

Hochschule Neubrandenburg  
Studiengang Geoinformatik

**Entwicklung eines generischen Prozessors zur  
Bewertung der Datennutzbarkeit verschiedener  
optischer Satellitenfernerkundungssensoren**

**Bachelorarbeit**

vorgelegt von: Martin Kriegel

Zum Erlangen des akademischen Grades  
„Bachelor of Engineering“ (B.Eng.)



Hochschule Neubrandenburg  
University of Applied Sciences



Erstprüfer: Prof. Dr.-Ing. A. Wehrenpfennig  
Zweitprüfer: Dr. rer. nat E. Borg

Bearbeitungszeitraum: 16.07.2010 – 10.09.2010

urn:nbn:de:gbv:519-thesis2010-0490-3

## **Danksagung**

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich bei der Erstellung dieser Arbeit mit Korrekturlesen, Hinweisen und Diskussionen unterstützt haben.

Besonderer Dank richtet sich an Prof. Dr.-Ing. A. Wehrenpfennig von der Hochschule Neubrandenburg und Dr. rer. nat. Erik Borg vom DLR Neustrelitz für die fachliche und organisatorische Unterstützung und die somit vorbildliche Betreuung dieser Arbeit und des vorangegangenen Praktikums beim DLR Neustrelitz.

Ich danke vor allem meiner Freundin und meinen Eltern für die großartige Unterstützung während des gesamten Studiums.

Neubrandenburg, 30. August 2010

## **Kurzfassung**

Mit fortschreitender technischer Entwicklung von Fernerkundungssensoren und Satellitensystemen nimmt die Nutzung optischer Satellitenfernerkundungsprodukte für die Erdbeobachtung in vielen Bereichen stetig zu. Das Ausmaß der Informationsausbeute dieser Daten hängt jedoch nur sekundär von der Wahl der optimalen Aufnahmetechnik ab. Vielmehr sind es die Wetterbedingungen wie zum Beispiel der Grad der Wolkenbedeckung, die über die Qualität der Daten entscheiden. Somit müssen Fernerkundungsdaten hinsichtlich deren Nutzbarkeit klassifiziert werden, um eine optimale Nutzung derer zu gewährleisten.

Diese Arbeit beschreibt die Entwicklung eines performanten generischen Prozessors, welcher auf Grundlage des von BORG (2007) entwickelten Algorithmus die automatisierte Bewertung der Datennutzbarkeit optischer Satellitenfernerkundungsprodukte durch die Detektion von Wolken erlaubt. Der generische Aspekt umfasst dabei die Fähigkeit des Prozessors ein großes Spektrum optisch-thermalen Satellitenfernerkundungsdaten verarbeiten zu können. Im Anschluss der Beschreibung der verschiedenen Phasen der Entwicklung wird die Funktionalität des aktuellen Entwicklungsstandes demonstriert.

## **Abstract**

With increasing technological development of remote sensors and satellite systems, the use of optical satellite remote sensing data for Earth observation is steadily increasing in many fields. The magnitude of the information yield of this data depends only secondarily on the choice of the optimal recording technique. Rather it is the weather conditions such as the amount and size of cloud cover, which will determine the quality of the data. Thus, remote sensing data must be classified as to their usefulness in order to ensure optimal use of those.

This paper describes the development of a generic high-performance processor which allows the automated assessment of usefulness of optical remote sensing products through the detection of clouds, based on the BORG (2007) developed algorithm. The generic aspect comprises the processing capability of the processor to handle a large range of thermal-optical satellite remote sensing data. After describing the various stages of development, the functionality of current state of development is demonstrated.

# Inhaltsverzeichnis

<b>1. Einführung</b>	<b>1</b>
1.1. Zielstellung der Arbeit . . . . .	1
1.2. Aufbau der Arbeit . . . . .	2
<b>2. Grundlagen</b>	<b>3</b>
2.1. Grundlagen der optischen Fernerkundung . . . . .	3
2.2. Anwendungsbereiche der optischen Fernerkundung . . . . .	5
2.3. Datennutzbarkeit . . . . .	5
2.4. Algorithmus zur Ermittlung der Datennutzbarkeit . . . . .	6
2.4.1. Detektion von Bildfehlern . . . . .	6
2.4.2. Kalibrierung der Satellitendaten . . . . .	7
2.4.3. Erzeugung einer Land-Wasser-Maske . . . . .	9
2.4.4. Die Wolkendetektion . . . . .	10
2.4.5. Die Bewertung der Datennutzbarkeit . . . . .	11
2.5. Der MACCA-Prozessor . . . . .	12
2.6. Schlussfolgerungen und Optimierungsansätze . . . . .	15
<b>3. Konzeption und Design</b>	<b>19</b>
3.1. Integration in das DIMS . . . . .	19
3.2. Grundlegender Aufbau des Prozessors . . . . .	22
3.3. Generische Komponenten des Prozessors . . . . .	27
3.4. Einbindung neuer Sensoren . . . . .	28
<b>4. Umsetzung</b>	<b>31</b>
4.1. Entwicklungsumgebung . . . . .	31
4.2. Erweiterung des JobOrderFiles . . . . .	32
4.3. Konfiguration des Prozessors . . . . .	33
4.4. Implementierung der Module . . . . .	34
4.4.1. Einlesen der Konfiguration . . . . .	35
4.4.2. Einlesen der Prozessierungsumgebung . . . . .	35
4.4.3. Check der Prozessierungsumgebung . . . . .	36
4.4.4. Generierung der Szenenliste . . . . .	36
4.4.5. Einlesen der Szenen . . . . .	37
4.4.6. Bildfehlerdetektion . . . . .	37
4.4.7. Einlesen der Metadaten . . . . .	37
4.4.8. Wolkendetektion . . . . .	38
4.4.9. Bewertung der Daten . . . . .	39
4.4.10. Ausgabe der Ergebnisse . . . . .	39

4.5. Demonstrative Anwendung des Prozessors . . . . .	41
<b>5. Ergebnisse und Ausblick</b>	<b>43</b>
<b>Anhang</b>	<b>45</b>
<b>A. Ergebnisse des Testdurchganges</b>	<b>45</b>

# Abbildungsverzeichnis

1.	Darstellung des elektromagnetischen Spektrums in Bezug zur atmosphärischen Opazität . . . . .	3
2.	Repräsentation spektraler Signaturen durch einen Sensor eines Landsat-Satelliten . . . . .	4
3.	Vereinfachte Darstellung der wesentlichen Schritte zur Ableitung der Datennutzbarkeit optisch-thermalen Satellitenfernerkundungsdaten . . . . .	7
4.	Darstellung einer in Satellitenprojektion transformierten Land-Wasser-Maske . . . . .	10
5.	Darstellung einer Wolkenmaske und der abgeleiteten Distanzmatrix . . . . .	12
6.	Darstellung des Arbeitsablaufes für das Profiling des MACCA-Prozessors . . . . .	14
7.	Verlauf der Nutzung des physikalischen Arbeitsspeichers im ersten Testdurchgang der Analyse des MACCA-Prozessors . . . . .	16
8.	Vergleichende Darstellung der Küstenlinien durch die CIA World Map und SWBD . . . . .	18
9.	Integration des GMACCA-Prozessors in das DIMS . . . . .	21
10.	Darstellung der Module des GMACCA-Prozessors in oberster Ebene . . . . .	23
11.	Darstellung der verschiedenen Komponenten des GMACCA-Prozessors im Hinblick auf die Detektion der Wolken . . . . .	26
12.	Prinzip der Auswahl spezifischer Module . . . . .	30
13.	Konfigurationsdatei <code>gmacca.conf</code> des GMACCA Prozessors . . . . .	34
14.	Das Modul <code>ConfigParser</code> . . . . .	35
15.	Das Modul <code>Parser_JobOrderFile</code> . . . . .	36
16.	Module zum Auslesen der Metadaten eines Quicklook-Produktes des Satelliten Landsat 7 . . . . .	38
17.	Darstellung bereits implementierter Module zur Wolkendetektion . . . . .	39
18.	Ergebnisse der Wolkendetektion einer Szene in verschiedener Darstellung . . . . .	40
19.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 22 Track 182 . . . . .	47
20.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 23 Track 182 . . . . .	48
21.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 24 Track 182 . . . . .	49
22.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 25 Track 182 . . . . .	50
23.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 26 Track 182 . . . . .	51
24.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 27 Track 182 . . . . .	52
25.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 28 Track 182 . . . . .	53
26.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 29 Track 182 . . . . .	54
27.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 30 Track 182 . . . . .	55
28.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 31 Track 182 . . . . .	56
29.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 32 Track 182 . . . . .	57
30.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 33 Track 182 . . . . .	58
31.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 34 Track 182 . . . . .	59
32.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 35 Track 182 . . . . .	60
33.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 36 Track 182 . . . . .	61
34.	Ergebnisse der Datennutzbarkeitsbewertung der Szene 37 Track 182 . . . . .	62

35. Aufbau der Modulbeschreibung in der Entwicklerdokumentation. . . . .	63
--	----

## Tabellenverzeichnis

1. Auflistung der dem JobOrderFile hinzugefügten Parameter im Teilbaum ProcessingParameters . . . . .	32
2. Zusammenfassung der Verarbeitungszeiten der Testdatensätze. . . . .	46

## Listings

1. Die Ergebnisse der Datennutzbarkeitsbewertung eines Quicklooks des Satelliten Landsat 7. . . . .	40
---	----

## Abkürzungsverzeichnis

<b>ACRES</b>	Australian Centre for Remote Sensing
<b>AE</b>	Astronomische Einheit
<b>CIA</b>	Central Intelligence Agency
<b>DFD</b>	Deutsches Fernerkundungsdatenzentrum
<b>DIMS</b>	Data and Information Management System
<b>DLR</b>	Deutsches Zentrum für Luft- und Raumfahrt e.V.
<b>DN</b>	Digital Number (Grauwert)
<b>DOM</b>	Document Object Model
<b>ECT</b>	Equator Crossing Time
<b>ETM<sup>+</sup></b>	Enhanced Thematic Mapper
<b>FST</b>	FAST-Landsat 7 Format
<b>GeoTIFF</b>	Tagged Image File Format (inkl. geografischer Informationen)
<b>GMACCA</b>	Generic Modular Automated Cloud Cover Assessment
<b>GSHHS</b>	Global Self-consistent, Hierarchical, High-resolution Shoreline
<b>GUI</b>	Graphical User Interface
<b>IDL</b>	Interactive Data Language
<b>JPEG</b>	Joint Photographers Experts Group (Bildformat)
<b>MACCA</b>	Modular Automated Cloud Cover Assessment
<b>NASA</b>	National Aeronautics and Space Administration
<b>NDSI</b>	Normalized Difference Soil Index
<b>PNG</b>	Portable Network Graphics (Bildformat)
<b>PSM</b>	Processing System Management
<b>RGB</b>	Rot Grün Blau (additiver Farbraum)
<b>SRTM</b>	Shuttle Radar Topography Mission
<b>SWBD</b>	SRTM Water Body Data
<b>XML</b>	eXtensible Markup Language



## 1. Einführung

Schon seit einigen Jahrzehnten nutzen Wissenschaftler aus den verschiedensten Fachbereichen die Fernerkundung, um die auf der Erde ablaufenden Prozesse zu untersuchen und zu überwachen. Dazu wurde eine große Zahl von Fernerkundungssatelliten in verschiedene Orbits gebracht. Diese erfassen mithilfe unterschiedlichster Sensoren elektromagnetische Strahlung. Daraus lassen sich Informationen über das System Erde ableiten.

Das Deutsche Fernerkundungsdatenzentrum (DFD) – ein Institut des Deutschen Zentrums für Luft- und Raumfahrt e.V. (DLR) – betreibt zu diesem Zweck sowohl Empfangs- als auch Kontrollstationen. Diese Stationen befinden sich in Oberpfaffenhofen bei München und Neustrelitz in Mecklenburg-Vorpommern. Neben dem Empfang der Daten nationaler und internationaler Fernerkundungssatelliten, werden die Daten durch das DFD zu Informationsprodukten weiterverarbeitet und verschiedensten Nutzern zur Verfügung gestellt beziehungsweise archiviert (DECH, 2010).

Die Verarbeitung dieser Daten erfolgt durch sogenannte *Prozessoren*. Dabei handelt es sich entgegen der allgemeinen Bedeutung um Softwareapplikationen zur automatisierten Verarbeitung von Fernerkundungsdaten und nicht um den eigentlichen Prozessor aus Hardwaresicht. In diesem Zusammenhang wird unter *Prozessierung* die automatisierte Verarbeitung von Fernerkundungsdaten verstanden.

### 1.1. Zielstellung der Arbeit

Diese Arbeit behandelt die Weiterentwicklung des in BORG (2007) entwickelten Prozessors mit der Bezeichnung *Modular Automated Cloud Cover Assessment (MACCA)*. Dieser Prozessor erlaubt die automatisierte Bewertung der Datennutzbarkeit von Quicklooks des Satelliten *Landsat 7*. Die Weiterentwicklung dieses Prozessors verfolgt verschiedene Ziele. Das wichtigste Ziel ist die Generalisierung des Prozessors bezüglich der Fähigkeit ein großes Spektrum optisch-thermalen Satellitenfernerkundungsdaten verschiedener Sensoren verarbeiten zu können. Aufgrund dieses Zieles wird die Weiterentwicklung des Prozessors mit *Generic Modular Automated Cloud Cover Assessment (GMACCA)* bezeichnet. Von untergeordneter, aber dennoch hoher Priorität ist die Steigerung der Performance zu nennen. Dies ist erforderlich, da die Verarbeitung des ständig wachsenden Datenvolumens moderner Satellitensysteme mitunter strengen Zeitanforderungen genügen muss.

## 1.2. Aufbau der Arbeit

Nach dieser kurzen Einleitung werden in Kapitel 2 zunächst wichtige Grundlagen der optischen Fernerkundung erläutert. Anschließend wird der durch BORG (2007) entwickelte Algorithmus zur Bewertung der Datennutzbarkeit beschrieben. Darauf aufbauend erfolgt eine kurze Beschreibung der durchgeführten Analyse der Implementierung des Algorithmus, dem MACCA-Prozessor. Die aus der Analyse abgeleiteten Informationen fließen schließlich mit in die Optimierungsansätze der Entwicklung des GMACCA-Prozessors ein.

Kapitel 3 beschreibt die Softwarearchitektur des Prozessors. Dabei wird neben der Integration des Prozessors in das *Data and Information Management System (DIMS)* des DLR der Fokus auf die Spezifizierung der Prozessorkomponenten gesetzt. Darüber hinaus wird das Prinzip der Einbindung verschiedener Sensoren erläutert.

Das Kapitel 4 beschreibt im Wesentlichen die Entwicklungsumgebung, sowie verschiedene Kernkomponenten und Module aus der Sicht der Implementierung. Im Anschluss daran wird die Funktionsfähigkeit der Entwicklung anhand der Verarbeitung verschiedener Testdatensätze demonstriert und deren Ergebnisse analysiert.

Zum Abschluss dieser Arbeit schafft das Kapitel 5 einen Überblick zum aktuellen Stand der Entwicklung. Darüber hinausgehend wird ein kurzer Ausblick hinsichtlich weiterzuführender Entwicklungsarbeiten gegeben.

Im Anhang findet sich unter anderem eine detaillierte Darstellung der Ergebnisse der vorangegangenen Prozessierung.

## 2. Grundlagen

### 2.1. Grundlagen der optischen Fernerkundung

Dieser Abschnitt gibt eine kurze Einführung in die optische Fernerkundung und stützt sich im Wesentlichen auf die Literatur von ALBERTZ und WIGGENHAGEN (2009); BALDENHOFER (2010) und KRAUS und SCHNEIDER (1988).

Wie in der Einführung bereits angedeutet, erfassen im All befindliche Sensoren elektromagnetische Strahlung. Aufgrund der Zusammensetzung der Atmosphäre der Erde lässt sich nicht die gesamte Bandbreite des elektromagnetischen Spektrums für die Fernerkundung durch Satelliten nutzen. In Abbildung 1 wird die atmosphärische Opazität in Bezug zum elektromagnetischen Spektrum dargestellt. Die für die Atmosphäre durchlässigen Bereiche des elektromagnetischen Spektrums werden auch als *Atmosphärische Fenster* bezeichnet. Nur in diesen Bereichen ist die Fernerkundung der Erdoberfläche durch Satelliten überhaupt möglich.

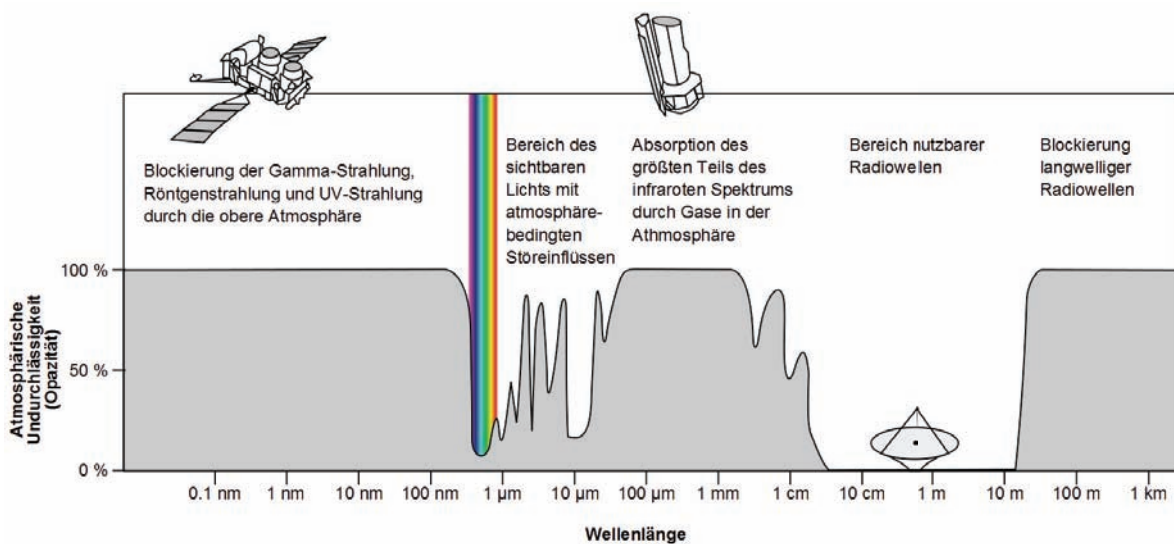


Abbildung 1: Darstellung des elektromagnetischen Spektrums in Bezug zur atmosphärischen Opazität (Undurchlässigkeit). Angelehnt an NASA (2008).

Die Erfassung der elektromagnetischen Strahlung erfolgt durch verschiedene Sensoren. Im Allgemeinen wird zwischen passiven und aktiven Sensoren unterschieden. Während passive Sensoren lediglich von der Erde ausgehende Strahlung messen, senden aktive Sensoren künstliche Strahlung und detektieren den reflektierten Anteil. Bei den passiven

Sensoren handelt es sich meist um Photosensoren, Thermalsensoren oder Mikrowellensysteme. Zu den aktiven Sensoren zählen vor allem Radar- und Lidarsysteme.

Die räumliche Auflösung der meisten Sensoren, also die Fähigkeit eng benachbarte Objekte zu differenzieren, reicht häufig nicht aus, um Objekte eindeutig zu identifizieren oder gar zu charakterisieren. So ist es Aufgabe sogenannter Multispektralsensoren die elektromagnetische Strahlung eines Gebietes in verschiedenen Spektralbereichen zu erfassen. Der aufzunehmende Spektralbereich reicht von etwa  $0,4 \mu\text{m}$  bis zu  $13 \mu\text{m}$  Wellenlänge des elektromagnetischen Spektrums (ALBERTZ, 2001, S. 15). Die spektrale Auflösung gibt dementsprechend an, wie viele Spektralkanäle (also Wellenlängenbereiche) gleichzeitig detektiert werden können. Je nach Aufgabenstellung lassen sich Objekte bei ausreichend hoher spektraler Auflösung anhand ihrer oftmals eindeutigen spektralen Signatur identifizieren und darüber hinaus auch charakterisieren. Unter der spektralen Signatur versteht man ein charakteristisches Frequenzmuster, welches vom Objekt reflektiert beziehungsweise emittiert wird. Dazu wird in vielen Fällen der Reflexionsgrad graphisch, als Funktion der Wellenlänge dargestellt. Abbildung 2 zeigt charakteristische spektrale Signaturen für Wasser, Boden und Vegetation und deren Repräsentation durch Grauwerte mithilfe eines Landsat-Sensors. Neben der räumlichen und spektralen

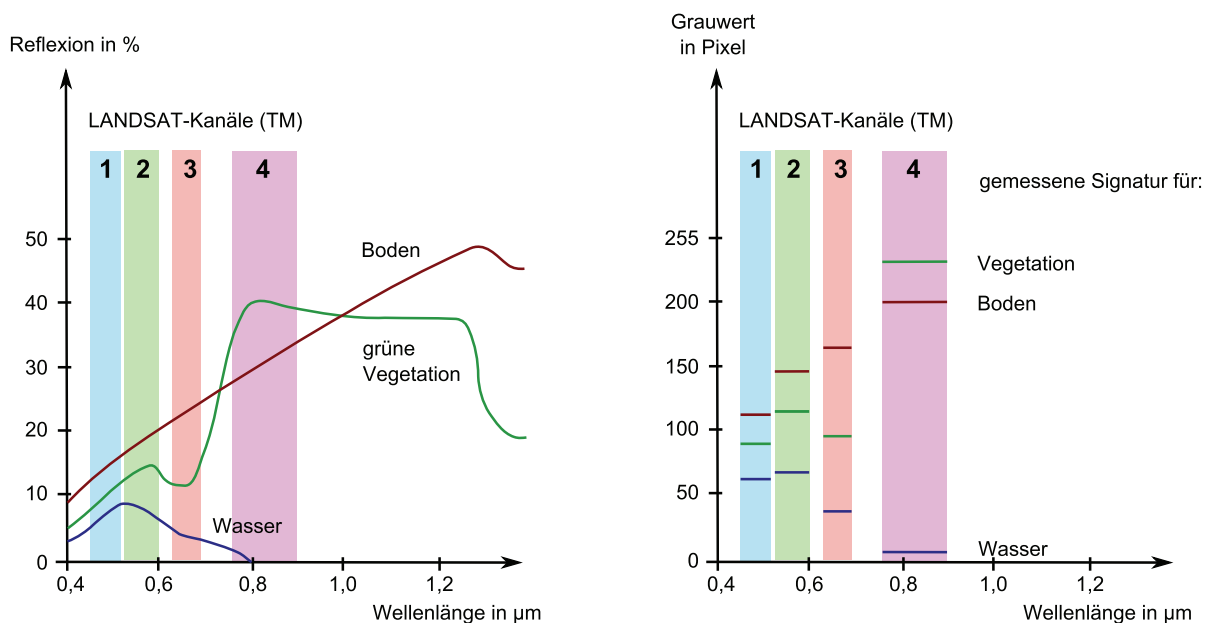


Abbildung 2: Diagramme der spektralen Signaturen von Wasser, Boden und Vegetation (links) und deren Repräsentation durch Grauwerte mithilfe eines Landsat-Sensors (rechts). Nach BALDENHOFER (2010, Grafiken bei Eintrag: spektrale Signatur).

Auflösung spielt die radiometrische Auflösung eine wesentliche Rolle. Sie kennzeichnet die minimal messbare Strahlungsdifferenz eines Sensors. Diese Auflösung wird oft in Bit angegeben, also in Anzahl möglicher Grauwerte pro Pixel. Die derzeit häufig genutzte radiometrische Auflösung beträgt 8 bis sogar 24 Bit. Durch Steigerung der beschriebenen Auflösungen lässt sich die Informationsausbeute wesentlich erweitern. Im gleichen Zuge ist jedoch der technische Aufwand, die erfassten Daten zur Erde zu übertragen und weiterzuverarbeiten immens (ALBERTZ und WIGGENHAGEN, 2009; BALDENHOFER, 2010).

## 2.2. Anwendungsbereiche der optischen Fernerkundung

Die optische Satellitenfernerkundung hat vielfältige Anwendungsgebiete. Klassische Anwendungsgebiete sind nach KRAUS und SCHNEIDER (1988, S. 8 – 9) zum Beispiel die globale Wetterbeobachtung, das Klima- und Umweltmonitoring und die Erfassung und Kartierung der Flächennutzung verschiedenster Regionen auf der Erdoberfläche. Dies umfasst neben der Herstellung und Laufendhaltung unter anderem topografischer und thematischer Karten auch die Erstellung digitaler Geländemodelle. Fernerkundungsdaten sind auch zunehmend Datengrundlage für das Katastrophenmanagement.

## 2.3. Datennutzbarkeit

Der Informationsgehalt von Fernerkundungsdaten hängt von vielen Faktoren ab. Diese sind zum Beispiel neben der thematischen Orientierung, die Wahl des Sensors (geometrische, spektrale und radiometrische Auflösung), der Trägerplattform (zum Beispiel Flugzeug oder Satellit) und die genutzte Aufnahmetechnik.

Die optisch-thermale Satellitenfernerkundung kann jedoch – trotz optimaler Auswahl der Aufnahmesysteme – keine interpretierbaren Daten liefern, wenn das zu observierende Gebiet durch Dunst oder Wolken maskiert wird. Die Hauptparameter stellen das Maß der Dunst- beziehungsweise Wolkenbedeckung und die strukturelle Verteilung der Wolken.

Dieser Umstand erfordert eine Bewertung dieser optisch-thermalen Fernerkundungsdaten hinsichtlich deren Nutzbarkeit aufgrund detektierter Wolken vor der Auslieferung der Daten an den Nutzer. Auf Anwenderseite können Fehlinvestitionen durch die Kenntnis der Datenqualität vermieden werden. Darüber hinaus kann durch die frühzeitige Sondierung der Fernerkundungsprodukte nach Datennutzbarkeit eine Entlastung automatischer Produktgenerierungssysteme auf Seiten des Bodensegmentes erreicht werden.

Nach BORG (2007) wird die Datennutzbarkeitsbewertung optisch-thermalen Satellitenfernerkundungsszenen oftmals manuell – durch sogenannte Operateure – vorgenommen.

Dies birgt sowohl Vorteile als auch Nachteile. Zum Vorteil wirkt sich die Tatsache aus, dass ein Operator zumeist auf fundiertes Basiswissen in Bezug auf die Bildinterpretation zurückgreifen kann und in Grenzfällen nach verschiedensten ihm bekannten Kriterien und Erfahrungswerten entscheiden und sondieren kann. Eher nachteilig wirkt sich die Subjektivität bei der Bewertung aus. So lassen sich Klassifizierungsfehler oftmals nicht erkennen, da sie keine Systematik aufweisen. Findet die Bewertung der Daten durch einen Automaten statt, so lässt sich die Subjektivität nahezu ausschalten. Treten bei der Klassifizierung durch einen Automaten Fehler auf, so sind diese systematisch, da der Automat nach einem fest definierten Schema arbeitet. Systematische Fehler sind gut zu entdecken und zu behandeln. Aus unter anderen diesen Gründen wird in BORG (2007) ein Algorithmus vorgestellt, welcher eine automatisierte Datennutzbarkeitsbewertung aufgrund einer Wolkendetektion vornimmt. Dazu wurde ein Automat – im folgenden Text als Prozessor bezeichnet – entwickelt, der die Datennutzbarkeit von Quicklook-Daten des Satelliten Landsat 7 ermittelt, indem er durch geeignete Klassifikationsstrategien Wolken detektiert und deren Verteilung analysiert.

