



Hochschule
München
University of
Applied Sciences

Bachelorarbeit

Räumlich hyperkomplexe Basen zur Steigerung der
Klassifikationsgenauigkeit in TerraSAR-X-Aufnahmen über Lac
Bam

angefertigt von

Richard Wolff

vorgelegt bei

Prof. Dr.-Ing. Andreas Schmitt

München, den 21.12.2022

Danksagung

Mit dieser Seite möchte ich mich bei allen Personen herzlich bedanken, die mich bei der Erstellung dieser Bachelorarbeit unterstützt haben.

An erster Stelle gebührt mein Dank Herrn Prof. Dr.-Ing. Andreas Schmitt, der meine Bachelorarbeit betreut und begutachtet hat. Auch die engagierte Betreuung, die hilfreichen Anregungen und die konstruktive Kritik haben zu dem Gelingen dieser Bachelorarbeit beigetragen.

Ebenfalls möchte ich dem Deutschen Zentrum für Luft- und Raumfahrt für die Bereitstellung der Daten danken. Diese wurden im Rahmen des Proposals "*Wetland monitoring and water stress in sub-Saharan West Africa*" (LAN2000) zur Verfügung gestellt

Auch möchte ich meinen Kommilitonen aus dem Jour Fixe bedanken, die mir mit einem anderen Blickwinkel neue Ideen bezüglich meiner Arbeit gegeben haben.

Zu guter Letzt danke ich noch meinen Freunden und meiner Familie, die mich während der gesamten Zeit motiviert und unterstützt haben.

Kurzfassung

Feuchtgebiete sind im halbtrockenen Afrika lebenswichtige Wasserressourcen für die lokale Bevölkerung und für die Artenvielfalt. Der sich im Norden von Burkina Faso befindende Lac Bam See ist der größte natürliche Süßwassersee in dieser Region. Die Feuchtgebiete weisen verschiedene Merkmale auf. Das sehr trübe und sedimentreiche Wasser ist von unterfluteter und schwimmender Vegetation bedeckt.

Trocken- und Regenzeit tragen neben dem sehr unberechenbaren Wetter zu saisonalen Flächenveränderungen in den Feuchtgebieten bei. Während der Trockenzeit ist der See neben einzelnen Brunnen der einzige Süßwasserspeicher in der Bam Region. Zusätzlich werden immer mehr Wasserpumpen für die Bewässerung von Felder errichtet. Diese bedeuten einen weiteren Eingriff in das Habitat des Sees.

Die Aufnahme des Lac Bam 2009 in die *List of Wetlands of International Importance*, der Ramsar-Konvention, trägt dazu bei, dass das Hauptaugenmerk der Fernerkundung seit 2013 auf dieses Gebiet gelenkt wird.

Die Studie verwendet dual-co-polarized HH-VV StripMap Daten des Terra-SAR-X Satelliten. Die Radarszenen wurden am DLR mit dem MultiSAR-System vorprozessiert. Durch die Vorprozessierung erhielt man Kennaugh-Elemente, die mit räumlichen-hyperkomplexen Basen in weitere Pixelstufen erweitert wurden.

Es gilt das Verfahren der überwachten-Klassifikation mithilfe der aus vorhergehenden Studien erhaltenen Referenzdaten zu automatisieren und eine Genauigkeitssteigerung der räumlich-hyperkomplexen Basen gegenüber dem herkömmlichen Multilooking nachzuweisen. Der trainierte Algorithmus soll auf zwölf neue Aufnahmen, die in den Jahren 2019-2021 von TerraSAR-X aufgenommen wurden, angewendet werden. Dadurch sollen Veränderungen und jahreszeitliche Trends des Lac Bam Sees festgestellt werden.

Abstract

Wetlands are vital water resources for the local population and for biodiversity in semi-arid Africa. Lake Lac Bam, located in the north of Burkina Faso, is the largest natural freshwater lake in this region. The wetlands have several characteristics. The very turbid and sediment-rich water is covered by flooded and floating vegetation.

Dry and wet seasons contribute to seasonal area changes in the wetlands in addition to the highly unpredictable weather. During the dry season, the lake is the only freshwater reservoir in the Bam region besides individual wells. In addition, more and more water pumps are being installed for the irrigation of fields. These represent a further encroachment on the lake's habitat.

The 2009 inclusion of Lac Bam in the List of Wetlands of International Importance, the Ramsar convention, contributes to the focus of remote sensing on this area since 2013.

The study uses dual-co-polarized HH-VV StripMap data from the Terra-SAR-X satellite. The radar scenes were pre-processed at DLR using the MultiSAR system. The pre-processing yielded Kennaugh elements that were extended into additional pixel levels using spatial-hypercomplex bases.

The goal is to automate the supervised classification procedure using reference data obtained from previous studies and to demonstrate an increase in accuracy of spatial-hypercomplex bases over conventional multiLooking. The trained algorithm will be applied to twelve new images acquired by TerraSAR-X in 2019-2021. This will be used to detect changes and seasonal trends in Lac Bam Lake.

