verspricht. Auch die Tatsache,

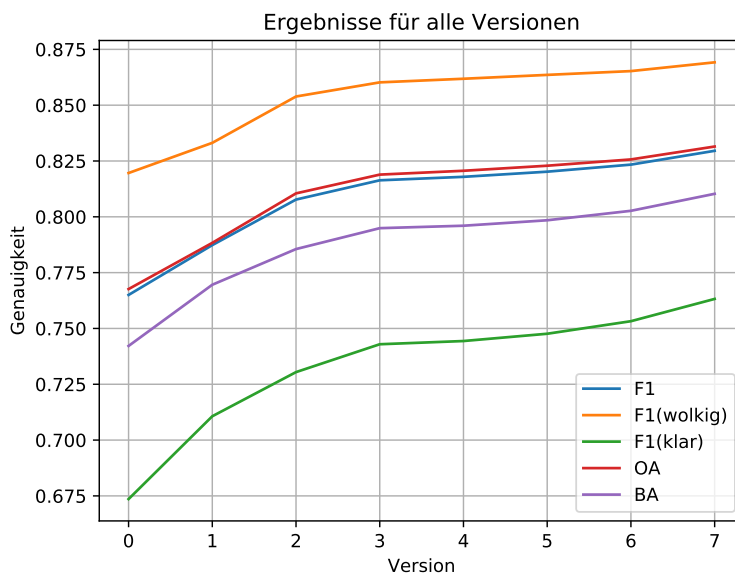


Abbildung 3.13: Vergleich der Ergebnisse der semantischen Segmentierung von Wolken für verschiedene Versionen des finalen Merkmalsvektors. Version 0 verwendet nur ein Merkmal, Version 7 den vollständigen Merkmalsvektor. Die jeweils verwendeten Merkmalsvektoren werden im Text definiert.

### 3. Methodik

dass bei der Merkmalsauswahl in der besten der getesteten Kombinationen aus Einzelbildmerkmalen und Bildserienmerkmalen nur zwei der Bildserienmerkmale verwendet wurden, scheint dieser These zu widersprechen. Dieses Verhalten ist in Anbetracht des theoretisch großen Potenzials solcher Merkmale überraschend und wird in Kapitel 4.2.3 auf Seite 68 genauer analysiert.

#### 3.3.4. Fenstergröße

Die Berechnung der nachbarschaftsbasierten Merkmale beruht auf der Betrachtung aller Nachbarpixel innerhalb eines Fensters um das betrachtete Pixel. Für die Merkmale Mittelwert, Varianz und maximale Korrelation muss die Größe dieses Fensters festgelegt werden, beim Laplace-Merkmal und den Haralick-Merkmalen wurde bereits eine feste Fenstergröße gewählt (vgl. Kapitel 3.2).

Je nach Größe des Fensters wird ein mehr oder weniger großes Gebiet mit in die Berechnung einbezogen. Die am besten geeignete Gebietsgröße ist abhängig von der Ausdehnung der gesuchten Strukturen. Im Falle der Wolkenerkennung in DNB-Daten wurden aufgrund der relativ großen Bodenpixelgröße die beiden kleinst möglichen Fenstergrößen  $3 \times 3$  und  $5 \times 5$  als die geeignetsten Optionen getestet. Bei einer Bodenpixelgröße von ungefähr 750 m entspricht dies bereits einer Kantenlänge der Nachbarschaft von 2250 m beziehungsweise 3750 m. Bei noch größeren Nachbarschaften ist das Risiko groß, dass benachbarte Strukturen das Ergebnis negativ beeinflussen.

Bei der Evaluierung der erzielbaren Ergebnisse unter Verwendung der anhand beider Fenstergrößen berechneten Nachbarschaftsmerkmale hat sich eindeutig die Fenstergröße  $5 \times 5$  als die geeignetere herausgestellt. Die in Abschnitt 3.3.3 bestimmte beste Merkmalskombination bezieht sich also für Mittelwert, Varianz und maximale Korrelation auf die Betrachtung einer  $5 \times 5$  Pixelnachbarschaft.

#### 3.3.5. Ground Truth

Neben der Bestimmung der idealen Hyperparameter für den Random Forest und der geeignetsten Merkmalskombination wurde zudem evaluiert, welche Form der Repräsentation der Ground Truth zu den besten Ergebnissen führt. Die originäre Repräsentation der ECM-Wolkenmaske ist die Darstellung in Form von kontinuierlichen Wolkenwahrscheinlichkeiten pro Pixel. Sie dient als Basis für die Ableitung der Zwei- und Vier-Klassen Darstellungen (vgl. Kapitel 2.1.2).

Die für das Ziel des Labelns eines Pixels als *wolkig* oder *klar* direkteste Vorgehensweise ist das Training anhand der Zwei-Klassen-Wolkenmaske, die genau diese Klassen repräsentiert. Da jedoch bei dieser Repräsentation der Wolkenmaske auch Pixel, bei denen der ECM-Algorithmus ein sehr unsicheres Ergebnis erzielt, eindeutig einer der beiden Klassen zugeordnet werden, befinden sich in den Trainingsdaten der beiden Klassen auch einige

### 3.3. Entwicklung des Verfahrens zur semantischen Segmentierung von Wolken

Daten, die potenziell der falschen Klasse zugeordnet worden sind. Deshalb wurde alternativ das Training anhand der Vier-Klassen-Wolkenmaske getestet. Hier werden die vom ECM-Algorithmus unsicher klassifizierten Pixel in den beiden zusätzlichen Klassen *wahrscheinlich klar* und *wahrscheinlich wolkig* zusammengefasst und verunreinigen somit nicht die Klassen *klar* und *wolkig*. Zur Auswertung der Ergebnisse wurde das so erzielte Ergebnis wieder zu den zwei Klassen *klar* und *wolkig* zusammengefasst.

Als dritte Option wurde die direkte Verwendung der Wolkenwahrscheinlichkeiten als Ground Truth getestet. Dies ist kein Klassifikationsproblem, sondern ein Regressionsproblem, welches ebenfalls mit dem Random Forest Ansatz gelöst werden kann.

Mit allen drei Varianten konnten im getesteten Fall ähnliche Ergebnisse erzielt werden, weshalb mit der binären Wolkenmaske die einfachste und direkteste Repräsentation zur Verwendung für die weiteren Untersuchungen ausgewählt wurde. Im Hinblick auf die möglicherweise nötige Erzeugung von Ground Truth Daten für ein zukünftiges Training auf andere Bilddaten, wäre eine binäre Darstellung außerdem am unkompliziertesten zu generieren und somit am praktikabelsten.



## 4. Analyse der Ergebnisse

Das Verfahren zur semantischen Segmentierung von Wolken wurde anhand von zwei exemplarischen Testgebieten (*muenchen* und *stuttgart*) entwickelt, wie im vorherigen Kapitel beschrieben wurde. Die so bestimmte Merkmalskombination und die Hyperparameter des Random Forest werden im Folgenden festgehalten, um die Qualität der Ergebnisse sowie deren Beeinflussung durch weitere Faktoren zu evaluieren.

Nach der Definition der betrachteten Genauigkeitsmaße im folgenden Abschnitt werden in Kapitel 4.2 zunächst die mit der Version aus Kapitel 3.3 erzielbaren Ergebnisse vorgestellt. Sie dienen als Grundlage der Betrachtungen. In den darauf folgenden Analysen werden zunächst die verwendeten Trainingsdaten variiert, um die Abhängigkeit der Ergebnisse von der Wahl der Trainingsdaten zu bestimmen. Anschließend wird analysiert, wie robust das Verfahren gegenüber einer Veränderung der radiometrischen Sensitivität des Sensors ist. Hierbei ist anzumerken, dass es selbstverständlich möglich ist, dass unter diesen veränderlichen Bedingungen die Verwendung anderer Merkmalskombinationen oder Random-Forest-Hyperparameter ein besseres Ergebnis erzielen würde. Die folgenden Ergebnisse sind also als Mindestmaß der erreichbaren Qualität zu verstehen und zeigen die Größenordnung der erreichbaren Genauigkeiten auf.

**Testdatensatz** Als Testdatensatz wird für alle durchgeführten Analysen eine für verschiedene Situationen repräsentative Auswahl der in Kapitel 3.1.1 vorgestellten Testgebiete verwendet. Diese ausgewählten Gebiete sind *stuttgart*, *belgien* und *neworleans* als Beispiele für verschiedene Ausprägungen urbaner Beleuchtung und *petronius*, *nil* und *alpen* als Beispiele für die besonderen Landbedeckungen Ozean, Wüste und Schnee (vgl. 3.1.1). Von jedem dieser Gebiete wurden die ersten 30 Bilder des Jahres 2018 verwendet, die nach der Vorverarbeitung übrig bleiben (das heißt die Daten können auch einen Zeitraum von mehr als 30 Tagen abdecken, falls manche Tage fehlen). Dies stellt sicher, dass alle möglichen Mondbeleuchtungs-Situationen mindestens einmal vorliegen. Tabelle 4.1 zeigt die Klassenverteilung dieses Testdatensatzes pro Gebiet auf.

	<i>stuttgart</i>	<i>belgien</i>	<i>neworleans</i>	<i>petronius</i>	<i>nil</i>	<i>alpen</i>	gesamt
Anteil <i>klar</i>	20.6%	35.0%	53.9%	33.7%	81.9%	22.6%	41.3%
Anteil <i>wolkig</i>	79.4%	65.0%	46.1%	66.3%	18.1%	77.4%	58.7%

Tabelle 4.1: Klassenverteilung des Testdatensatzes pro Gebiet.

### 4.1. Definition der betrachteten Genauigkeitsmaße

Das in dieser Arbeit beschriebene Verfahren führt eine semantische Segmentierung der Bilder in die beiden Klassen *wolkig* und *klar* durch. Der daraus resultierende Aufbau der Konfusionsmatrix, die das Klassifikationsergebnis aller Pixel beschreibt, wird in Tabelle 4.2

#### 4. Analyse der Ergebnisse

dargestellt. Die beiden aus der Konfusionsmatrix ableitbaren Größen Recall und Precision werden dort ebenfalls definiert [48].

		Prädiktion		Recall
		<i>klar</i>	<i>wolkig</i>	
GT	<i>klar</i>	TN	FP	$\frac{TN}{TN+FP}$
	<i>wolkig</i>	FN	TP	$\frac{TP}{FN+TP}$
Precision		$\frac{TN}{TN+FN}$	$\frac{TP}{FP+TP}$	

Tabelle 4.2: Definition der Konfusionsmatrix für den Fall der binären Klassifikation mit den Klassen *klar* und *wolkig*. Die beiden Größen Recall und Precision können direkt aus der Konfusionsmatrix berechnet werden.

Um Precision und Recall gemeinsam zu betrachten, kann aus ihnen das harmonische Mittel gebildet werden. Dieses wird bei gleicher Gewichtung von Precision und Recall  $F_1$ -Maß genannt [49].

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.1)$$

Das  $F_1$ -Maß kann sowohl pro Klasse als auch für beide Klassen gemeinsam berechnet werden. Im Fall der Ableitung eines Maßes für beide Klassen gemeinsam wurde in dieser Arbeit das mit dem Support beider Klassen gewichtete Mittel der  $F_1$ -Maße der beiden Klassen verwendet.

Neben diesem mittleren  $F_1$ -Maß werden zwei weitere Maße zur Gesamtbeurteilung des Klassifikationsergebnisses eingesetzt. Die Overall Accuracy (kurz OA, siehe Gleichung (4.2)) ist der Anteil der korrekt klassifizierten Pixel an den insgesamt betrachteten Pixeln. Die Balanced Accuracy (BA) entspricht dem Mittelwert des Recalls der beiden Klassen und wird verwendet, um die Beurteilung des Klassifikators weniger abhängig von der in den Daten vorliegenden Klassenverteilung zu machen.

$$\text{OA} = \frac{TP + TN}{P + N} \quad \text{mit} \quad P = TP + FN, N = FP + TN \quad (4.2)$$

$$\text{BA} = \frac{1}{2} \cdot \left( \frac{TP}{P} + \frac{TN}{N} \right) \quad (4.3)$$

Als Ground Truth (GT), also als Repräsentation der wahren Gegebenheit, wird in dieser Arbeit die ECM-Wolkenmaske verwendet (vgl. Kapitel 2.1.2). Dies entspricht natürlich nicht den tatsächlichen Eigenschaften der ECM, da diese den in Kapitel 2.1.2 angesprochenen Ungenauigkeiten unterliegt. Alle im Folgenden durchgeführten Analysen und Berechnungen von Klassifikationsgenauigkeiten basieren jedoch auf dieser Annahme der ECM als perfekte Wolkenmaske und treffen dadurch keine direkte Aussage darüber, wie gut die tatsächlich vorhandene Bewölkung erkannt wird, sondern nur, wie gut die von ECM erkannte Bewölkung erkannt wird. Dies wird als Näherung der tatsächlichen Genauigkeiten betrachtet.

Der Fall, dass das in dieser Arbeit entwickelte Verfahren an einer Stelle, wo ECM falsch liegt, eine richtige Entscheidung trifft, spiegelt sich nicht in den auf Basis von ECM abgeleiteten Genauigkeiten wider. Solche Fälle können darum nur anhand einer visuellen Beurteilung identifiziert werden.

## 4.2. Ergebnisse für verschiedene Testgebiete

Mit dem in Kapitel 3.3 definierten Verfahren wird für den gesamten Testdatensatz (vgl. Tabelle 4.1) eine Overall Accuracy von 70.54% und eine Balanced Accuracy von 67.20% erreicht. Tabelle 4.3 zeigt die entsprechende Konfusionsmatrix inklusive Recall und Precision. Die Ergebnisse der Klasse *wolkig* sind deutlich besser als die der Klasse *klar*, was vor

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
G	<i>klar</i>	803262	868699	0.48	0.57
	<i>wolkig</i>	324269	2053770	0.86	0.77
Precision		0.71	0.70		

Tabelle 4.3: Konfusionsmatrix für die semantische Segmentierung des gesamten Testdatensatzes (alle 6 Testgebiete) mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

allem durch den hohen Recall der Klasse *wolkig* verursacht wird. Die Precision ist für beide Klassen ähnlich. Dies entspricht dem erwarteten Verhalten, da bewusst die Ergebnisse für die Klasse *wolkig* optimiert wurden.

Dieses Gesamtergebnis ist aufgrund der geringen Precision hinsichtlich einer praktischen Anwendung nicht zufriedenstellend, da im Ergebnis beider Klassen jeweils ca. 30% der Pixel falsch zugeordnet sind.

### 4.2.1. Ergebnisse pro Gebiet

Betrachtet man dahingegen die nach den einzelnen Testgebieten aufgeschlüsselten Ergebnisse in Abbildung 4.1, ergibt sich ein anderes Bild. Es zeigt sich, dass die Performance des Klassifikators für verschiedene Gebiete sehr unterschiedlich ist. Die Betrachtung des Gesamtergebnisses vernachlässigt diese große Verschiedenheit der Testgebiete und besitzt eine relativ schwache Aussagekraft, da unter den Testgebieten bewusst auch solche sind, bei denen Schwierigkeiten erwartet werden.

Insgesamt schneiden die drei urbanen Testgebiete bei Betrachtung der Balanced Accuracy am besten ab, was aufgrund des ausschließlichen Trainings mit urbanen Daten und des bei der Entwicklung des Verfahrens auf solche Daten gesetzten Fokus der Erwartung entspricht. Mit Abstand am schlechtesten sind die Ergebnisse für das Gebiet *nil*.

#### 4. Analyse der Ergebnisse

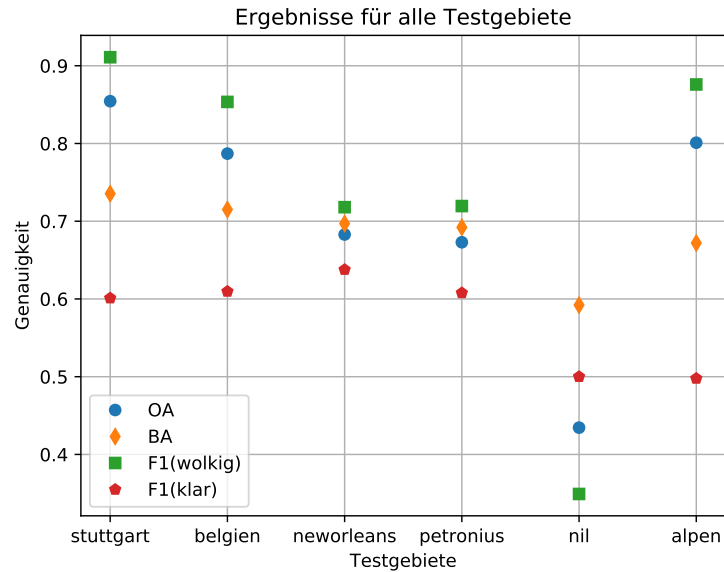


Abbildung 4.1: Ergebnisse pro Testgebiet unter Verwendung von *muenchen* (100 Bilder) zum Training. Dargestellt werden die Overall Accuracy (OA), Balanced Accuracy (BA) und die  $F_1$ -Maße der beiden Klassen, die jeweils für alle Bilder eines Gebietes berechnet wurden.

Betrachtet man die nach Klassen getrennten Ergebnisse, also  $F_1(wolkig)$  und  $F_1(klar)$ , gibt es noch größere Unterschiede zwischen den Gebieten. Zum einen fallen die großen Unterschiede in der Differenz der beiden Maße pro Gebiet auf. So sind diese für *neworleans* mit einer Differenz von weniger als 10% relativ ähnlich und für *stuttgart* und *alpen* mit einer Differenz von über 30% sehr verschieden. Dieser Effekt ist eine Auswirkung der unterschiedlichen Klassenverteilungen in den Testdaten (vgl. Tabelle 4.1). So haben *stuttgart* und *alpen* von allen Testgebieten den kleinsten Anteil an Pixeln der Klasse *klar*, wodurch es für den auf die Erkennung von Wolken optimierten Algorithmus einfacher ist, einen hohen Wert für  $F_1(wolkig)$  zu erreichen, als bei der relativ ausgeglichenen Klassenverteilung von *alpen*. Durch den kleinen Anteil an klaren Pixeln wirken sich Fehler bei deren Zuordnung viel stärker auf das  $F_1$ -Maß für die Klasse *klar* aus, als für die Klasse *wolkig*.

Zum anderen fällt der große Ausreißer von  $F_1(wolkig)$  beim Testgebiet *nil* auf, der bei  $F_1(klar)$  nicht auftritt. Dieses Verhalten lässt sich anhand der im Folgenden durchgeführten genaueren Betrachtung der Ergebnisse pro Bild erklären.