Im Zuge der Weiterentwicklung des MACCA-Prozessors ist die Phase der Analyse der bestehenden Methodiken äußerst wichtig. Dazu wird zunächst der in BORG (2007) entwickelte Algorithmus untersucht und beschrieben. Im Anschluss daran wird die programmtechnische Umsetzung des Algorithmus – also der MACCA-Prozessor – analysiert. Daraus lassen sich im nächsten Schritt Schlussfolgerungen und Optimierungsansätze für die Entwicklung des generischen Prozessors gewinnen.

## 2.4. Algorithmus zur Ermittlung der Datennutzbarkeit

Die Datennutzbarkeit lässt sich nach BORG (2007) wie in Abbildung 3 auf Seite 7 dargestellt ermitteln.

### 2.4.1. Detektion von Bildfehlern

Der erste Schritt nach dem Einlesen der Metadaten und Bilddaten ist die Detektion von Bildfehlern. Bildfehler können sowohl durch Übertragungsverluste, Komprimierungsfehler als auch durch technische Defekte am Sensor hervorgerufen werden. Oftmals sind ganze Zeilen oder Spalten fehlerhaft und somit das gesamte Bild für die weitere Auswertung unbrauchbar. Ist ein Bild fehlerhaft, so wird es dementsprechend schlecht bewertet. Ist ein Bild frei von Fehlern, so erfolgt die weitere Verarbeitung der Daten.

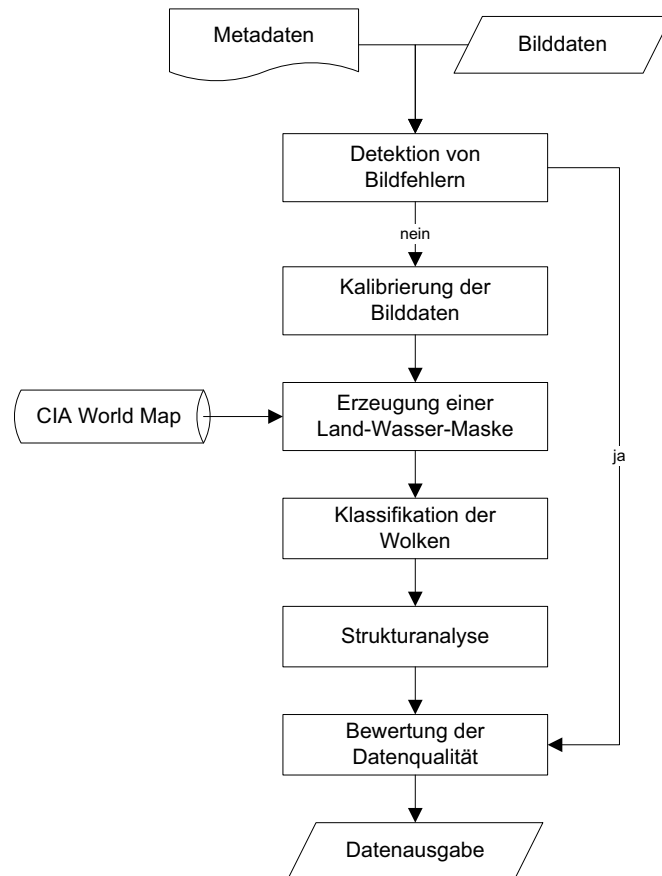


Abbildung 3: Vereinfachte Darstellung der wesentlichen Schritte zur Ableitung der Datennutzbarkeit optisch-thermaler Satellitenfernerkundungsdaten nach BORG (2007).

### 2.4.2. Kalibrierung der Satellitendaten

Bevor eine Wolkenklassifikation stattfinden kann, müssen die aufgenommenen Bilddaten kalibriert werden. Dazu werden die Grauwerte der einzelnen Kanäle zurück in am Sensor gemessene Strahldichten  $L_\lambda$  beziehungsweise am Sensor gemessene Temperaturen  $T$  transformiert. Für die beispielhaft prozessierten Quicklook-Daten ergeben sich nach NASA (2009, Kapitel 11) und BORG (2007) die folgenden Transformationsgleichungen.

Die Umwandlung der Grauwerte in am Sensor gemessene Strahldichten erfolgt durch die Anwendung der Gleichungen (1), (2) und (3) auf Seite 8. Die in diesen Gleichungen enthaltenen Parameter  $L_{MIN_\lambda}$ ,  $L_{MAX_\lambda}$ ,  $Q_{calmin}$  und  $Q_{calmax}$  sind NASA (2009, Kapitel 11, Tabelle 11.2) zu entnehmen.

$$L_\lambda = G_{rescale} \cdot Q_{cal} + B_{rescale} \quad (1)$$

$$G_{rescale} = \frac{LMAX_\lambda - LMIN_\lambda}{Q_{calmax} - Q_{calmin}} \quad (2)$$

$$B_{rescale} = LMIN_\lambda - \left( \frac{LMAX_\lambda - LMIN_\lambda}{Q_{calmax} - Q_{calmin}} \right) \quad (3)$$

mit

$L_\lambda$  = spektrale Strahldichte am Sensor [ $W/(m^2 sr \mu m)$ ]

$Q_{cal}$  = Grauwert [ $DN$ ]

$Q_{calmin}$  = minimaler Grauwert entsprechend  $LMIN_\lambda$  [ $DN$ ]

$Q_{calmax}$  = maximaler Grauwert entsprechend  $LMAX_\lambda$  [ $DN$ ]

$LMIN_\lambda$  = spektrale Strahldichte am Sensor skaliert zu  $Q_{calmin}$  [ $W/(m^2 sr \mu m)$ ]

$LMAX_\lambda$  = spektrale Strahldichte am Sensor skaliert zu  $Q_{calmax}$  [ $W/(m^2 sr \mu m)$ ]

$G_{rescale}$  = bandspezifischer Kalibrierungskoeffizient für den Gain [ $W/(m^2 sr \mu m)/DN$ ]

$B_{rescale}$  = bandspezifischer Kalibrierungskoeffizient für den Bias [ $W/(m^2 sr \mu m)$ ]

Die Umwandlung der Grauwerte der Infrarotbilder in Temperaturen erfolgt nach Gleichung (4). Die für die Berechnung notwendigen Kalibrierungsparameter  $K1$  und  $K2$  sind sensorspezifisch und ebenfalls in NASA (2009, Kapitel 11, Tabelle 11.5) zu finden.

$$T = \frac{K2}{\ln\left(\frac{K1}{L_\lambda} + 1\right)} \quad (4)$$

mit

$T$  = am Sensor gemessene Temperatur [ $K$ ]

$K1$  = Kalibrierungskonstante 1 [ $K$ ]

$K2$  = Kalibrierungskonstante 2 [ $K$ ]



Nach der Ermittlung der am Sensor gemessenen Strahldichte und Temperaturen kann die planetare Albedo mithilfe der Gleichung (5) bestimmt werden. Der bandspezifische Parameter  $ESUN_\lambda$  ist dem Handbuch NASA (2009, Kapitel 11, Tabelle 11.3) zu entnehmen.

$$\rho_\lambda = \frac{\pi \cdot L_\lambda \cdot d^2}{ESUN_\lambda \cdot \cos \theta_S} \quad (5)$$

mit

$\rho_\lambda$  = planetare Albedo

$L(\lambda)$  = gemessene Strahldichte am Sensor [ $W/(m^2 sr \mu m)$ ]

$d$  = Distanz zwischen Erde und Sonne [ $AE$ ]

$ESUN_\lambda$  = extraterrestrische Solarstrahlung [ $W/(m^2 \mu m)$ ]

$\theta_S$  = Sonnenzenitwinkel [ $Grad$ ]

Nachdem die Bilddaten kalibriert wurden, lassen sich verschiedene Verfahren der Wolkenklassifikation anwenden.

### 2.4.3. Erzeugung einer Land-Wasser-Maske

Nach Abschluss der Kalibrierung wird im Hinblick auf die Wolkenklassifikation eine sogenannte Land-Wasser-Maske generiert. Dazu wird die Kartenvorlage nach einem von FICHTELMANN et al. (2004) entwickelten Verfahren in die Satellitenprojektion transformiert. Als Datengrundlage dienen dabei die Daten der CIA World Map aus dem Jahre 1993. Aufgrund des unterschiedlichen Maßstabs innerhalb der Daten müssen die niedriger aufgelösten Basisdaten mit den höher aufgelösten Daten der Küstenlinien verschnitten werden, um eine Kartenvorlage mit bestmöglichem Maßstab zu generieren. Abbildung 4 auf Seite 10 stellt eine Land-Wasser-Maske grafisch dar.

Als Ergebnis der Land-Wasser-Masken-Generierung entsteht eine Kartenvorlage in den Dimensionen und der Projektion der Satellitenszene. Diese Kartenvorlage repräsentiert zwei verschiedene Klassen; die Klasse *Land* und die Klasse *Wasser*. Aufgrund der Transformation der Kartenvorlage in Satellitenprojektion lassen sich die Klassen *Land* und *Wasser* der kalibrierten Originalszene durch einfache Überlagerung pixelgenau zuweisen. Die Kenntnis welchen Pixeln in der Satellitenszene die Klasse *Land* und welchen die



Abbildung 4: Darstellung eine Land-Wasser-Maske. Die linke Abbildung stellt die niedrig aufgelösten Basisdaten dar, die rechte Abbildung das Ergebnis der Verschneidung mit den höher aufgelösten Daten der Küstenlinien. Graue Pixel entsprechen der Klasse *Land*, blaue Pixel der Klasse *Wasser*.

Klasse *Wasser* zugewiesen wurde, ermöglicht nach BORG (2007) eine wesentlich genauere Detektion von Wolken und vor allem von Dunst.

#### 2.4.4. Die Wolkendetektion

Der in BORG (2007) entwickelte Algorithmus verfügt über drei verschiedene Module zur Detektion von Wolken. Das erste Teilmodul ist ein Schwellwertverfahren. Zuerst wird mithilfe von Testspektren auf gesättigte Pixel getestet. Sowohl Schnee als auch Wolken führen zur Sättigung des Sensors. Zur Differenzierung werden sowohl der NDSI-Wert als auch der NDSL<sub>2</sub>-Wert gegen einen Schwellwert getestet. Die beiden NDSI-Werte ergeben sich aus den Gleichungen (6) und (7) auf Seite 11. Die Schwellwerttests gegen die NDSI-Werte lassen eine gute Unterscheidung zwischen Schnee und Wolken oder Dunst zu. Im Anschluss daran finden verschiedene Tests bezüglich des Thermalverhaltens der verbliebenen Pixel statt. Wurde die Überprüfung auf das Thermalverhalten abgeschlossen, findet in den verschiedenen Spektralkanälen eine erneute Prüfung des NDSI-Wertes gegen einen Schwellwert statt.

Das zweite Wolkenerkennungsmodul nutzt zur Klassifikation der Wolken definierte Trainingspektren verschiedener Landbedeckungsklassen. Mithilfe der Pearsonkorrelation werden die zu untersuchenden Spektren mit den Testspektren hinsichtlich spektraler

Ähnlichkeit analysiert und entsprechend klassifiziert. Das dritte Wolkenerkennungsmodul nutzt die vorher generierte Kartenvorlage zur Klassifizierung der Wolken. Aufgrund der Land-Wasser-Verteilung werden dafür verschiedene Klassifikationsalgorithmen genutzt.

$$\text{NDSI} = \frac{\text{Kanal 2} - \text{Kanal 5}}{\text{Kanal 2} + \text{Kanal 5}} \quad (6)$$

$$\text{NDSL}_2 = \frac{\text{Kanal 4} - \text{Kanal 7}}{\text{Kanal 4} + \text{Kanal 7}} \quad (7)$$

Durch die Nutzung verschiedener Wolkendetektionsalgorithmen werden auch verschiedene Resultate erzeugt. Dabei handelt es sich nach NISCHWITZ et al. (2007, S. 350) um Zweipegelebilder, welche die Klassen *Wolke* und *Nicht-Wolke* repräsentieren. Bei kombinierter Nutzung der verschiedenen Wolkendetektionsmodule müssen die einzelnen Ergebnisse – nachfolgend Wolkenmasken genannt – auf geeignete Art und Weise fusioniert werden, um ein Gesamtergebnis zu erhalten. Die Fusion der einzelnen Pixel der verschiedenen Wolkenmasken erfolgt im Prinzip durch einen Mehrheitsvergleich. Letztendlich wird die Klassifizierung eines Pixels in das Gesamtergebnis übernommen, die am häufigsten in den Teilresultaten auftritt. Eine Gewichtung der Teilresultate wird nicht vorgenommen. Detailliertere Informationen zur Funktionsweise der verschiedenen Wolkendetektionsalgorithmen sowie der Fusion der Wolkenmasken sind BORG (2007, S. 42 – 47) zu entnehmen.

#### 2.4.5. Die Bewertung der Datennutzbarkeit

Im Anschluss an die Wolkenklassifizierung sieht der Algorithmus die Bewertung der Datennutzbarkeit vor. Die Grundlage dafür ist die Berechnung mehrerer sogenannter Distanzmatrizen.

Dazu werden jeweils pro Pixel Distanzen sowohl zwischen Bildrand und Pixel der Klasse *Wolke* (Wolkenpixel) als auch zwischen Wolkenrand und Wolkenpixel berechnet. Diese Berechnung erfolgt ausgehend von acht Richtungen, entsprechend der 8er-Nachbarschaft eines Pixels. Im Anschluss daran wird das pixelweise Minimum der acht Matrizen in die endgültige Distanzmatrix übernommen. Diese Matrix enthält für jedes Pixel die minimale Distanz entweder zum Bildrand oder zum nächsten Wolkenpixel. Abbildung 5 auf Seite 12 stellt ein Resultat der Wolkendetektion und der daraus abgeleiteten Distanzmatrix zur Bewertung der Datennutzbarkeit dar.

Die Berechnung des Maßes der Datennutzbarkeit erfolgt dann durch einen Formel-

apparat unter Nutzung spezieller Parameter. Diese Parameter werden aus der vorher generierten Distanzmatrix abgeleitet. Eine genaue Darstellung der Parameter sowie des gesamten Formelapparates zur Ableitung der Bewertung der Datennutzbarkeit findet sich in BORG (2007, S. 47 – 51). Dort wird auch das Verfahren zur Ableitung der Distanzmatrix abgebildet. Nachdem die Satellitenszene hinsichtlich ihrer Nutzbarkeit bewertet wurde erfolgt die Ausgabe der Ergebnisse.

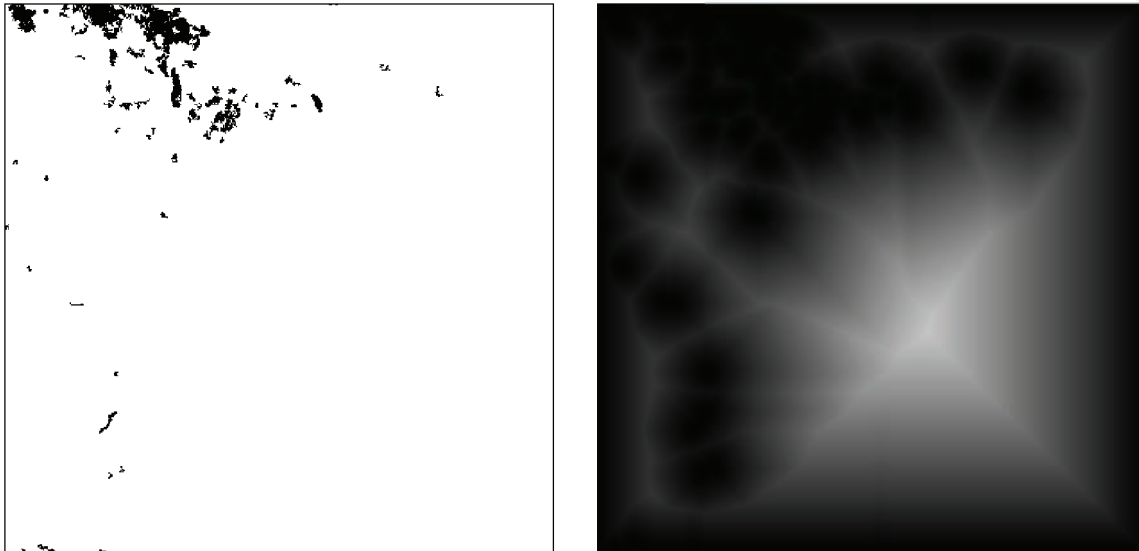


Abbildung 5: Darstellung der Wolkenmaske auf der linken Seite. Diese Maske kann sowohl ein Teilprodukt als auch ein fusioniertes Gesamtergebnis der Wolken-detektion sein. Detektierte Wolkenpixel sind hier schwarz gekennzeichnet. Auf der rechten Seite ist die dazugehörige Distanzmatrix dargestellt. Umso heller die Pixel in der rechten Abbildung dargestellt sind, umso größer ist die Distanz dieser Pixel zum Bildrand oder zu detektierten Wolken. Diese Distanzmatrix ist die Grundlage der Bewertung der Datennutzbarkeit.

## 2.5. Der MACCA-Prozessor

Nachdem der Algorithmus zur Ermittlung der Datennutzbarkeit aufgrund der Detektion von Wolken beschrieben wurde, liegt das Augenmerk in diesem Abschnitt auf der Analyse des MACCA-Prozessors. Dieser Prozessor ist die Grundlage der im Rahmen dieser Arbeit durchgeführten Entwicklungstätigkeiten.

Der MACCA-Prozessor wurde in der Interactive Data Language (IDL) implementiert. Diese Programmiersprache wird durch ITT VISUAL INFORMATION SOLUTIONS entwi-

ckelt und eignet sich aufgrund ihrer einfachen und kompakten Syntax für die schnelle Erstellung von Software zur Datenauswertung und -visualisierung. Darüber hinaus ist die Sprache plattformunabhängig und somit auf einer Vielzahl von Systemen einsetzbar. Als Entwicklungsumgebung dient die IDL-Workbench, eine auf dem Eclipse-Framework basierende Entwicklungsumgebung. Neben der kostenpflichtigen Entwicklerversion wird auch eine kostenlose, plattformunabhängige Lösung zur Ausführung bereits kompilierter Software bereitgestellt. Dabei handelt es sich um die *IDL Virtual Machine* (ITT VIS, 2009a, 2010).

Für die Analyse der Performance des MACCA-Prozessors wurde eigens eine Umgebung entwickelt, mithilfe derer sich die Erfassung verschiedener Messdaten automatisieren lässt. Die Implementierung dieser Umgebung erfolgte ebenfalls in der Programmiersprache IDL. Bei den erfassten Messdaten handelt es sich um die Anzahl der Aufrufe sämtlicher Funktionen, deren Ausführungszeit und der Bedarf an Arbeitsspeicher, den der Prozessor während der Ausführung für sich beansprucht. Eine Analyse einer Softwareapplikation hinsichtlich dieser Messdaten wird auch als *Profiling* bezeichnet. Die Abbildung 6 auf Seite 14 stellt den Arbeitsablauf des durchgeführten Profilings dar. Das Modul `PROFILING_CONTROL` implementiert die durch IDL zur Verfügung gestellten Funktionalitäten für ein Profiling. Dabei handelt es sich um die sinnvolle und zielführende Anwendung der IDL-Funktionen `PROFILER` und `MEMORY`. Parallel zu diesem Modul läuft ein Shellskript, welches den Arbeitsspeicherbedarf protokolliert.

Während der Analyse einer Anwendung wird diese mehrmals durchlaufen. Die Anzahl der Testdurchläufe ist den Testbedingungen anzupassen und frei wählbar. Nach Durchlaufen aller Testdurchgänge werden die erzeugten Ergebnisse durch die Tabellenkalkulation `CALC` des Programmpaketes `OPEN OFFICE` analysiert und visualisiert. Dieser Schritt ließe sich auch mit IDL bewerkstelligen und in die Analyseumgebung integrieren. Der Aufwand der Implementierung dieser Möglichkeiten, war im Hinblick auf die Zielvorstellung – der schnellen Analyse des MACCA-Prozessors – jedoch nicht zu rechtfertigen.

Um aussagekräftige Messwerte zu erhalten, wurde ein Testdatensatz fünfzig mal prozessiert. Darüber hinaus wurden sämtliche Funktionen des MACCA-Prozessors dokumentiert und es wurde eine Funktionsaufrufhierarchie erstellt. Es wurde also untersucht, welche Funktion, wann, unter welchen Voraussetzungen, durch welche andere Funktion aufgerufen wird und welche Aufgaben die Funktion löst. Eine detailliertere Darstellung der Analyse des MACCA-Prozessors und der dafür genutzten Methodiken findet sich in KRIEGEL (2010, S. 29 – 62).

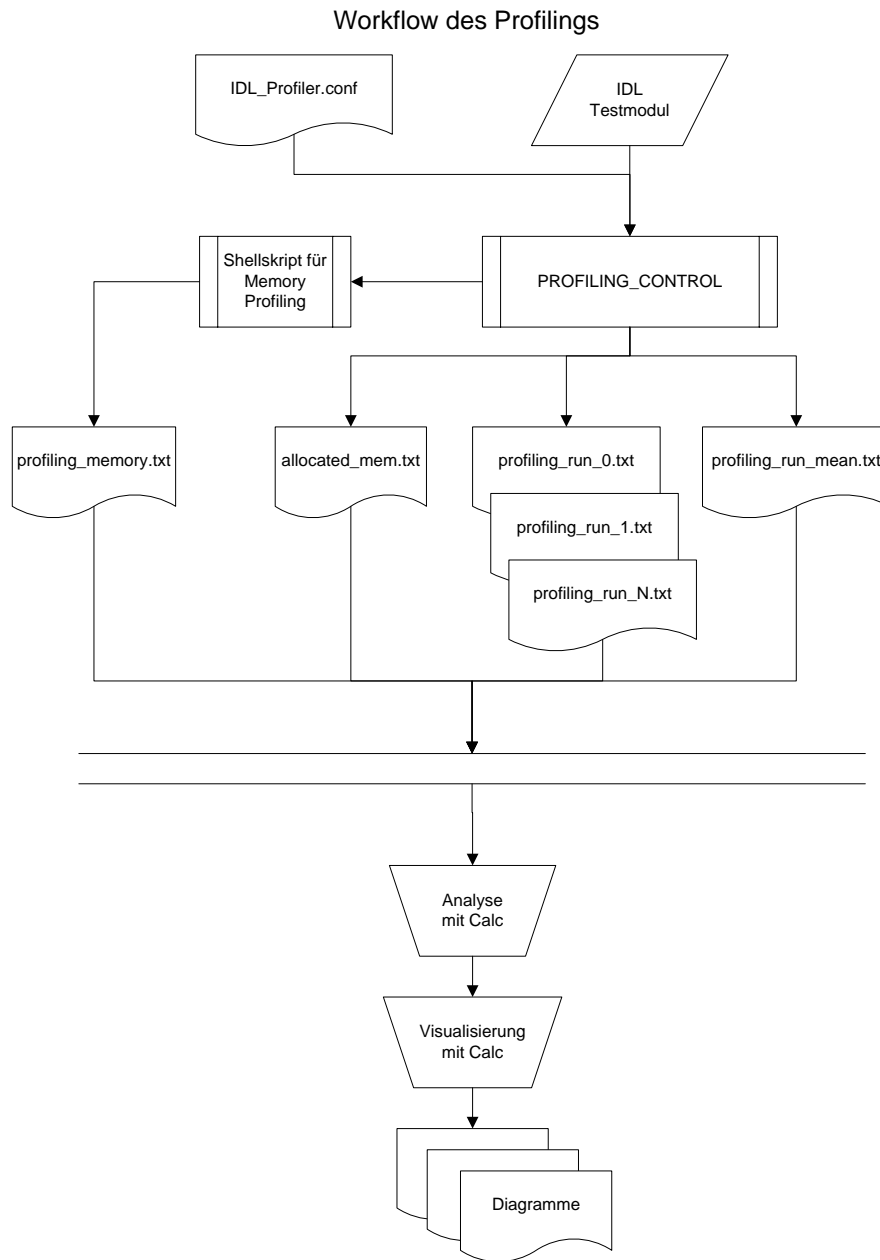


Abbildung 6: Darstellung des Arbeitsablaufes für die Erfassung der Performance des MACCA-Prozessors.

## 2.6. Schlussfolgerungen und Optimierungsansätze

Die dieser Arbeit zugrunde liegende Version des MACCA-Prozessors verwendet zur Verarbeitung von Quicklooks des Satelliten Landsat 7 den in Abschnitt 2.4 beschriebenen Algorithmus. Allerdings wurde der Prozessor einzig und allein für die Bearbeitung von Quicklooks des Satelliten Landsat 7 konzipiert. Bei der Implementierung des Prozessors handelt es sich um einen proprietären Forschungsprozessor. Dementsprechend ausgeprägt ist die Qualität und Dokumentation des Prozessors. Für die Bearbeitung einer einzelnen Quicklook-Szene – bestehend aus acht 1000 mal 1000 Pixel großen JPEG-Dateien zuzüglich einiger ASCII-Metadatendateien – werden durchschnittlich 57 Sekunden Verarbeitungszeit benötigt. Bei gleichzeitiger Nutzung der grafischen Oberfläche des Prozessors werden zusätzlich 55 Sekunden benötigt, so dass der Gesamtzeitaufwand der Verarbeitung einer Quicklook-Szene durchschnittlich 112 Sekunden beträgt. Die grafische Oberfläche ist dabei nur für die Visualisierung der Ergebnisse konzipiert und besitzt keinerlei interaktive Elemente. Die Abbildung 7 auf Seite 16 zeigt die Nutzung des Arbeitsspeichers während des ersten der fünfzig Testdurchgänge. Demnach werden im Durchschnitt etwa 53 MB des physikalischen Arbeitsspeichers genutzt. Die maximale Allokation (Speicherreservierung) beträgt kurzzeitig etwa 124 MB. Das Profiling des MACCA-Prozessors fand auf dem im Abschnitt 4.1 beschriebenen System statt.

Der Prozessor verfügt im Grunde über keine Möglichkeiten der externen Konfiguration. Darüber hinaus sind sämtliche Parameter, wie zum Beispiel wichtige Sensorparameter oder die Dimensionen der Bilddaten fest und unstrukturiert im Quelltext verankert. Dies erschwert sowohl die Wartung als auch die Erweiterung des Prozessors.

Aufgrund der genannten Probleme findet eine komplette Neuentwicklung beziehungsweise Reimplementierung des Prozessors statt. Dabei werden verschiedene Ziele verfolgt. Wie in Abschnitt 1.1 auf Seite 1 bereits angedeutet, hat die Generalisierung des Prozessors im Hinblick auf die Verarbeitung verschiedenster optisch-thermaler Satellitenfernerkundungsprodukte höchste Priorität. Dies erfordert einen streng modularen Aufbau des Prozessors, so dass neu entwickelte Komponenten an feste Schnittstellen andocken und somit problemlos und ohne Aufwand in den Prozessor integriert werden können. Darüber hinaus erleichtert eine stark ausgeprägte Modularisierung auch die Wartung des Prozessors. Fehler lassen sich einfacher lokalisieren, fehlerhafte Module können auf einfache Art und Weise ausgetauscht werden, einzelne Module lassen sich separat testen und schrittweise optimieren. Auf diese Weise kann die Robustheit des Prozessors erhöht und die Performance schrittweise verbessert werden.

Neben der strikten Modularisierung muss der GMACCA-Prozessor von außen konfigu-

rierbar sein, ohne dass dafür Quelltextmanipulationen erforderlich sind. Darüber hinaus soll der GMACCA-Prozessor in das DIMS des DLR eingebunden werden können. Durch diese Integration wird eine zentralisierte Prozessorsteuerung und ein einheitliches Datenmanagement gewährleistet, so dass der GMACCA-Prozessor optimal in die Geschäftsprozesse des DLR integriert werden kann. Großen Wert wird auch auf die Dokumentation des Prozessors gelegt. Dazu gehört neben der Anfertigung einer HTML-Dokumentation auch gut strukturierter und dokumentierter Quellcode.

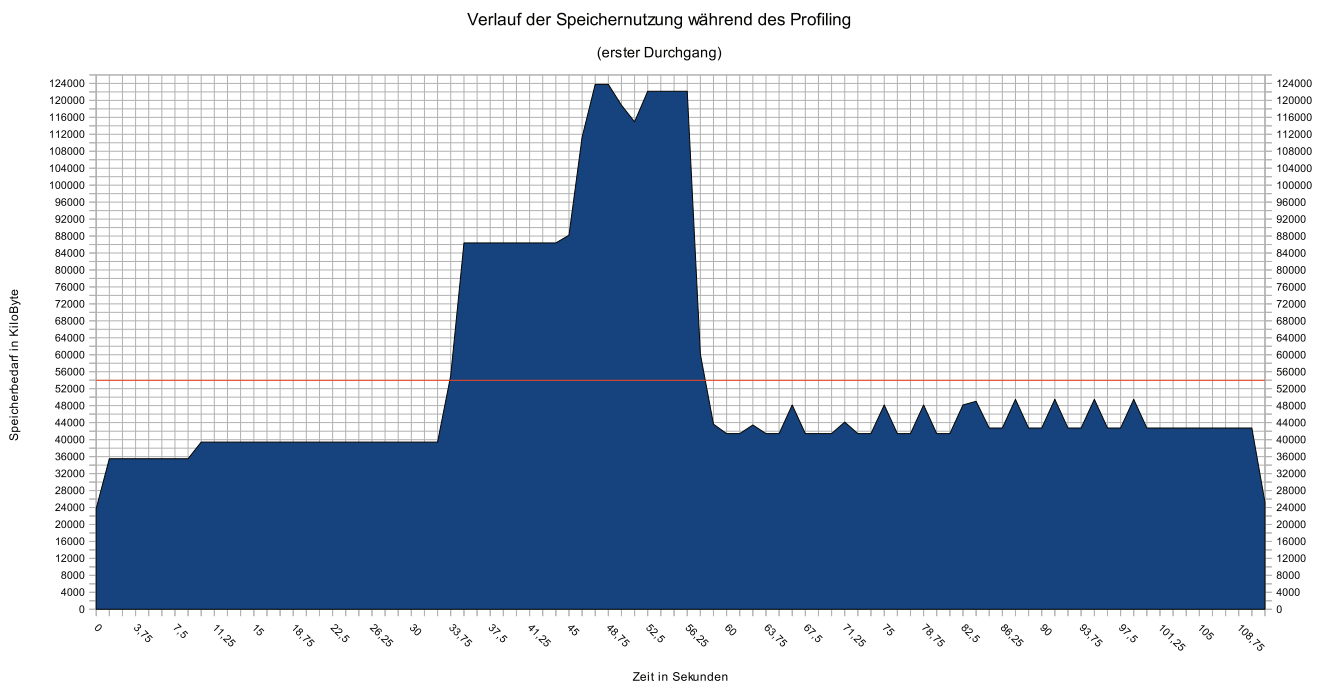


Abbildung 7: Verlauf der Nutzung des physikalischen Arbeitsspeichers im ersten Testdurchgang der Analyse des MACCA-Prozessors. Die rote Linie kennzeichnet den durchschnittlichen Verbrauch von Arbeitsspeicher.

In Abschnitt 2.4.3 wurde die automatisierte Generierung einer Land-Wasser-Maske auf Grundlage des Datenbestandes der CIA World Map beschrieben. Da sich die Klassifizierung der Wolken – wie in Abschnitt 2.4.4 dargestellt – unter anderem auf die Ergebnisse der Generierung einer Land-Wasser-Maske des betroffenen Gebietes stützt, lassen sich



umso genauere Klassifizierungsergebnisse erzielen, je höher die räumliche Auflösung der Datengrundlage zur Generierung einer Land-Wasser-Maske ist.

Während der 11-tägigen *Shuttle Radar Topography Mission (SRTM)* im Jahre 2000 wurde der Großteil der Erdoberfläche höhenmäßig erfasst. Mithilfe der dort erfassten Daten konnte erstmals ein hochaufgelöstes dreidimensionales digitales Höhenmodell der Landflächen zwischen 60° nördlicher Breite und 58° südlicher Breite erstellt werden (RAMIREZ, 2009). Als Nebenprodukt entstanden die *SRTM Water Body Data (SWBD)*-Produkte. Dabei handelt es sich um Vektordaten zur Beschreibung des Wasserkörpers der Erde im aufgenommenen Gebiet. Die räumliche Auflösung dieser Daten beträgt etwa 20 m. Die Daten werden in Form von ESRI 3D Shapefiles zur Verfügung gestellt. Pro Shapefile wird eine 1° mal 1°-Kachel des aufgenommenen Gebietes abgedeckt, so dass der gesamte Datenbestand 12 229 Shapefiles umfasst. Dabei ist zu beachten, dass für den 1° mal 1°-Ausschnitt der Erdoberfläche nur dann ein Shapefile erzeugt wurde, sofern dieser Bereich Wasser enthält beziehungsweise durch die Shuttle Mission erfasst wurde. Der Zugriff sowohl auf die SWBD-Daten als auch auf deren Spezifikation ist kostenfrei über das Internet<sup>1</sup> möglich.

Im Zuge der Entwicklung des GMACCA-Prozessors wurde die Nutzung dieser Daten getestet und hinsichtlich deren Nutzbarkeit bewertet. Dafür wurde eigens eine Applikation entwickelt, welche sowohl die SWBD-Daten für die Generierung einer Land-Wasser-Maske nutzt als auch Daten der CIA World Map als Vergleichskriterium visualisiert. Abbildung 8 auf Seite 18 zeigt die Visualisierung beider Datensätze. Während die Küsten- beziehungsweise Uferlinien auf Datengrundlage der CIA World Map rot eingefärbt sind, sind die Küstenlinien auf Grundlage der SWBD-Daten grün eingefärbt. Der Abbildung 8 folgend ist eindeutig festzustellen, dass die Klassifizierung der Gewässer mithilfe der SWBD-Daten wesentlich detaillierter ist, als die unter Nutzung der Daten der CIA World Map. Dies umfasst neben der detaillierteren, realitätsgetreueren Darstellung der Küsten- beziehungsweise Uferlinien auch die Erfassung wesentlich kleinerer Wasserkörper.

Für Bereiche auf der Erdoberfläche, für die keine SWBD-Datensätze verfügbar sind, werden weiterhin die Daten der CIA World Map zur Generierung der Land-Wasser-Maske genutzt. Eine Szene kann jedoch zu Teilen sowohl im Datenbestand der SWBD-Daten als auch im Bestand der CIA World Map liegen. Dies erfordert die optimale Nutzung beider Datenquellen zur Generierung einer gemeinsamen und in sich stimmigen Land-Wasser-Maske mit höchstmöglicher Qualität.

---

<sup>1</sup>[http://dds.cr.usgs.gov/srtm/version2\\_1/SWBD/](http://dds.cr.usgs.gov/srtm/version2_1/SWBD/)

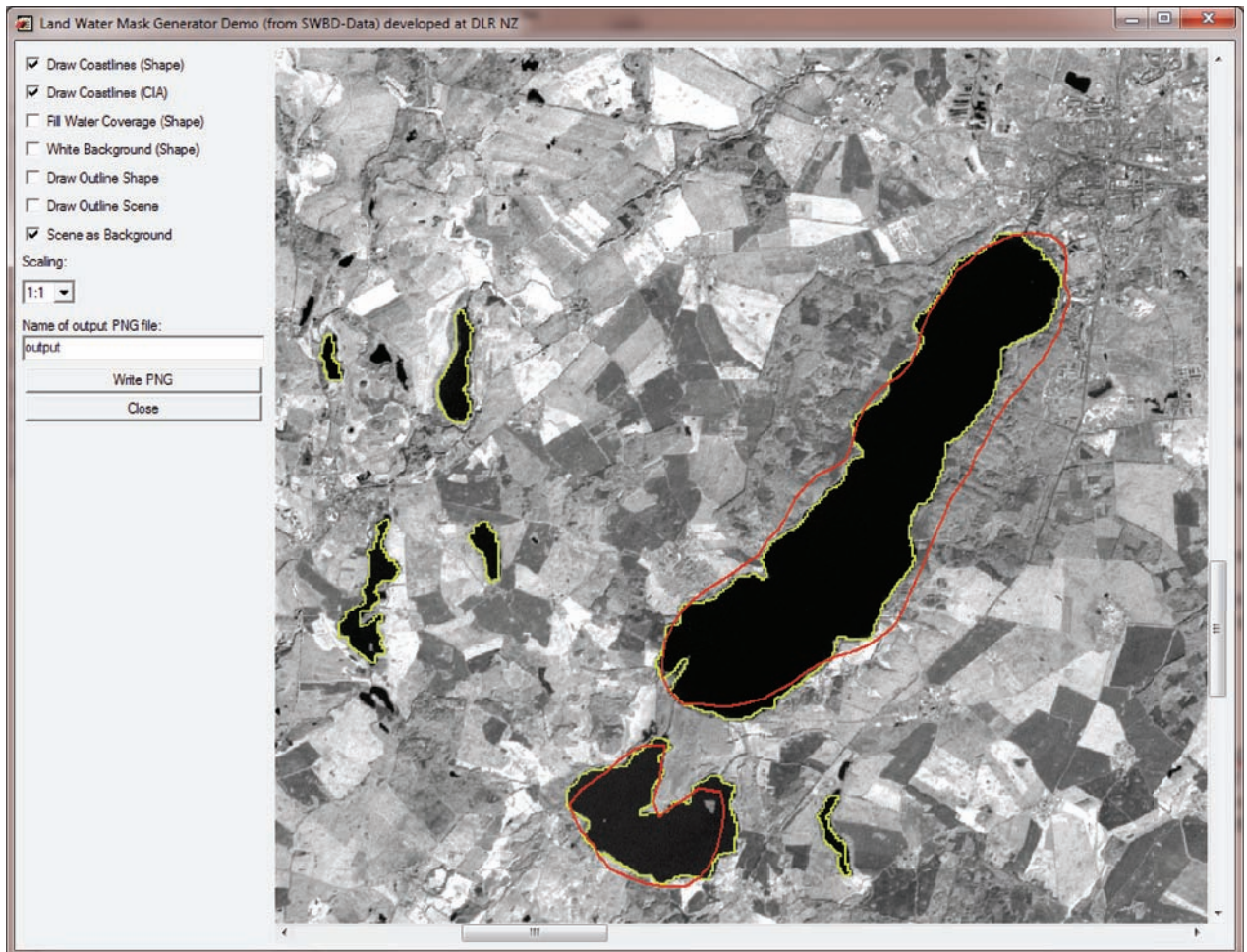


Abbildung 8: Darstellung der Küstenlinien auf Datengrundlage der SWBD-Daten (grün eingefärbt) im Vergleich zu den Küstenlinien, die auf Basis der Daten der CIA World Map (rot eingefärbt) erzeugt wurden am Beispiel der Region um den Tollensesee. Aufgrund der höheren räumlichen Auflösung der SWBD-Daten lassen sich qualitativ höherwertige Land-Wasser-Masken zur genaueren Wolkendetektion erzeugen.

## 3. Konzeption und Design

Nach einer kurzen Einführung in die Grundlagen der optischen Satellitenfernerkundung wurde der durch BORG (2007) entwickelte Algorithmus zur Ermittlung der Datennutzbarkeit erläutert. Im Anschluss daran fand eine Analyse des MACCA-Prozessors statt. Aus den Ergebnissen der Analyse heraus erfolgte die Entwicklung verschiedener Optimierungsansätze für die Entwicklung des GMACCA-Prozessors. Dieser Abschnitt ist der Architektur des GMACCA-Prozessors und der Beschreibung wichtiger Prozessorkomponenten gewidmet.

### 3.1. Integration in das DIMS

Nach BRANDT-POOK und KOLLMEIER (2008) beschreibt die Software-Architektur eines Systems sowohl die Strukturen als auch die wesentlichen Bausteine eines Systems. Dies beinhaltet auch die Beschreibung der Eigenschaften und Beziehungen der Bausteine zueinander. Architekturentscheidungen sind demnach die grundlegendsten Entscheidungen im Software-Entwicklungsprozess. Bereits hier entstehende Fehlentscheidungen haben im späteren Verlauf der Entwicklung gravierende Auswirkungen. Die Wahl der Architektur hat also unmittelbaren Einfluss auf die Wartung und die Erweiterung – und somit auch auf die Lebensdauer – einer Anwendung (BRANDT-POOK und KOLLMEIER, 2008; VOGEL et al., 2009).

Einer der Optimierungsansätze aus Abschnitt 2.6 war die Möglichkeit der Integration des GMACCA-Prozessors in das *Data and Information Management System (DIMS)*. DIMS ist beim Deutschen Fernerkundungsdatenzentrum (DFD) für Produkte der Erdbeobachtung erfolgreich im Einsatz und bietet umfassende Funktionalität zum Management großer heterogener digitaler Datenbestände mit Raum-Zeit-Bezug. DIMS besteht aus mehreren Kernkomponenten. So beinhaltet es neben den Komponenten für die Produktion und Produktgenerierung auch Komponenten für die Katalogisierung und Langzeitarchivierung der Produkte. Darüber hinaus umfasst das DIMS Komponenten für die Abwicklung von Kundenaufträgen und die Auslieferung der Daten. Dies umfasst auch Komponenten für den Online Zugriff auf den Datenkatalog über das Internet (KIEMLE, 2010a).

Um dem Anspruch eines Multi-Mission-Systems gerecht zu werden, wurde das DIMS so konzipiert, dass es von der internen Struktur und den jeweiligen Prozessierungsbedingungen einer Mission völlig entkoppelt ist. Somit ermöglicht die offene Architektur die Unterstützung neuer Produkttypen und Anwendungen (KIEMLE, 2010a). Das Data and

Information Management System ist als modulares, komponentenbasiertes System konzipiert. Die für die Produktgenerierung verantwortliche Komponente ist das sogenannte *Processing System*. Mit der Integration des GMACCA-Prozessors in das DIMS wird eine neue Instanz dieses sogenannten Processing Systems erzeugt. In der Abbildung 9 auf Seite 21 wird diese Instanz als *GMACCA Processing System* bezeichnet.

Der GMACCA-Prozessor stellt den eigentlichen Datenprozessor da. Dieser verfügt über eine eigene Konfigurationsdatei, sowie über eine eigene Wissensbasis. In der Wissensbasis sind wichtige plattform-, sensorspezifische und zielobjektorientierte Informationen, wie zum Beispiel Kalibrierungsparameter oder Bahnparameter der verschiedenen Sensoren enthalten, die der GMACCA-Prozessor verarbeiten kann. Für die Generierung der Land-Wasser-Masken benötigt der Prozessor Direktzugriff sowohl auf die Daten der CIA World Map als auch auf die SWBD-Daten.

Entsprechend Abschnitt 1.1 sind die Produkte verschiedener optisch-thermalen Satellitenfernerkundungssensoren Datengrundlage des generischen Prozessors. Im Speziellen handelt es sich dabei um die Kombination von Bilddaten und dazugehöriger Metadaten. Diese werden dem Prozessor über einen *Cache* entweder aus externen Quellen oder aus der *Product Library* übergeben. Die *Product Library* ist ein zentraler Bestandteil des DIMS, die als Quelle und Ziel für die Produktgenerierung dient und die Vollständigkeit und Konsistenz der Produkte gewährleistet (KIEMLE, 2010b).

Im Zuge der Prozessierung der Satellitenprodukte werden verschiedene Produkte generiert. Dabei handelt es sich primär um die Ergebnisse der Datennutzbarkeitsbewertung in Form von Textdateien. Darüber hinaus entstehen weitere Nebenprodukte. So wird ein Produkt erzeugt, welches alle klassifizierten Wolken in einem RGB-Bild darstellt. Darüber hinaus werden verschiedene temporäre Dateien erzeugt. Diese enthalten zum Beispiel die pixelweise berechneten Koordinaten beziehungsweise pixelweise berechnete Sonnenzenit- und Sonnenazimutwinkel. Es besteht die Möglichkeit sämtliche Ergebnisse verschiedener Prozessierungsschritte durch eine grafische Oberfläche zu visualisieren. Diese Oberfläche kann – im Gegensatz zum MACCA-Prozessor – auch interaktive Elemente, wie zum Beispiel das dynamische Zooming und Panning in der Szene oder die dynamische Anpassung verschiedener Schwellwerte für die Wolkenklassifikation einer Szene enthalten. Die Darstellung einer grafischen Oberfläche ist jedoch nur für den Einzelbetrieb im Zuge der Entwicklung des Prozessors wirklich sinnvoll, da die Darstellung grafischer Oberflächen der Datenprozessoren für den operationellen Betrieb im DIMS keine beziehungsweise eine untergeordnete Rolle spielt. Dort ist der Prozessor nur als Prozess-

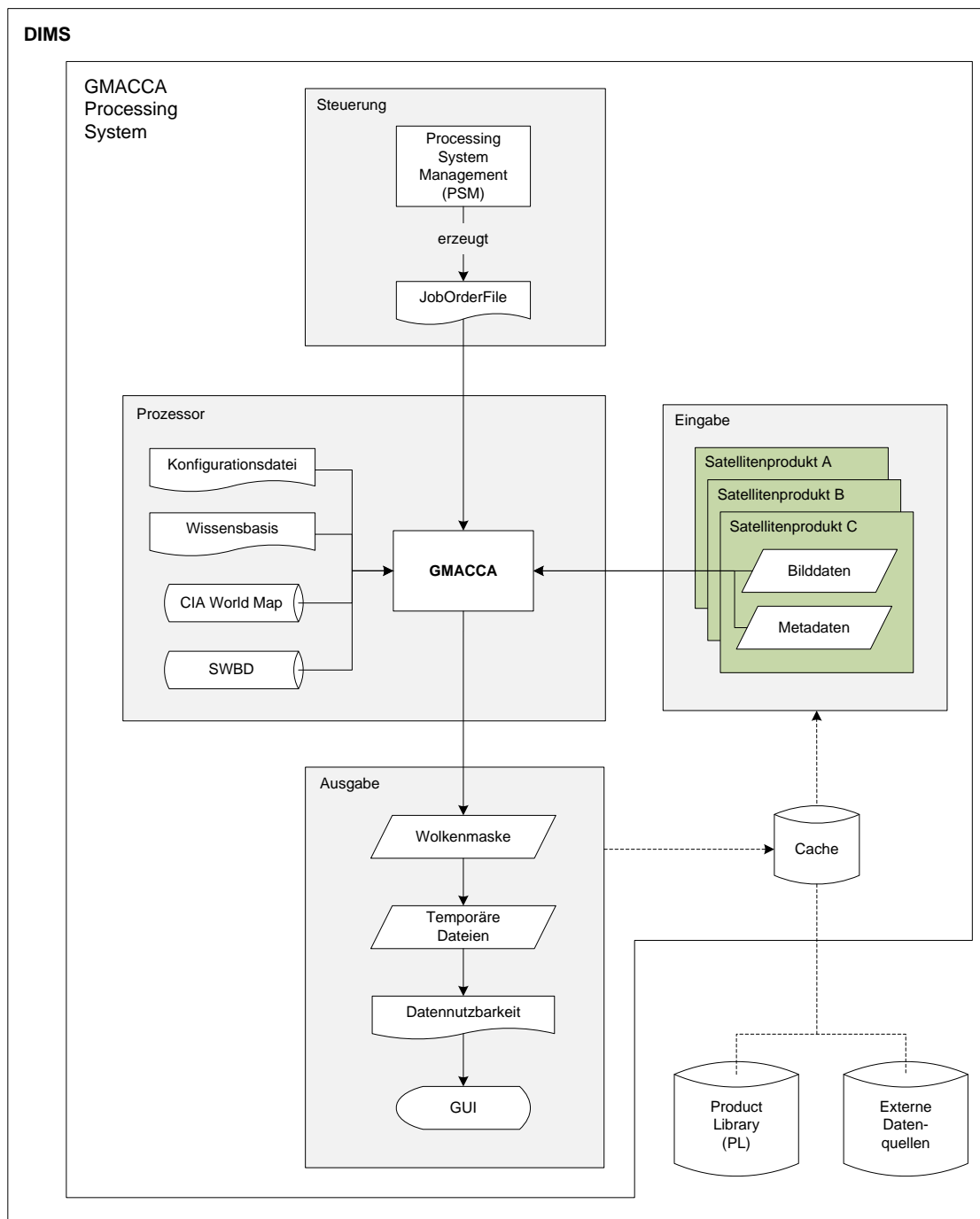


Abbildung 9: Darstellung der mit dem GMACCA-Prozessor interagierenden Komponenten und der Ein- und Ausgabeprodukte bei Integration des Prozessors in das Data and Information Management System des DLR.

instanz in einer Produktionslinie vom Datenempfang über die Datenverarbeitung zur Datenarchivierung und -auslieferung zu verstehen.

Die Steuerung der verschiedenen Komponenten eines *Processing Systems* übernimmt die Komponente *Processing System Management (PSM)*. Hier werden auf Grundlage fest definierbarer Regeln verschiedene Requests – also Prozessierungsinstruktionen – behandelt. Diese Instruktionen spezifizieren die Art der Ein- und Ausgabedaten und verschiedenste Parameter. Unter anderem erzeugt die PSM ein sogenanntes JobOrderFile.

Das JobOrderFile ist ein XML-Dokument. Es wird dem Prozessor beim Start der Prozessierung übergeben und beinhaltet verschiedene Konfigurations- oder Modellparameter und die Pfade zu den zu verarbeitenden Daten. Im Wesentlichen besteht die XML-Datei aus 3 wichtigen Hauptknoten. Der Knoten *Ipfcnf* enthält allgemeine Konfigurationsparameter, wie zum Beispiel den Namen des Prozessors, seine Version und verschiedene Level für das Logging und die Ein- und Ausgabe. Im Knoten *ProcessingParameters* werden die verschiedenen Parameter, die für die Prozessierung notwendig sind, abgelegt. Der Knoten *List\_of\_Ipf\_Procs* enthält sowohl weitere Parameter als auch weitere Informationen der für die Prozessierung notwendigen Eingabedateien. Des Weiteren werden dort auch die Bezeichnung und der Zielort des zu erzeugenden Produktes festgelegt. Zusammenfassend lässt sich also feststellen, dass durch die Übergabe des JobOrderFiles eine dynamische Zuweisung der Prozessierungsumgebung an den Prozessor stattfindet. Weitere Informationen zur Integration von Datenprozessoren in das DIMS finden sich in WERUM SOFTWARE & SYSTEMS AG (2009).

## 3.2. Grundlegender Aufbau des Prozessors

Das Hauptziel der Entwicklung war die Generalisierung des Prozessors hinsichtlich der Verarbeitung verschiedener optisch-thermaler Satellitenfernerkundungsprodukte. Eng damit verknüpft ist die Forderung nach strenger Modularisierung. In diesem Abschnitt sollen die Kernkomponenten des GMACCA-Prozessors vorgestellt werden. Bezugnehmend auf die Abbildung 10 auf Seite 23 werden aus Gründen der Übersichtlichkeit und des Umfangs dieser Arbeit nur ausgewählte Module des Prozessors beschrieben. Eine vollständige Dokumentation der einzelnen Module, inklusiver grafischer Modellabbildungen findet sich in der HTML-Dokumentation des GMACCA-Prozessors. Der Aufbau der HTML-Dokumentation wird im Anhang beschrieben.