Betrachtet man die Ergebnisse der einzelnen Testgebiete genauer (Konfusionsmatrizen in den Tabellen 4.4 bis 4.9), werden die Stärken und Schwächen des Verfahrens deutlich. Für *stuttgart* (Tabelle 4.4) und *belgien* (Tabelle 4.5) sind die Ergebnisse deutlich besser als das zuvor vorgestellte Gesamtergebnis. Insbesondere der Recall der Klasse *wolkig* ist hier mit über 90% sehr gut und auch die Precision ist höher als beim Gesamtergebnis. Für die von

## 4.2. Ergebnisse für verschiedene Testgebiete

urbaner Beleuchtung geprägten Gebiete, die Ähnlichkeiten zum Trainingsgebiet *muenchen* aufweisen, funktioniert das Verfahren also sehr gut.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
CT	<i>klar</i>	74036	64770	0.53	0.60
	<i>wolkig</i>	33503	502691	0.94	0.91
Precision		0.69	0.89		

Tabelle 4.4: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *stuttgart* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
CT	<i>klar</i>	112311	123683	0.48	0.61
	<i>wolkig</i>	20136	418870	0.95	0.85
Precision		0.85	0.77		

Tabelle 4.5: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *belgien* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

Für die beiden Gebiete *neworleans* und *petronius* konnten laut Abbildung 4.1 sehr ähnliche Ergebnisse erzielt werden, die für die Klasse *wolkig* deutlich schlechter als für *stuttgart* und *belgien* sind. Dies ist zunächst überraschend, da für *neworleans*, das ebenfalls ein Beispiel für urbane Beleuchtung darstellt, ähnliche Ergebnisse wie für die anderen urbanen Gebiete erwartet werden könnten.

Die entsprechenden Konfusionsmatrizen (Tabellen 4.6 und 4.7) zeigen jedoch, dass sich die Performance für *neworleans* und *petronius* doch stark unterscheidet. Während bei *neworleans* der Klassifikator weiterhin eher dazu tendiert, zu viele Pixel als *wolkig* zu klassifizieren (hoher Recall der Klasse *wolkig* bei relativ geringer Precision), ist bei *petronius* genau das Gegenteil der Fall. Eine Erklärung hierfür ist, dass Wolken über dem dunklen Ozean ohne Mondbeleuchtung schlicht unsichtbar sind, da sie zum einen nicht durch von der Erdoberfläche emittiertes Licht beleuchtet werden, wie es bei urbanen Gebieten der Fall ist.

Zum anderen tritt der Effekt des Verschwimmens der (beleuchteten) Strukturen der Erdoberfläche durch Bewölkung, auf dem einige verwendete Merkmale basieren, in diesem Fall nicht auf. Eine Wolke ist also ohne Mondbeleuchtung über dem Ozean für den Klassifikator nicht von klaren Sichtverhältnissen unterscheidbar (siehe hierzu auch Bild 12 in Abbildung A.6 im Anhang). Bestätigt wird diese Überlegung dadurch, dass die besten Ergebnisse für *petronius* alle bei Bildern mit starker Mondbeleuchtung auftreten (vgl. Abbildung 4.5).

Die Konfusionsmatrix des Testgebietes *nil* (Tabelle 4.8) zeigt sehr deutlich das Zusammenspiel von Recall und Precision. Es werden zwar 84% der Wolken gefunden, aber nur gut ein Fünftel der Pixel in der resultierenden Wolkenklasse sind tatsächlich wolkig, wodurch

#### 4. Analyse der Ergebnisse

der hohe Recall für die Anwendung wertlos wird. Für *alpen* sind wiederum sowohl Recall als auch Precision der Klasse *wolkig* sehr gut, siehe Tabelle 4.9.

Im Folgenden werden zur Veranschaulichung der Auswirkungen des beschriebenen Verhaltens des Klassifikators für die verschiedenen Testgebiete beispielhaft einige Ergebnisbilder betrachtet.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
GT	<i>klar</i>	188407	175629	0.52	0.64
	<i>wolkig</i>	38406	272558	0.88	0.72
Precision		0.83	0.61		

Tabelle 4.6: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *neworleans* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
GT	<i>klar</i>	171065	56618	0.75	0.61
	<i>wolkig</i>	164164	283153	0.63	0.72
Precision		0.51	0.83		

Tabelle 4.7: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *petronius* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
GT	<i>klar</i>	190885	362090	0.35	0.50
	<i>wolkig</i>	19641	102384	0.84	0.35
Precision		0.91	0.22		

Tabelle 4.8: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *nil* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

		Prädiktion		Recall	$F_1$
		<i>klar</i>	<i>wolkig</i>		
GT	<i>klar</i>	66558	85909	0.44	0.50
	<i>wolkig</i>	48419	474114	0.91	0.88
Precision		0.58	0.85		

Tabelle 4.9: Konfusionsmatrix für die semantische Segmentierung des Testgebietes *alpen* mit dem an *muenchen* (100 Bilder) trainierten Klassifikator.

### 4.2.2. Ergebnisse pro Bild

Nicht nur zwischen den verschiedenen Testgebieten gibt es große Unterschiede in der Qualität der Ergebnisse, sondern auch zwischen den einzelnen Bildern jedes Gebietes. Im Folgenden werden deshalb beispielhaft für einige Testgebiete besonders gute und schlechte Ergebnisbilder betrachtet, um die Ursachen für diese Schwankungen abzuleiten. Ein besonderes Augenmerk wird hierbei ab Seite 63 auf den Einfluss der Mondbeleuchtung gelegt.

Zunächst werden die Ergebnisse des Testgebietes *stuttgart* evaluiert, für das die Gesamtergebnisse am besten sind. Die Abbildungen 4.2 und 4.3 zeigen diejenigen Bilder, für die das Verfahren die besten und die schlechtesten Ergebnisse erzielt hat. Die besten Ergebnisse werden für stark bewölkte Bilder erzielt, bei denen die Struktur der Stadt sehr stark verschwommen ist. Dies entspricht genau der Eigenschaft, die durch viele der Merkmale charakterisiert wird.

Bei den schlechtesten Ergebnissen fällt auf, dass die Masken der prädizierten Klassen mitunter besser zum visuellen Eindruck des DNB-Bildes passen, als die entsprechenden ECM-Masken (vgl. Bilder Nr. 19 und 23 in Abbildung 4.3). Vor allem bei Bild 19 ist in der oberen linken Bildecke des DNB-Bildes eindeutig eine Wolke sichtbar, die nicht Teil der ECM-Maske ist, vom in dieser Arbeit entwickelten Verfahren aber richtig erkannt wurde. In diesen Fällen ist also nur die Übereinstimmung mit der ECM-Maske schlecht, aber nicht zwingend die Übereinstimmung mit der Realität. Die Performance des entwickelten Verfahrens bei Anwendung auf das Testgebiet *stuttgart* ist also insgesamt als sehr gut zu bewerten.

Für die Gebiete *belgien*, *neworleans* und *alpen* werden die besten Ergebnisse genau wie bei *stuttgart* für Bilder mit starker Bewölkung erzielt, die den ganzen Bildbereich überdeckt. Die entsprechenden Abbildungen befinden sich im Anhang (Abb. A.1, A.3, A.8). Die visuelle Analyse der schlechtesten Ergebnisse offenbart verschiedene Ursachen.

Für *belgien* (Abb. A.2) tritt in drei der vier als am schlechtesten bewerteten Bilder wie bei *stuttgart* der Fall auf, dass die prädizierte Wolkenmaske besser zum visuellen Eindruck der DNB-Daten passt, als die ECM-Maske, bei der offensichtlich erkennbare Wolken fehlen. Diese Ergebnisse sind somit unter Betrachtung der DNB-Bilder als Referenz als sehr gut zu bewerten. Im Bild Nr. 4 wird tatsächlich eine Schwäche des Verfahrens deutlich. Bei sehr dünner Bewölkung, die nur zu einem leichten Verschwimmen der städtischen Beleuchtung führt, trifft der Algorithmus sehr unsichere und dadurch oft falsche Entscheidungen.

Für *neworleans* und *alpen* treten trotz der unterschiedlichen Charakteristika der beiden Gebiete sehr ähnliche Fehler auf. Bei den schlechtesten Ergebnissen (Abb. A.4, A.9) werden alle künstlich beleuchteten Strukturen der Klasse *klar* zugeordnet und das gesamte Umland als *wolkig* klassifiziert. Lediglich bei einem Bildausschnitt von *alpen* ist die Ursache hierfür eindeutig erkennbar, da hier die vom Mond beleuchteten schneebedeckten Berge mit Wolken verwechselt werden (vgl. Bild 22 in Abbildung A.9). Bei den restlichen Bildern in den beiden Abbildungen ist gerade das Gegenteil der Fall. Das falsch klassifizierte Umland ist sehr dunkel und wirkt nur teilweise leicht inhomogen oder verrauscht. Beim Trainingsgebiet *muenchen* treten bei klarem Himmel aufgrund der urbanen Strukturen

#### 4. Analyse der Ergebnisse

deutlich größere Inhomogenitäten auf, es gibt keine solchen ausgedehnten, relativ homogenen dunklen Flächen. Betrachtet man nur die Homogenität der Textur kann die falsche Zuordnung also damit erklärt werden, dass in den Trainingsdaten für klaren Himmel solche Flächen nicht repräsentiert werden. Dadurch besitzt das falsch klassifizierte Umland eine höhere Ähnlichkeit zu Bewölkung, denn diese ist in den Trainingsdaten im Vergleich zu den urbanen Strukturen bei wolkenfreien Bedingungen deutlich homogener.

Laut dieser Annahme müssten bei *petronius* die gleichen Fehler auftreten, da dieses das Testgebiet mit der homogensten Landoberfläche ist. Die Abbildungen A.6 und A.7 zeigen, dass dies jedoch nur manchmal der Fall ist. In vielen Fällen wird eine homogene dunkle Fläche durchaus als *klar* klassifiziert. Das Verhalten des Klassifikators bezüglich homogener dunkler Flächen ist also aufgrund der visuellen Beurteilung der Ergebnisbilder nicht abschließend erklärbar. Für solche Bereiche funktioniert das Verfahren somit insgesamt nicht zuverlässig.

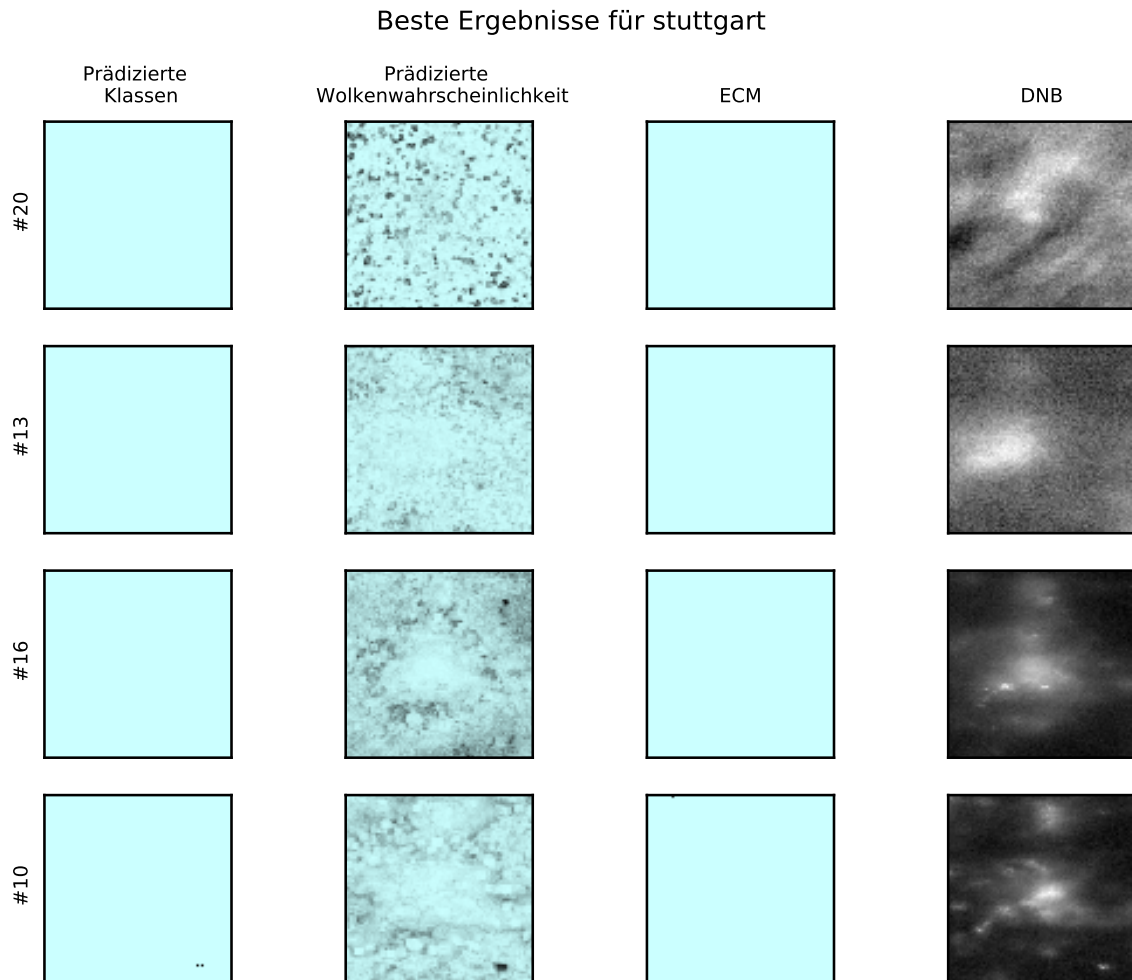


Abbildung 4.2: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *stuttgart* mit den besten Ergebnissen. Blau ist die Klasse *wolzig*, schwarz die Klasse *klar*. In den DNB-Bildern ist die in Kapitel 2.2.4 angesprochene geometrische Verschiebung von beleuchteten Städten durch die Aufnahme einer von Bewölkung überdeckten Stadt in Schrägsicht gut erkennbar.

Legende (gilt für alle Abbildungen dieses Typs):

Bedeutung der Spalten (von rechts nach links): DNB=originaler DNB-Bildausschnitt; ECM=Ground Truth; Prädizierte Wolkenwahrscheinlichkeit=Vom Random Forest bestimmte Wahrsch. für die Klasse *wolzig*; Prädizierte Klassen=Vom Random Forest bestimmte Klassenzugehörigkeit (*wolzig*, wenn prädizierte Wolkenwahrscheinlichkeit  $> 0.5$ )

Weitere Informationen zu den Bildern können anhand der Bildnummer am Zeilenanfang dem Anhang entnommen werden (Tabellen B.1 - B.6).

#### 4. Analyse der Ergebnisse

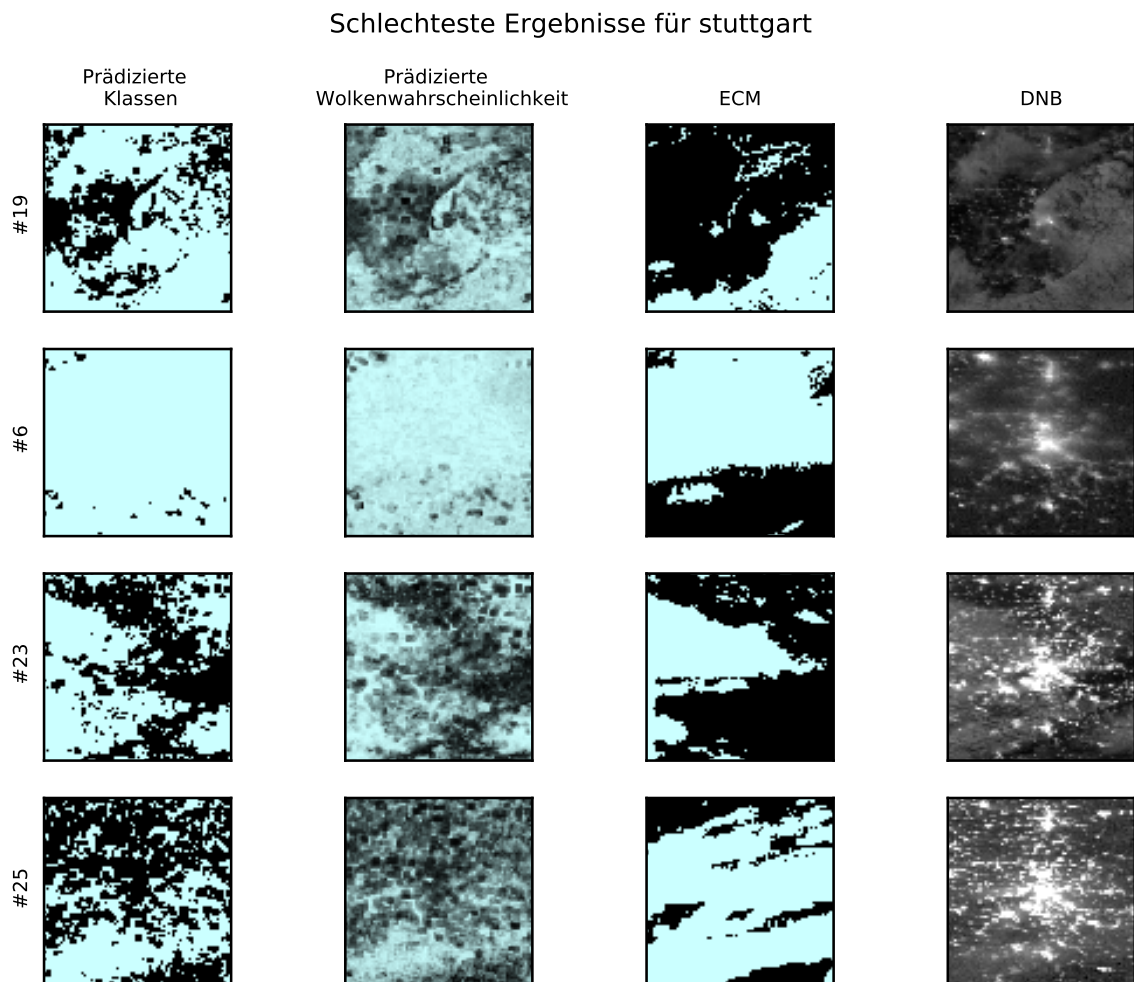


Abbildung 4.3: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *stuttgart* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolzig*, schwarz die Klasse *klar*.

In den Bildern 19 und 23 passen die Klassengrenzen der prädizierten Wolkenmaske im Gegensatz zur ECM-Maske sehr gut zum visuellen Eindruck in den DNB-Bildern.