Der von BORG (2007) entwickelte Algorithmus zur Bewertung der Datennutzbarkeit ist patentiert. Dies macht eine Zugriffsbeschränkung des Prozessors notwendig, so dass der GMACCA-Prozessor nur auf Hardware legitimer Personen oder Institutionen

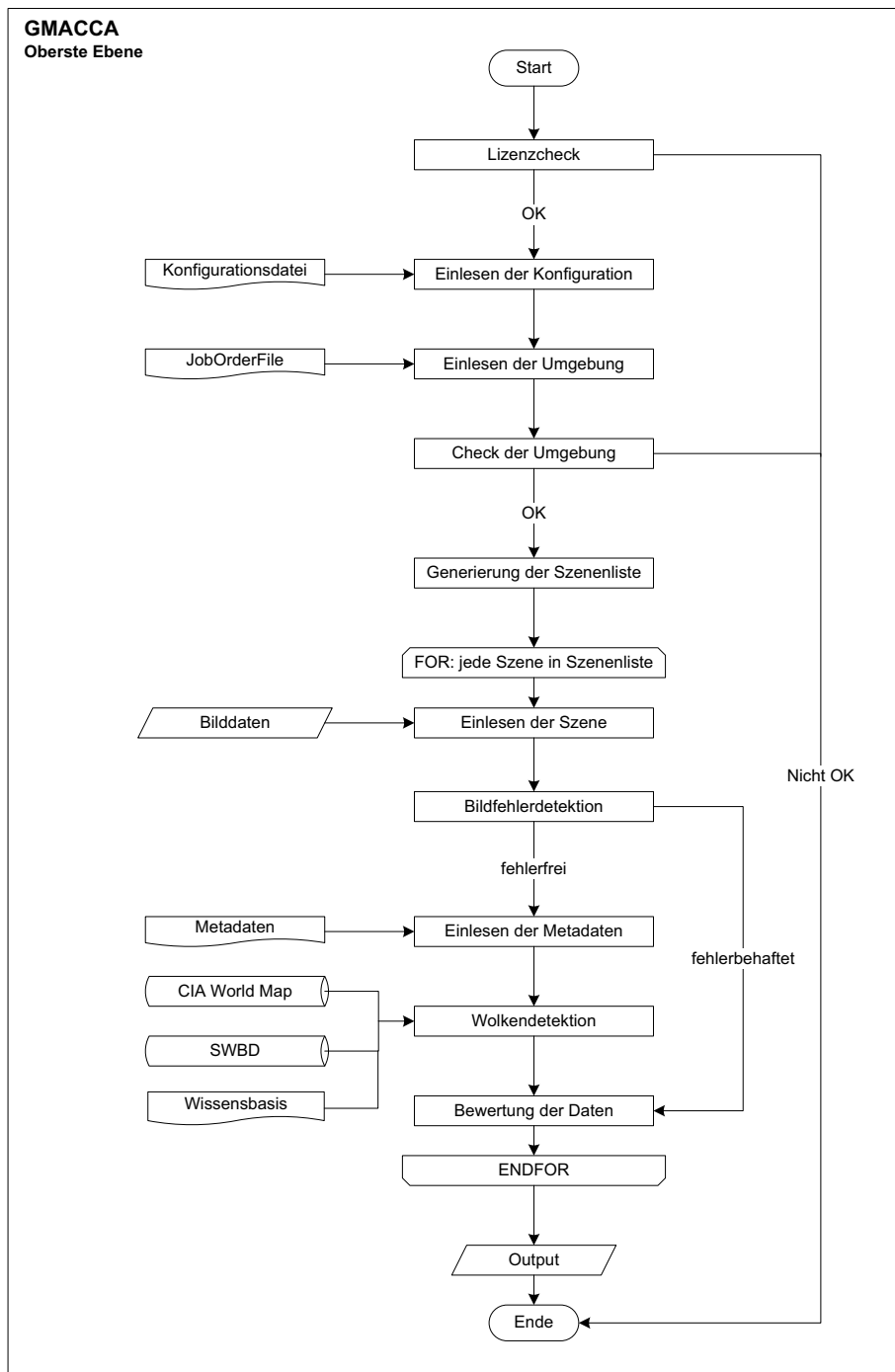


Abbildung 10: Sinngemäße Darstellung der Module des GMACCA-Prozessors in oberster Ebene.

lauffähig ist. Somit ist der erste Schritt die Feststellung einer gültigen Lizenz. Nur bei korrekter Authentifizierung kann mit der Prozessierung fortgefahren werden. Nachdem die autorisierte Nutzung des Prozessors sichergestellt wurde, erfolgt das Einlesen der eigenen Konfiguration des Prozessors. Wurde dieser Vorgang abgeschlossen, so erfolgt die Verarbeitung des JobOrderFiles. Der Prozessor verarbeitet also seine ihm dynamisch zugewiesene Prozessierungsumgebung. Anschließend überprüft der Prozessor, ob alle zur Verarbeitung der Daten notwendigen Pfade existieren und das Rechte-Management (Schreib- und Leserechte) der verschiedenen Pfade korrekt ist. Ist dies der Fall, so wird anhand der übergebenen Pfade zu den Bilddaten eine sogenannte Szenenliste erstellt. Dadurch werden die durch das JobOrderFile definierten Eingabedaten sortiert und in Strukturen transformiert mit denen der Prozessor arbeiten kann. Im Anschluss daran wird die soeben erstellte Szenenliste durchlaufen und es werden sequenziell mehrere Module aufgerufen. Zuerst werden alle Bilddaten einer Szene eingelesen und dann auf Bildfehler geprüft. Wurden keine Bildfehler detektiert, so werden im nächsten Schritt die Metadaten dieser Szene eingelesen.

Nach Durchlaufen aller bisher genannten Schritte – eher organisatorischer Natur – findet die eigentliche Wolkenerkennung statt. Der Prozess der Wolkendetektion erfordert verschiedene, teilweise komplexe Schritte. Entsprechend des von BORG (2007) entwickelten Algorithmus erfolgt die Wolkenklassifikation für bereits kalibrierte Bilddaten. Somit ist die Kalibrierung der Bilddaten entsprechend der in Abschnitt 2.4.2 beschriebenen Formelapparate der erste wesentliche Schritt zur Wolkendetektion. Für die Berechnung der planetaren Albedo  $\rho_\lambda$  nach Gleichung (5) auf Seite 9 werden verschiedene Parameter verwendet. Diese Parameter werden dynamisch berechnet, so dass für jede Berechnung jeweils ein Modul verantwortlich ist. Dabei handelt es sich im Speziellen um die Berechnung am Sensor gemessener Strahldichten  $L(\lambda)$ , der Distanz zwischen Erde und Sonne  $d$  und der Berechnung der Zeitgleichung, also der Differenz zwischen wahrer Ortszeit und mittlerer Ortszeit (BENNETT et al., 2009, S. 136). Diese Parameter fließen dann in die Berechnung der Sonnenzenitwinkel  $\theta_S$  mit ein. Obwohl die Sonnenzenitwinkel für die Zentrumsordinate einer Szene bei Produkten des Satelliten Landsat 7 in den Metadaten enthalten sind, werden sie in diesem Fall pixelgenau berechnet, da die Berechnung der planetaren Albedo ebenfalls pixelweise stattfindet. Sowohl die Beachtung der Zeitgleichung als auch die pixelweise Berechnung der Sonnenzenitwinkel erlauben nach BORG (2007) eine präzise Kalibrierung der Bilddaten und führen somit zu exakteren Klassifikationsergebnissen. Die für die Ermittlung der am Sensor gemessenen Strahldichten  $L(\lambda)$  notwendigen band- und sensorspezifischen Kalibrierungskoeffizienten – wie zum Beispiel



zur Berechnung von *Bias* oder *Gain* – werden der zuvor angesprochenen Wissensbasis entnommen. Für die Klassifikation der Wolkenpixel werden vier verschiedene Wolkendetektionsalgorithmen verwendet. Dabei handelt es sich um einen an die hiesigen regionalen Bedingungen angepassten ACRES-Klassifikator, sowie um die in BORG (2007, S. 58 – 65) beschriebenen drei Wolkendetektionsmodule des MACCA-Prozessors. Die verschiedenen Detektoren sind durch das Setzen von Parametern im JobOrderFile dynamisch einsetzbar. Sofern Wolkendetektionsmodule des MACCA-Processors verwendet werden, wird eine Land-Wasser-Maske des entsprechenden Gebietes der Szene erzeugt. Dazu werden sowohl die Daten der CIA World Map als auch die SWBD-Daten verwendet, um den Optimierungsansätzen aus Abschnitt 2.6 gerecht zu werden. Nach Abschluss der Klassifikation kann das Klassifikationsergebnis optional als Bild exportiert werden. Dazu wird mithilfe der Spektralkanäle der Satellitenszene ein RGB-Bild erzeugt. Da die spektralen Bandbreiten pro Kanal je nach Sensor variieren, werden stets die Kanäle ausgewählt, die am ehesten den RGB-Kanälen entsprechen. Im so erzeugten RGB-Bild werden detektierte Wolken auf verschiedene Art und Weise gekennzeichnet. So kann zwischen der einfachen farbigen Umrandung und der zusätzlichen farbigen Füllung der detektierten Wolkenbereiche gewählt werden. Diese Ausgabeoptionen werden in der Konfigurationsdatei des Prozessors festgelegt (siehe Abschnitt 4.3).

Die erzeugten Wolkenmasken dienen als direkte Grundlage zur Bewertung der Datennutzbarkeit. Die Bewertung der Datennutzbarkeit erfolgt stets für Teilbereiche der Satellitenszene. So werden die Quicklooks des Satelliten Landsat 7 beispielsweise in Quadranten geteilt, welche unabhängig voneinander bewertet werden. Dies hat verschiedene Gründe. Zum einen verringert sich der Berechnungsaufwand zur Erzeugung einer Distanzmatrix mit schwindender Bildgröße. Zum anderen lassen getrennte Analysen eine differenziertere Auswertung zu. So ist es möglich, dass zwar drei Quadranten einer Szene vollständig mit Wolken bedeckt sind, jedoch der Quadrant, der die zu interpretierenden Daten enthält, weniger oder gar keine Wolkenbedeckung aufweist. Somit wäre ein solcher Quadrant unter Umständen noch nutz- und auswertbar. Dieser Aspekt kommt vor allem dann zum Tragen, wenn die Erfordernis einer Auswertung hochaktueller Daten besteht und keine alternativen aktuellen Daten genau dieser Region verfügbar sind.

Die eigentliche Bewertung der einzelnen Quadranten erfolgt wie in Abschnitt 2.4.5 dargestellt. Die Ergebnisse, sowohl der Wolkenklassifikation als auch der Datennutzbarkeitsanalyse, können persistent in Form von Bild- oder Textdateien gespeichert werden. Diese Optionen werden ebenfalls durch das Setzen verschiedener Parameter im JobOr-

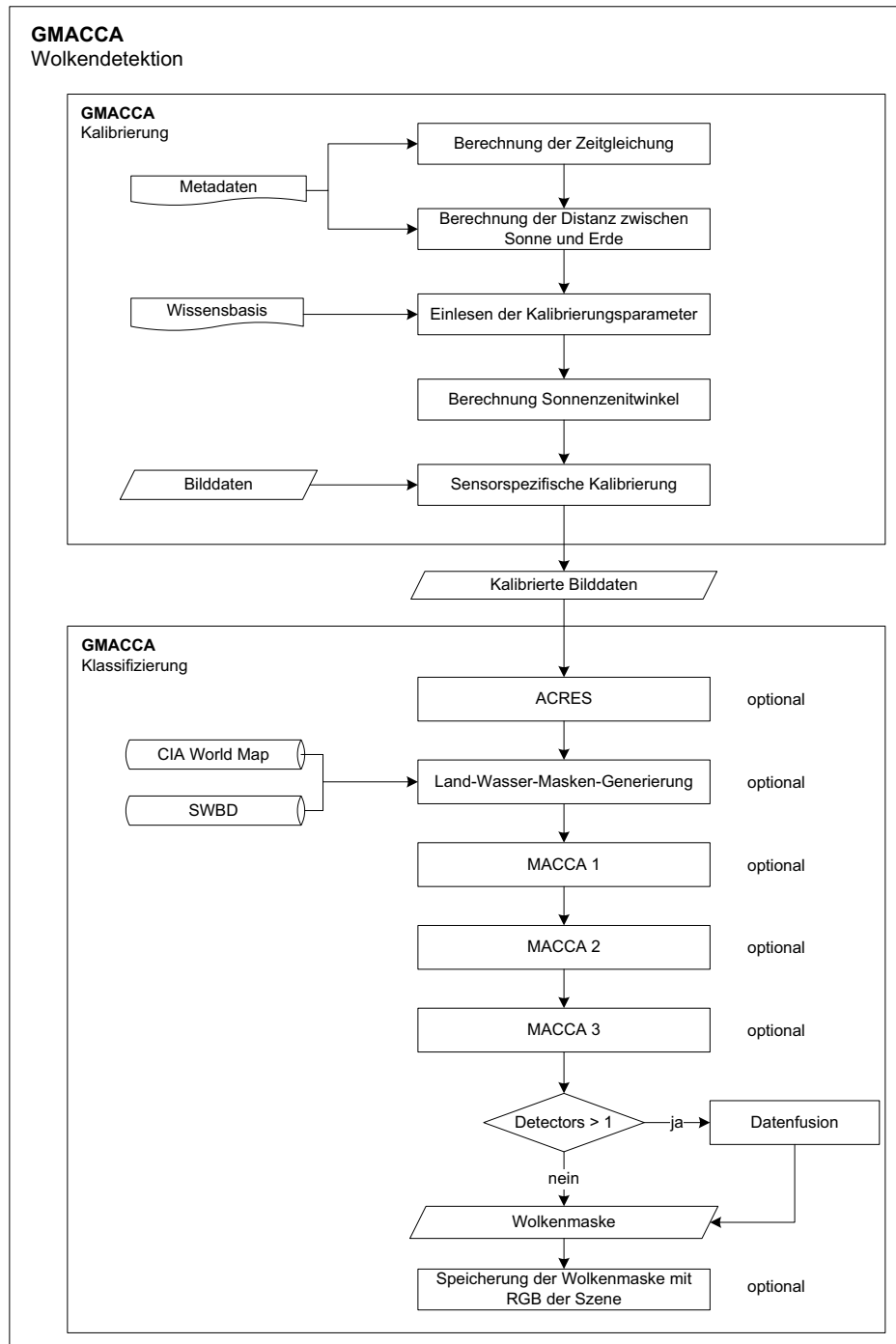


Abbildung 11: Darstellung der verschiedenen Komponenten des GMACCA-Prozessors im Hinblick auf die Detektion von Wolken. Die Wolkendetektion gliedert sich in die Kalibrierung der Bilddaten und die Klassifizierung der kalibrierten Bilddaten. Die Auswahl geeigneter Klassifikatoren erfolgt dynamisch durch das Setzen verschiedener Parameter im JobOrderFile. Bei der Nutzung eines MACCA-Klassifikators erfolgt die Generierung der Land-Wasser-Maske des Gebietes der aktuellen Szene. Werden mehrere Klassifikatoren genutzt, so erfolgt die Fusion der Wolkenmasken wie in Abschnitt 2.4.4 erläutert.

derFile festgelegt. Abbildung 11 auf Seite 26 veranschaulicht den Verfahrensablauf der Wolkendetektion des GMACCA-Prozessors.

### 3.3. Generische Komponenten des Prozessors

Verschiedene optisch-thermale Satellitenfernerkundungsprodukte unterscheiden sich aus der Sicht des Prozessors hinsichtlich verschiedener Kriterien. Ein Unterscheidungskriterium ist die Auflösung der Satellitendaten. Die Auflösungen – vor allem die geometrischen und spektralen Auflösungen – der im All befindlichen Sensoren variieren mitunter stark. Dies beeinflusst vor allem die Dimensionen der Szenen und deren Dateigröße. Während Unterschiede der radiometrischen und spektralen Auflösung vor allem sensor-spezifisch sind, so variiert die geometrische Auflösung gar pro Band eines Sensors. So ist die geometrische Auflösung des panchromatischen Kanals in der Regel höher als die der Multispektralkanäle des gleichen Sensors. Die Auflösungen einer Satellitenszene haben direkten Einfluss auf die Verarbeitungszeit und erfordern unter Umständen verschiedene Verarbeitungstechniken.

Neben der unterschiedlichen Auflösung werden Satellitenfernerkundungsprodukte in den unterschiedlichsten Datenformaten ausgeliefert. Dabei kann es sich zum Beispiel um JPEG-Bilder für Quicklooks, um GeoTIFFs oder um eigens für den Sensor entwickelte Formate handeln. Beispielsweise werden die Produkte des Sensors *Enhanced Thematic Mapper (ETM<sup>+</sup>)* des Satelliten Landsat 7 unter anderem im exotischen FST-Format ausgeliefert. Somit verfügt der GMACCA-Prozessor sowohl über Module für die Eingabe als auch für die Ausgabe verschiedener Bildformate.

Neben den unterschiedlichen Bildformaten unterscheiden sich auch der Aufbau und der Umfang der Metadaten je nach Satellitenprodukt. Insbesondere das Auslesen von Zeitinformationen, der Koordinaten einer Szene und verschiedener Parameter für die Kalibrierung der Bilddaten sind für den Algorithmus der Datennutzbarkeitsbewertung ausschlaggebend. Darüber hinaus besitzen verschiedene Satellitenprodukte auch unterschiedliche Namenskonventionen. Dort sind zum Beispiel Meta-Angaben (Bezeichnung des Satelliten, Sensor, Band . . .), räumliche Angaben (Orbit, Pass, Szene . . .) und zeitliche Angaben (Aufnahmezeitpunkt) kodiert. Da diese Informationen in der Prozessierung genutzt werden, um beispielsweise Szenenlisten zu erzeugen, werden dafür produktspezifische Module entwickelt. Nachdem dargestellt wurde, welche Unterscheidungskriterien für die Entwicklung dieses Prozessors relevant sind, wird im folgenden Absatz das Prinzip der sensorspezifischen Modulauswahl erläutert.

Damit der GMACCA-Prozessor verschiedenste optisch-thermale Satellitenfernerkun-

dungsprodukte verarbeiten kann, müssen diese durch den Prozessor hinsichtlich verschiedener Kriterien differenzierbar sein. So muss beim Start der Prozessierung der Bild- und Metadaten zunächst überprüft werden, welches Aufnahmesystem die Daten detektiert hat. Dies geschieht durch einfache Abfragen. Die relevanten Informationen über das Aufnahmesystem erhält der Prozessor aus dem JobOrderFile, welches zur eindeutigen Identifizierung der zu verarbeitenden Daten durch verschiedene Prozessierungsparameter erweitert wurde (siehe folgenden Abschnitt 4.2). Im Speziellen handelt es sich dabei um die Bezeichnung des Satelliten, die Bezeichnung des Sensors und die Bezeichnung des Produkttyps. Nachdem der Prozessor seine Umgebung – also das JobOrderFile – eingelesen hat, können die entsprechenden Module für das Verarbeiten dieser Daten über verschiedene Abfragen geladen werden. Wird ein Aufnahmesystem vom Prozessor nicht erkannt beziehungsweise kann er die Daten eines Aufnahmesystems nicht verarbeiten, so stoppt die Verarbeitung der Daten und es wird eine entsprechende Fehlermeldung erzeugt. Die Abbildung 12 auf Seite 30 zeigt die spezifische Auswahl eines Moduls exemplarisch anhand des Moduls zum Einlesen der Metadaten.

### 3.4. Einbindung neuer Sensoren

Aufgrund des modularen Aufbaus des gesamten Prozessors wird die Einbindung neuer Produkttypen in die Prozessierung wesentlich erleichtert und erfordert kaum Quelltextanpassungen bereits bestehenden Codes. Von Seiten der PSM-Komponente des DIMS müssen bei der Generierung des JobOrderFiles lediglich die Parameter Satelliten-, Sensor- und Produktbezeichnung neu definiert werden. Durch einfache Erweiterung der Abfragen nach genau diesen Parametern auf Prozessorebene lassen sich die entsprechenden spezifischen Module direkt integrieren. Die Entwicklung der spezifischen Module kann extern erfolgen. Folgende spezifische Module müssen bei der Integration neuer Sensoren zusätzlich entwickelt werden:

- Modul zum Erstellen der Szenenliste
- Modul zum Auslesen der relevanten Metadaten
  - Modul zum Ermitteln der Koordinaten der Länge der Ecken einer Szene
  - Modul zum Ermitteln der Koordinaten der Breite der Ecken einer Szene
  - Modul zum Ermitteln der Datums- und Zeitinformationen
  - Modul zum Ermitteln der Kalibrierungsparameter wie zum Beispiel des *Gain*-Zustandes diverser Multispektralkanäle

- Modul zur Kalibrierung der aufgenommenen Bilddaten
- Modul für die Einteilung der Szene in Bewertungsbereiche für die separate Daten-nutzbarkeitsanalyse

Weiterführende Informationen und die genauen Schnittstellendefinitionen finden sich in der HTML-Dokumentation. Darüber hinaus muss die Wissensbasis des Prozessors um bestimmte Parameter des neu hinzuzufügenden Systems erweitert werden. Dies umfasst verschiedene Bahnparameter (zum Beispiel Inklination, Equator Crossing Time (ECT)) sowie Parameter zur Charakteristik der physikalischen Eigenschaften des Sensorsystems (zum Beispiel spektrale Bandbreiten der einzelnen Multispektralkanäle, Konstanten).

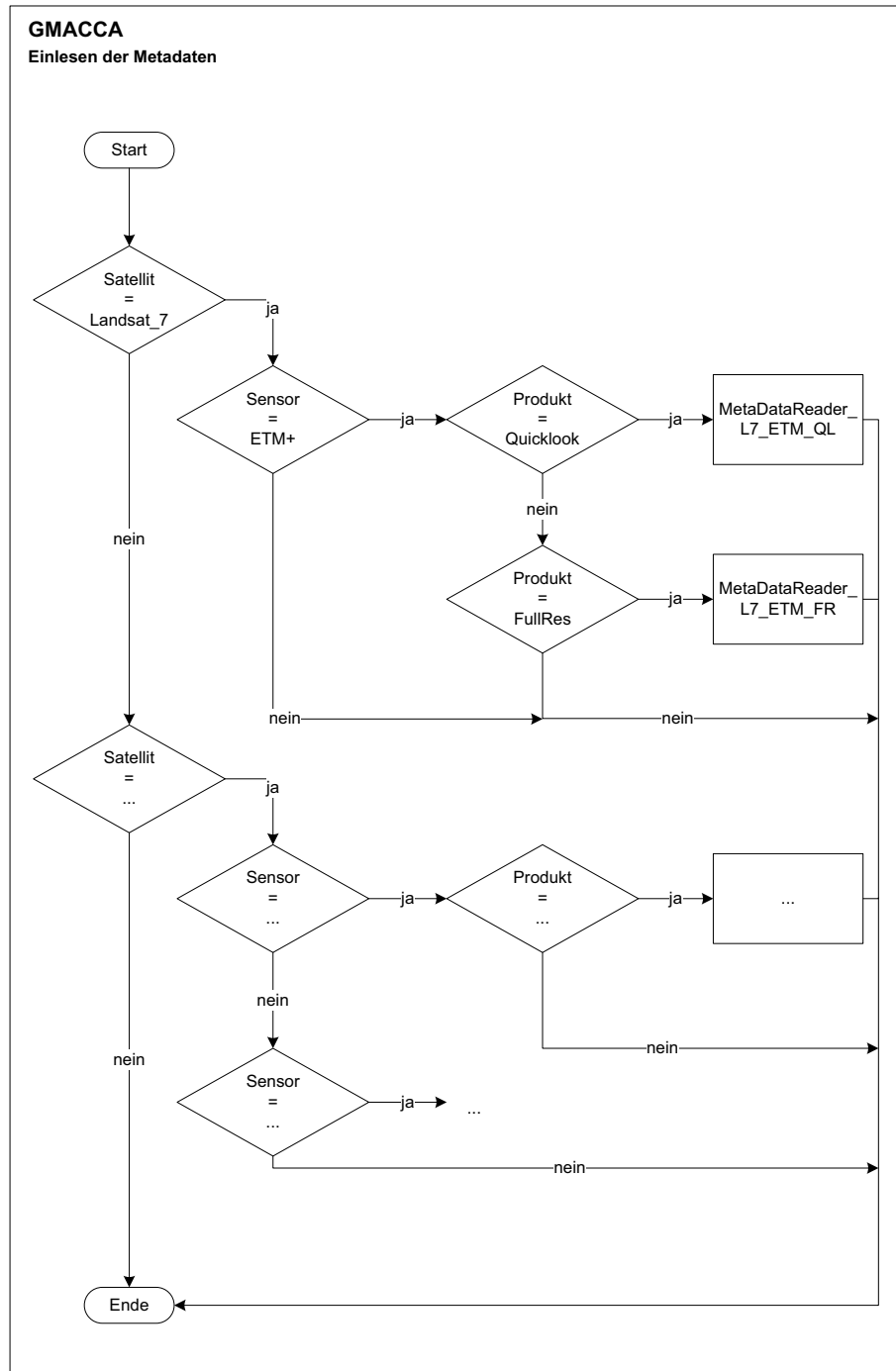


Abbildung 12: Prinzipielle Darstellung der Auswahl spezifischer Module durch einfache Abfragen spezifischer Angaben zu Satellit, Sensor und Produkt anhand des Moduls zum Einlesen der Metadaten eines optisch-thermalen Satellitenfernerkundungsproduktes des Satelliten Landsat 7.

## 4. Umsetzung

Nachdem im letzten Kapitel wesentliche Design und Architekturentscheidungen beschrieben wurden, gibt dieses Kapitel einen Überblick über die bisherige Implementierung des GMACCA-Prozessors. Im Prinzip stellt der GMACCA-Prozessor die Weiterentwicklung des von BORG (2007) entwickelten MACCA-Prozessors dar. Damit der hier entwickelte Prozessor im Zuge der Implementierung stets validiert werden kann, werden die ersten vom Prozessor zu verarbeitenden Produkte ebenfalls Quicklooks des Satelliten Landsat 7 sein. Eine ausführliche Beschreibung der Daten findet sich in BORG (2007). Durch die prozessorübergreifende Verarbeitung gleicher Datensätze können sowohl die korrekte Funktionsweise festgestellt als auch konzeptbedingte Vorteile der Neuentwicklung offenbart werden.

### 4.1. Entwicklungsumgebung

Die Entwicklung des Prozessors erfolgt auf einem Notebook, welches über einen 1,6 GHz CENTRINO DUO-Prozessor von INTEL, sowie 3 GB Arbeitsspeicher und 320 GB Festplattenkapazität verfügt. Obwohl der GMACCA-Prozessor konzeptionell plattformübergreifend arbeiten wird, findet die Entwicklung unter dem Betriebssystem WINDOWS 7 in aktueller Version statt.

Die Implementierung erfolgt in der Programmiersprache IDL in Version 7.0.6. Dies hat mehrere Gründe. Zum einen wurde bereits der MACCA-Prozessor in dieser Sprache implementiert. So wird die Adaption und Optimierung bereits entwickelter Komponenten wesentlich erleichtert. Darüber hinaus wurde die Programmiersprache primär für die Verarbeitung von Bilddaten entwickelt und verfügt über Konzepte (wie zum Beispiel Pointer, Threadpools, Ausnahmebehandlung, GUI-Programmierung), deren Einsatz auch die Entwicklung komplexer und gleichzeitig performanter Softwareapplikationen ermöglicht. Ein weiterer Vorteil der Nutzung dieser Sprache ist der relativ einfache Umgang mit Kartenprojektionen und der damit verknüpften Georeferenzierung der Daten. Schließlich lassen sich entwickelte Applikationen auch lizenzieren und distribuieren. So werden sämtliche zur Ausführung notwendigen Komponenten der IDL-Installation zusammen mit der kompilierten Applikation in einer Verzeichnisstruktur zusammengefasst. Dadurch lässt sich die Applikation auf einfache Art und Weise verteilen und durch die Nutzung der *IDL Virtual Machine* auf legitimierten Rechnern kostenlos ausführen, sogar wenn dort keine Installation von IDL vorgenommen wurde.

Für die Erstellung der Dokumentation des Prozessors werden zwei weitere Tools ver-

wendet. Die HTML-Dokumentation wird mithilfe von NATURALDOCS in Version 1.4 erstellt. In dieser Dokumentation sind grafische Darstellungen der Abläufe der verschiedenen Module eingebettet. Diese werden mithilfe des Tools PAPDESIGNER in Version 2.2.0.8 erstellt.

## 4.2. Erweiterung des JobOrderFiles

Bezugnehmend auf Abschnitt 3.1 erhält der GMACCA-Prozessors seine Prozessierumgebung in Form eines dynamisch erzeugten JobOrderFiles. Dieses JobOrderFile wurde um verschiedene Parameter erweitert. Diese werden in Tabelle 1 dargestellt. Die Werte der Parameter `Satellite`, `Sensor` und `ProductName` dienen dem Prozessor als Entscheidungsgrundlage für die Auswahl spezifischer Module (siehe Abschnitt 3.3). Die Parameter `Channels`, `ColumnsIn` und `LinesIn` spezifizieren die Auflösung der Eingabebilddaten. Durch den Parameter `CCA-Output` wird festgelegt, ob als Nebenprodukt ein RGB-Bild der Klassifikationsergebnisse erzeugt werden soll. Die Dimensionen des RGB-Bildes werden durch die Parameter `ColumnsOut` und `LinesOut` festgelegt. Der Parameter `DQA-Output` steuert die Ausgabe der Ergebnisse der Datennutzbarkeitsbewertung. Darauf Einfluss nehmend wird durch `ReadOperatorVote` festgelegt, ob eventuell vorliegende Datennutzbarkeitsbewertungen aus den Metadaten mit eingelesen und somit vergleichend wieder ausgegeben werden. Der Parameter `ReadGain` beeinflusst das Einlesen der Metadaten und ist nur bei Sensoren mit dynamischer Gain-Anpassung anwendbar. Die Parameter `UseDetector1`, `UseDetector2`, `UseDetector3` und `UseModACRES` erlauben eine skalierbare Anwendung der zur Verfügung stehenden Klassifikatoren.