**Abhängigkeit der Ergebnisse von der Mondbeleuchtung** Wie in Kapitel 2.2 bereits erläutert, ist das Erscheinungsbild von Wolken in den DNB-Bildern stark von der Beleuchtung durch den Mond abhängig. Dies legt nahe, dass bei der semantischen Segmentierung von Wolken eine Abhängigkeit der Qualität der Ergebnisse von der Mondbeleuchtung auftreten könnte.

Abbildung 4.4 stellt diesen Zusammenhang zwischen der Overall Accuracy und der modellierten Mondbeleuchtung pro Bild für die verschiedenen Testgebiete her. Insgesamt ist keine sehr deutliche Abhängigkeit ablesbar, lediglich eine  $OA < 0.4$  kommt bei fast allen Testgebieten nur bei fehlender Mondbeleuchtung vor. Betrachtet man die Testgebiete einzeln, ist jedoch bei *petronius* und *nil* eine besonders deutliche Abhängigkeit der beiden Größen erkennbar. Für *petronius* wurde die Ursache für diese Abhängigkeit bereits

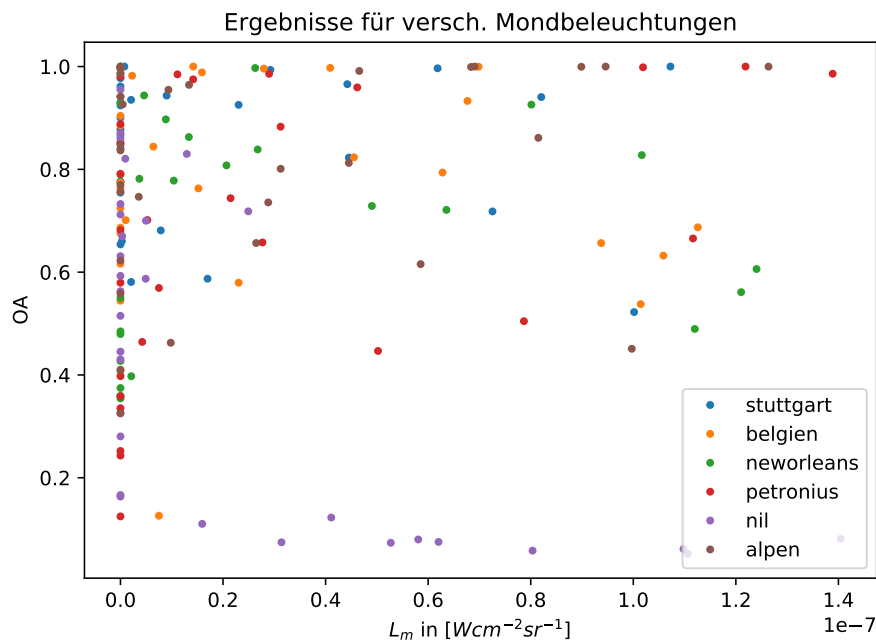


Abbildung 4.4: Abhängigkeit der Ergebnisse von der Mondbeleuchtung für alle Testgebiete. Jeder Datenpunkt beschreibt den Zusammenhang von Mondbeleuchtung  $L_m$  und OA für ein Bild.

bei der gebietsweisen Betrachtung der Ergebnisse in Kapitel 4.2.1 erläutert. Umso stärker die Mondbeleuchtung, umso seltener treten hier schlechte Ergebnisse auf bzw. umso besser werden die Ergebnisse im Durchschnitt, da Wolken erst dann vom Untergrund unterscheidbar werden. Abbildung 4.5 verdeutlicht diesen Zusammenhang. Die Bilder mit den Ergebnissen befinden sich im Anhang in den Abbildungen A.5, A.6 und A.7. In Abbildung A.7 (Bilder 17 und 3) werden einige weitere Ergebnisse für *petronius* dargestellt, die verdeutlichen, dass die Erkennung von Wolken bei vorhandener Mondbeleuchtung auch für nur leicht bewölkte Gebiete gut funktioniert.

#### 4. Analyse der Ergebnisse

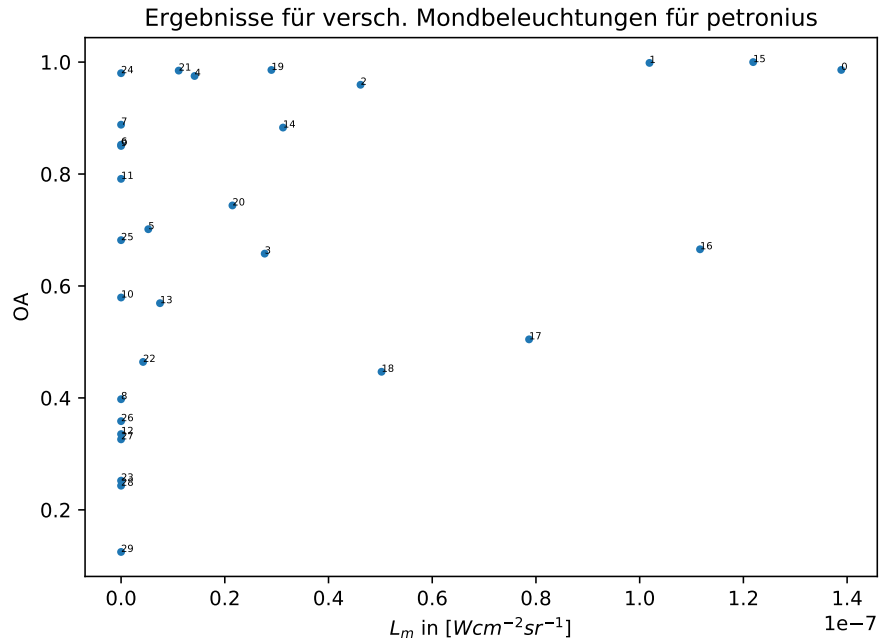


Abbildung 4.5: Abhängigkeit der Ergebnisse von der Mondbeleuchtung für *petronius*. Jeder Datenpunkt beschreibt den Zusammenhang von Mondbeleuchtung und OA für ein Bild, die Bildnummer wird durch die zugehörige Zahl gekennzeichnet. Die schlechtesten Ergebnisse treten alle bei Bildern ohne Beleuchtung durch den Mond auf.

Noch stärker ist die Abhängigkeit der Ergebnisse von der Mondbeleuchtung für das Testgebiet *nil*. Abbildung 4.6 zeigt, dass hier für keines der unter stärkerer Mondbeleuchtung aufgenommenen Bilder gute Ergebnisse erzielt werden. Die visuelle Betrachtung der schlechtesten Ergebnisse für das Gebiet *nil* in Abbildung 4.8 auf Seite 67 erklärt dieses Verhalten. Mit Ausnahme des beleuchteten Nil-Ufers werden alle Pixel als wolkeig klassifiziert, obwohl die Bilder wolkenfrei sind. Bei den besten Ergebnissen in Abbildung 4.7 ist dies seltener der Fall. Der Unterschied zwischen den wolkenfreien Bildern mit den besten und den schlechtesten Ergebnissen ist das bei Mondbeleuchtung durch die hohe Albedo von Wüstenboden veränderte Erscheinungsbild des Testgebietes. Es bestätigt sich also die in Kapitel 3.1.1 erläuterte Vermutung, dass vom Mond beleuchteter Wüstensand mit Bewölkung verwechselt werden könnte.

Insgesamt kann somit je nach den Eigenschaften des Gebietes eine positive oder negative Abhängigkeit der Qualität der Ergebnisse von der Mondbeleuchtung auftreten.

## 4.2. Ergebnisse für verschiedene Testgebiete

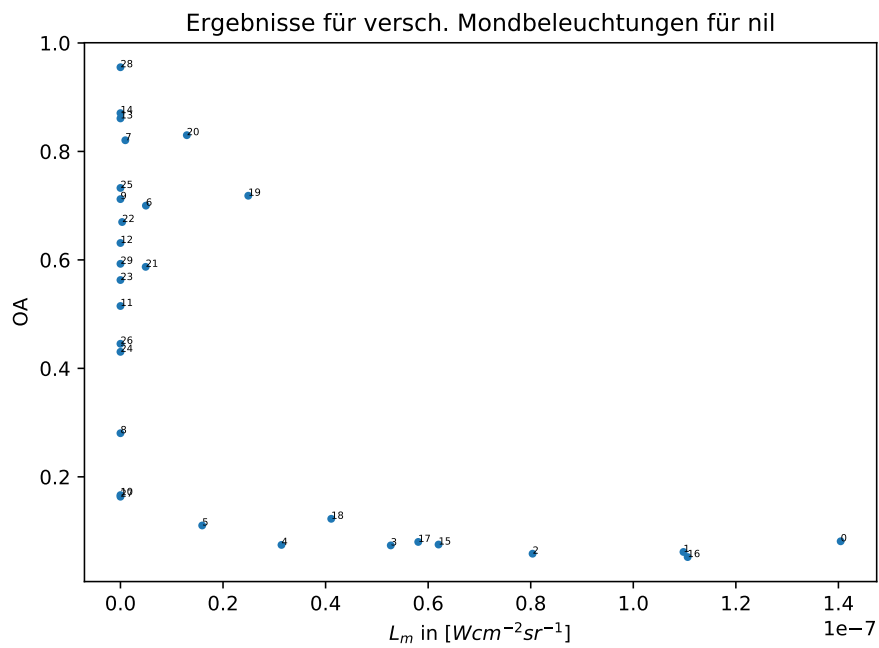


Abbildung 4.6: Abhängigkeit der Ergebnisse von der Mondbeleuchtung für *nil*. Jeder Datenpunkt beschreibt den Zusammenhang von Mondbeleuchtung und OA für ein Bild, die Bildnummer wird durch die zugehörige Zahl gekennzeichnet. Bei stärkerer Beleuchtung durch den Mond werden nur noch sehr schlechte Ergebnisse erzielt.

#### 4. Analyse der Ergebnisse

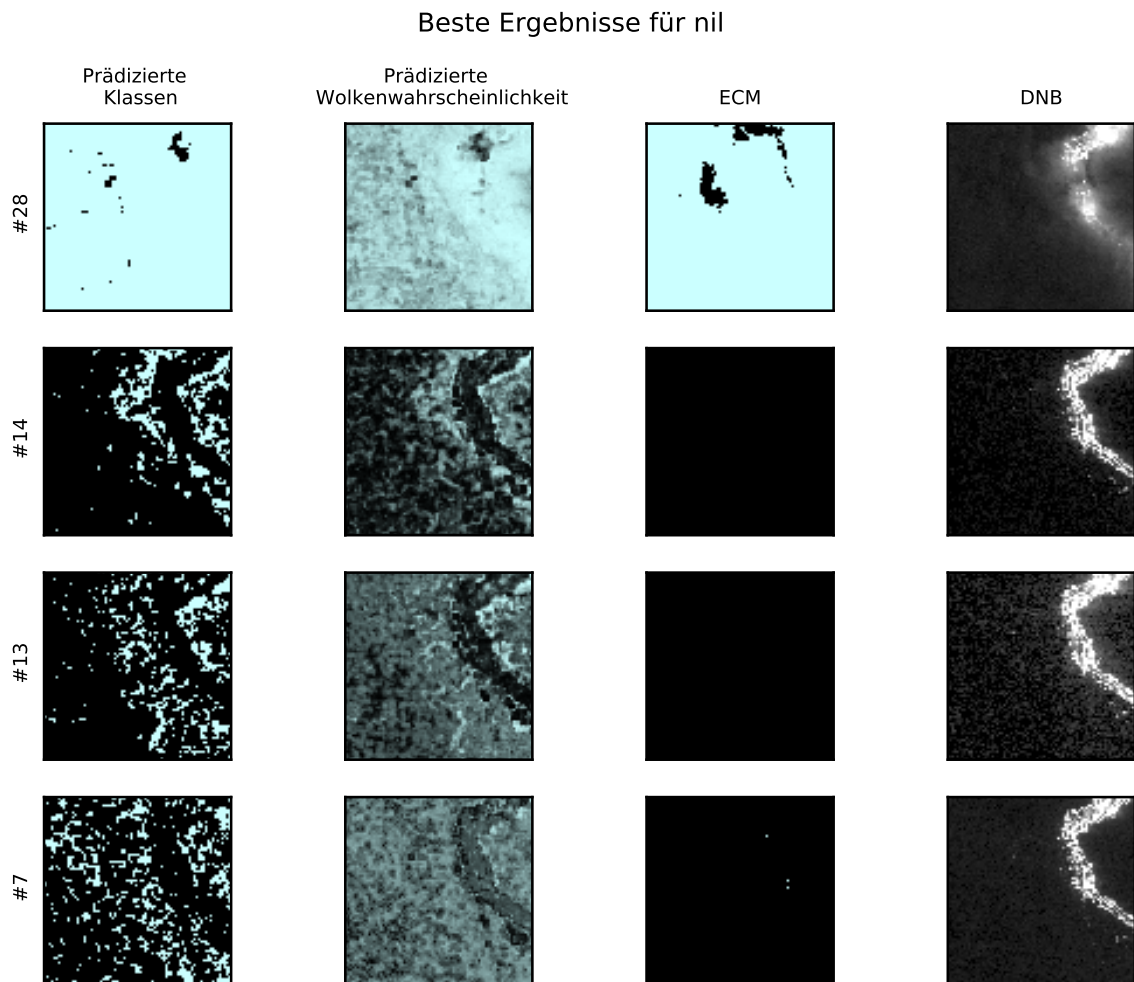


Abbildung 4.7: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *nil* mit den besten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

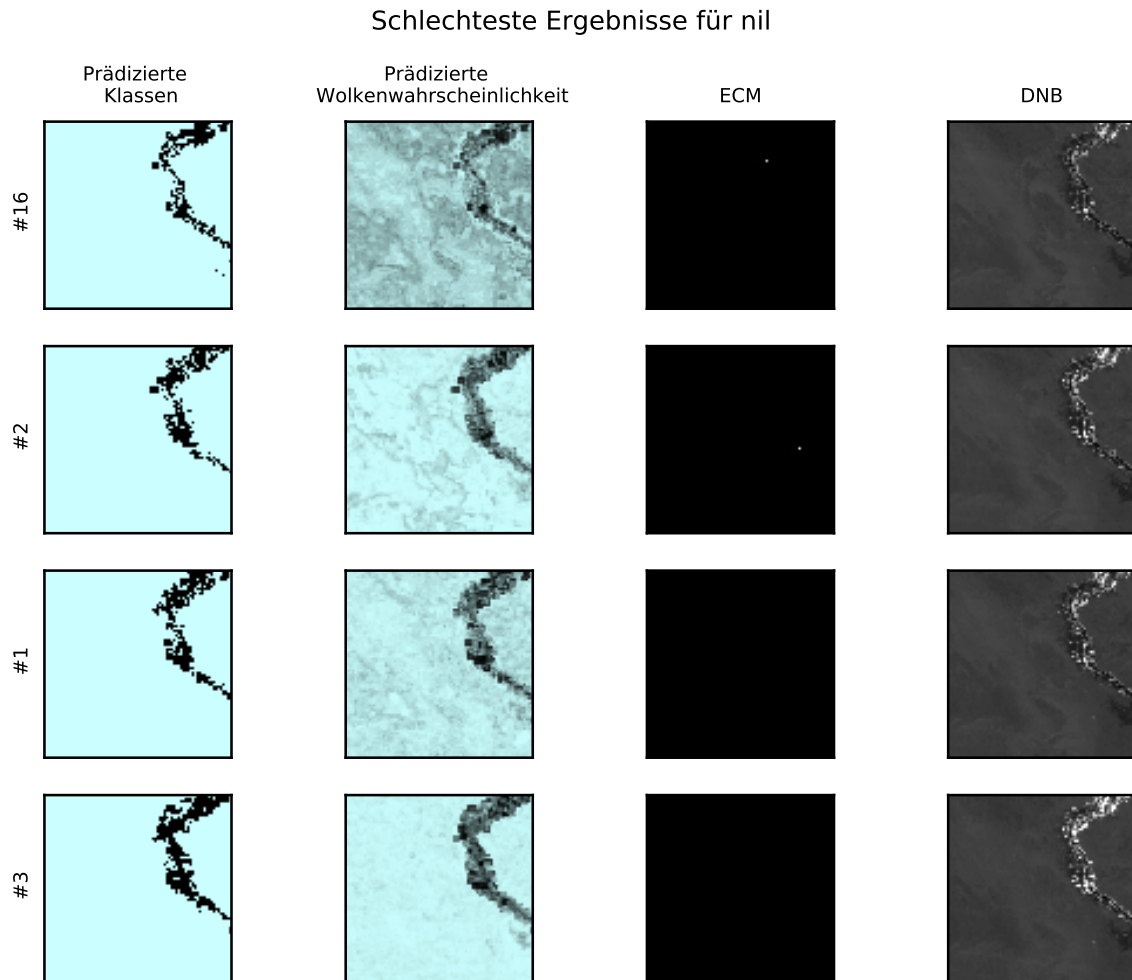


Abbildung 4.8: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *nil* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

In den DNB-Bildern ist die Struktur des Wüstenbodens durch die Beleuchtung durch den Mond im Vergleich zu Abbildung 4.7 eindeutig erkennbar. Sie wird bei den Ergebnissen mit Bewölkung verwechselt.

## 4. Analyse der Ergebnisse

### 4.2.3. Einfluss von Bildserienmerkmalen auf die Ergebnisse

In Kapitel 3.3.3 wurde bereits erwähnt, dass bei der Auswahl der besten Merkmalskombination mit Bildserienmerkmalen bei der Betrachtung von *stuttgart* als Testgebiet die Bildserienmerkmale einen deutlich kleineren positiven Einfluss auf die Ergebnisse haben, als erwartet. Im Folgenden wird die Ursache dieses Verhaltens analysiert.

Vergleicht man die Ergebnisse für die Merkmalsvektoren ohne und mit Bildserienmerkmalen (Tabelle 4.10) für die weiteren Testgebiete, zeigt sich, dass die Bildserienmerkmale bei den Testgebieten *petronius* und *alpen* durchaus zu deutlichen Verbesserungen führen, vor allem bei *petronius*. Der Unterschied dieser Gebiete zu *stuttgart* und den anderen Testge-

	<i>stuttgart</i>	<i>belgien</i>	<i>neworleans</i>	<i>petronius</i>	<i>nil</i>	<i>alpen</i>
$F_1(klar) : V_0$	0.5832	0.6108	0.6484	<b>0.5650</b>	0.5308	<b>0.4495</b>
$F_1(klar) : V_1$	0.6010	0.6097	0.6378	<b>0.6078</b>	0.5000	<b>0.4977</b>
$F_1(wolkig) : V_0$	0.9088	0.8535	0.7019	<b>0.6314</b>	0.3570	0.8673
$F_1(wolkig) : V_1$	0.9110	0.8535	0.7181	<b>0.7195</b>	0.3491	0.8759

Tabelle 4.10: Vergleich der Ergebnisse für die Testgebiete für die beiden Klassifikations-Versionen  $V_0$  (Merkmalsvektor  $M_{Einzelbild}$ ) und  $V_1$  (Merkmalsvektor  $M_{Bildserie}$ ).

bieten, bei denen die Bildserienmerkmale kaum einen Einfluss haben, ist das Fehlen bzw. das deutlich geringere Ausmaß an künstlicher Beleuchtung. Es liegt deshalb die Vermutung nahe, dass die zeitliche Betrachtung lediglich bei künstlich beleuchteten urbanen Gebieten nicht die theoretisch erwarteten Verbesserungen bewirkt.