Name	Wert	Name	Wert
Satellite	Landsat_7	CCAOutput	true
Sensor	ETM+	DQAOutput	true
Channels	8	UseDetector1	false
ProductName	Quicklook	UseDetector2	false
ColumnsIn	1000	UseDetector3	false
LinesIn	1000	UseModACRES	true
ColumnsOut	1000	ReadGain	true
LinesOut	1000	ReadOperatorVote	true

Tabelle 1: Auflistung der dem JobOrderFile hinzugefügten Parameter im Teilbaum `ProcessingParameters`.



### 4.3. Konfiguration des Prozessors

Der hier entwickelte GMACCA-Prozessor ist konfigurierbar. Dafür wurde eine XML-Konfigurationsdatei definiert. Deren Inhalt wird in Abbildung 13 auf Seite 34 dargestellt. Dem Hauptknoten `<Configuration>` folgen drei wesentliche Unterknoten.

Der Knoten `<General_Information>` enthält Informationen, die den Prozessor beschreiben. Somit handelt es sich in gewisser Art und Weise um Metainformationen. Neben der Prozessorbezeichnung (`<Processor_Name>`) und der Version (`<Version>`) wird hier die Anzahl der verfügbaren Wolkendetektionsmodule (`<Number_Of_Detectors>`) festgelegt. Der Prozessor verfügt konzeptionell über vier Wolkendetektionsmodule. Neben dem bereits implementierten ACRES-Klassifikator wurde die zukünftige Verwendung der drei Klassifikationsmodule des MACCA-Prozessors konzipiert.

Im Knoten `<Parameters>` werden verschiedene Parameter definiert, die den Prozessor in seiner Ausführung beziehungsweise in seiner Kommunikationsfähigkeit beeinflussen. Der GMACCA-Prozessor verfügt konzeptionell über eine modulare grafische Oberfläche. Diese lässt sich durch das Setzen des Parameters `<GUI_Enabled>` aktivieren oder deaktivieren. Über den Parameter `<Fill_Clouds>` lässt sich die visuelle Ausgabe der detektierten Wolken beeinflussen. Wird dieser Parameter auf `true` gesetzt werden die klassifizierten Wolken gelb gefüllt und rot umrandet. Bei Zuweisung von `false` werden die detektierten Wolken lediglich rot umrandet. Der Parameter `<Verbose>` skaliert die Anzahl der Ausgaben auf die Standardausgabe. Es werden drei verschiedene Szenarien unterstützt. Um die Ausgabe von Informationen so weit wie möglich einzuschränken, ist dem Parameter `Verbose` der Wert 0 zuzuweisen. Diese Einstellung ist vor allem dann sinnvoll, wenn der Prozessor operativ arbeitet und maximale Performance notwendig ist. Gegenteilig kann dem Parameter der maximale Wert 2 zugewiesen werden. In diesem Fall werden sämtliche Ausschriften angezeigt. Dieser Modus ist vor allem bei der Entwicklung – wie zum Beispiel beim Debugging – sinnvoll. Die Zuweisung des Wertes 1 stellt einen Kompromiss zwischen optimaler Laufzeit und Informationsdichte dar.

Der letzte der drei Unterknoten von `<Configuration>` ist der Knoten `<Paths>`. Hier werden die zur Ausführung notwendigen Pfade gesetzt. Der Knoten `<Temp>` verweist auf das Verzeichnis in dem temporäre Dateien abgelegt werden. Zur Kalibrierung des grafischen Systems und der korrekten Darstellung der Ergebnisse ist eine definierte Farbübersetzungstabelle (Color Translation Table) notwendig. Deren Pfad wird durch den Parameter `<Color_Table>` festgelegt. Wie in den Abschnitten 2.6 und 3.2 bereits beschrieben, werden für die Generierung einer Land-Wasser-Maske des aufgenommenen Ge-

```
1  <?xml version="1.0" encoding="ISO-8859-1"?>
2  <Configuration>
3      <!--> allgemeine Informationen für den generischen Prozessor <!-->
4      <General_Information>
5          <Processor_Name>Generic_MACCA</Processor_Name>
6          <Version>prototype</Version>
7          <Number_Of_Detectors>4</Number_Of_Detectors>
8      </General_Information>
9      <!--> verschiedene Schalter zur Konfiguration des Prozessors <!-->
10     <Parameters>
11         <GUI_Enabled>false</GUI_Enabled>
12         <Fill_Clouds>false</Fill_Clouds>
13         <Verbose>0</Verbose>
14     </Parameters>
15     <!--> verschiedene Pfade zu notwendigen Dateien des Prozessors <!-->
16     <Paths>
17         <Temp>C:\IDL_Projects\GMACCA\temp</Temp>
18         <Color_Table>C:\IDL_Projects\GMACCA\ini\idlsave.dat</Color_Table>
19         <SRTM_Data>C:\IDL_Projects\GMACCA\ini\SRTM_SWDB_2_1</SRTM_Data>
20     </Paths>
21 </Configuration>
```

Abbildung 13: Konfigurationsdatei gmacca.conf des GMACCA Prozessors.

bietet einer Szene unter anderem SWBD-Daten verwendet. Der Parameter `<SRTM_Data>` verweist auf die SWBD-Daten.

Der eigentliche Vorgang der Konfiguration kann im ersten Schritt durch manuelles Editieren des XML-Dokumentes und im nächsten Schritt durch die konzipierte grafische Oberfläche erfolgen. Darüber hinaus kann der Umfang der Parameter im Zuge weiterer Entwicklungsarbeit durchaus erweitert werden.

## 4.4. Implementierung der Module

Im Zuge der Entwicklung wurde der GMACCA-Prozessor soweit implementiert, dass der Funktionsumfang des MACCA-Prozessors nahezu abdeckt wird. Das heißt, der neu entwickelte GMACCA-Prozessor ist in der Lage Quicklooks des Satelliten Landsat 7 zu verarbeiten. Die Wolkendetektion erfolgt dabei einzig und allein auf Grundlage der Nutzung des ACRES-Prozessors. Die nun folgenden Unterabschnitte orientieren sich an der Reihenfolge der Prozessierung wie sie in Abbildung 10 auf Seite 23 dargestellt wurde. Aus unter anderem patentrechtlichen Gründen erfolgt die Beschreibung der bisher implementierten Module auf einem Niveau, welches dem Leser zwar erlaubt den grundlegenden Ablauf der Prozessierung nachzuvollziehen, jedoch verschiedene Implementierungsdetails verbirgt.

#### 4.4.1. Einlesen der Konfiguration

Im Abschnitt 4.3 wurde die Konfigurationsdatei des GMACCA-Prozessors beschrieben. Das Modul `ConfigParser` extrahiert die Informationen aus dieser Konfigurationsdatei unter Verwendung verschiedener Untermodule. Abbildung 14 zeigt das dafür verantwortliche Modul einschließlich der Untermodule. Das Untermodul `getDOMXML` erzeugt das *Document Object Model (DOM)* der XML-Datei. Die anderen drei Module dienen dazu, die drei Hauptknoten des DOMs rekursiv zu durchlaufen und deren Parameter einzulesen. Die Daten der einzelnen Teilbäume werden in Form zweidimensionaler Arrays abgelegt. Während die erste Spalte die Bezeichnung des Parameters enthält, wird in der zweiten Spalte der zugehörigen Wert abgelegt. Wurden alle drei Teilbäume eingelesen, so werden die drei Arrays in eine Struktur übernommen. Der Zeiger auf diese Struktur wird an das aufrufende Modul zurückgegeben.

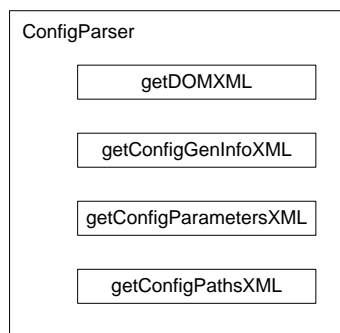


Abbildung 14: Das Modul `ConfigParser` extrahiert mithilfe verschiedener Untermodule Informationen aus der Konfigurationsdatei des GMACCA-Prozessors.

#### 4.4.2. Einlesen der Prozessierungsumgebung

Das Einlesen des `JobOrderFiles` geschieht prinzipiell nach dem gleichen Schema wie das Einlesen der Konfigurationsdatei. Entsprechend der Abbildung 15 auf Seite 36 sind dafür mehrere Module verantwortlich. Wie bereits beim Einlesen der Konfigurationsparameter wird das Modul `getDOMXML` für die Erzeugung des DOM genutzt. Anschließend werden die Parameter aus den verschiedenen Teilbäumen gelesen und in Arrays abgelegt. Diese Arrays werden wiederum in eine definierte Struktur übernommen, dessen Zeiger vom Modul `Parser_JobOrderFile` zurückgegeben wird.

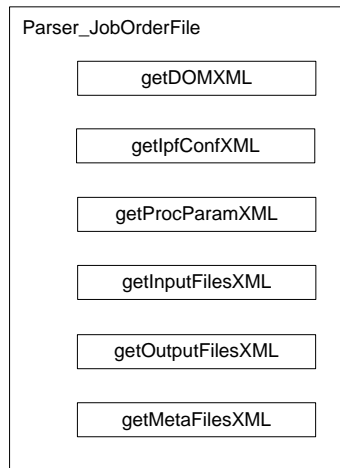


Abbildung 15: Durch die Verwendung verschiedener Untermodule werden die Parameter aus dem JobOrderFile eingelesen.

#### 4.4.3. Check der Prozessierungsumgebung

Nachdem das Einlesen der Konfigurationsdatei und des JobOrderFiles abgeschlossen ist, werden durch das Modul `FileSystemCheck` sämtliche Verzeichnisse hinsichtlich deren Existenz überprüft. Dazu werden die IDL-Funktionen `FILE_TEST` und `FILE_SEARCH` aufgerufen. Sollten Ausgabeverzeichnisse nicht existieren, so werden diese neu angelegt und mit Schreibrechten versehen. Leseverzeichnisse, die zum Beispiel die Bild- und Metadaten enthalten, werden hinsichtlich richtig gesetzter Leserechte geprüft. Sofern alle Verzeichnisse vorhanden und deren Rechte korrekt gesetzt sind, kann die Prozessierung fortgesetzt werden.

#### 4.4.4. Generierung der Szenenliste

Das Modul `SceneListBuilder` ist für die Erzeugung einer Struktur zuständig, in der sämtliche zu verarbeitende Bild- und Metadaten sortiert enthalten sind. Diese Struktur enthält eine Liste der sortierten Bilddaten, eine Liste sortierter Metadaten, Information über die Anzahl der Szenen und Informationen über die Anzahl der Bilder pro Szene, also der Anzahl der Kanäle. Da die Namenskonventionen pro Produkt variieren, werden zur Erzeugung der Struktur produktspezifische Module eingesetzt. Diese zerteilen die Pfade der durch das JobOrderFile definierten Bild- und Metadaten nach den entsprechenden Regeln und erzeugen die entsprechenden Listen. Bei den Listen handelt es sich wiederum um zweidimensionale Arrays.

### 4.4.5. Einlesen der Szenen

Das Einlesen der Szenen erfolgt pro Band durch das Modul `SceneReader`. Dieses versucht im ersten Schritt die Informationen über das Bildformat und die Dimensionen des Bildes durch die Anwendung der IDL-Funktion `QUERY_IMAGE` zu ermitteln. Können auf diese Art und Weise keine hinreichenden Informationen über das Bildformat und die Dimensionen gesammelt werden, so werden die Informationen über das Format aus dem Dateipfad und die Informationen über die Dimensionen aus den Parametern im `JobOrderFile` (`ColumnsIn`, `LinesIn`) ermittelt. Anschließend werden die Bilder eingelesen. Dazu werden entsprechend des Bildformates geeignete IDL-Funktionen aufgerufen. Konzeptionell kann der Prozessor sämtliche durch IDL unterstützte Bildformate lesen und schreiben. Eine Aufzählung der durch IDL nativ unterstützten Bild- und Datenformate findet sich in ITT VIS (2009b, S. 10-12). Bisher implementiert wurde das Einlesen von Quicklooks des Satelliten Landsat 7 im Bildformat JPEG.

### 4.4.6. Bildfehlerdetektion

Nachdem die Bilder eingelesen wurden, erfolgt die Bildfehlerdetektion durch das Modul `DataErrorAssessment`. Dazu wird ein Bild einer Szene zeilenweise durchlaufen. Eine Zeile wird dann als fehlerhaft eingestuft, wenn diese eine bestimmte Anzahl zusammenhängender Spalten mit einem hohen Anteil schwarzer Pixel aufweist.

### 4.4.7. Einlesen der Metadaten

Das Modul `MetaDataReader` realisiert das Einlesen der Metadaten. Dazu werden dem Modul die Daten aus dem `JobOrderFile` und die Szenenliste übergeben. Da der Aufbau und Inhalt der Metadaten produktspezifisch ist, werden auch hier spezifische Module geladen. Die Abbildung 12 auf Seite 30 stellte dieses Modul bereits in Abschnitt 3.4 grafisch dar.

Somit ist das Modul `MetaDataReader_L7_ETM_QL` für das Einlesen der Metadaten für die Quicklook-Produkte des Satelliten Landsat 7 verantwortlich. In Abbildung 16 auf Seite 38 werden die bereits implementierten Module zum Auslesen dieser Daten dargestellt. Die Metadaten der Quicklooks bestehen im Grunde aus einer MTA-Datei pro Szene und einer TXT-Datei pro Band. Während die TXT-Datei vor allem Orbitparameter und Eckkoordinaten enthält, so finden sich bandspezifische Parameter wie zum Beispiel der *Gain*-Zustand in der MTA-Datei. Darüber hinaus sind in der MTA-Datei die Bewertungen der Operateure hinsichtlich der Datennutzbarkeit der einzelnen Quadranten hinterlegt.

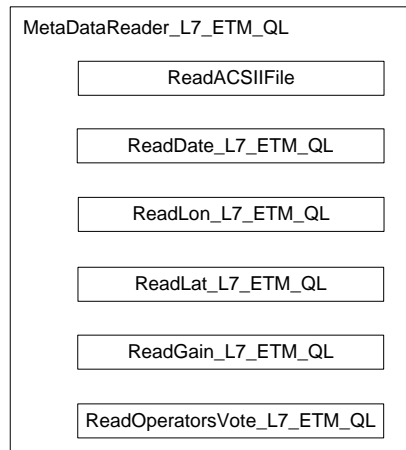


Abbildung 16: Module zum Auslesen der Metadaten eines Quicklook-Produktes des Satelliten Landsat 7.

#### 4.4.8. Wolkendetektion

Die Wolkendetektion beruht auf der Verarbeitung kalibrierter Bilddaten. Das Verfahren der Kalibrierung wurde bereits im Abschnitt 2.4.2 beschrieben. Entgegen der konzeptionellen Abbildung 11 auf Seite 26 wurde der Wolkendetektionsalgorithmus bisher nur rudimentär implementiert. Abbildung 17 auf Seite 39 zeigt die bereits implementierten Module. Demnach wurde die Kalibrierung der Bilddaten bereits vollständig implementiert. Dies umfasst auch die Implementierung des Moduls `CalibrateImages_L7_ETM` für die Kalibrierung der Quicklooks des Satelliten Landsat 7. Die dafür erforderlichen Kalibrierungsparameter werden bis zur endgültigen Spezifizierung und Nutzung der Wissensbasis durch das Dummy-Modul `ReadCalibrationData` zur Verfügung gestellt.

Im Hinblick auf die Module zur Klassifikation der Wolken wurde bisher nur der modifizierte ACRES-Klassifikator implementiert. Die Implementierung der Klassifikatoren des MACCA-Prozessors erfolgt in der weiteren Entwicklung. Da zur Zeit auch nur ein Klassifikator einsatzfähig ist, wurde im gleichen Zuge auf die bisherige Implementierung des Moduls für die Datenfusion der verschiedenen Wolkenmasken verzichtet. Das Modul zur Ableitung der Land-Wasser-Maske des betroffenen Gebietes wurde bereits implementiert. Dies umfasst sowohl die Nutzung der Daten der CIA World Map als auch der SWBD-Daten. Sofern die Ausgabe der Ergebnisse der Wolkendetektion gewünscht ist, erzeugt das Modul `ClassifiedSceneBuilder` ein entsprechendes RGB-Bild in dem die Wolken hervorgehoben sind.

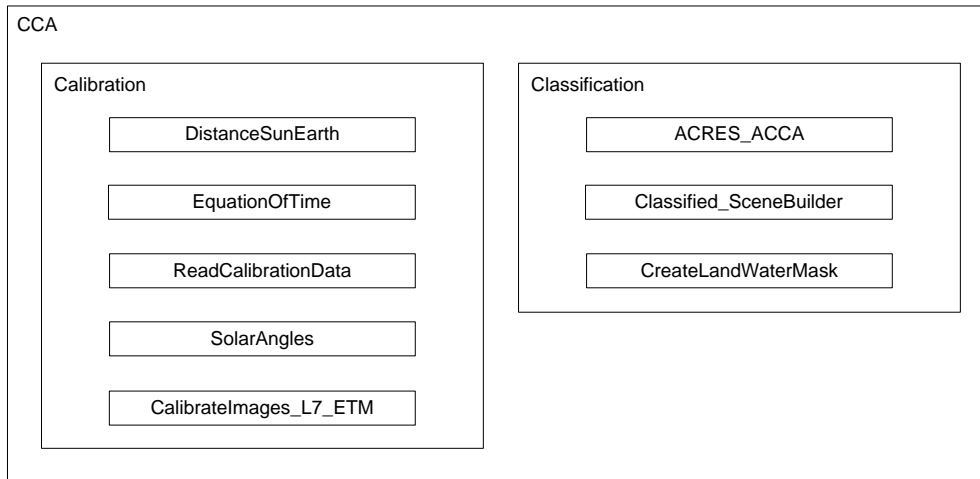


Abbildung 17: Darstellung bereits implementierter Module zur Wolkendetektion.

#### 4.4.9. Bewertung der Daten

Die Bewertung der Datennutzbarkeit erfolgt durch das Modul `DataQualityAssessment`. Bevor es zur eigentlichen Bewertung kommt, wird die Wolkenmaske für die weitere Verarbeitung modifiziert. So werden nur Wolkengebiete ab einer bestimmten Größe für die Bewertung der Datennutzbarkeit herangezogen. Anschließend wird die vorbereitete Wolkenmaske in verschiedene Teilbereiche eingeteilt. Die Art der Teilung wird durch den Produkttyp festgelegt. Entsprechend den Vorschriften für die Quicklooks des Satelliten Landsat 7 ermittelt das Untermodul `getSections_L7_ETM_QL` die Eckkoordinaten der vier zu bewertenden Quadranten. Die Weiterverarbeitung der Quadranten erfolgt durch das Modul `SectionQualityAssessment`. Hier wird die entsprechende Distanzmatrix erzeugt. Im Gegensatz zum MACCA-Prozessor wird dazu die IDL-Funktion `MORPH_DISTANCE` genutzt, um Rechenzeit zu sparen. Auf Grundlage der so erzeugten Distanzmatrizen wird die Datennutzbarkeit ermittelt.

#### 4.4.10. Ausgabe der Ergebnisse

Für die Ausgabe der Ergebnisse wurden mehrere Module implementiert. Die visuellen Ergebnisse der Wolkendetektion werden durch das Modul `ClassifiedSceneBuilder` erzeugt und durch das Modul `SceneWriter` als Bilddatei gespeichert. Zu Testzwecken unterstützt das Modul `SceneWriter` bisher das PNG-Format als Ausgabeformat. Abbildung 18 auf Seite 40 zeigt die Ergebnisse der Wolkendetektion eines Quicklooks des Satelliten Landsat 7. Die Ergebnisse der Datennutzbarkeitsbewertung werden durch An-

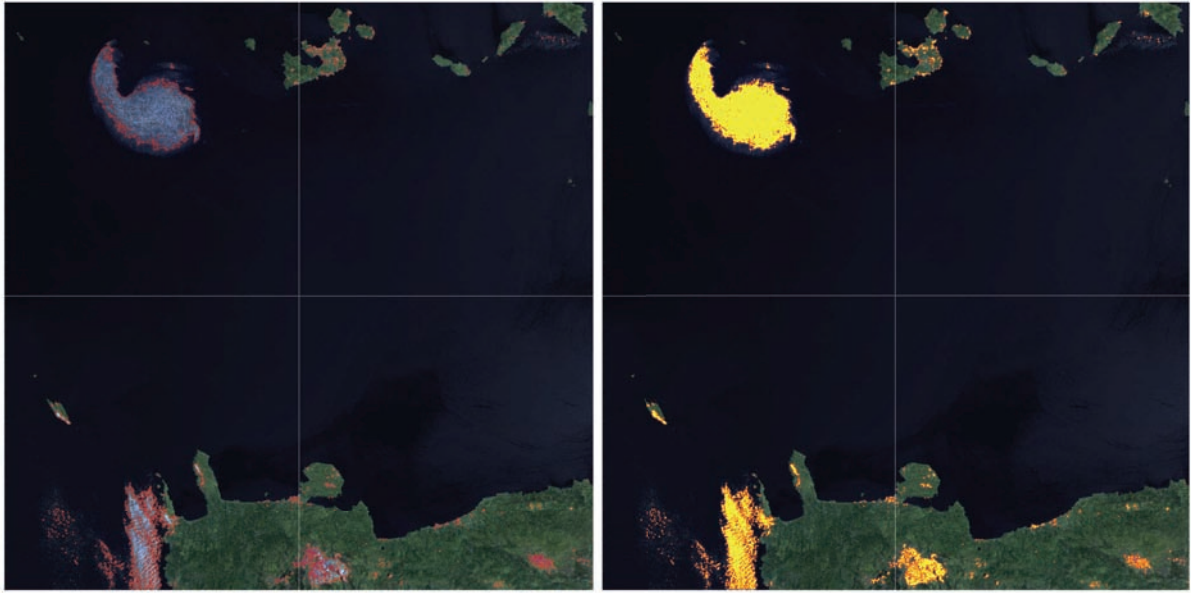


Abbildung 18: Ergebnisse der Wolkendetektion der Szene aus Listing 1 in verschiedener Darstellung.

wendung des Moduls `WriteDQAOutput` in einer Textdatei abgelegt. Diese enthält die einzelnen Bewertungen der entsprechenden Quadranten. Erhält ein Quadrant die Bewertung 0, dann lassen sich die Daten optimal nutzen da keine Wolkenbedeckung oder Dunst detektiert wurde. Eine Bewertung mit 90 entspricht einer fehlerhaften oder vollständig durch Wolken bedeckten Szene. An erster Stelle steht die mittlere Bewertung aller Quadranten, stellvertretend für die Datenqualität der gesamten Szene. Nachfolgend erscheinen die Bewertungen des oberen, rechten, unteren und linken Quadranten (beginnend oben links). Listing 1 stellt einen Auszug aus der Ergebnisdatei grafisch dar.

```

1 Scene:          LN7T-_016452_0182_0035_NZ_LRS--SJ-001-QLUC.jpg
2 Operator:      15      20      0      30      10
3 Automat:       30      40     10     40     30

```

Listing 1: Die Ergebnisse der Datennutzbarkeitsbewertung eines Quicklooks des Satelliten Landsat 7.



## 4.5. Demonstrative Anwendung des Prozessors

In diesem Abschnitt soll die Funktionsfähigkeit des GMACCA-Prozessors demonstriert werden. Dazu wird ein Testdatensatz prozessiert. Bei den Testdaten handelt es sich um Quicklooks des Satelliten Landsat 7. Die Quicklooks sind im Grunde vorprozessierte Szenen in verringerter geometrischer Auflösung (1000 mal 1000 Pixel) im Format JPEG. Pro Kanal wurde ein Bild erzeugt, sodass eine Szene aus acht Einzelbildern besteht. Die Szenen wurden am 19. Mai 2002 vom Sensor detektiert. Die Generierung der Quicklooks erfolgte am 22. Mai 2002 durch die Bodenstation Neustrelitz. Im Anschluss an die Prozessierung der Quicklooks wurde die Datennutzbarkeit dieser Szenen durch sogenannte Operateure visuell bestimmt. Die Ergebnisse dieser Analysen wurden in den Metadaten abgelegt.

Bei der Auswahl der Szenen wurden verschiedene Sachverhalte berücksichtigt, um den Prozessor eingehend zu testen. Zum einen sollen fehlerhafte Szenen prozessiert werden. Dadurch soll überprüft werden, ob der Prozessor fehlerhafte Szenen erkennt und deren weitere Verarbeitung versagt. Neben der Prozessierung fehlerhafter Szenen, soll eine Szene eine Gain-Umschaltung des Sensor beinhalten. Durch die Umschaltung des Gain-Zustandes kann die maximale Sensitivität des Sensors – beispielsweise beim Überflug von schneebedeckten Gebieten – beeinflusst werden. Die Veränderung der maximalen Sensitivität eines Sensor kann somit mitten in einer Szene stattfinden. Diese Umschaltung ist bei der Kalibrierung der Bilddaten zu beachten, damit eine korrekte Wolkendetektion erreicht werden kann. Bei der Auswahl der Szenen wurde zusätzlich darauf geachtet, dass sowohl die Wolken- als auch die Landbedeckung variiert, um möglichst variable Testbedingungen zu schaffen. Für die Prozessierung der Szenen wurde ein JobOrderFile generiert. Dieses enthält neben den Pfaden zu den Bild- und Metadaten auch die in Abschnitt 4.2 vorgestellten Parameter. Die Werte der Parameter entsprechen denen aus Tabelle 1 auf Seite 32.

Der GMACCA-Prozessor verarbeitete insgesamt 16 verschiedene Szenen. Die Tabelle 2 auf Seite 46 fasst die Verarbeitungszeiten der 16 Szenen zusammen. Die Gesamtverarbeitungszeit betrug 136,32 Sekunden. Dies entspricht einer mittleren Verarbeitungszeit pro Szene von etwa 8,52 Sekunden. Die mittlere Verarbeitungszeit ist jedoch wenig aussagekräftig, da die Verarbeitungszeit sowohl von der Bildqualität als auch vom Ausmaß der zu detektierenden Wolkenbedeckung abhängt. Fehlerhafte Szenen werden vom Prozessor übersprungen und weisen somit eine sehr geringe Verarbeitungszeit auf. Darüber hinaus haben sowohl die Wolkendichte als auch die Verteilung der Wolken direkten Einfluss auf die Verarbeitungszeit. Diese Faktoren bestimmen unter anderem die Komplexität der

Berechnung der Distanzmatrix. Darüber hinaus werden für die Berechnung der Datennutzbarkeit lediglich Wolken einer bestimmten Größe zugelassen. Bei Vernachlässigung des begünstigenden Einflusses fehlerhafter Szenen beträgt die bereinigte mittlere Verarbeitungszeit einer Szene etwa 9,69 Sekunden. Somit ist die Verarbeitungszeit einer Szene im Vergleich zum MACCA-Prozessor mehr als 5 mal geringer. Hierbei werden beide Prozessoren in gleicher Testumgebung und ohne Anwendung einer grafischen Benutzeroberfläche zum Vergleich herangezogen. Der Bedarf an Arbeitsspeicher während der Anwendung des GMACCA-Prozessors beträgt maximal etwa 100 MB und ist nur minimal geringer (24 MB) als der Bedarf des MACCA-Prozessors.