Eine Ursache hierfür kann sein, dass bei urbanen Gebieten eine exakte Georeferenzierung deutlich wichtiger für den Erfolg einer solchen zeitlichen Betrachtung ist, als bei nicht urbanen Gebieten. Sobald das Pixelraster verschoben ist, sind die Aussagen des Vergleichs des Pixels mit seinen zeitlichen Vorgängern nicht mehr valide, da sich der Vergleich auf einen falschen Teil der Erdoberfläche bezieht. Bei künstlich beleuchteten urbanen Gebieten sind die Intensitäten benachbarter Pixel aufgrund der urbanen Strukturen oft sehr unterschiedlich, wodurch bereits eine kleine Verschiebung zu einer großen Abweichung des fälschlicherweise betrachteten Pixels vom richtigen Vergleichspixel führt. Bei Gebieten ohne künstliche Beleuchtung sind die Flächen insgesamt homogener, wodurch eine Verschiebung des Pixelrasters beim zeitlichen Vergleich einen deutlich kleineren Einfluss auf die Abweichung des betrachteten Pixels vom korrekten Vergleichspixel hat. Dadurch wird der zeitliche Vergleich weniger stark verfälscht, als bei künstlich beleuchteten Gebieten. Aufgrund der Problematiken der Georeferenzierung am Bildrand und bei Bewölkung (vgl. Kapitel 3.1.2 und 2.1.1) muss davon ausgegangen werden, dass diese nicht in allen Fällen ideal ist und die soeben erläuterten Effekte eintreten können.

Der negative Einfluss der Bildserienmerkmale auf die Ergebnisse des Testgebietes *nil* weist auf einen weiteren Faktor hin, der den Erfolg einer zeitlichen Betrachtung beeinflusst. Wie

bereits erläutert, besteht für dieses Gebiet eine starke Abhängigkeit des Erscheinungsbildes von der Mondbeleuchtung. Vergleicht man ein Pixel nun mit zeitlichen Vorgängern, die bei einer abweichenden Mondbeleuchtung aufgenommen wurden, ergeben sich Abweichungen, die nicht durch Bewölkung verursacht, jedoch als solche interpretiert werden. Dieser Effekt ließe sich durch die ausschließliche Verwendung von Bildern mit ähnlicher Mondbeleuchtung bei der Berechnung der Bildserienmerkmale reduzieren.

### 4.3. Variation der Trainingsdaten

Der erste Faktor, dessen Einfluss auf die Ergebnisse im Rahmen dieser Analysen betrachtet wird, ist die Wahl der Trainingsdaten. Zur Entwicklung des Klassifikators in Kapitel 3.3 wurden lediglich 100 Bilder von *muenchen* als Trainingsdaten verwendet, der Klassifikator wurde also nur auf ein einziges urbanes Gebiet trainiert. Dies entspricht im Allgemeinen nicht der idealen Konfiguration für Trainingsdaten, da generell ein für alle Situationen repräsentativer Datensatz mit ausreichend großem Umfang angestrebt wird. Ziel der folgenden Betrachtungen ist es deshalb, abzuleiten, wie ein solcher idealer Trainingsdatensatz für das in dieser Arbeit entwickelte Verfahren aufgebaut sein sollte.

Im Folgenden wird deshalb zum einen analysiert, inwiefern sich die Art beziehungsweise die Diversität der Trainingsdaten bei gleichbleibender Anzahl an Bildern auf die Qualität der Ergebnisse auswirkt und zum anderen, wie diese durch die Anzahl der Trainingsbilder beeinflusst wird. Der Klassifikator wurde hierfür mit vier weiteren Varianten von Trainingsdatensätzen trainiert, die sich aus 1-4 Gebieten zusammensetzen. Die in diesen Varianten verwendeten Gebiete wurden so gewählt, dass die Diversität der durch sie repräsentierten Landbedeckungen stetig erhöht wird. So wurde als Kontrast zum Training mit *muenchen* das Training mit einem Gebiet ohne urbane künstliche Beleuchtung (*indozean*) getestet. Außerdem wurde *muenchen* um weitere Gebiete ergänzt, zunächst mit einem weiteren ähnlichen urbanen Gebiet (*mailand*), anschließend mit *indien*, das auch ländlichere Regionen sowie Ozean enthält und schließlich mit *indozean*, um eine Mischung aus Gebieten mit und ohne künstliche Beleuchtung zu erhalten. Die letzte Kombination ist also diejenige mit der höchsten Diversität der Trainingsdaten.

Für die Gebietskombinationen, die *muenchen* beinhalten, wurde zusätzlich jeweils die Verwendung von 100, 200 und 300 Bildern getestet. Tabelle 4.11 fasst die Eigenschaften dieser verschiedenen Versionen von Trainingsdaten zusammen.

In Abbildung 4.9 wird die Balanced Accuracy der semantischen Segmentierung des Testdatensatzes mit diesen Trainingsdatenversionen aufgeschlüsselt nach den einzelnen Testgebieten dargestellt. Da zuvor bereits festgestellt wurde, dass sich die Qualität der Ergebnisse für die beiden Klassen oft sehr unterschiedlich verhält, werden zusätzlich die beiden  $F_1$ -Maße betrachtet. Abbildung 4.10 zeigt diese für die drei urbanen Testgebiete, Abbildung 4.11 für die restlichen, nicht urbanen Gebiete. Es zeigt sich, dass die Qualität der Ergebnisse für die Klasse *klar* im Allgemeinen von der Variation der Trainingsdaten deutlich stärker beeinflusst wird, als die der Klasse *wolkig*.

#### 4. Analyse der Ergebnisse

Tabelle 4.11: Übersicht über die verschiedenen getesteten Trainingsdatenversionen.

	0	1	2	3	4
Trainings- gebiete (Bilder)	<i>muenchen</i> (100)	<i>indozean</i> (100)	<i>muenchen</i> (50) <i>mailand</i> (50)	<i>muenchen</i> (34) <i>mailand</i> (33) <i>indien</i> (33)	<i>muenchen</i> (25) <i>mailand</i> (25) <i>indien</i> (25) <i>indozean</i> (25)
Anteil <i>klar</i>	31.0%	40.4%	23.0%	35.8%	43.3%
” <i>wolkig</i>	69.0%	59.6%	77.0%	64.2%	56.7%

(a) Trainingsdatenversionen mit 100 Bildern.

	0 <sub>200</sub>	2 <sub>200</sub>	3 <sub>200</sub>	4 <sub>200</sub>
Trainings- gebiete (Bilder)	<i>muenchen</i> (200)	<i>muenchen</i> (100) <i>mailand</i> (100)	<i>muenchen</i> (67) <i>mailand</i> (66) <i>indien</i> (66)	<i>muenchen</i> (50) <i>mailand</i> (50) <i>indien</i> (50) <i>indozean</i> (50)
Anteil <i>klar</i>	43.2%	30.8%	31.5%	39.4%
” <i>wolkig</i>	56.8%	69.2%	68.5%	60.6%

(b) Trainingsdatenversionen mit 200 Bildern.

	0 <sub>300</sub>	2 <sub>300</sub>	3 <sub>300</sub>	4 <sub>300</sub>
Trainings- gebiete (Bilder)	<i>muenchen</i> (264)	<i>muenchen</i> (150) <i>mailand</i> (150)	<i>muenchen</i> (100) <i>mailand</i> (100) <i>indien</i> (100)	<i>muenchen</i> (75) <i>mailand</i> (75) <i>indien</i> (75) <i>indozean</i> (75)
Anteil <i>klar</i>	46.8%	37.5%	31.2%	34.9%
” <i>wolkig</i>	53.2%	62.5%	68.8%	65.1%

(c) Trainingsdatenversionen mit 300 Bildern.

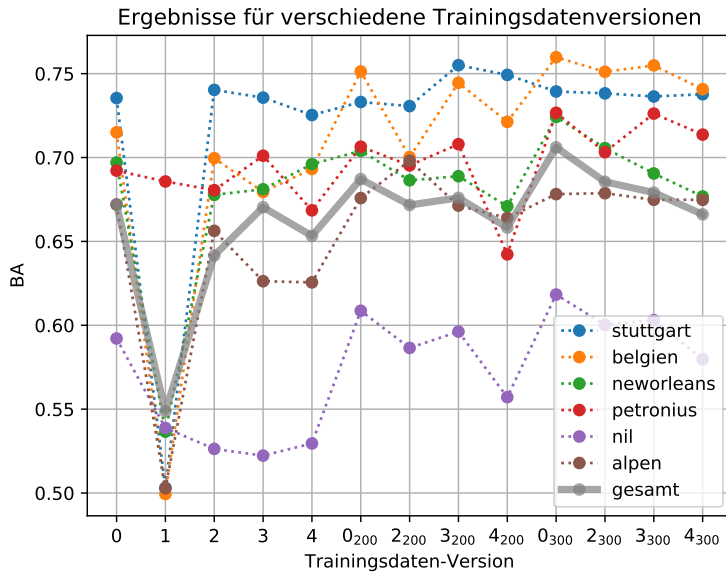


Abbildung 4.9: Ergebnisse (BA) für verschiedene Trainingsdatenversionen, aufgeschlüsselt nach Testgebieten.

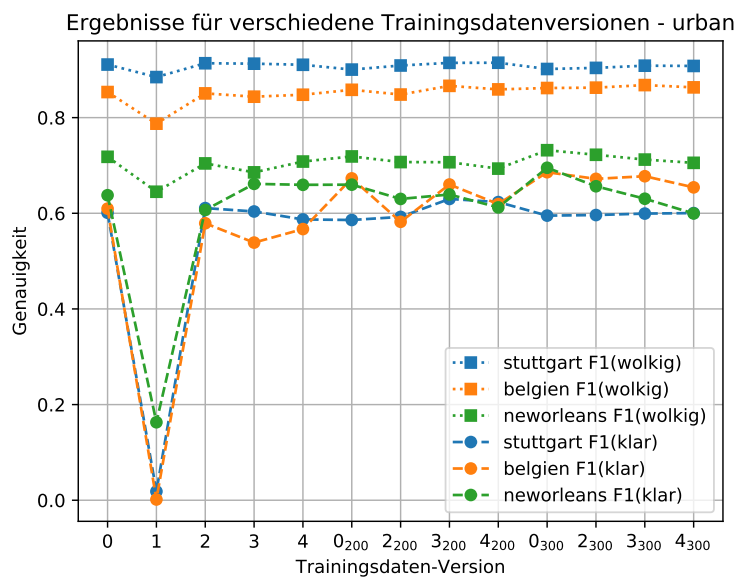


Abbildung 4.10: Ergebnisse ( $F_1(klar)$  und  $F_1(wolkig)$ ) für verschiedene Trainingsdatenversionen für alle urbanen Testgebiete.

#### 4. Analyse der Ergebnisse

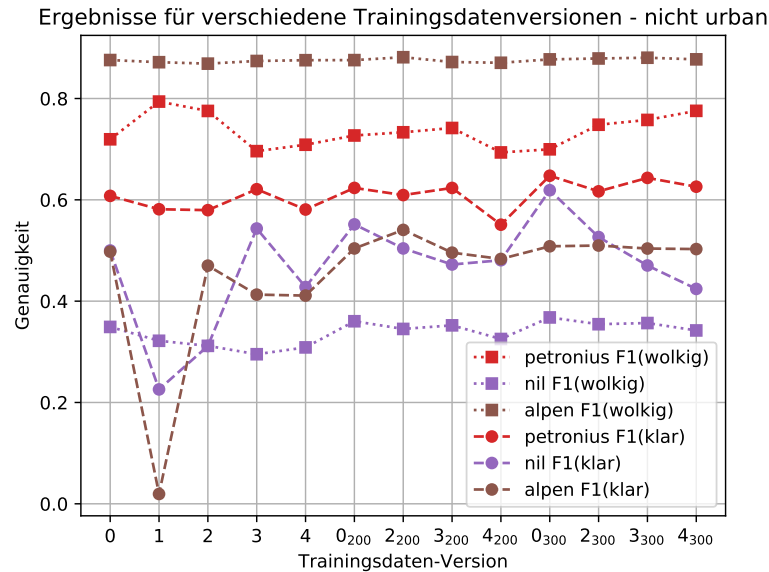


Abbildung 4.11: Ergebnisse ( $F_1(klar)$  und  $F_1(wolkig)$ ) für verschiedene Trainingsdatenversionen für alle nicht urbanen Testgebiete.

**Art und Diversität der Trainingsgebiete** Vergleicht man die Ergebnisse der verschiedenen Kombinationen aus Trainingsgebieten mit jeweils konstanter Bildanzahl miteinander (Abb. 4.9 - 4.11), ist das auffälligste Ergebnis, dass das ausschließliche Training anhand eines nicht urbanen Gebietes (Version 1) für alle Gebiete außer *petronius*, welches ebenfalls einen Ozean zeigt, nicht funktioniert. Für *petronius* liefert diese Trainingsdatenversion durchschnittlich gute Ergebnisse, für  $F_1(wolkig)$  wird sogar der beste Wert der Tests erreicht. Zur semantischen Segmentierung von urbanen Gebieten sollten diese somit auf jeden Fall auch in den Trainingsdaten repräsentiert sein. Umgekehrt funktioniert jedoch die semantische Segmentierung eines Gebietes ohne künstliche Beleuchtung wie *petronius* auch mit einem Klassifikator, der ausschließlich anhand urbaner Trainingsdaten trainiert wurde.

Hinsichtlich der Diversität der Trainingsgebiete lässt sich anhand dieses Tests bezüglich der Testgebiete keine pauschale Aussage darüber ableiten, ob die Verwendung von mehreren verschiedenen Trainingsgebieten (unabhängig davon, ob ausschließlich urban oder gemischt) vorteilhaft gegenüber der Verwendung eines einzigen Trainingsgebietes ist. Für jedes Gebiet ist die Entwicklung der Ergebnisse durch das Hinzufügen von Gebieten zum Trainingsdatensatz (also pro Bildanzahl jeweils die Versionen 0-4) verschieden.

Beim Gesamtergebnis gibt es jedoch mit Ausnahme der Versionen mit 100 Trainingsbildern eine klare Tendenz dazu, dass mit weniger verschiedenen Gebieten bessere Ergebnisse erzielt werden. Die BA sinkt hier jeweils von den Versionen 0 nach 4. Eine mögliche Ursache hierfür kann sein, dass die Zeitreihen pro Gebiet bei Erhöhung der Gebietsanzahl

und gleichbleibender Gesamtanzahl der Bilder kürzer werden und die zeitlichen Merkmale dadurch weniger aussagekräftig sind.

Es ist zudem zu erwähnen, dass nicht nur die in den Trainingsdaten repräsentierte Landbedeckung, sondern auch die Klassenverteilung einen Einfluss auf die Ergebnisse hat. Soll ein Gebiet mit hohem Anteil an Bewölkung klassifiziert werden, ist es ein Nachteil, wenn in den Trainingsdaten nur sehr wenige Beispiele für Bewölkung vorhanden sind, da die gelernten Klassengrenzen nur von sehr wenigen Punkten beeinflusst werden – und zwar unabhängig von der Landbedeckung.

Die Art und die Anzahl der verwendeten Trainingsgebiete haben also generell einen großen Einfluss auf die Ergebnisse. Die Analysen zeigen, dass es vorteilhaft ist, wenn die Trainingsgebiete den Testgebieten ähneln – sowohl in der Art der Landbedeckung, als auch in der Klassenverteilung.

**Anzahl der Trainingsbilder** Hinsichtlich der Anzahl der verwendeten Trainingsbilder ist für das Gesamtergebnis der Trend erkennbar, dass durch eine Erhöhung der Anzahl der Trainingsdaten eine Verbesserung der Ergebnisse erreicht werden kann, siehe Abbildung 4.12. Das Ausmaß der Auswirkungen ist jedoch stark vom Testgebiet abhängig, für *stuttgart* verschlechtern sich die Ergebnisse bei Erhöhung der Anzahl der Trainingsbilder sogar teilweise.

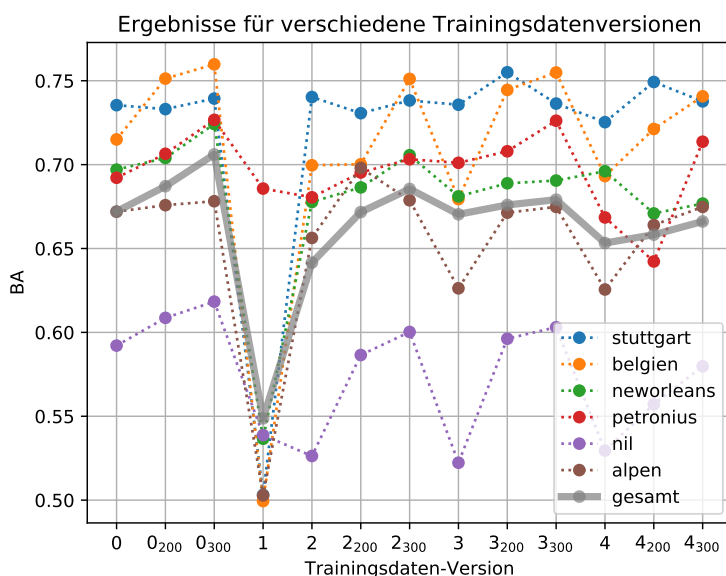


Abbildung 4.12: Ergebnisse (BA) für verschiedene Trainingsdatenversionen, aufgeschlüsselt nach Testgebieten. Um den Fokus auf die Veränderungen durch die verwendete Anzahl an Trainingsbildern zu legen, wurde die Anordnung der Versionen auf der x-Achse ggü. Abb. 4.9 verändert.

#### 4. Analyse der Ergebnisse

Insgesamt haben somit sowohl die Art der verwendeten Trainingsbilder als auch die Menge der Bilder einen großen Einfluss auf die Ergebnisse. Es ist also zunächst eine sorgfältige Auswahl der Art der Trainingsdaten erforderlich. Um optimierte Ergebnisse zu erhalten, ist hierbei eine Anpassung der Trainingsdaten an die Eigenschaften der Gebiete notwendig, auf die das Verfahren angewandt werden soll. Ist dies erfolgt, kann durch die Erhöhung der Menge der Bilder eine weitere Verbesserung erreicht werden.