Die Abbildungen 27 und 28 auf Seite 55 und 56 zeigen einige Klassifizierungsfehler des verwendeten Moduls für die Wolkendetektion. Diese Fehler treten jedoch sowohl beim MACCA-Prozessor als auch beim GMACCA-Prozessor auf, da der ACRES-Klassifikator in beiden Prozessoren auf die gleiche Art und Weise implementiert wurde. Bis auf die beschriebenen Klassifizierungsfehler erfolgt die Detektion der Wolken durch den modifizierten ACRES-Prozessor zufriedenstellend. Dasselbe trifft auf den Algorithmus der Ermittlung der Datennutzbarkeit zu. Dieser wurde in seinen Verfahrensschritten ebenfalls nicht modifiziert, sondern lediglich auf Codeebene optimiert. Dies umfasst beispielsweise die Nutzung optimierter Routinen für die Ableitung der Distanzmatrix (siehe Abschnitt 4.4.9). Somit decken sich die Ergebnisse der Datennutzbarkeit mit denen des MACCA-Prozessors. Eine Einschätzung beziehungsweise Analyse der durch die Automaten ermittelten Ergebnisse wird in BORG (2007, S. 75ff.) vorgenommen.

## 5. Ergebnisse und Ausblick

Der GMACCA-Prozessor wurde soweit implementiert, dass eine automatisierte Bewertung der Datennutzbarkeit von Quicklooks des Satelliten Landsat 7 möglich ist. Somit wurden bereits viele der bereits in Abschnitt 2.6 aufgezählten Optimierungsansätze realisiert. Es wurde ein Prozessor entwickelt, dessen modularer Aufbau die Verarbeitung verschiedener optisch-thermalen Satellitenfernerkundungsprodukten erlaubt. Die Integration neuer Module kann konzeptionell auf einfache Art und Weise erreicht werden. Durch die Fähigkeit JobOrderFiles zu verarbeiten wurde außerdem ein wichtiger Grundstein für die Integration des Prozessors in das Data and Information Management System (DIMS) des DLR gelegt. Im Gegensatz zum MACCA-Prozessor lässt sich die Neuentwicklung konfigurieren, ohne dass dafür Quelltextanpassungen notwendig sind. Dies fördert sowohl die Robustheit als auch die Wartbarkeit des Prozessors. Darüber hinaus ist der neu entwickelte GMACCA-Prozessor performanter in seiner Ausführung als der MACCA-Prozessor.

Die Entwicklung des Prozessors ist noch nicht abgeschlossen. Nachdem die Funktionalität des Prozessors anhand der Prozessierung von ersten Fernerkundungsprodukten sichergestellt wurde, ist weitere Entwicklungsarbeit notwendig. Neben der Implementierung der anderen Wolkendetektionsmodule umfasst die weitere Entwicklungsarbeit die Einbindung neuer Produkttypen. Mit der voranschreitenden Modulentwicklung ist auch die Laufendhaltung der Entwicklerdokumentation verbunden.

In Bezug auf die Verbesserung der Ergebnisse der Wolkendetektion besteht die Möglichkeit neben der Nutzung der CIA World Map und der SWBD-Daten andere Datenbestände zu verwenden, um die Qualität der Land-Wasser-Maske zu steigern. Dabei kann es sich beispielsweise um *Global Self-consistent, Hierarchical, High-resolution Shoreline (GSHHS)*-Daten<sup>2</sup> handeln. Vor allem in Bereichen in denen keine SWBD-Daten vorliegen sind Daten höherer räumlicher Auflösung von Vorteil.

Konzeptionell verfügt der Prozessor über eine Wissensbasis. In dieser sind unter anderem wichtige sensorspezifische Kalibrierungskoeffizienten enthalten. Die Implementierung der Wissensbasis erfordert jedoch weiteren Entwicklungsaufwand. Zum einen muss eine Spezifizierung der Art der Datenspeicherung erarbeitet werden. So können diese Daten beispielsweise in einer XML-Datenstruktur verwaltet werden. Bei steigendem Datenvolumen wäre auch der Einsatz eines Datenbanksystems denkbar. Diese Entscheidungen

---

<sup>2</sup><http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>

bedürfen jedoch weiterführenden Analysen hinsichtlich des zu verwaltenden Informationsbedarfes.

Der GMACCA-Prozessor benötigt schon in diesem Stadium der Entwicklung wesentlich weniger Verarbeitungszeit als der Vorgängerprozessor. Diese Effizienz lässt sich unter Umständen weiter steigern. Dies kann zum einen durch die Parallelisierung verschiedener Teilkomponenten und zum anderen durch die Nutzung der Grafikkarte<sup>3</sup> erreicht werden.

---

<sup>3</sup>[http://www.nvidia.de/object/cuda\\_home\\_new\\_de.html](http://www.nvidia.de/object/cuda_home_new_de.html)

---

# Anhang

## A. Ergebnisse des Testdurchganges

Die folgenden Seiten enthalten die Ergebnisse der Prozessierung der Testdatensätze durch den GMACCA-Prozessor. Jede Abbildung enthält zwei visuelle Darstellungen der Ergebnisse der Wolkendetektion. Dabei handelt es sich im Speziellen um RGB-Bilder in denen die detektierten Wolkenpixel auf verschiedene Weise kenntlich gemacht wurden. In der oberen Abbildung werden detektierte Wolken lediglich umrandet. Dadurch lässt sich ein Großteil der Originaldaten in Kombination mit den Ergebnissen der Wolkendetektion darstellen. In der unteren Abbildung werden Wolkengebiete zusätzlich mit gelber Farbe gefüllt. Durch die Hervorhebung wird zum einen die Erkennbarkeit kleiner Wolkengruppen verbessert, zum anderen lassen sich Fehler in der Wolkendetektion – vor allem *Löcher* – besser darstellen.

Unter den beiden Abbildungen finden sich Informationen zur Bezeichnung der Szene. Repräsentativ wird die Bezeichnung des ersten Bandes dargestellt. Darüber hinaus werden die Ergebnisse der Datennutzbarkeitsanalyse sowohl der Operateure als auch des GMACCA-Prozessors dargestellt. Entsprechend Abschnitt 4.4.10 erhält ein Quadrant die Bewertung 0, wenn keine Wolkenbedeckung detektiert wurde. Eine Bewertung mit 90 entspricht einer fehlerhaften oder vollständig durch Wolken bedeckten Szene. An erster Stelle steht die mittlere Bewertung aller Quadranten, stellvertretend für die Datenqualität der gesamten Szene. Nachfolgend erscheinen die Bewertungen des oberen, rechten, unteren und linken Quadranten (beginnend oben links). Abschließend wird die Verarbeitungszeit der Szene in Sekunden angegeben. Abbildung 19 auf Seite 47 und Abbildung 34 auf Seite 62 wurden nur zur Darstellung generiert. Im normalen Betrieb werden fehlerhafte Szenen lediglich mit 90 bewertet und nicht weiter prozessiert.

Im Zuge der Entwicklung des GMACCA-Prozessors wird eine Entwicklerdokumentation erstellt. Dazu wurde – wie bereits beschrieben – das Tool NATURALDOCS verwendet. Die Dokumentation folgt einem einheitlichen Aufbau und ist in verschiedene Rubriken unterteilt. Dazu zählen Informationen zum Ziel des Projektes, zu den Entwicklern und zum aktuellen Projektstatus. Der wohl wichtigste Teil der Dokumentation ist der Teil der Modulbeschreibungen. Hierzu werden die Module hierarchisch gegliedert erläutert. Die Erläuterungen umfassen neben einer Beschreibung der Funktionsweise auch die Übergabeparameter und Rückgabewerte der Module. Zum besseren Verständnis wer-

den die Programmabläufe der Module grafisch dargestellt. Um die Navigation zwischen zusammenhängenden Modulen zu vereinfachen, wurden zusammenhängende Modulbeschreibungen untereinander verlinkt. Die Abbildung 35 auf Seite 63 stellt den Aufbau einer Modulbeschreibung grafisch dar.

Track	Frame	Zeit [s]	Abbildung	Anmerkung
182	22	0.375	Abb. 19 auf S. 47	fehlerhafte Bilddaten
182	23	10.134	Abb. 20 auf S. 48	hoher Bewölkungsgrad
182	24	10.423	Abb. 21 auf S. 49	hoher Bewölkungsgrad
182	25	10.705	Abb. 22 auf S. 50	hoher Bewölkungsgrad
182	26	11.371	Abb. 23 auf S. 51	hoher Bewölkungsgrad
182	27	10.666	Abb. 24 auf S. 52	hoher Bewölkungsgrad
182	28	11.706	Abb. 25 auf S. 53	hoher Bewölkungsgrad
182	29	10.245	Abb. 26 auf S. 54	Gain-Change in Szene
182	30	9.852	Abb. 27 auf S. 55	Detektionsfehler
182	31	9.643	Abb. 28 auf S. 56	Detektionsfehler
182	32	8.257	Abb. 29 auf S. 57	geringer Bewölkungsgrad
182	33	7.677	Abb. 30 auf S. 58	geringer Bewölkungsgrad
182	34	8.006	Abb. 31 auf S. 59	geringer Bewölkungsgrad
182	35	8.148	Abb. 32 auf S. 60	geringer Bewölkungsgrad
182	36	8.760	Abb. 33 auf S. 61	geringer Bewölkungsgrad
182	37	0.249	Abb. 34 auf S. 62	fehlerhafte Bilddaten

Tabelle 2: Zusammenfassung der Verarbeitungszeiten der Testdatensätze.

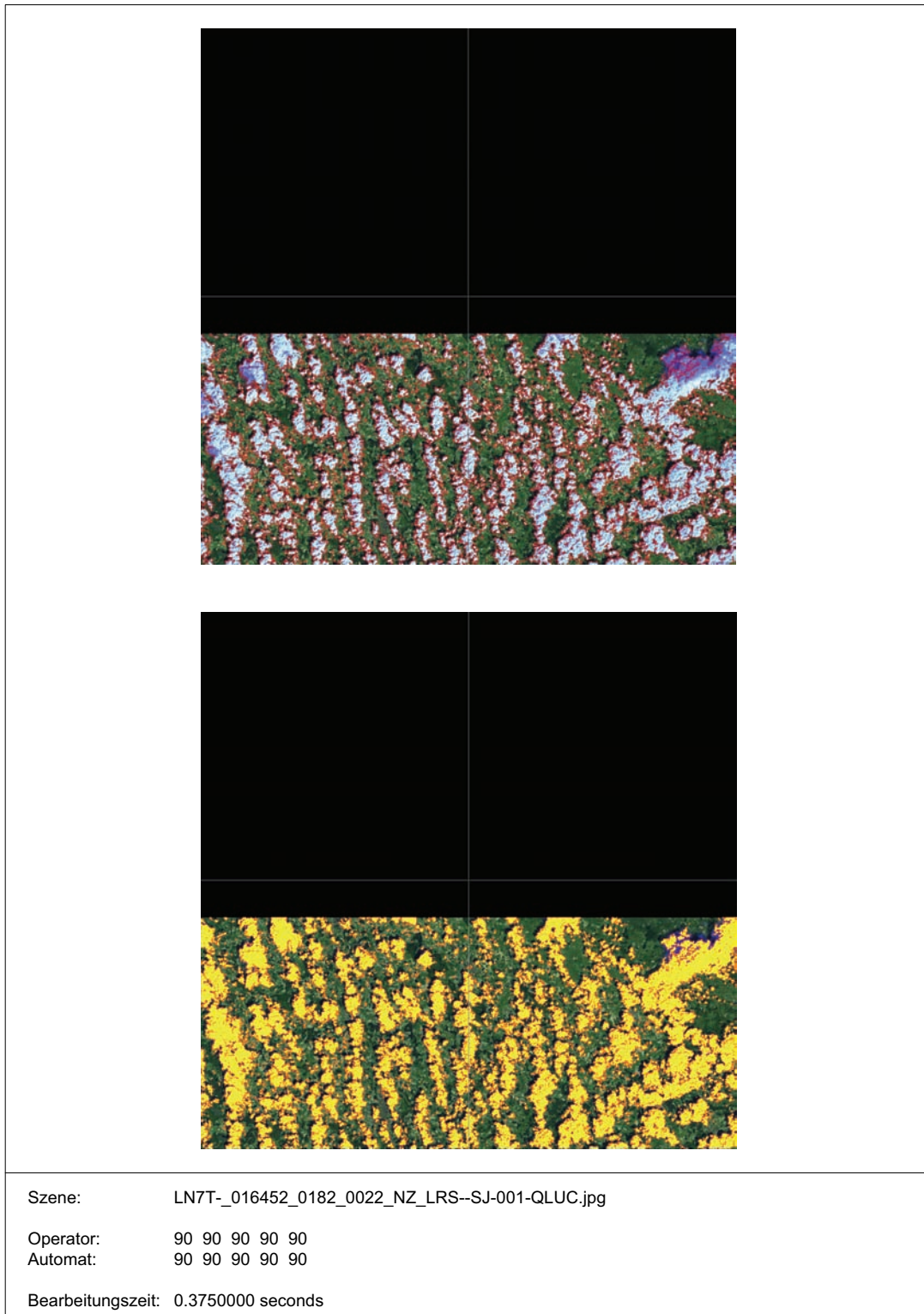


Abbildung 19: Ergebnisse der Datennutzbarkeitsbewertung der Szene 22 im Track 182 durch den GMACCA-Prozessor.

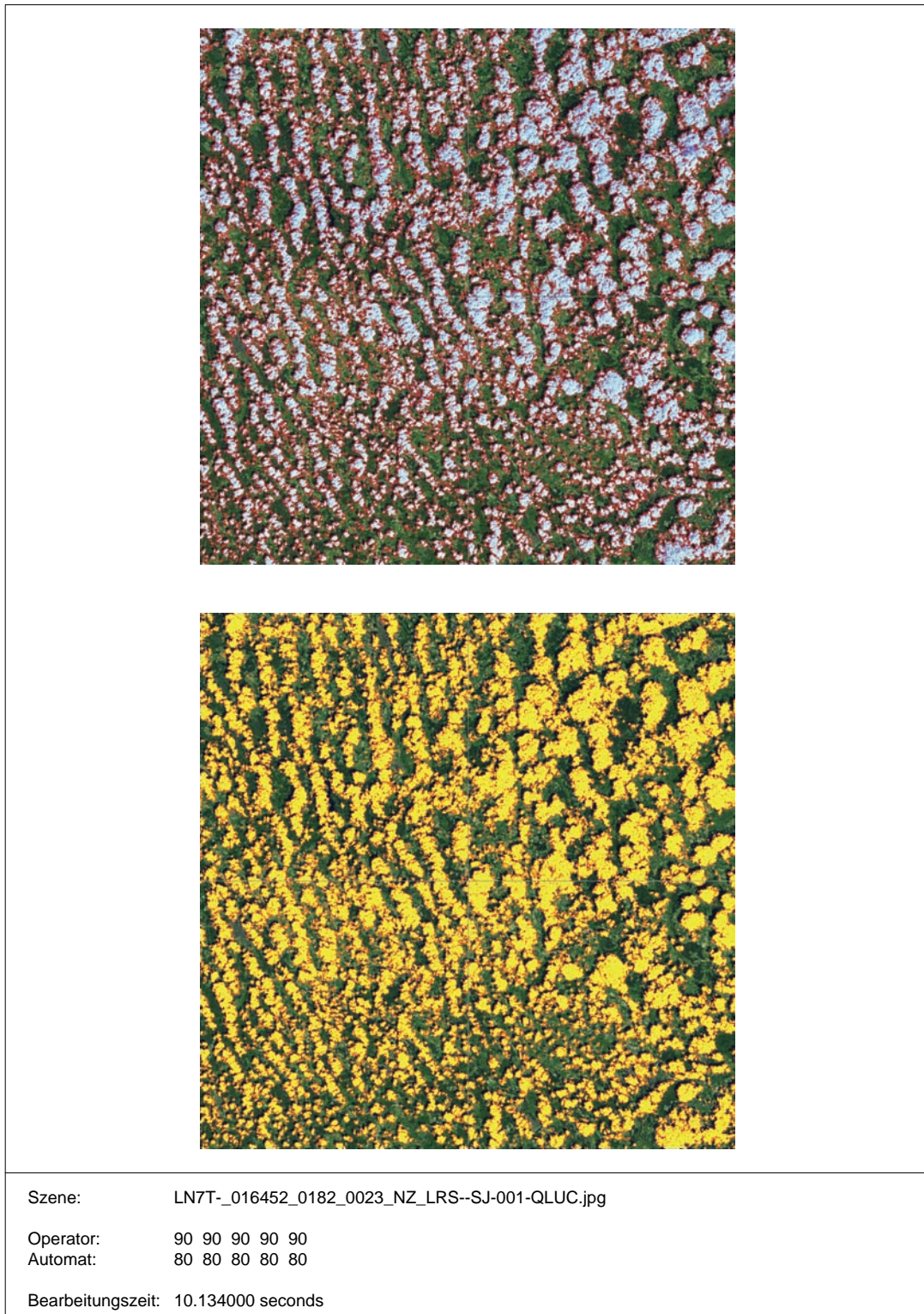


Abbildung 20: Ergebnisse der Datennutzbarkeitsbewertung der Szene 23 im Track 182 durch den GMACCA-Prozessor.



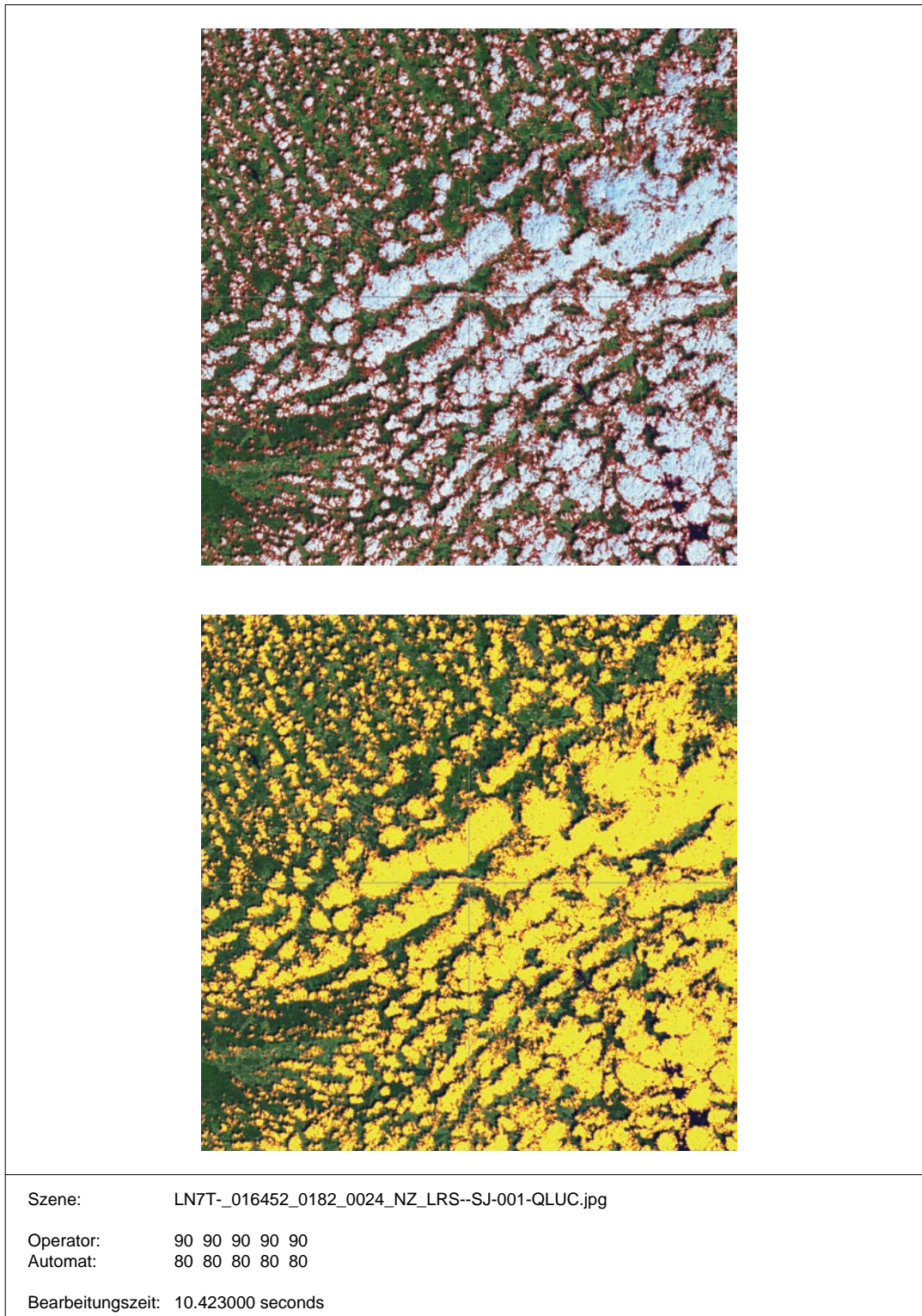


Abbildung 21: Ergebnisse der Datennutzbarkeitsbewertung der Szene 24 im Track 182 durch den GMACCA-Prozessor.

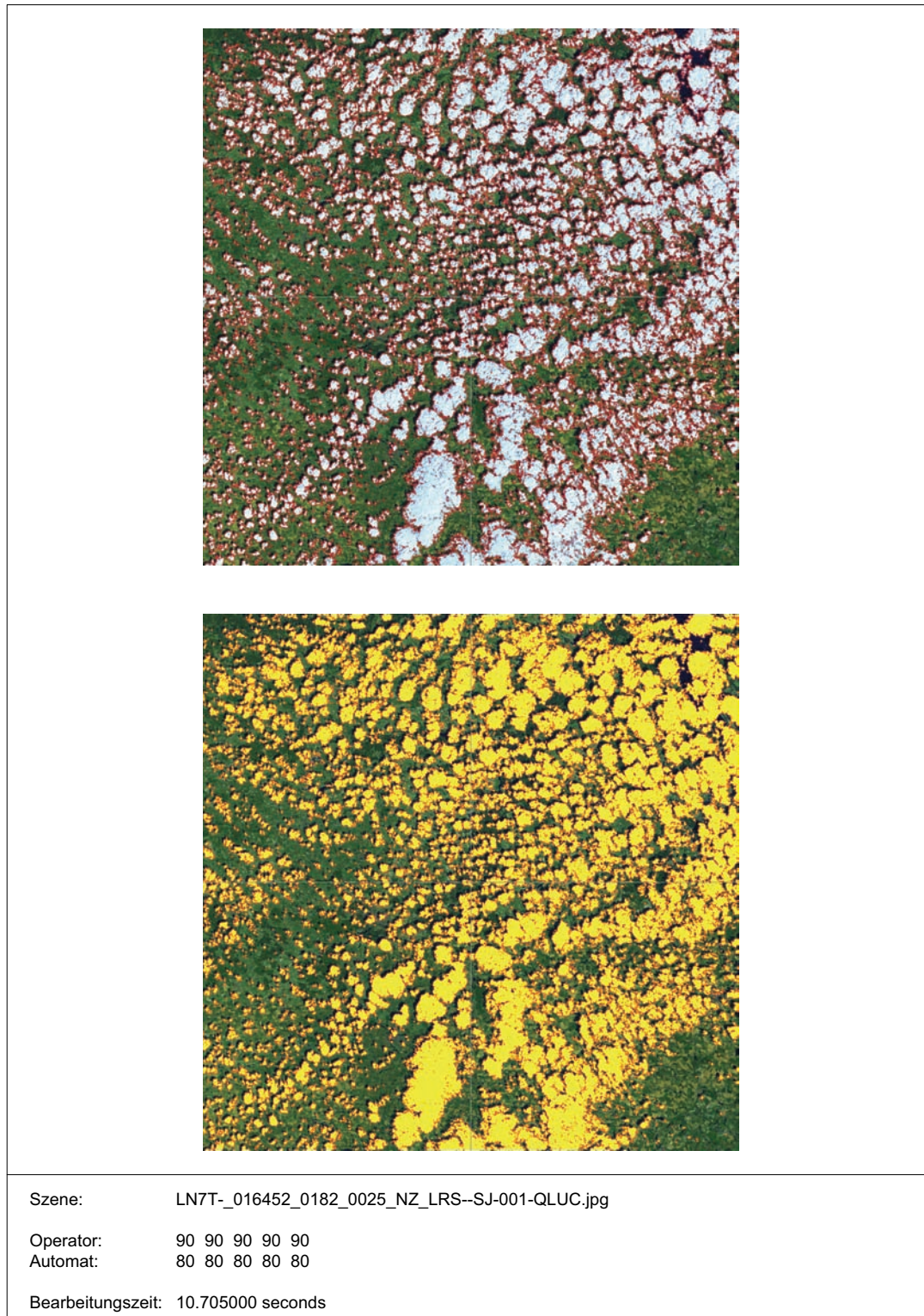


Abbildung 22: Ergebnisse der Datennutzbarkeitsbewertung der Szene 25 im Track 182 durch den GMACCA-Prozessor.

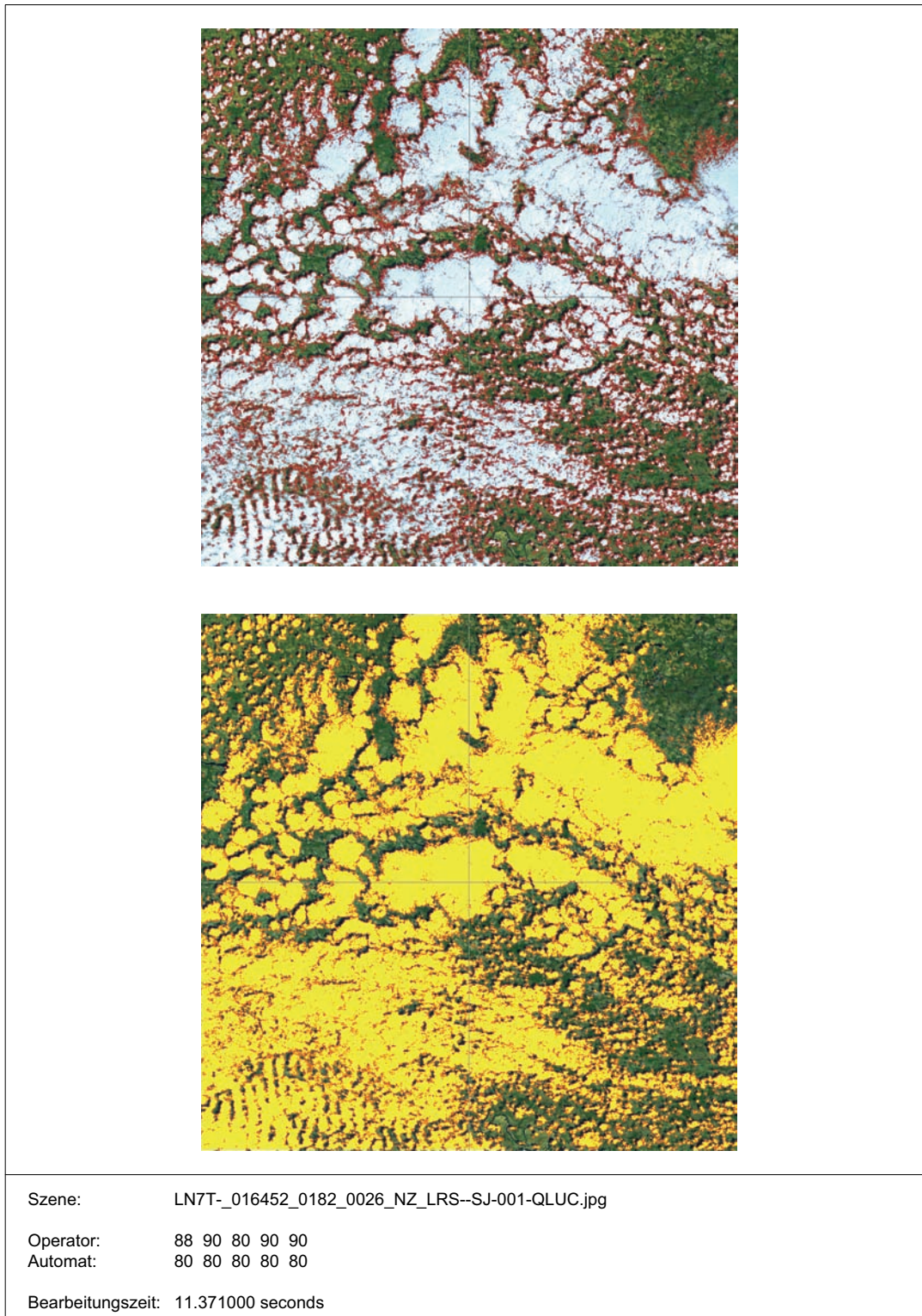


Abbildung 23: Ergebnisse der Datennutzbarkeitsbewertung der Szene 26 im Track 182 durch den GMACCA-Prozessor.

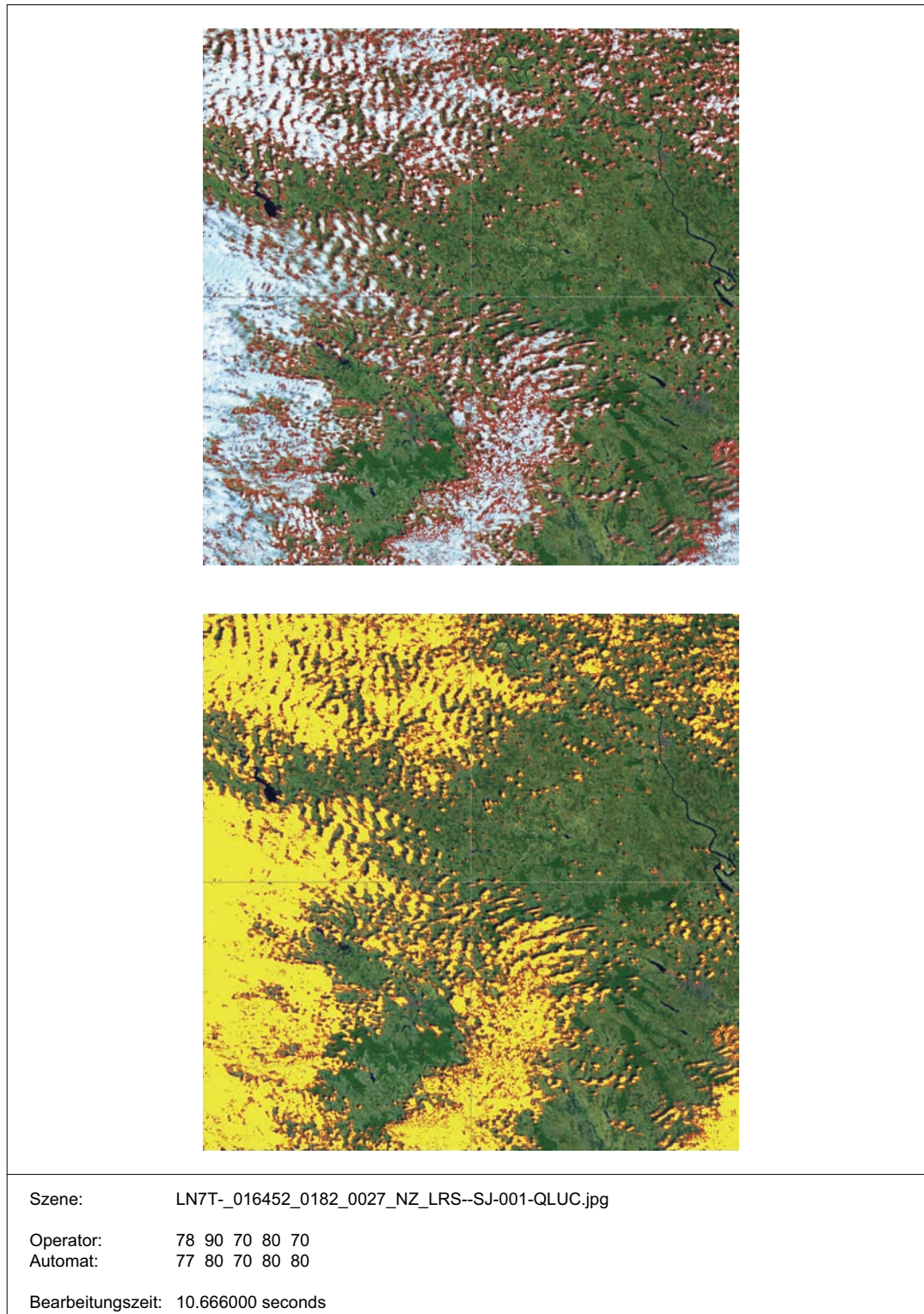


Abbildung 24: Ergebnisse der Datennutzbarkeitsbewertung der Szene 27 im Track 182 durch den GMACCA-Prozessor.

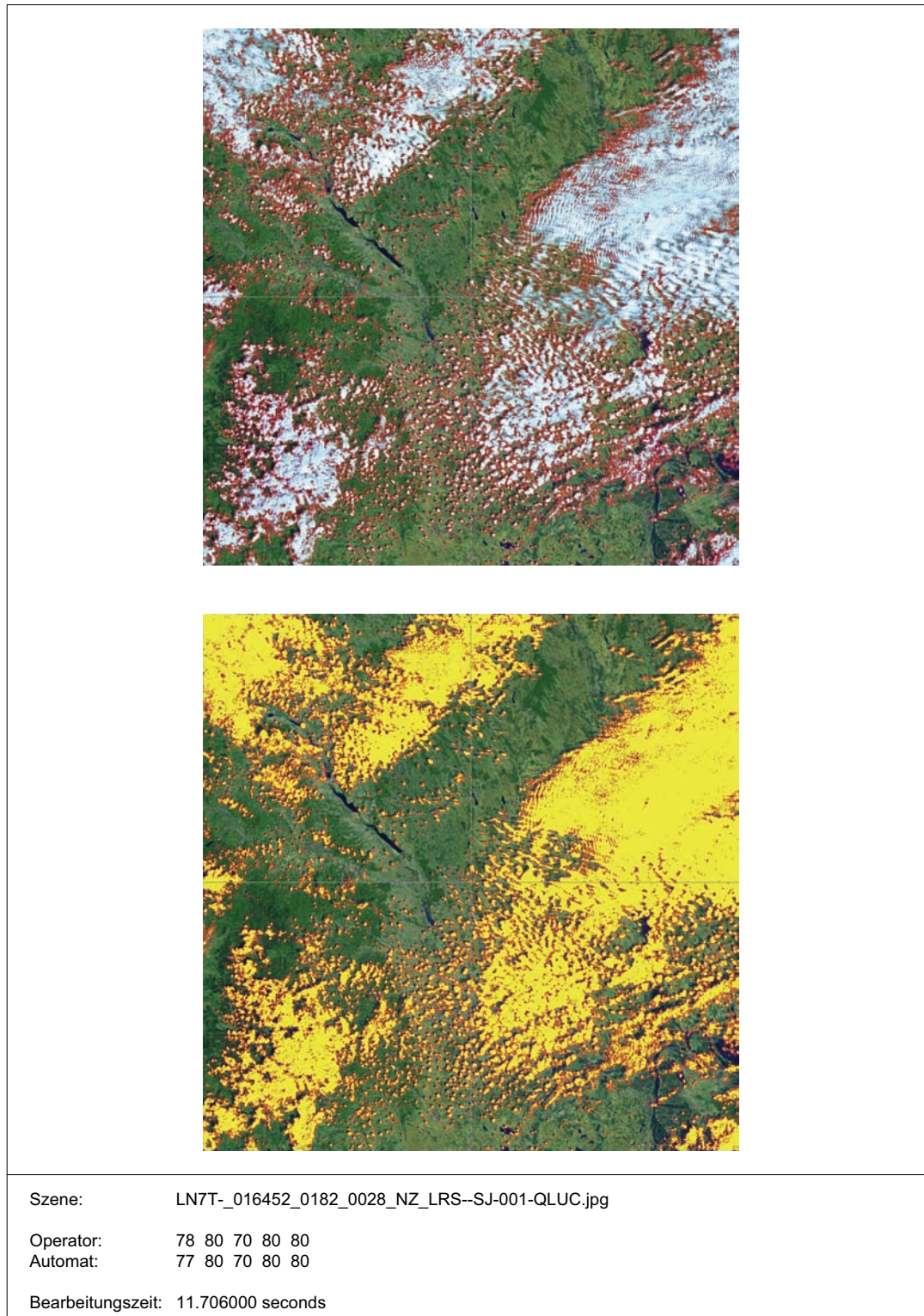


Abbildung 25: Ergebnisse der Datennutzbarkeitsbewertung der Szene 28 im Track 182 durch den GMACCA-Prozessor.

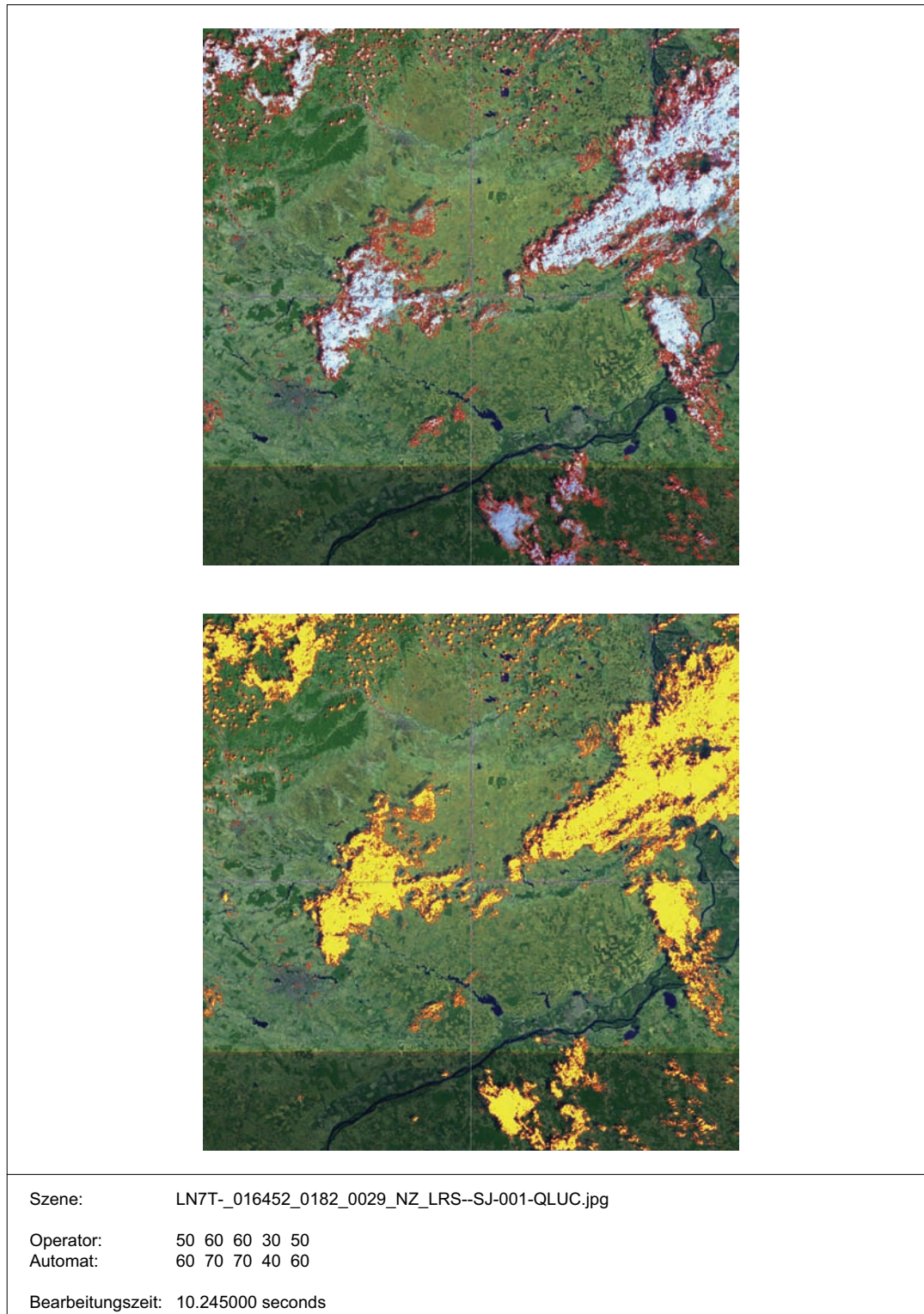


Abbildung 26: Ergebnisse der Datennutzbarkeitsbewertung der Szene 29 im Track 182 durch den GMACCA-Prozessor.

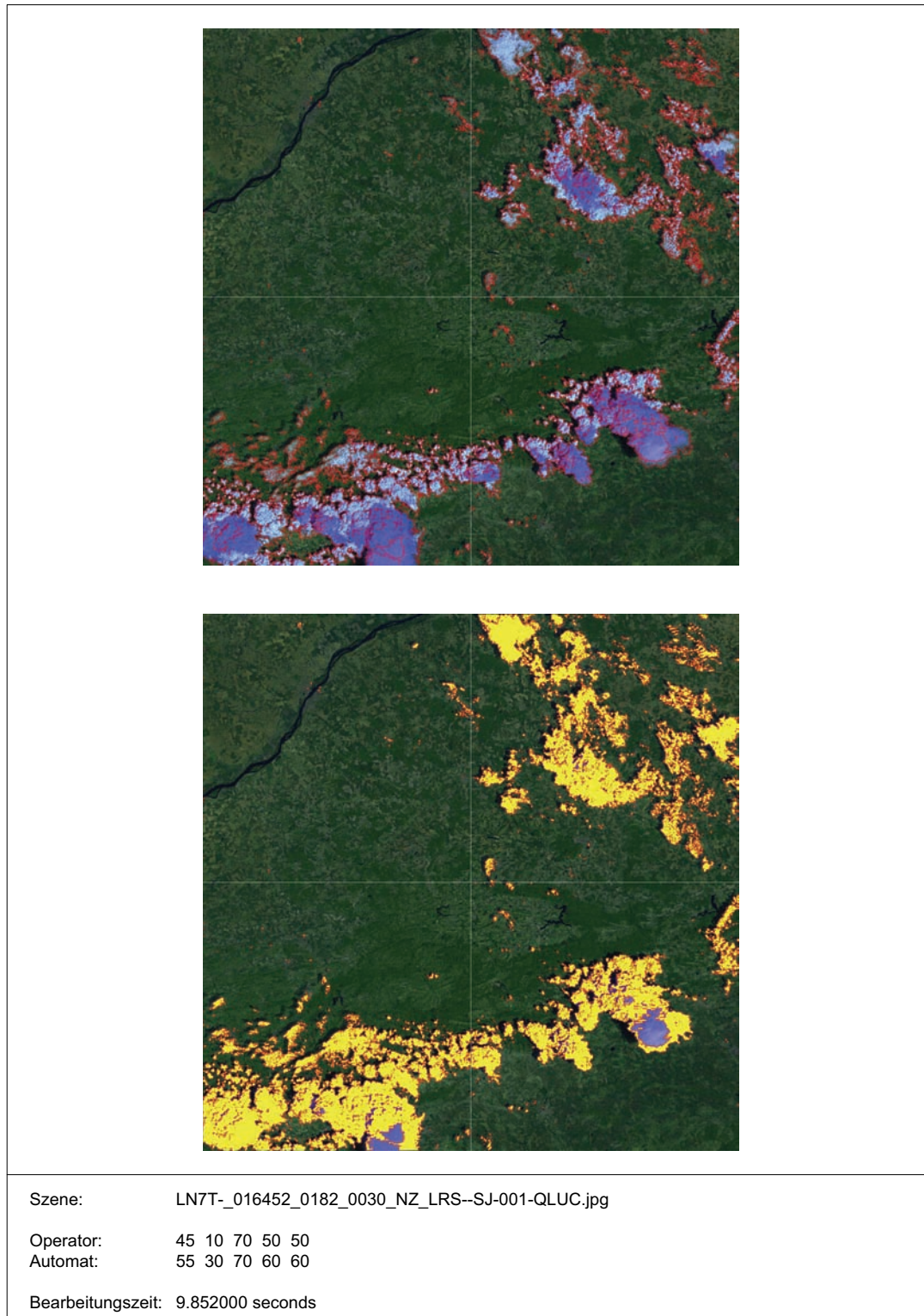


Abbildung 27: Ergebnisse der Datennutzbarkeitsbewertung der Szene 30 im Track 182 durch den GMACCA-Prozessor.

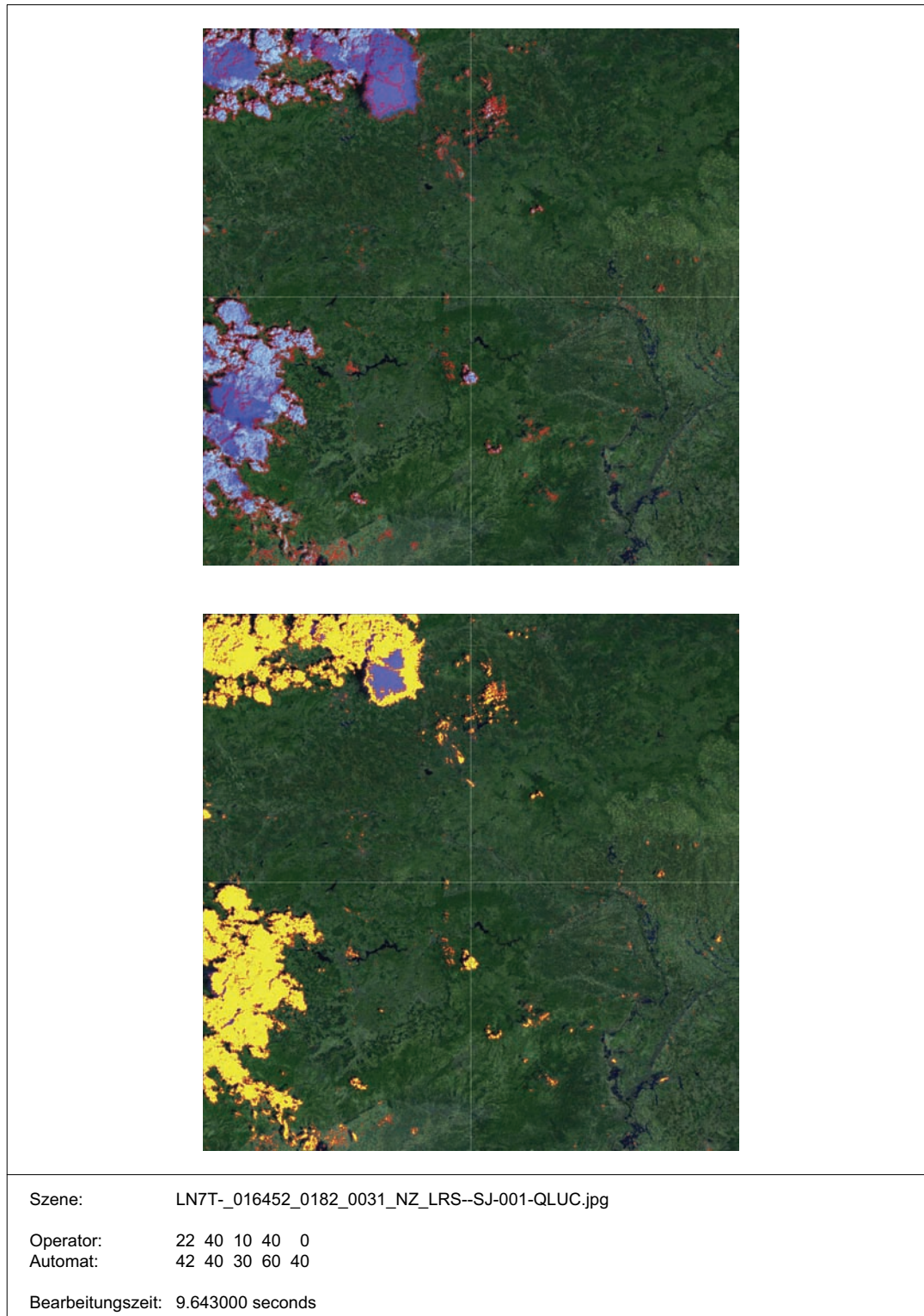


Abbildung 28: Ergebnisse der Datennutzbarkeitsbewertung der Szene 31 im Track 182 durch den GMACCA-Prozessor.



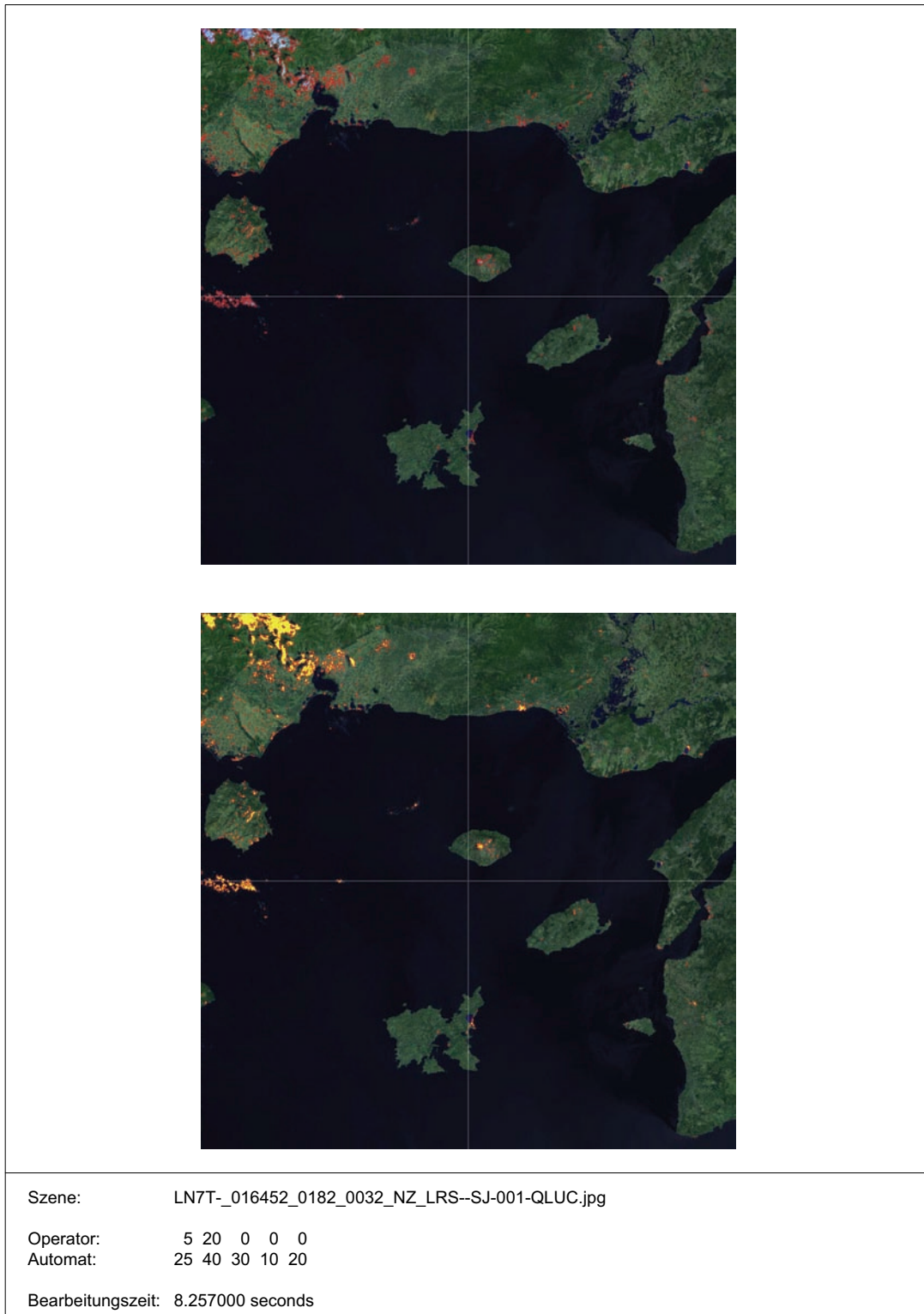


Abbildung 29: Ergebnisse der Datennutzbarkeitsbewertung der Szene 32 im Track 182 durch den GMACCA-Prozessor.