#### 4.4. Variation der Sensitivität des Sensors

Für DNB wird nominell ein minimal detektierbares Signal von  $3 \times 10^{-9} \text{ Wcm}^{-2}\text{sr}^{-1}$  angegeben, in der Praxis werden aber auch noch deutlich kleinere Strahldichten von bis zu  $2 \times 10^{-10} \text{ Wcm}^{-2}\text{sr}^{-1}$  detektiert [31]. Sollen wie bei N8 Bilder mit einer deutlich höheren räumlichen Auflösung erzeugt werden, geht dies mit Einbußen bei der Sensitivität einher. Eine realistische Sensitivität für eine solche Mission liegt in der Größenordnung von  $2.5 \times 10^{-8} \text{ Wcm}^{-2}\text{sr}^{-1}$ , siehe hierzu auch [13]. Dies entspricht in etwa der Größenordnung des von der Erdoberfläche reflektierten Mondlichtes, sodass Strukturen von schwach reflektierenden Flächen nicht detektiert werden. Generell gehen durch die Erhöhung des minimal detektierbaren Signals (Textur-)Informationen verloren und feine Unterschiede in der Strahldichte können nicht mehr aufgelöst werden.

Da das in dieser Arbeit entwickelte Verfahren sehr stark auf der Analyse der Grauwertunterschiede der lokalen Nachbarschaft basiert, stellt sich die Frage, inwiefern es bei Anwendung auf Daten mit einer geringeren Sensitivität noch funktioniert.

Um dies zu überprüfen, wurden durch Rundung der originalen DNB-Strahldichten des Trainings- und Testdatensatzes fünf verschiedene DNB-Datensätze mit unterschiedlichen Sensitivitäten simuliert. Die Ergebnisse der semantischen Segmentierung anhand dieser simulierten DNB-Daten werden in Abbildung 4.13 dargestellt.

Bei einer Rundung der beobachteten Strahldichten auf  $10^{-9} \text{ Wcm}^{-2}\text{sr}^{-1}$  wird für das Gesamtergebnis sowie für fast alle Testgebiete das beste Ergebnis erzielt, wobei die Unterschiede zur Originalversion klein sind. Eine Ursache für diese Verbesserung gegenüber der Verwendung der Originaldaten kann sein, dass durch die Rundung auch störendes Rauschen eliminiert wurde, das sonst zum Beispiel durch seine Textur leicht mit Wolken verwechselt werden kann. Hierfür spricht ebenfalls, dass bei dieser Version insbesondere  $F_1(\textit{klar})$  (in dieser Grafik nicht dargestellt) verbessert wird.

Sinkt die Sensitivität um eine weitere Größenordnung, werden die Ergebnisse für alle Gebiete deutlich schlechter, da zu viele Details verloren gehen. Für  $10^{-7}$  und  $10^{-6} \text{ Wcm}^{-2}\text{sr}^{-1}$  wird schließlich kein reflektiertes Mondlicht mehr detektiert. Lediglich bei sehr hellen, künstlich beleuchteten Pixeln bleibt hier Bildinformation vorhanden, alle anderen Strahldichten sind konstant. Da diese hellen Pixel jedoch nur einen sehr kleinen Anteil an der Gesamtmenge der Pixel ausmachen, treten in der Menge der Trainings- und Testdaten fast ausschließlich konstante Merkmals-Werte auf. Deshalb sind die beiden Klassen nicht mehr voneinander trennbar. In diesen Fällen ist  $BA = 0.5$ , da alle Pixel als *wolkig* klassifiziert

#### 4.4. Variation der Sensitivität des Sensors

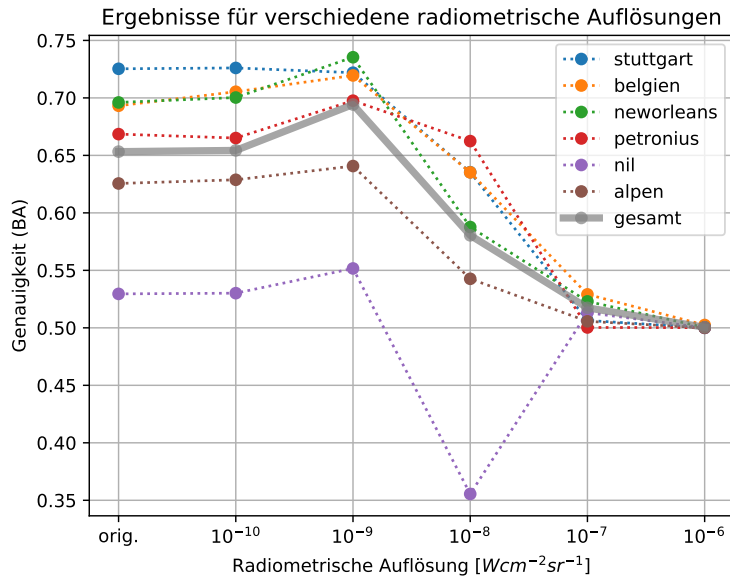


Abbildung 4.13: Ergebnisse (BA) für verschiedene radiometrische Sensitivitäten, aufgeschlüsselt nach Testgebieten.

werden, weil dies der Mehrheit der Pixel in den Trainingsdaten entspricht.

Die zur Anwendung des Verfahrens mindestens notwendige Sensitivität liegt also zwischen  $10^{-9}$  und  $10^{-8} Wcm^{-2}sr^{-1}$ .



## 5. Fazit und Ausblick

Die in dieser Arbeit vorgestellten Ergebnisse zeigen, dass Wolkenerkennung in nächtlichen Satellitenbildern des VNIR-Bereiches generell möglich ist. Für urbane Gebiete wie *stuttgart* und *belgien* können mit einer solchen Methode sehr gute Ergebnisse erzielt werden, die nahe an die für operationelle Verfahren wie ECM spezifizierten Genauigkeiten herankommen (bestes Ergebnis für *stuttgart* mit 85.4% Gesamtgenauigkeit im Vergleich zur ECM-Spezifikation von 88% über Land [19]). In manchen Fällen passt die in dieser Arbeit erzeugte Wolkenmaske sogar besser zum DNB-Bild als die ECM-Wolkenmaske. Für den Fokus der N8-Mission, nämlich die Analyse urbaner Gebiete, hat die Wolkenerkennung in VNIR-Aufnahmen also definitiv Potenzial.

Generell liegt die Stärke des Verfahrens bei der Erkennung dicker, großflächiger Bewölkung. Fehler treten vermehrt bei sehr dünner Bewölkung auf, diese ist jedoch generell schwieriger zu erkennen, weshalb hier auch die operationellen Verfahren weniger zuverlässig arbeiten [19].

Bei der Wolkenerkennung für Testgebiete ohne künstliche Beleuchtung oder mit besonderen Landbedeckungen (Wüste, Schnee, Ozean) treten größtenteils unabhängig von der Art der Bewölkung einige Schwierigkeiten auf. Eine Ursache für die schlechtere Performance des Verfahrens für solche Gebiete ist die fehlende künstliche Beleuchtung, wodurch insbesondere die verwendeten Kontrastmerkmale weniger aussagekräftig sind. Lediglich dicke, großflächige Bewölkung wird auch hier sehr zuverlässig erkannt, zumindest bei vorhandener Mondbeleuchtung. In diesen Fällen spielt der Effekt der künstlichen Beleuchtung auch bei urbanen Gebieten keine Rolle bei der Wolkenerkennung, da die Bewölkung diese komplett verdeckt. Generell sind die Ergebnisse bei den getesteten nicht-urbanen Gebieten oft stärker von der Mondbeleuchtung abhängig, als es bei urbanen Gebieten der Fall ist.

Unabhängig von der Genauigkeit der Ergebnisse hat das vorgestellte Verfahren zwei Vorteile gegenüber auf TIR basierten Verfahren wie ECM. Zum einen werden durch die Unabhängigkeit von Zusatzdaten auch einige mögliche Fehlerquellen eliminiert. Zum anderen könnte das Verfahren als Ergänzung von TIR-basierten Verfahren dienen, um zum Beispiel tief liegende Wolken besser zu erkennen. Für diese ist TIR nicht sensitiv, da ihre Temperatur sich zu wenig von der Temperatur der Erdoberfläche unterscheidet [37]. Eine Ergänzung durch ein Verfahren, das VNIR berücksichtigt, könnte hier zu Verbesserungen führen.

Zusammenfassend lässt sich sagen, dass das in dieser Arbeit vorgestellte Verfahren bisher für eine globale Anwendung zu wenig generalisiert ist, aber für bestimmte Interessensgebiete verwendet werden kann, indem es speziell auf diese trainiert wird (z.B. auf mittelgroße europäische Städte). Um das Ziel eines global anwendbaren VNIR-Wolkenerkennungsverfahrens zu erreichen, muss das in dieser Arbeit vorgeschlagene Verfahren noch weiterentwickelt werden, wofür im Folgenden einige mögliche Ansätze vorgestellt werden.

### 5.1. Ausblick

Die einfachste Möglichkeit zur Verbesserung der Ergebnisse für diejenigen Testgebiete, bei denen das Verfahren schon relativ gut funktioniert, ist die Integration einer Nachverarbeitung der Wolkenmasken. [21] und [24] implementieren hierfür ein Verfahren, das das Ergebnis für einzelne Pixel, die einer anderen Klasse zugeordnet wurden, als der Großteil ihrer Nachbarn, an ihre Nachbarschaft anpasst. Die Wolkenmasken werden also geglättet und Ausreißer entfernt. Zu erwägen wäre hierbei, ob ein solches Verfahren nur für isolierte Pixel der Klasse *klar* angewandt wird, da kleine, absolut wolkenfreie Lücken in Wolken deutlich unwahrscheinlicher sind, als das Auftreten von einzelnen kleinen Wolken mit einer Ausdehnung in der Größenordnung eines Pixels.

In einen solchen Nachverarbeitungsschritt könnte zudem die vom Random Forest ausgegebene prädizierte Wolkenwahrscheinlichkeit integriert werden, um die Entscheidung über die Anpassung der Klasse eines Pixels an dessen Nachbarschaft zu stützen. Liegt die Wolkenwahrscheinlichkeit nahe an 50%, erfolgte die Klassifikation nur sehr unsicher und eine Anpassung an die mit höherer Sicherheit klassifizierte Nachbarschaft ist wahrscheinlich sinnvoll. Umso höher die Sicherheit bei der Klassifikation des isolierten Pixels ist, umso größer ist wiederum die Wahrscheinlichkeit, dass das Beibehalten der ursprünglichen Klasse die bessere Entscheidung ist.

Um die Ergebnisse für alle getesteten Gebiete zu verbessern, muss die Generalisierung des Verfahrens erhöht werden. Dies betrifft zum einen die verwendeten Merkmale, die in dieser Arbeit mit dem Fokus auf die Beschreibung der durch Wolken verursachten Effekte über urbanen Gebieten ausgewählt wurden. Hier ist die Entwicklung neuer Merkmale nötig, die auch für andere Landbedeckungen aussagekräftig sind. Zum anderen ist ein generalisierteres Training notwendig, was vor allem die Auswahl der Trainingsdaten, aber auch das Design des Random Forest betrifft. Besonders aufgrund der hohen Abhängigkeit der Ergebnisse von der Art der Trainingsdaten bedarf es genauerer Untersuchungen zum idealen Aufbau des Trainingsdatensatzes.

Neben diesen Aspekten, die basierend auf der Grundstruktur des vorgeschlagenen Verfahrens zu Verbesserungen führen können, haben auch grundlegende Veränderungen an der Struktur des Verfahrens Potenzial. So erfordert die hohe Diversität der Daten gegebenenfalls ein komplexeres Klassifikator-Design.

Ein Ansatz hierfür ist, nicht alle Daten mit dem gleichen Klassifikator zu behandeln, sondern das Verfahren aus spezialisierteren Teil-Klassifikatoren aufzubauen, die jeweils mit einer bestimmten Untermenge der Daten trainiert werden und gegebenenfalls auch verschiedene, spezialisierte Merkmale berücksichtigen. Anhand von Fallunterscheidungen kann anschließend bei der Anwendung des Verfahrens der jeweils geeignetste Klassifikator verwendet werden. Eine mögliche Fallunterscheidung ist eine Trennung nach der Stärke der Mondbeleuchtung, im einfachsten Fall durch jeweils einen Klassifikator für vorhandene und nicht vorhandene Mondbeleuchtung. Dies kann den Einfluss der Mondbeleuchtung auf die Ergebnisse verringern, da die in Abhängigkeit der Mondbeleuchtung teilweise gegensätzlichen Effekte von Wolken nicht mehr durch einen einzigen Klassifikator gemeinsam

berücksichtigt werden müssen. Insbesondere bei den Bildserienmerkmalen ist es vorteilhaft, wenn beim Vergleich mit vorangegangenen Bildern ausschließlich Bilder ähnlicher Mondbeleuchtung berücksichtigt werden, um eine möglichst ähnliche Vergleichsgrundlage zu erhalten (vgl. Kapitel 4.2.3).

Eine weitere mögliche Fallunterscheidung ist die Trennung nach urbanen und nicht-urbanen Gebieten in Anlehnung an die Berücksichtigung der Landbedeckung beim ECM-Algorithmus [19]. Hierdurch können zum Beispiel Kontrastmerkmale nur für urbane Gebiete eingesetzt werden und für die nicht-urbanen Gebiete kann der Fokus stärker auf Merkmale, die zeitliche Veränderungen berücksichtigen, gelegt werden, da diese dort besser funktionieren. Voraussetzung hierfür ist, dass Informationen über die Urbanität eines Gebietes entweder direkt aus den Daten abgeleitet werden können oder aus Zusatzdaten zur Verfügung stehen.

Generell kann die Verwendung von zusätzlichen Informationen zur Verbesserung des Verfahrens eingesetzt werden, wobei dabei jedoch der Vorteil der Unabhängigkeit von weiteren Daten verloren geht. So können Informationen zur Landbedeckung nicht nur zur Unterscheidung verschiedener Klassifikatoren verwendet werden, sondern auch durch eine Verwendung bei der Merkmalsberechnung berücksichtigt werden. Eine weitere zusätzliche Datenquelle können globale Wetter- bzw. Bewölkungsdaten sein, die je nach verfügbarer Auflösung zumindest Tendenzen über die Wahrscheinlichkeit von Bewölkung im betrachteten Gebiet liefern. Solche Informationen könnten auch in der Nachverarbeitung als Entscheidungshilfe für unsicher klassifizierte Pixel dienen. Wetter- und Landbedeckungsdaten werden unter anderem auch beim ECM-Algorithmus [19] berücksichtigt.

Es besteht somit noch einiges an Potenzial, um die Wolkenerkennung anhand der derzeit verfügbaren panchromatischen VNIR-Daten zu verbessern. Mit N8 stünden jedoch zusätzlich zu panchromatischen Bildern auch spektrale Informationen zur Verfügung (z.B. 5 zusätzliche Kanäle im VNIR-Bereich [35]), die ebenfalls zur Wolkenerkennung genutzt werden können. So können zum Beispiel bei urbanen Gebieten die Unterschiede der Beobachtungen der verschiedenen spektralen Kanäle als Merkmale herangezogen werden, da das Spektrum von an Wolken reflektiertem Mondlicht im gesamten VNIR-Bereich relativ gleichverteilt ist, wohingegen das von künstlicher Beleuchtung emittierte Licht sehr spezifische Spektren besitzt. Bei dünner Bewölkung müssen hierbei jedoch Dämpfungseffekte und die Mischung von künstlichem Licht und reflektiertem Mondlicht beachtet werden. Generell sollte insbesondere auch für nicht-urbane Gebiete untersucht werden, inwiefern für die Anwendung bei Tag gängige Wolkenerkennungsverfahren, die auf spektralen Informationen im VNIR-Bereich basieren, auf die Anwendung bei Nacht übertragen werden können.

Insgesamt ist das in dieser Arbeit vorgestellte Verfahren also ein einfaches, auf urbane Gebiete optimiertes Basis-Verfahren, das durch die angesprochenen Ergänzungen weiter verbessert und generalisiert werden kann. Die bereits mit diesem Verfahren erzielbaren Ergebnisse zeigen, dass die Entwicklung eines Verfahrens zur Wolkenerkennung anhand nächtlicher VNIR-Aufnahmen definitiv Potenzial hat und weiter verfolgt werden sollte.



## Literatur

- [1] Ackerman, Steve, Kathleen Strabala, Paul Menzel, Richard Frey, Chris Moeller, Liam Gumley, Bryan Baum, Crystal Schaaf und George Riggs (2010). *Discriminating clear-sky from cloud with MODIS. Algorithm theoretical basis document (MOD35)*. Techn. Ber. Version 6.1.
- [2] Baker, Neal (2017). *Joint Polar Satellite System (JPSS) VIIRS Geolocation Algorithm Theoretical Basis Document (ATBD)*. Techn. Ber. Revision A. NOAA NESDIS Center for Satellite Applications und Research.
- [3] Baugh, Kimberly, Feng-Chi Hsu, Christopher Elvidge und Mikhail Zhizhin (2013). „Nighttime Lights Compositing Using the VIIRS Day-Night Band: Preliminary Results“. In: *Proceedings of the Asia-Pacific Advanced Network* 35, S. 70–86. DOI: 10.7125/APAN.35.8.
- [4] Bergstra, James und Y. Bengio (2012). „Random Search for Hyper-Parameter Optimization“. In: *The Journal of Machine Learning Research* 13, S. 281–305.
- [5] Bishop, Christopher M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc. ISBN: 0387310738.
- [6] Breiman, Leo (1984). *Classification and Regression Trees*. New York: Routledge.
- [7] Breiman, Leo (1996). „Bagging predictors“. In: *Machine Learning* 24.2, S. 123–140. ISSN: 1573-0565. DOI: 10.1007/BF00058655.
- [8] Breiman, Leo (2001). „Random Forests“. In: *Machine Learning* 45.1, S. 5–32. ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.
- [9] Center for Satellite Applications and Research (STAR). *NOAA-20 Visible Infrared Imaging Radiometer Suite (VIIRS)*. <https://ncc.nesdis.noaa.gov/NOAA-20/NOAA20VIIRS.php>. Zuletzt besucht: 20.10.2019.
- [10] Changyong, Cao, (Jack) Xiong Xiaoxiong, Wolfe Robert, DeLuccia Frank, (Mark) Liu Qianhua, Blonski Slawomir, (Gary) Lin Guoqing, Nishihama Masahiro, Pogorzala Dave, Oudrari Hassan und Don Hillger (2017). *Visible Infrared Imaging Radiometer Suite (VIIRS) Sensor Data Record (SDR) User’s Guide*. Techn. Ber. Version 1.3. Washington, DC: U.S. Department of Commerce, National Oceanic and Atmospheric Administration.
- [11] Derrien, M. und H. Le Gléau (2010). „Improvement of cloud detection near sunrise and sunset by temporal-differencing and region-growing techniques with real-time SEVI-

## Literatur

- RI“. In: *International Journal of Remote Sensing* 31.7, S. 1765–1780. DOI: 10.1080/01431160902926632. eprint: <https://doi.org/10.1080/01431160902926632>.
- [12] Elvidge, Christopher D, Kimberly Baugh, Mikhail Zhizhin, Feng Chi Hsu und Tilotama Ghosh (2017). „VIIRS night-time lights“. In: *International Journal of Remote Sensing* 38.21, S. 5860–5879. DOI: 10.1080/01431161.2017.1342050. eprint: <https://doi.org/10.1080/01431161.2017.1342050>.
- [13] Elvidge, Christopher, Pierantonio Cinzano, D. Pettit, J. Arvesen, Paul Sutton, Christopher Small, Ramakrishna Nemani, Travis Longcore, C. Rich, J. Safran und Steeve Ebenner (2007). „The Nightsat mission concept“. In: *International Journal of Remote Sensing - INT J REMOTE SENS* 28, S. 2645–2670. DOI: 10.1080/01431160600981525.
- [14] Frey, Richard, Steve Ackermann, Robert Holz und Steve Dutcher (2019). *The Continuity MODIS-VIIRS Cloud Mask (MVCN) User Guide*. Techn. Ber. Version 1.0. Space Science und Engineering Center, University of Wisconsin-Madison.
- [15] Godin, Ray (2014). *Joint Polar Satellite System (JPSS) VIIRS Cloud Mask (VCM) Algorithm Theoretical Basis Document (ATBD)*. Techn. Ber. Version Revision E. NASA/Goddard Space Flight Center, Greenbelt, Maryland.
- [16] Goward, Samuel N. (2005). „Albedo and Reflectivity“. In: *Encyclopedia of World Climatology*. Hrsg. von John E. Oliver. Dordrecht: Springer Netherlands, S. 32–35. ISBN: 9781402032660. DOI: 10.1007/1-4020-3266-8\_8.
- [17] Hänsch, Ronny und Olaf Hellwich (2017). „Random Forests“. In: *Photogrammetrie und Fernerkundung: Handbuch der Geodäsie, herausgegeben von Willi Freuden und Reiner Rummel*. Christian Heipke, S. 603–643. ISBN: 9783662470930. DOI: 10.1007/978-3-662-47094-7\_46.
- [18] Haralick, Robert M., K. Sam Shanmugam und Its’hak Dinstein (1973). „Textural Features for Image Classification“. In: *IEEE Trans. Systems, Man, and Cybernetics* 3, S. 610–621.
- [19] Heidinger, Andrew, Denis Botambekov und Andi Walther (2016). *A Naive Bayesian Cloud Mask Delivered to NOAA Enterprise, Algorithm Theoretical Basis Document*. Techn. Ber. Version 1.1. NOAA NESDIS Center for Satellite Applications und Research.
- [20] Heidinger, Andrew, Thomas Kopp und Denis Botambekov (2016). „Enterprise Cloud Mask (ECM)“.
- [21] Hu, Shensen, Shuo Ma, Wei Yan, Jun Jiang und Yunxian Huang (2017). „A new multichannel threshold algorithm based on radiative transfer characteristics for detecting fog/low stratus using night-time NPP/VIIRS data“. In: *International Journal of Re-*

- mote Sensing* 38.21, S. 5919–5933. DOI: 10.1080/01431161.2016.1265691. eprint: <https://doi.org/10.1080/01431161.2016.1265691>.
- [22] Hunt, G. E. (1973). „Radiative properties of terrestrial clouds at visible and infrared thermal window wavelengths“. In: *Quarterly Journal of the Royal Meteorological Society* 99.420, S. 346–369. DOI: 10.1002/qj.49709942013. eprint: <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49709942013>.
- [23] Jedlovec, Gary (2009). „Automated Detection of Clouds in Satellite Imagery“. In: ISBN: 9789533070056. DOI: 10.5772/8326.
- [24] Jiang, Jun, Wei Yan, Shuo Ma, Yangyang Jie, Xiarong Zhang, Shensen Hu, Lei Fan und Linyu Xia (2015). „Three Cases of a New Multichannel Threshold Technique to Detect Fog/Low Stratus during Nighttime Using SNPP Data“. In: *Weather and Forecasting* 30.6, S. 1763–1780. DOI: 10.1175/WAF-D-15-0050.1. eprint: <https://doi.org/10.1175/WAF-D-15-0050.1>.
- [25] Johnson, R. S., J. Zhang, E. J. Hyer, S. D. Miller und J. S. Reid (2013). „Preliminary investigations toward nighttime aerosol optical depth retrievals from the VIIRS Day/Night Band“. In: *Atmospheric Measurement Techniques* 6.5, S. 1245–1255. DOI: 10.5194/amt-6-1245-2013.
- [26] Kokhanovsky, A. (2004). „Optical properties of terrestrial clouds“. In: *Earth-Science Reviews* 64.3, S. 189–241. ISSN: 0012-8252. DOI: [https://doi.org/10.1016/S0012-8252\(03\)00042-4](https://doi.org/10.1016/S0012-8252(03)00042-4).
- [27] Kotarba, Andrzej Z. und Sebastian Aleksandrowicz (2016). „Impervious surface detection with nighttime photography from the International Space Station“. In: *Remote Sensing of Environment* 176, S. 295–307. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2016.02.009>.
- [28] Lee, Thomas E., Steven D. Miller, F. Joseph Turk, Carl Schueler, Richard Julian, Steve Deyo, Patrick Dills und Sherwood Wang (2006). „The NPOESS VIIRS Day/Night Visible Sensor“. In: *Bulletin of the American Meteorological Society* 87.2, S. 191–200. DOI: 10.1175/BAMS-87-2-191. eprint: <https://doi.org/10.1175/BAMS-87-2-191>.
- [29] Li, Xi, Christopher Elvidge, Yuyu Zhou, Changyong Cao und Timothy Warner (2017). „Remote sensing of night-time light“. In: *International Journal of Remote Sensing* 38.21, S. 5855–5859. DOI: 10.1080/01431161.2017.1351784. eprint: <https://doi.org/10.1080/01431161.2017.1351784>.
- [30] Li, Xi, Xiya Li, Deren Li, Xiaojun He und Michael Jendryke (2019). „A preliminary investigation of LuoJia-1 night-time light imagery“. In: *Remote Sensing Letters* 10.6, S. 526–535. DOI: 10.1080/2150704X.2019.1577573. eprint: <https://doi.org/10.1080/2150704X.2019.1577573>.

## Literatur

- [31] Liao, L., Stephanie Weiss, Steve Mills und Bruce Hauss (2013). „Suomi NPP VIIRS day-night band on-orbit performance“. In: *Journal of Geophysical Research: Atmospheres* 118. DOI: 10.1002/2013JD020475.
- [32] L.Imhoff, Marc, William T. Lawrence, David C. Stutzer und Christopher D. Elvidge (1997). „A technique for using composite DMSP/OLS City Lights satellite data to map urban area“. In: *Remote Sensing of Environment* 61.3, S. 361–370. ISSN: 0034-4257. DOI: [https://doi.org/10.1016/S0034-4257\(97\)00046-1](https://doi.org/10.1016/S0034-4257(97)00046-1).
- [33] Lyapustin, A., Y. Wang und R. Frey (2008). „An automatic cloud mask algorithm based on time series of MODIS measurements“. In: *Journal of Geophysical Research: Atmospheres* 113.D16. DOI: 10.1029/2007JD009641.
- [34] McCluney, William Ross (2014). *Introduction to radiometry and photometry*. 2. ed. Artech House applied photonics series. Boston u.a.: Artech House.
- [35] Meester, Jasper De (2019). „Simulation of representative nocturnal satellite imagery for urban areas with high spectral and high spatial resolution“. Magisterarb. Leibniz Universität Hannover.
- [36] Miller, S. D. und R. E. Turner (2009). „A Dynamic Lunar Spectral Irradiance Data Set for NPOESS/VIIRS Day/Night Band Nighttime Environmental Applications“. In: *IEEE Transactions on Geoscience and Remote Sensing* 47.7, S. 2316–2329. ISSN: 0196-2892. DOI: 10.1109/TGRS.2009.2012696.
- [37] Miller, Steven, Stephen Mills, Christopher Elvidge, Daniel Lindsey, Thomas Lee und Jeffrey Hawkins (2012). „Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities“. In: *Proceedings of the National Academy of Sciences of the United States of America* 109, S. 15706–11. DOI: 10.1073/pnas.1207034109.
- [38] Miller, Steven, William Straka, Stephen Mills, Christopher Elvidge, Thomas Lee, Jeremy Solbrig, Andi Walther, Andrew Heidinger und Stephanie Weiss (2013). „Illuminating the capabilities of the suomi national polar-orbiting partnership (NPP) visible infrared imaging radiometer suite (VIIRS) day/night band“. In: *Remote Sensing* 5.12, S. 6717–6766.
- [39] Pack, Dee W. und Brian S. Hardy (2016). „CUBESAT Nighttime Lights“. In: *Proceedings of the 30th annual AIAA/USU Conference on Small Satellites, Logan, UT, USA*. DigitalCommons, The Aerospace Corporation.
- [40] Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot und E. Duchesnay (2011). „Scikit-learn: Machine Learning in Python“. In: *Journal of Machine Learning Research* 12, S. 2825–2830.

- [41] Rees, Gareth [Verfasser], Hrsg. (2013). *Physical principles of remote sensing*. Englisch. 3. ed. Includes bibliographical references and index. Cambridge [u.a.]: Cambridge Univ. Pr., XVII, 441 Seiten. ISBN: 052118116X.
- [42] Román, Miguel O. und Eleanor C. Stokes (2015). „Holidays in lights: Tracking cultural patterns in demand for energy services“. In: *Earth's Future* 3.6, S. 182–205. DOI: 10.1002/2014EF000285. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014EF000285>.
- [43] Román, Miguel O., Zhuosen Wang, Qingsong Sun, Virginia Kalb, Steven D. Miller, Andrew Molthan, Lori Schultz, Jordan Bell, Eleanor C. Stokes, Bhartendu Pandey, Karen C. Seto, Dorothy Hall, Tomohiro Oda, Robert E. Wolfe, Gary Lin, Navid Gollpayegani, Sadashiva Devadiga, Carol Davidson, Sudipta Sarker, Cid Praderas, Jeffrey Schmaltz, Ryan Boller, Joshua Stevens, Olga M. Ramos González, Elizabeth Padilla, José Alonso, Yasmín Detrés, Roy Armstrong, Ismael Miranda, Yasmín Conte, Nitza Marrero, Kytta MacManus, Thomas Esch und Edward J. Masuoka (2018). „NASA's Black Marble nighttime lights product suite“. In: *Remote Sensing of Environment* 210, S. 113–143. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2018.03.017>.
- [44] Rossow, William B. und Leonid C. Garder (1993). „Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP“. In: *Journal of Climate* 6.12, S. 2341–2369. DOI: 10.1175/1520-0442(1993)006<2341:CDUSMO>2.0.CO;2. eprint: [https://doi.org/10.1175/1520-0442\(1993\)006<2341:CDUSMO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<2341:CDUSMO>2.0.CO;2).
- [45] Seaman, Curtis (2013). *Beginner's Guide to VIIRS Imagery Data*. [http://rammb.cira.colostate.edu/projects/npp/Beginner\\_Guide\\_to\\_VIIRS\\_Imagery\\_Data.pdf](http://rammb.cira.colostate.edu/projects/npp/Beginner_Guide_to_VIIRS_Imagery_Data.pdf). CIRA/Colorado State University, zuletzt besucht: 17.09.2019.
- [46] Seaman, Curtis J. und Steven D. Miller (2015). „A dynamic scaling algorithm for the optimized digital display of VIIRS Day/Night Band imagery“. In: *International Journal of Remote Sensing* 36.7, S. 1839–1854. DOI: 10.1080/01431161.2015.1029100. eprint: <https://doi.org/10.1080/01431161.2015.1029100>.
- [47] *SNAP-ESA Sentinel Application Platform*. <http://step.esa.int>. Version v6.0.0.
- [48] Ting, Kai Ming (2010a). „Confusion Matrix“. In: *Encyclopedia of Machine Learning*. Hrsg. von Claude Sammut und Geoffrey I. Webb. Boston, MA: Springer US, S. 209–209. ISBN: 9780387301648. DOI: 10.1007/978-0-387-30164-8\_157.
- [49] Ting, Kai Ming (2010b). „Precision and Recall“. In: *Encyclopedia of Machine Learning*. Hrsg. von Claude Sammut und Geoffrey I. Webb. Boston, MA: Springer US, S. 781–781. ISBN: 9780387301648. DOI: 10.1007/978-0-387-30164-8\_652.

## Literatur

- [50] Vaughan, M. A., D. M. Winker und K.A. Powell (2005). *CALIOP algorithm theoretical basis document. Part 2: Feature detection and layer properties algorithm*. Techn. Ber. Version 1.01.
- [51] Walt, Stéfan van der, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulougne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu und the scikit-image contributors (2014). „scikit-image: image processing in Python“. In: *PeerJ* 2, e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453.
- [52] Wang, Wenhui, Changyong Cao, Yan Bai, Slawomir Blonski und Mitchell A. Schull (2017). „Assessment of the NOAA S-NPP VIIRS Geolocation Reprocessing Improvements“. In: *Remote Sensing* 9, S. 974.
- [53] Xi Shao Changyong Cao, Sirish Uprety (2013). *Vicarious calibration of S-NPP/VIIRS day-night band*. DOI: 10.1117/12.2023412.
- [54] Zhao, Min, Yuyu Zhou, Xuecao Li, Wenting Cao, Chunyang He, Bailang Yu, Xi Li, Christopher D. Elvidge, Weiming Cheng und Chenghu Zhou (2019). „Applications of Satellite Remote Sensing of Nighttime Light Observations: Advances, Challenges, and Perspectives“. In: *Remote Sensing* 11.17. ISSN: 2072-4292. DOI: 10.3390/rs11171971.
- [55] Zheng, Qiming, Qihao Weng, Lingyan Huang, Ke Wang, Jinsong Deng, Ruowei Jiang, Ziran Ye und Muye Gan (2018). „A new source of multi-spectral high spatial resolution night-time light imagery—JL1-3B“. In: *Remote Sensing of Environment* 215, S. 300–312. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2018.06.016>.

## A. Weitere Ergebnisbilder

A. Weitere Ergebnisbilder

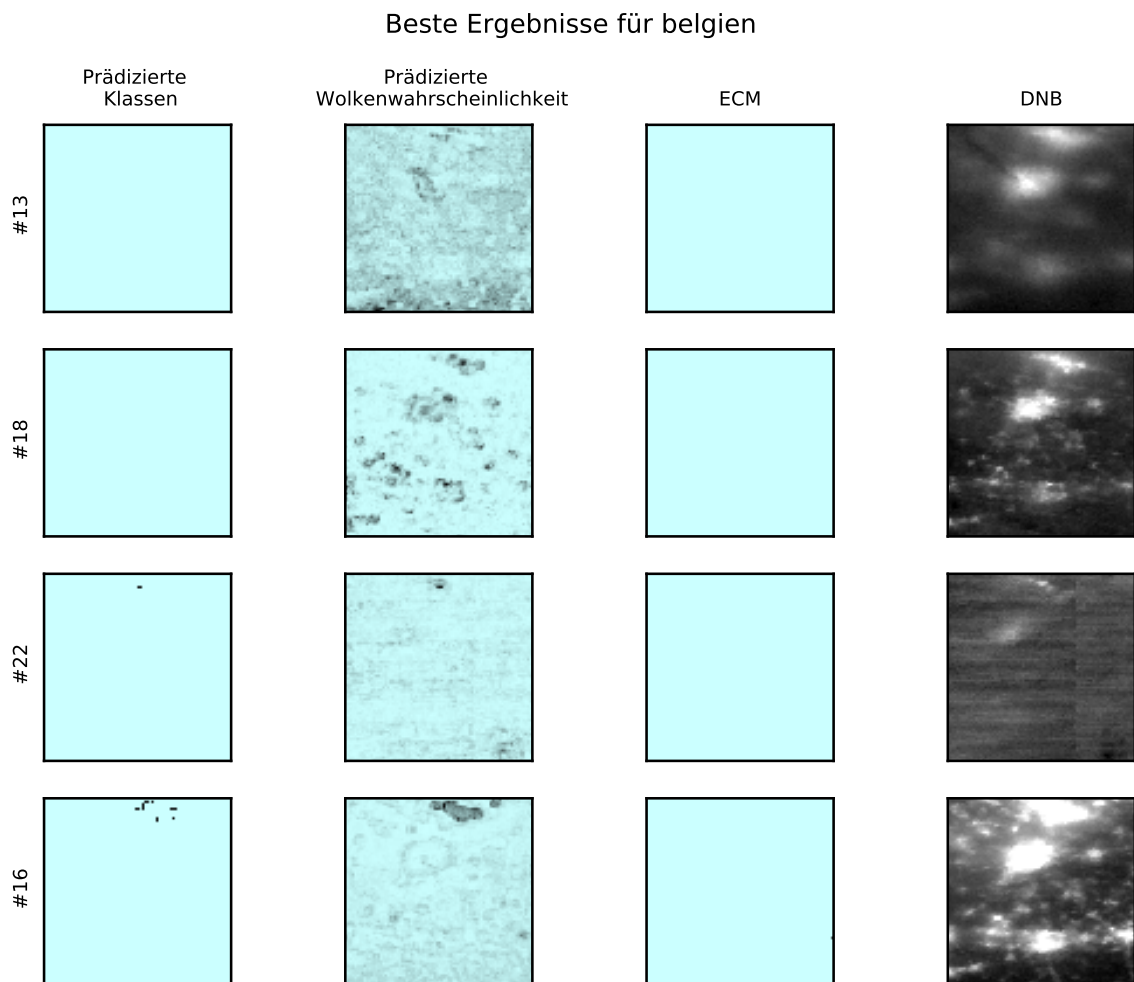


Abbildung A.1: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *belgien* mit den besten Ergebnissen. Blau ist die Klasse *wolzig*, schwarz die Klasse *klar*.