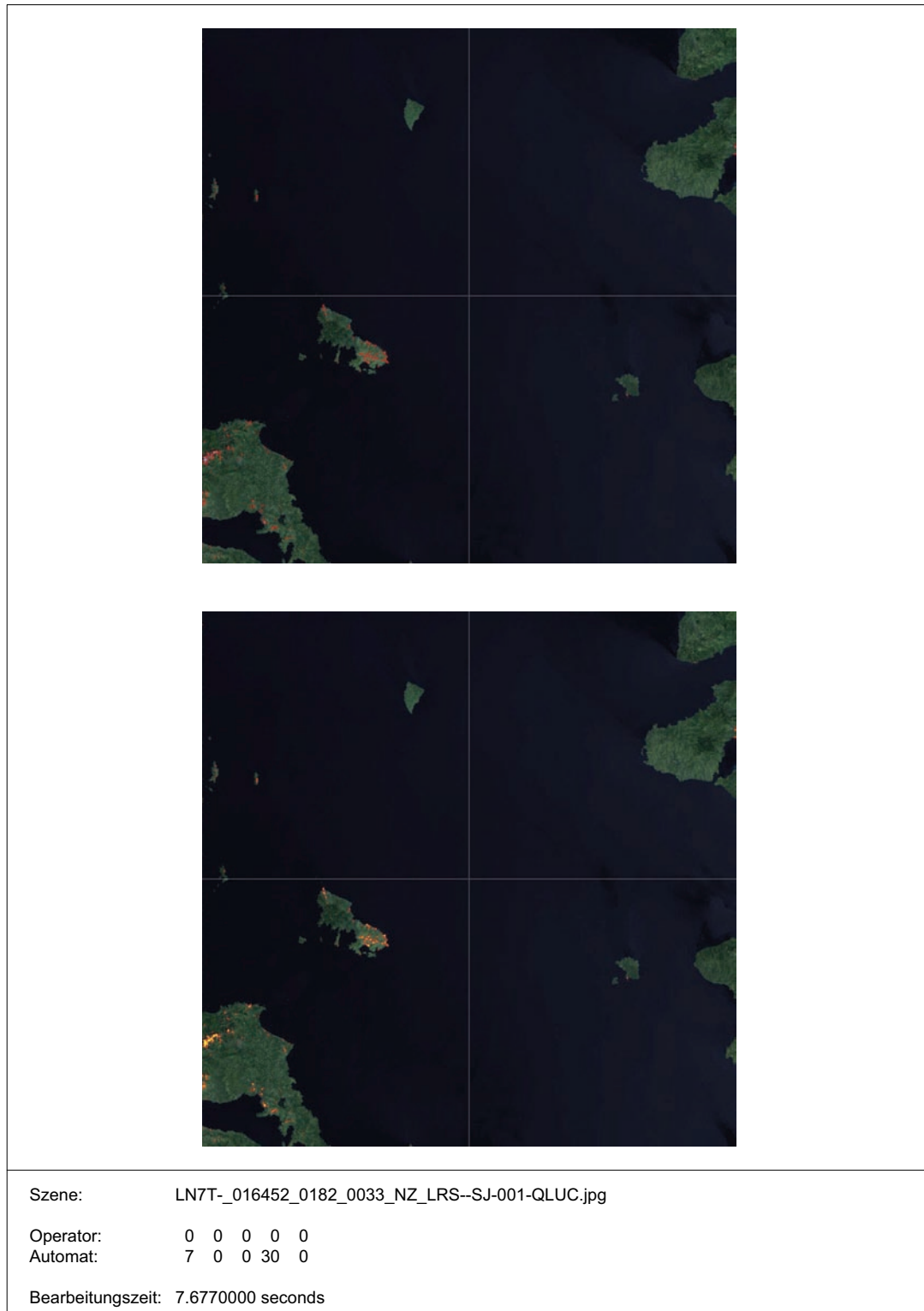


Abbildung 30: Ergebnisse der Datennutzbarkeitsbewertung der Szene 33 im Track 182 durch den GMACCA-Prozessor.

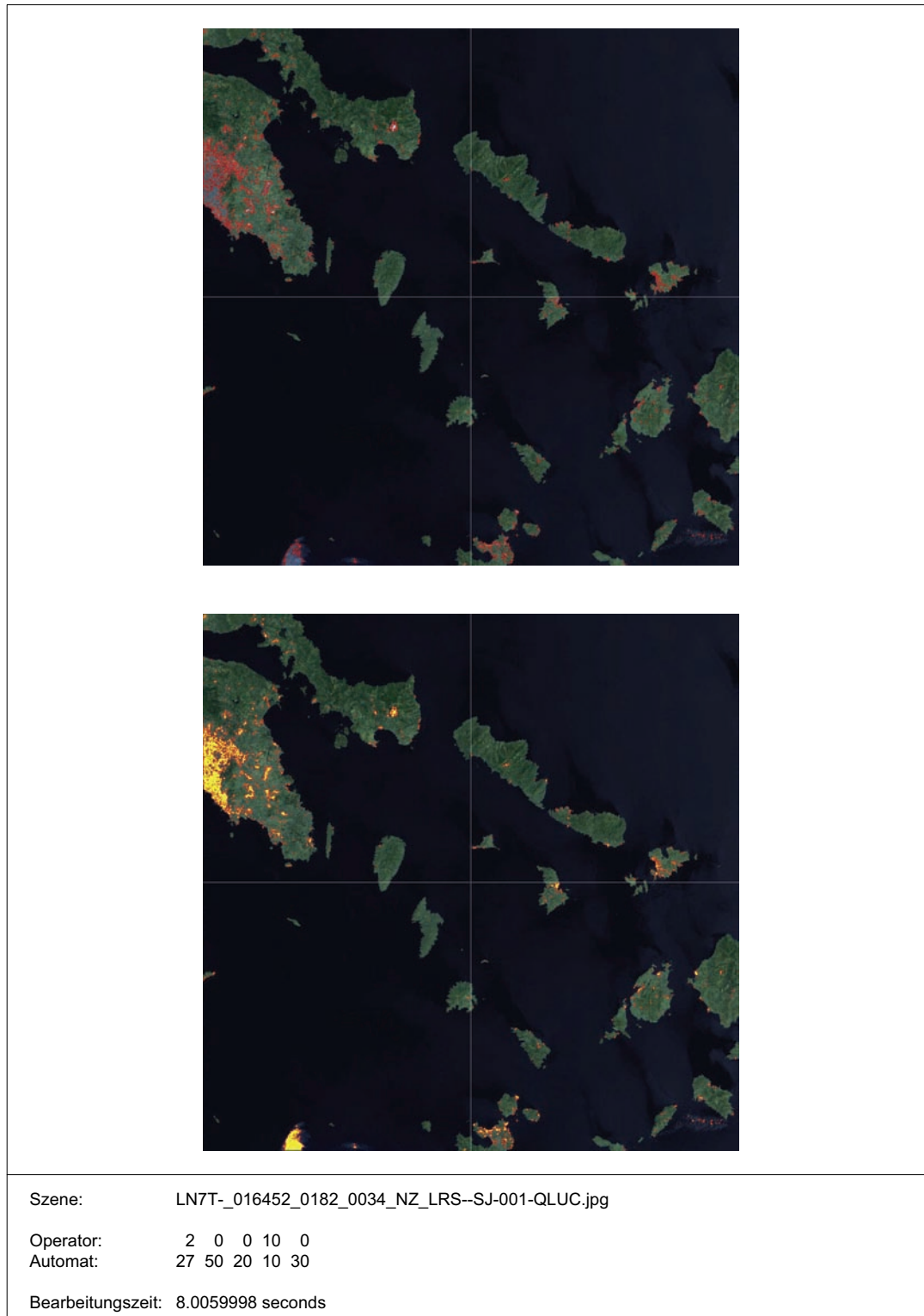


Abbildung 31: Ergebnisse der Datennutzbarkeitsbewertung der Szene 34 im Track 182 durch den GMACCA-Prozessor.

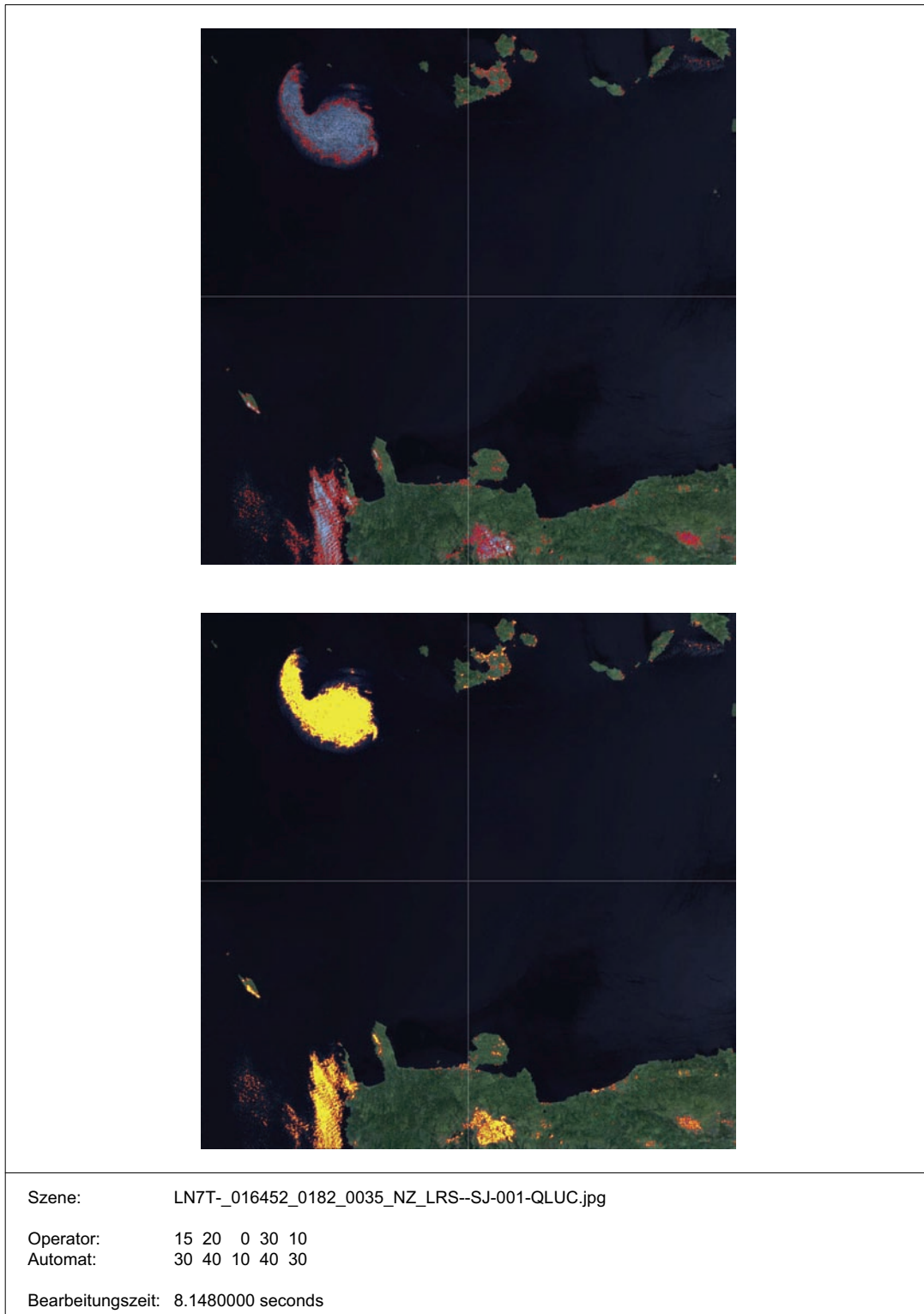


Abbildung 32: Ergebnisse der Datennutzbarkeitsbewertung der Szene 35 im Track 182 durch den GMACCA-Prozessor.

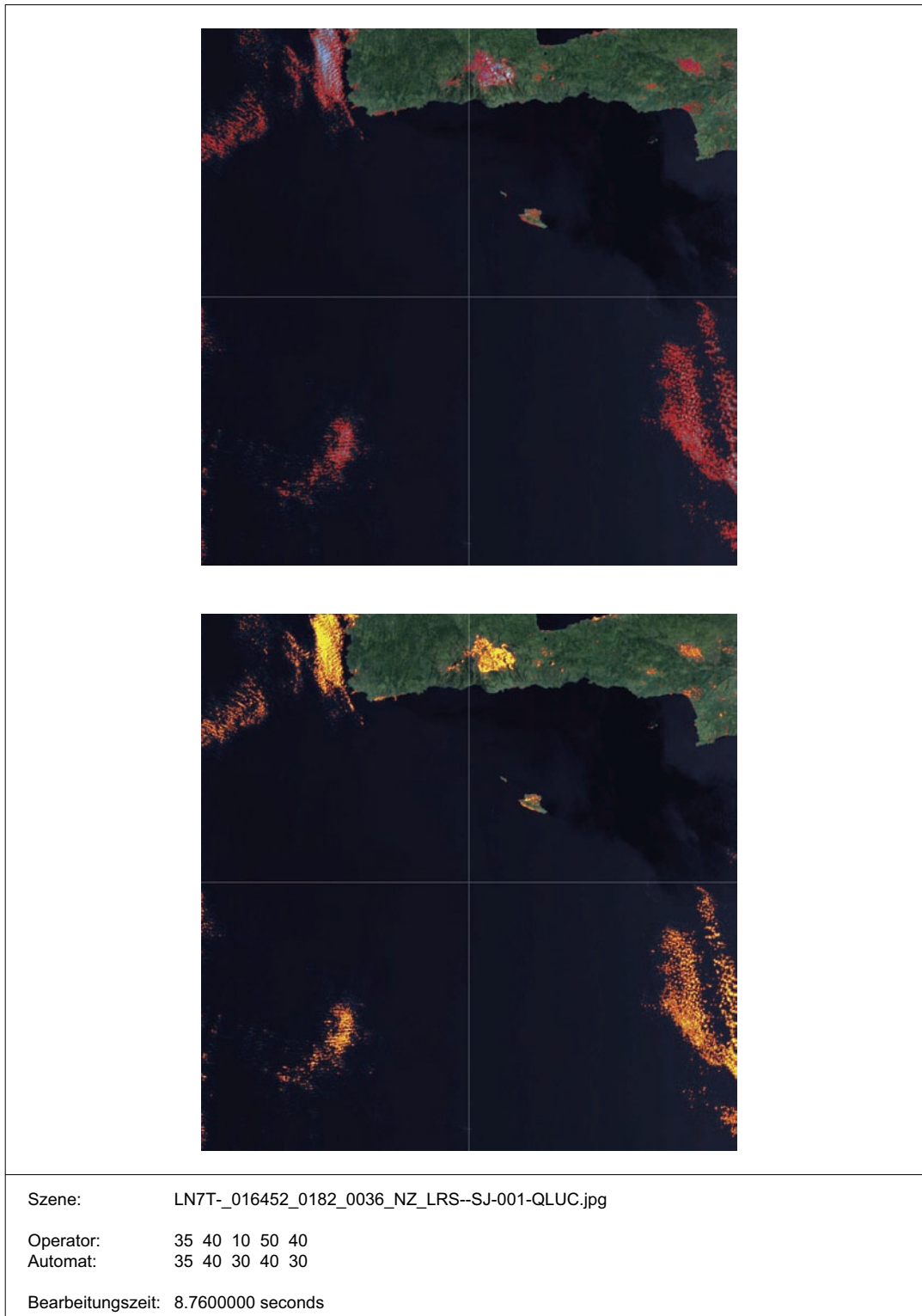


Abbildung 33: Ergebnisse der Datennutzbarkeitsbewertung der Szene 36 im Track 182 durch den GMACCA-Prozessor.

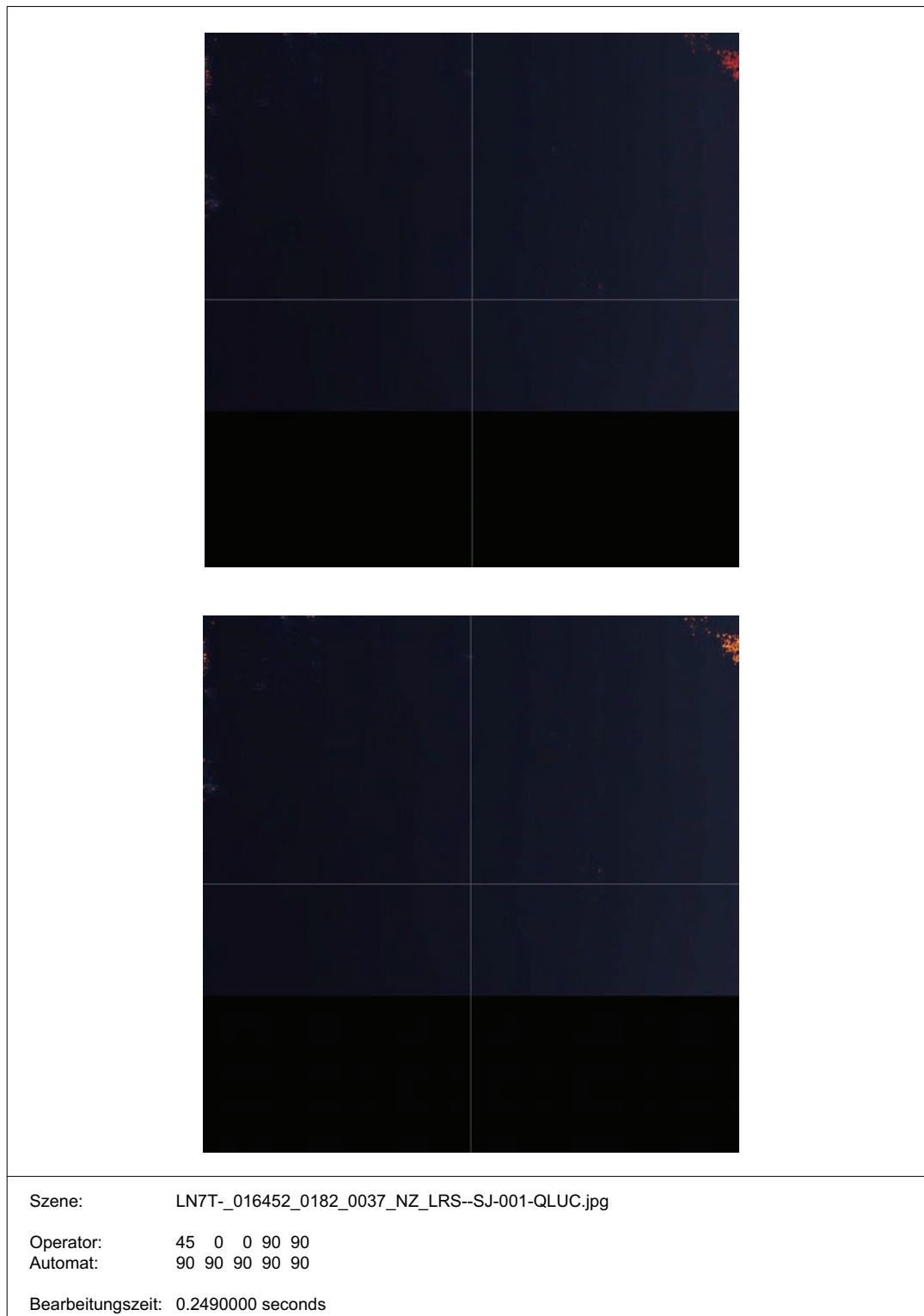


Abbildung 34: Ergebnisse der Datennutzbarkheitsbewertung der Szene 37 im Track 182 durch den GMACCA-Prozessor.

## MetaDataReader

FUNCTION MetaDataReader, PTR\_JobOrderStruct, PTR\_SceneList, SceneIndex

This function reads the metadata of all supported satellite products and stores them into a user defined structure.

Depending on the marking of the used satellite, sensor and product type different subroutines get called. This is because the metadata format changes according to the remote sensing system.

### Parameter

PTR\_JobOrderStruct pointer to structure containing information from JobOrderFile  
 PTR\_SceneList pointer to SceneList  
 Scene\_Index index of current scene

### Returns

PTR\_Metadata pointer to user defined structure with information read into

### The following information must be read into

Metadata.Longitude array with information about longitude per scene [UL, UR, LR, LL, CENTER]  
 Metadata.Latitude array with information about latitude per scene [UL, UR, LR, LL, CENTER]  
 Metadata.DateInformation userdefined structure with information about date per scene [Acquisition\_Date, Day, Month, Year]

### The following information is optional, depending on remote sensing system an parameters in JobOrderFile

Metadata.OperatorsAssessment array with votes of operators according cloud cover assessment [mean, UL, UR, LL, LR]  
 Metadata.GainInformation user defined structure with information about gain and gain change

### Following Products are currently readable

- LANDSAT 7, ETM, QUICKLOOK products (MetaDataReader\_L7\_ETM\_QL)

### Program Flow Chart

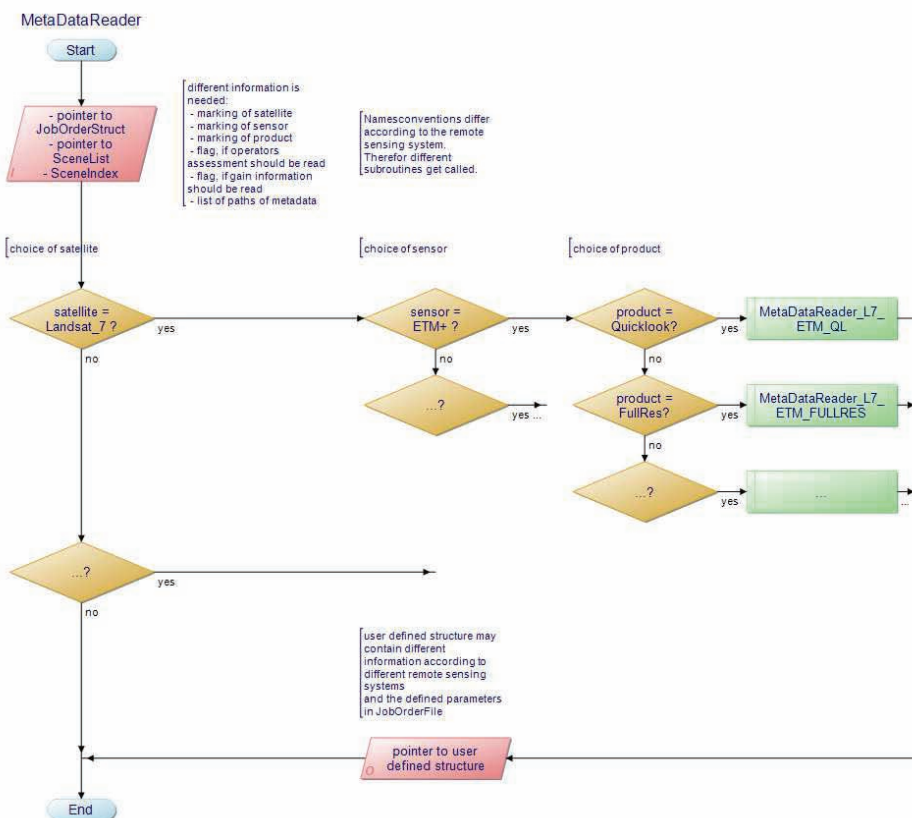


Abbildung 35: Aufbau der Modulbeschreibung in der Entwicklerdokumentation.

## Literatur

- ALBERTZ, JÖRG (2001). *Einführung in die Fernerkundung*. Wissenschaftliche Buchgesellschaft, 2. überarbeitete und erweiterte Aufl.
- ALBERTZ, JÖRG und M. WIGGENHAGEN (2009). *Taschenbuch zur Fernerkundung und Photogrammetrie*. Wichmann, 5. Aufl.
- BALDENHOFER, KURT (2010). *Lexikon der Fernerkundung*. Internet. <http://www.fe-lexikon.info/lexikon-s.htm>; letzter Zugriff: 15.07.2010 15:13 Uhr.
- BENNETT, JEFFREY, M. DONAHUE, N. SCHNEIDER und M. VOIT (2009). *Astronomie - Die kosmische Perspektive*. Pearson Education.
- BORG, ERIK (2007). *Entwicklung und Anwendung eines automatischen Prozessors zur Erfassung der Wolkenbedeckung und Datennutzbarkeit am Beispiel von LANDSAT 7 / ETM+-Daten*. Doktorarbeit, Mathematisch-Naturwissenschaftliche Fakultät der Universität Potsdam.
- BRANDT-POOK, HANS und R. KOLLMEIER (2008). *Softwareentwicklung kompakt und verständlich*. Vieweg + Teubner | GWV Fachverlage GmbH, Wiesbaden 2008, 1. Aufl.
- DECH, STEFAN (2010). *Deutsches Fernerkundungsdatenzentrum (DFD)*. Internet. [http://www.dlr.de/caf/desktopdefault.aspx/tabid-5278/8856\\_read-15911/](http://www.dlr.de/caf/desktopdefault.aspx/tabid-5278/8856_read-15911/); letzter Zugriff: 29.08.2010 16:32.
- FICHELTMANN, BERND, E. BORG und A. GÜNTHER (2004). *Automatische Georeferenzierung von Fernerkundungsdaten*. DE 10 2004 045 563 A1 2006.03.30.
- ITT VIS (2009a). *Getting Started with IDL*. IDL Version 7.1 May 2009 Edition.
- ITT VIS (2009b). *Using IDL*. IDL Version 7.1 May 2009 Edition.
- ITT VIS (2010). *Homepage ITT Visual Information Solutions*. Internet. <http://www.ittvis.com/ProductServices/IDL.aspx>; letzter Zugriff: 21.07.2010, 14:08.
- KIEMLE, STEPHAN (2010a). *Datenmanagement und Datenprozessierung (im Cluster für Angewandte Fernerkundung des DLR)*. Internet. [http://www.dlr.de/caf/desktopdefault.aspx/tabid-5335/8945\\_read-22872/](http://www.dlr.de/caf/desktopdefault.aspx/tabid-5335/8945_read-22872/); letzter Zugriff: 13.08.2010 09:30.



- KIEMLE, STEPHAN (2010b). *Digital Data Library (im Cluster für Angewandte Fernerkundung des DLR)*. Internet. [http://www.dlr.de/caf/desktopdefault.aspx/tabid-5351/10424\\_read-22884/](http://www.dlr.de/caf/desktopdefault.aspx/tabid-5351/10424_read-22884/); letzter Zugriff: 13.08.2010 09:30.
- KRAUS, KARL und W. SCHNEIDER (1988). *Fernerkundung: Band 1 Physikalische Grundlagen und Aufnahmetechniken*. Ferd. Dümmlers Verlag Bonn, 1. Aufl.
- KRIEGEL, MARTIN (2010). *Praktikumsbericht für das Praktikum am Deutschen Zentrum für Luft- und Raumfahrt e.V. vom 01.03.2010 bis 30.06.2010*.
- NASA (2008). *Atmospheric electromagnetic opacity*. Internet. [http://commons.wikimedia.org/wiki/File:Atmospheric\\_electromagnetic\\_transmittance\\_or\\_opacity.jpg](http://commons.wikimedia.org/wiki/File:Atmospheric_electromagnetic_transmittance_or_opacity.jpg); letzter Zugriff: 19.07.2010 10:46.
- NASA (2009). *Landsat 7 Science Data Users Handbook*. Internet. [http://landsathandbook.gsfc.nasa.gov/handbook/handbook\\_toc.html](http://landsathandbook.gsfc.nasa.gov/handbook/handbook_toc.html); letzter Zugriff: 20.07.2010 09:58.
- NISCHWITZ, ALFRED, M. FISCHER und P. HEBERÄCKER (2007). *Computergrafik und Bildbearbeitung*. Friedr. Vieweg & Sohn Verlag | GWV Fachverlage GmbH, Wiesbaden 2007, 2. verbesserte und erweiterte Aufl.
- RAMIREZ, ERIC (2009). *Shuttle Radar Topography Mission (SRTM) - The Mission to map the world*. Internet. <http://www2.jpl.nasa.gov/srtm/index.html>; NASA; letzter Zugriff: 26.07.2010 18:09.
- VOGEL, OLIVER, I. ARNOLD, A. CHUGHTAI, E. IHLER, T. KEHRER, U. MEHLIG und U. ZDUN (2009). *Software-Architektur: Grundlagen - Konzepte - Praxis*. Spektrum Akademischer Verlag Heidelberg 2009, 2. Aufl.
- WERUM SOFTWARE & SYSTEMS AG (2009). *DIMS Processing System Management - Integrator's Manual and ICD*. Werum Software & Systems AG. Release 2.0; Reg. No. 32.5.2; Date: 30.09.2009.

## **Eidesstattliche Erklärung**

Hiermit versichere ich, die vorliegende Bachelorarbeit ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus den Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Neubrandenburg, 30. August 2010

Martin Kriegel