### Schlechteste Ergebnisse für belgien

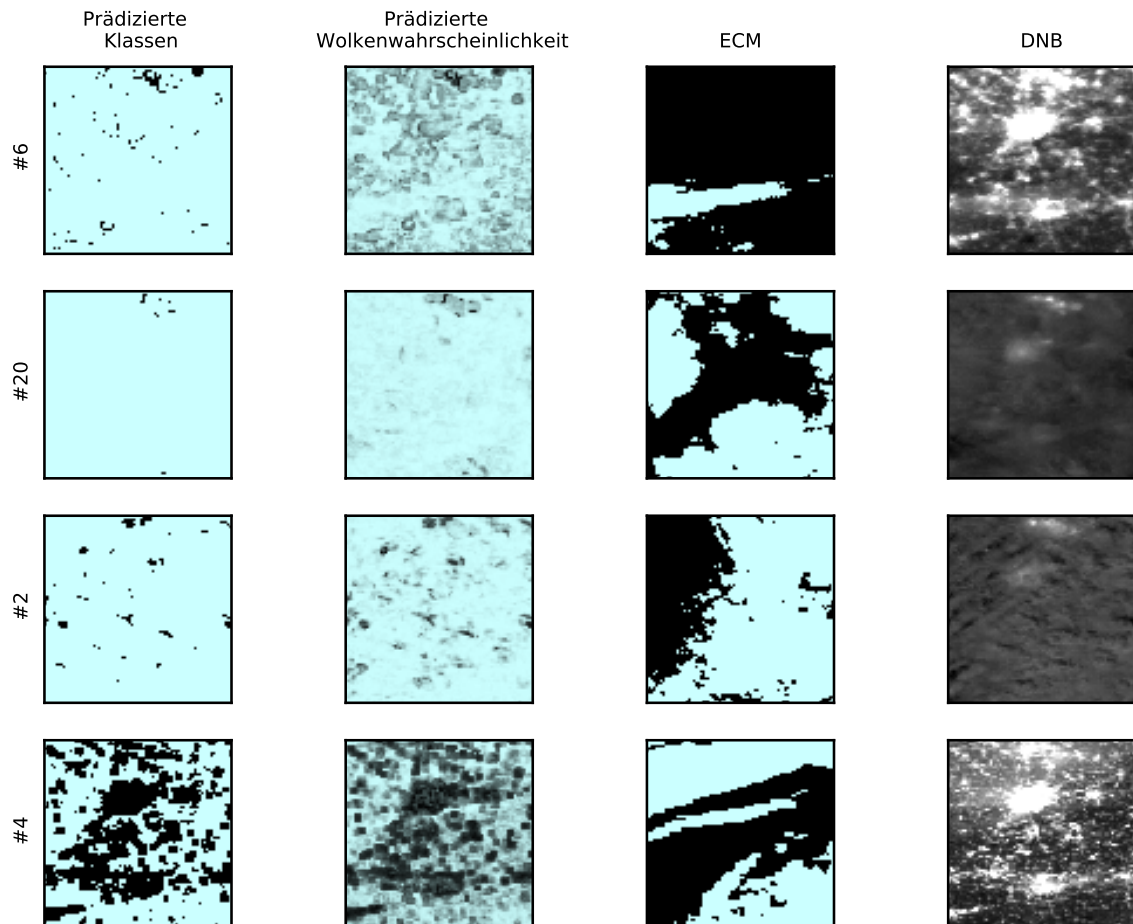


Abbildung A.2: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *belgien* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

A. Weitere Ergebnisbilder

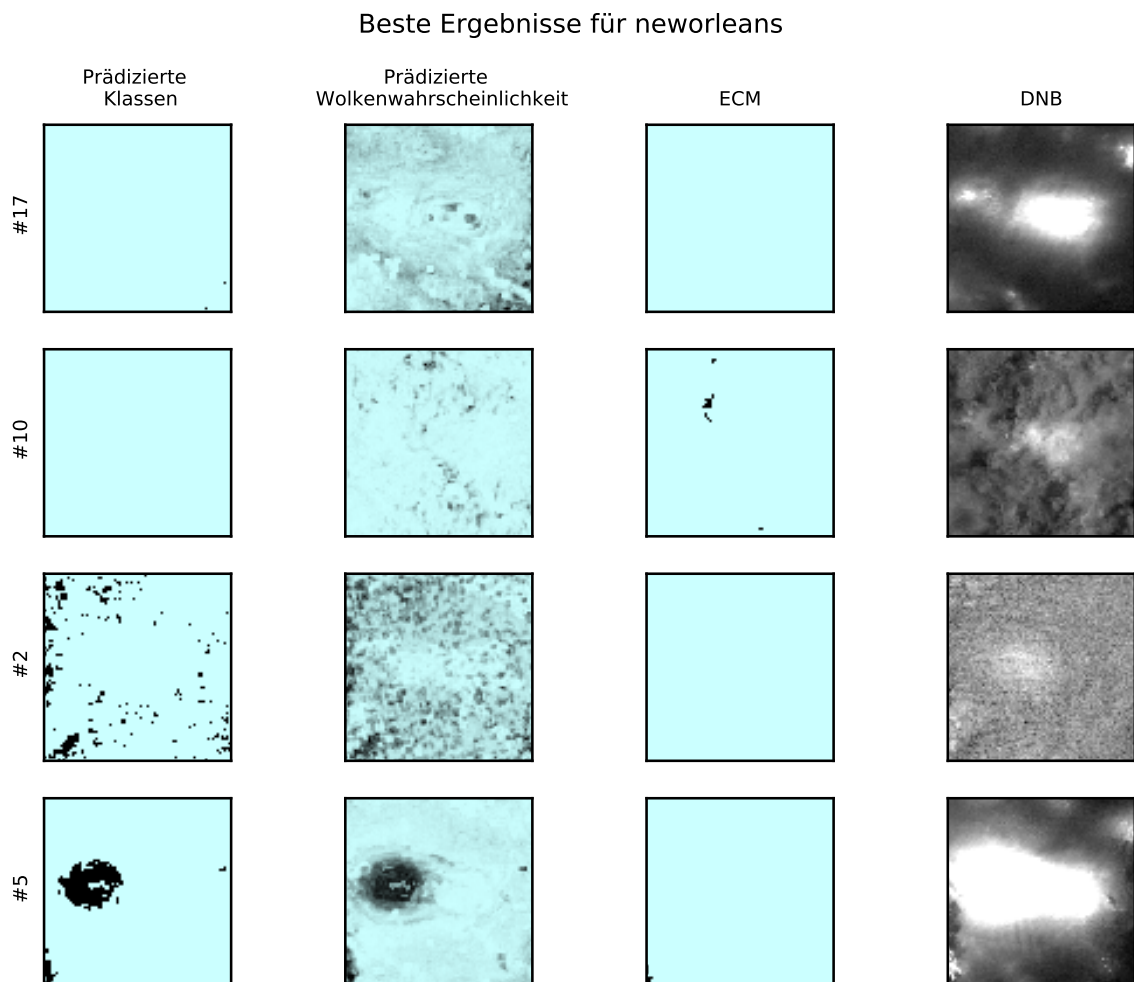


Abbildung A.3: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *neworleans* mit den besten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

Schlechteste Ergebnisse für neworleans

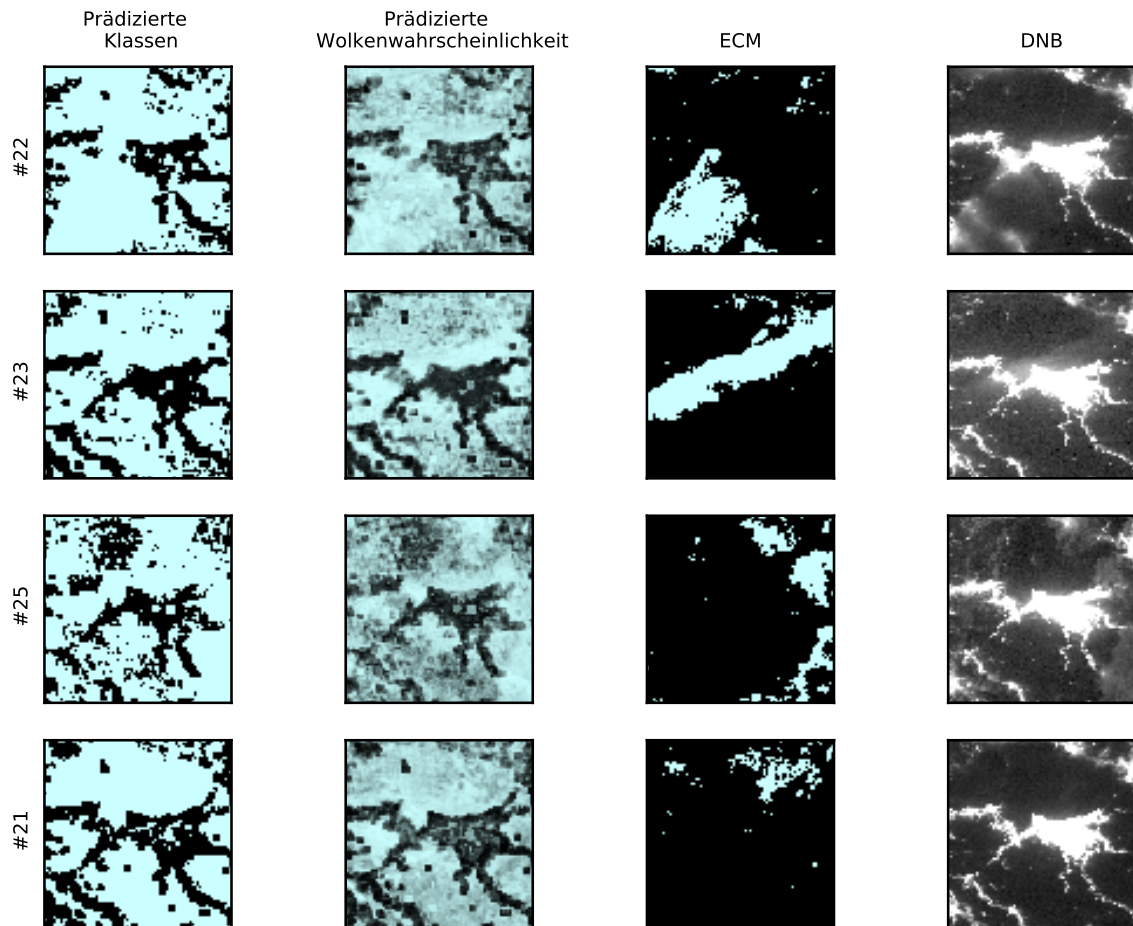


Abbildung A.4: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *neworleans* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

A. Weitere Ergebnisbilder

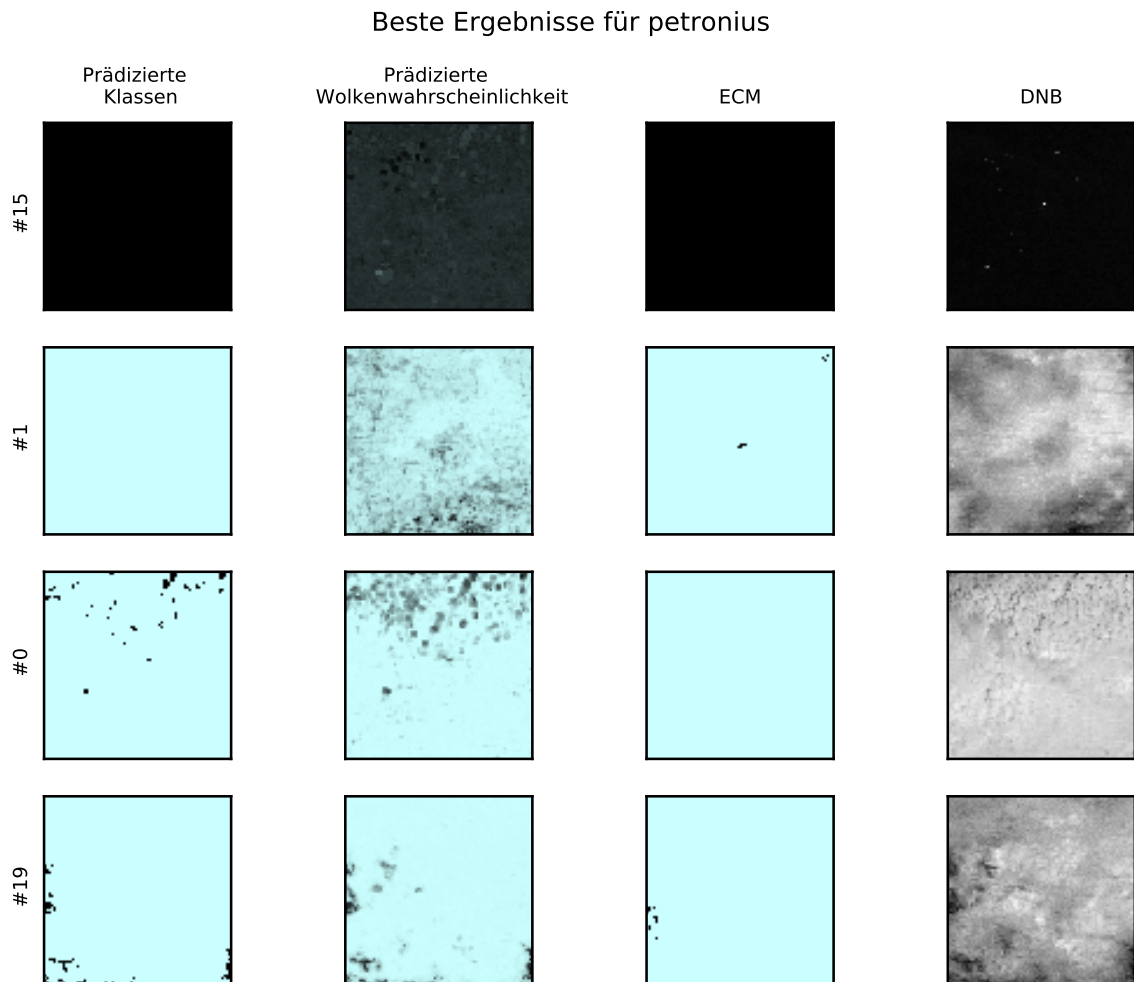


Abbildung A.5: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *petronius* mit den besten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

### Schlechteste Ergebnisse für petronius

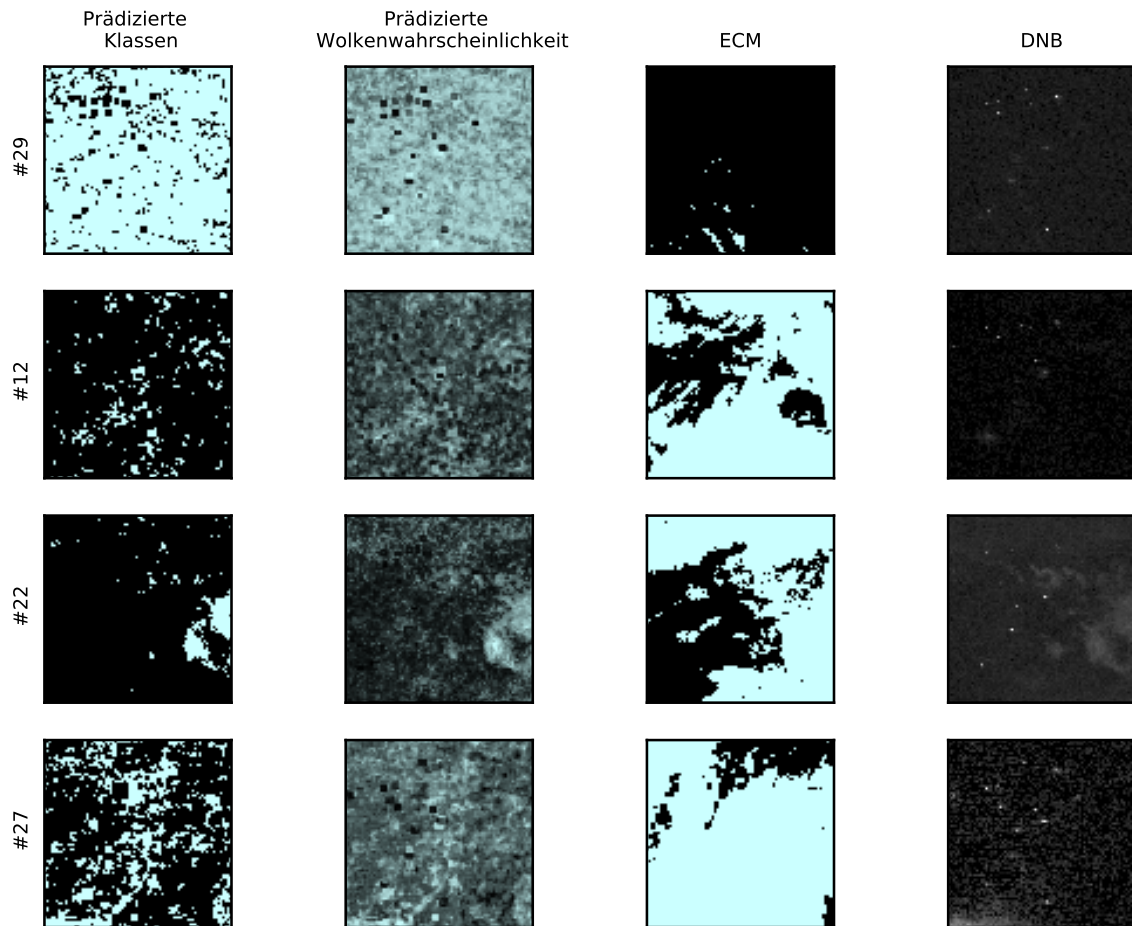


Abbildung A.6: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *petronius* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*. Die Wolken der ECM-Maske sind visuell größtenteils nicht in den DNB-Bildern erkennbar, da keine Mondbeleuchtung vorliegt.

## A. Weitere Ergebnisbilder

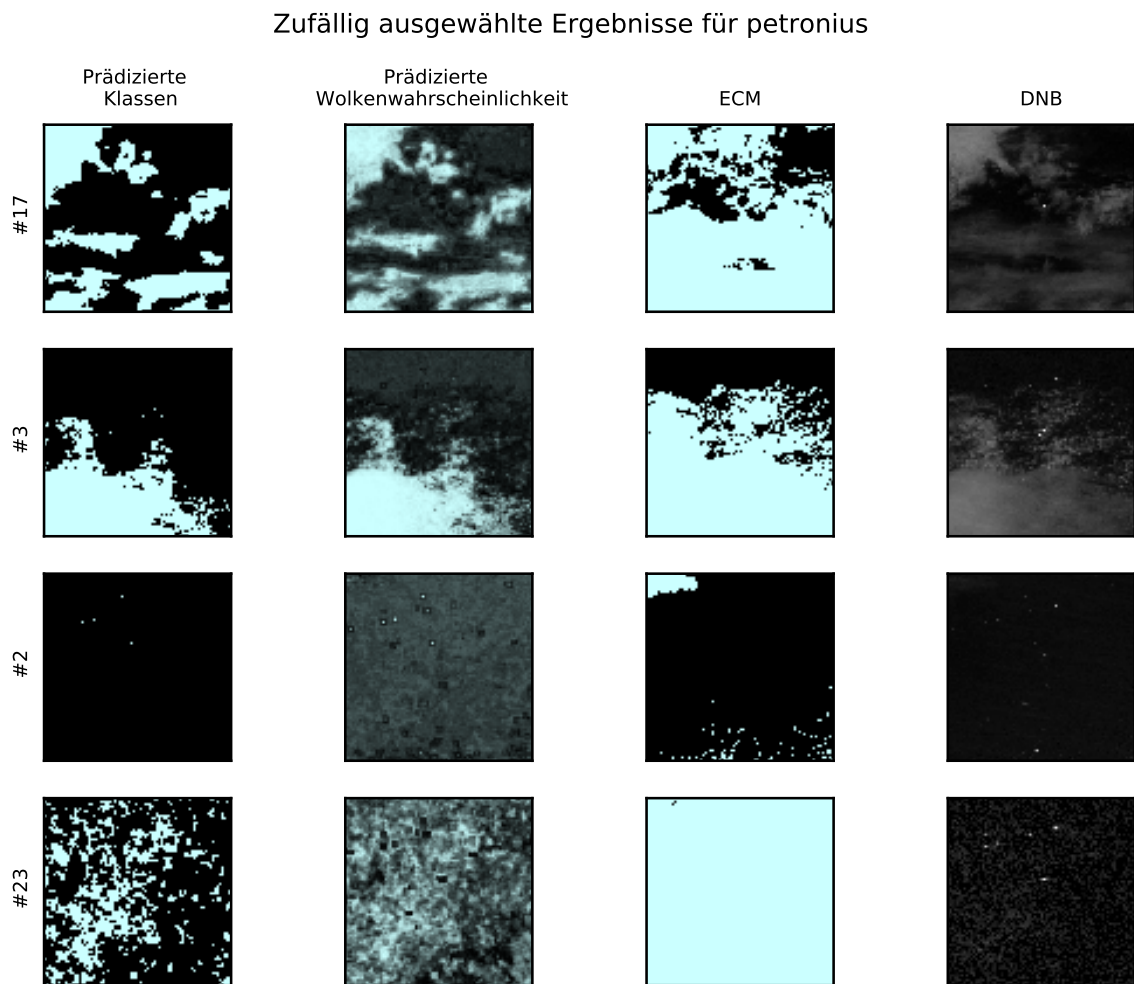


Abbildung A.7: Training mit *muenchen* (100 Bilder): Weitere Ergebnis-Bilder des Testgebietes *petronius*. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*. Bei den ersten beiden Bildern passt die Wolkenmaske sehr gut zu den im DNB-Bild sichtbaren Wolken.

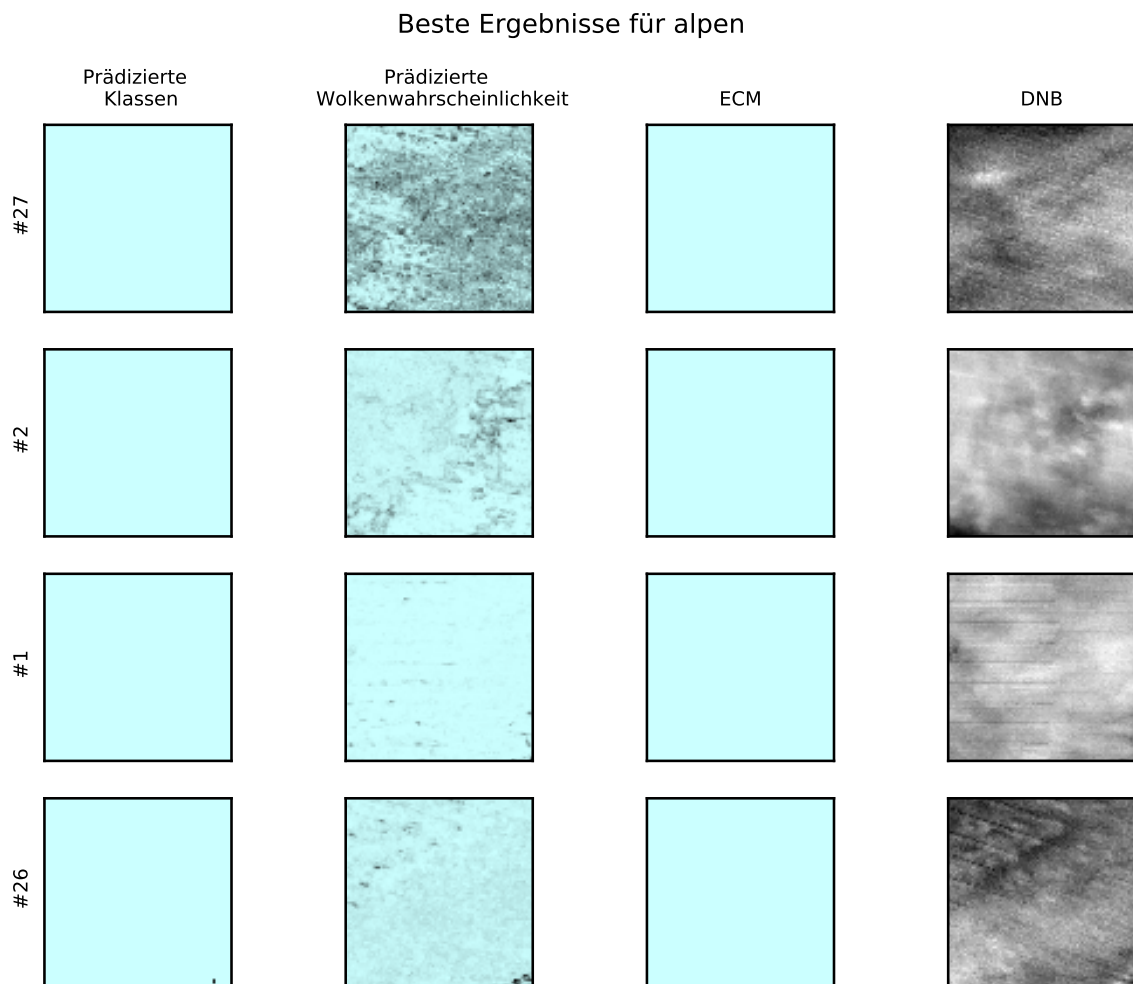


Abbildung A.8: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *alpen* mit den besten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

A. Weitere Ergebnisbilder

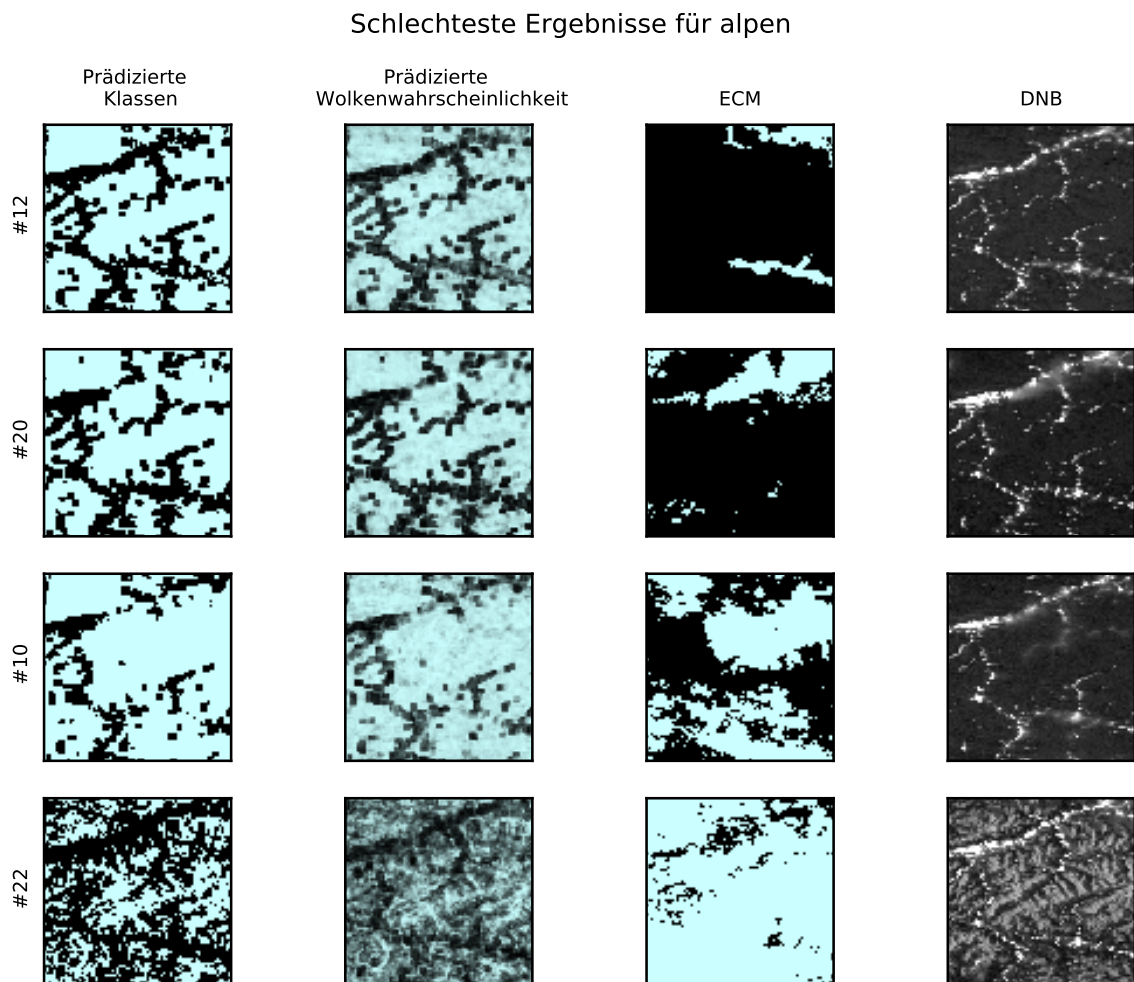


Abbildung A.9: Training mit *muenchen* (100 Bilder): Bilder des Testgebietes *alpen* mit den schlechtesten Ergebnissen. Blau ist die Klasse *wolkig*, schwarz die Klasse *klar*.

## B. Informationen über die verwendeten Bilder

B. Informationen über die verwendeten Bilder

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	01.01.2018	00:27	8,21E+07
1	04.01.2018	02:48	6,18E+07
2	05.01.2018	02:30	4,42E+07
3	06.01.2018	02:13	2,92E+08
4	10.01.2018	02:40	2,07E+07
5	11.01.2018	02:16	2,80E+06
6	12.01.2018	01:59	0
7	15.01.2018	02:43	0
8	17.01.2018	00:26	0
9	18.01.2018	00:08	0
10	19.01.2018	01:27	0
11	20.01.2018	01:10	0
12	21.01.2018	00:52	0
13	22.01.2018	00:29	0
14	23.01.2018	00:12	0
15	25.01.2018	01:13	0
16	26.01.2018	00:56	7,58E+05
17	27.01.2018	00:38	9,00E+06
18	28.01.2018	00:21	2,30E+08
19	31.01.2018	01:05	1,00E+09
20	01.02.2018	00:42	1,07E+06
21	02.02.2018	00:24	7,25E+06
22	03.02.2018	00:07	4,45E+07
23	05.02.2018	01:09	1,70E+08
24	06.02.2018	00:51	7,86E+06
25	07.02.2018	00:34	2,07E+05
26	08.02.2018	00:10	0
27	11.02.2018	00:55	0
28	12.02.2018	00:37	0
29	13.02.2018	00:20	0

Tabelle B.1: Bildinformationen für *stuttgart*.

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	01.01.2018	02:04	6,77E+07
1	02.01.2018	01:46	1,13E+09
2	03.01.2018	01:29	9,37E+07
3	05.01.2018	00:48	4,09E+08
4	06.01.2018	00:30	2,31E+08
5	07.01.2018	01:55	1,59E+08
6	08.01.2018	01:32	7,50E+05
7	09.01.2018	01:15	2,27E+06
8	11.01.2018	00:40	0
9	12.01.2018	01:59	0
10	13.01.2018	01:41	0
11	14.01.2018	01:24	0
12	17.01.2018	00:26	0
13	18.01.2018	01:45	0
14	19.01.2018	01:27	0
15	20.01.2018	01:10	0
16	24.01.2018	01:37	0
17	25.01.2018	01:13	0
18	28.01.2018	01:58	1,42E+08
19	30.01.2018	01:23	6,28E+07
20	31.01.2018	01:05	1,01E+09
21	01.02.2018	00:42	1,06E+09
22	02.02.2018	00:24	6,98E+07
23	03.02.2018	01:49	4,55E+07
24	04.02.2018	01:26	2,80E+08
25	05.02.2018	01:09	1,52E+08
26	06.02.2018	00:51	6,41E+06
27	07.02.2018	00:34	1,02E+06
28	09.02.2018	01:35	0
29	10.02.2018	01:18	0

Tabelle B.2: Bildinformationen für *belgien*.

B. Informationen über die verwendeten Bilder

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	06.01.2018	07:26	2,68E+08
1	07.01.2018	07:03	1,33E+08
2	08.01.2018	06:45	4,61E+07
3	20.01.2018	08:00	0
4	21.01.2018	07:42	0
5	22.01.2018	07:25	0
6	24.01.2018	08:26	0
7	25.01.2018	08:09	0
8	26.01.2018	07:46	0
9	27.01.2018	07:28	8,84E+06
10	28.01.2018	07:11	2,63E+08
11	31.01.2018	07:55	1,24E+09
12	01.02.2018	07:37	1,12E+09
13	03.02.2018	06:57	4,90E+07
14	05.02.2018	07:58	2,07E+07
15	06.02.2018	07:41	1,04E+06
16	07.02.2018	07:23	3,69E+07
17	12.02.2018	07:27	0
18	14.02.2018	06:52	0
19	15.02.2018	08:11	0
20	16.02.2018	07:54	0
21	17.02.2018	07:36	0
22	18.02.2018	07:19	0
23	19.02.2018	06:55	0
24	22.02.2018	07:40	0
25	24.02.2018	07:05	2,11E+07
26	28.02.2018	07:31	6,36E+07
27	01.03.2018	07:08	1,02E+09
28	02.03.2018	06:51	1,21E+08
29	03.03.2018	08:10	8,01E+07

Tabelle B.3: Bildinformationen für *neworleans*.

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	02.01.2018	06:59	1,39E+08
1	03.01.2018	06:41	1,02E+09
2	05.01.2018	07:43	4,62E+06
3	06.01.2018	07:26	2,77E+07
4	07.01.2018	07:03	1,42E+08
5	08.01.2018	06:45	5,25E+06
6	12.01.2018	07:12	0
7	13.01.2018	06:54	0
8	14.01.2018	06:31	0
9	20.01.2018	08:00	0
10	24.01.2018	08:26	0
11	25.01.2018	08:09	0
12	26.01.2018	07:46	0
13	27.01.2018	07:28	7,50E+06
14	29.01.2018	08:30	3,12E+07
15	31.01.2018	07:55	1,22E+09
16	01.02.2018	07:37	1,12E+09
17	02.02.2018	07:14	7,87E+07
18	03.02.2018	06:57	5,02E+07
19	04.02.2018	06:39	2,90E+08
20	05.02.2018	07:58	2,15E+07
21	06.02.2018	07:41	1,11E+08
22	07.02.2018	07:23	4,24E+06
23	09.02.2018	06:43	0
24	12.02.2018	07:27	0
25	13.02.2018	07:09	0
26	14.02.2018	06:52	0
27	15.02.2018	06:34	0
28	16.02.2018	07:54	0
29	17.02.2018	07:36	0

Tabelle B.4: Bildinformationen für *petronius*.

B. Informationen über die verwendeten Bilder

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	01.01.2018	22:33	1,40E+09
1	02.01.2018	23:52	1,10E+09
2	03.01.2018	23:34	8,04E+07
3	04.01.2018	23:17	5,27E+07
4	05.01.2018	22:59	3,14E+07
5	06.01.2018	22:36	1,59E+08
6	08.01.2018	23:44	4,96E+06
7	09.01.2018	23:20	9,47E+05
8	10.01.2018	23:03	0
9	11.01.2018	22:45	0
10	14.01.2018	23:30	0
11	15.01.2018	23:12	0
12	16.01.2018	22:49	0
13	17.01.2018	22:31	0
14	21.01.2018	22:58	0
15	29.01.2018	23:46	6,20E+07
16	30.01.2018	23:28	1,11E+09
17	02.02.2018	22:30	5,81E+08
18	03.02.2018	23:55	4,11E+08
19	04.02.2018	23:32	2,49E+07
20	05.02.2018	23:14	1,29E+08
21	06.02.2018	22:57	4,92E+06
22	07.02.2018	22:39	3,29E+06
23	09.02.2018	23:41	0
24	10.02.2018	23:24	0
25	11.02.2018	23:00	0
26	12.02.2018	22:43	0
27	14.02.2018	23:45	0
28	15.02.2018	23:27	0
29	16.02.2018	23:10	0

Tabelle B.5: Bildinformationen für *nil*.

Nr.	Datum	Uhrzeit (UTC)	$L_{MT2009}$ in $\text{Wcm}^{-2}\text{sr}^{-1}$
0	01.01.2018	00:27	8,15E+07
1	02.01.2018	00:09	1,26E+09
2	03.01.2018	01:29	9,46E+07
3	04.01.2018	01:11	6,84E+07
4	05.01.2018	00:54	4,45E+06
5	06.01.2018	00:30	2,65E+08
6	07.01.2018	00:13	1,34E+08
7	08.01.2018	01:38	9,35E+06
8	09.01.2018	01:15	3,57E+07
9	10.01.2018	00:57	4,44E+05
10	11.01.2018	00:40	0
11	12.01.2018	00:22	0
12	14.01.2018	01:24	0
13	15.01.2018	01:06	0
14	16.01.2018	00:43	0
15	17.01.2018	02:08	0
16	18.01.2018	01:45	0
17	21.01.2018	02:29	0
18	22.01.2018	02:12	0
19	23.01.2018	01:54	0
20	24.01.2018	01:37	0
21	27.01.2018	02:21	0
22	28.01.2018	01:58	9,80E+06
23	29.01.2018	01:40	2,88E+08
24	30.01.2018	01:23	5,85E+08
25	31.01.2018	01:05	9,97E+07
26	01.02.2018	02:24	8,99E+07
27	02.02.2018	02:07	6,91E+07
28	03.02.2018	00:07	4,66E+07
29	04.02.2018	01:26	3,12E+07

Tabelle B.6: Bildinformationen für *alpen*.