Erklärung

gemäß § 16 Abs. 10 APO in Zusammenhang mit § 35 Abs. 7 RaPO

Name: Wolff

Vorname: Richard

Geburtsdatum: 18.03.1991

Studiengang: Angewandte Geodäsie und Geoinformatik

Studiengruppe: 7W

Matrikel-Nr.: 20385417

Semester: Wintersemester 2022/23

Betreuer: Prof. Dr.-Ing. Andreas Schmitt

Hiermit erkläre ich, dass ich die Bachelorarbeit selbstständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt, sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

Ort, Datum Unterschrift

Datum 21.12.2022 München

Unterschrift Wolff

INHALTSVERZEICHNIS

Danksagung

Kurzfassung

Abstract

Erklärung

Tabellenverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

1.	Einleitung.....	1
1.1	Untersuchungsgebiet	2
1.1.1	Feuchtgebiete.....	3
1.1.2	Gesellschaft	5
1.2	Klima	5
1.2.1	Innertropische Konvergenzzone.....	6
1.3	Rolle der Fernerkundung.....	8
1.4	Vorarbeiten	9
1.5	Zielsetzung.....	11
2.	Verwendete Daten	13
2.1	TerraSAR-X Radardaten	13
2.1.1	Synthetic Aperture Radar (SAR)	14
2.2	Multi-SAR.....	15
2.2.1	Die Kennaugh Zerlegung	15
2.2.2	Die Komplexe Basis.....	17
2.2.3	Die räumlich-hyperkomplexe Basis	18
2.3	Spektraldata	20
2.3.1	Sentinel 2	20
2.3.2	Planet.....	20
2.4	Wetterdaten	21
3.	Verwendete Software	23

4.	Methodik	26
4.1	Aufbereiten der Referenzdaten	27
4.1.1	Erstellen der Klasse Land	27
4.1.2	Anpassung des Trainingsgebietes	27
4.2	Polarimetrisch-räumliche Signaturen.....	28
4.3	Klassifizierung.....	28
4.4	Validierung	29
4.5	Visualisieren der Klassifikationsergebnisse.....	30
4.6	Berechnung der Fläche.....	31
5.	Ergebnisse	33
5.1	Polarimetrisch-räumliche Signaturen.....	33
5.2	Ergebnisse Validierung	34
5.2.1	Ergebnisse Gesamtgenauigkeit	34
5.2.2	Ergebnis F1-Score	37
5.2.3	Ergebnis K-Nearest-Neighbor.....	38
5.3	Visualisieren der Klassifizierungsergebnis.....	39
5.4	Ergebnis Flächenberechnung	43
6.	Diskussion	45
6.1	Bewertung der Referenzdaten	45
6.2	Satellitendaten, Höhenmodell und Open Street Map	45
6.3	Bewertung der Wetterdaten	47
6.4	Bewertung der Validierungsergebnisse	48
6.4.1	Fehlklassifikationen	48
6.5	Bewertung der Klassen.....	49
6.6	Bewertung der Kartendarstellung.....	50
6.7	Bewertung der neuen Aufnahmen.....	50
6.8	Bewertung der Flächenberechnungen.....	51
6.9	Empfehlungen für weiterführende Forschung.....	52
7.	Zusammenfassung	54
8.	Literaturverzeichnis	56
	Anhang	59

TABELLENVERZEICHNIS

Tabelle 2.1 Einteilung Kennaugh Elemente.....	15
Tabelle 2.2 Übersicht der mit räumlichen-hyperkomplexen Basen erweiterten TerraSAR-X Daten. TSX_0 zeigt die Originalaufnahme aus dem MultiSAR-Prozessor mit den vier verfügbaren Kennaugh-Elementen. Pro Stufe werden räumliche Kennaugh- Elemente hinzugefügt.	19
Tabelle 2.3 Pixelstufen MultiLooking	19
Tabelle 5.1 F1-Score aller Klassen und die Gesamtgenauigkeit in der Pixelstufe 20m des K-Nearest-Neighbor Algorithmus	38

FORMELVERZEICHNIS

Formel 2.1 Kennaugh-Matrix K	16
Formel 2.2 K0 = Gesamtintensität von HH und VV	16
Formel 2.3 Intensitätsverhältnis zwischen Doppelreflexion und Oberflächenstreuung	16
Formel 2.4 Verhältnis zwischen HH- und VV-Intensität	16
Formel 2.5 Phasenverschiebung zwischen Doppelreflexion und Oberflächenstreuung	16
Formel 2.6 Die Komplexe Basis von 2,5m-Pixel zu 5m-Pixel.....	17
Formel 2.7 Quaternion Basis von 2,5m-Pixel zu 10m-Pixel.....	18
Formel 2.8 P als Patch (Bildausschnitt), die linke Multiplikation überführt die Zeilen, die rechte Multiplikation die Spalten auf räumlich-hyperkomplexe Basen.....	18
Formel 4.1 Precision.....	30
Formel 4.2 Recall	30
Formel 4.3 F1-Score.....	30
Formel 4.4 Accuracy	30

ABBILDUNGSVERZEICHNIS

1.1 Untersuchungsgebiet Lac Bam. Zugeschnittene TerraSAR-X Aufnahme, Planet, Sentinel 2	1
1.2 Typische Feuchtgebietslandschaft in Burkina Faso am Beispiel des Lac Bam: (a) trübes offenes Wasser; (b) schwimmende Seerosen; (c) auftauchende Gräser; (d) natürliche Küstenvegetation; (e) freiliegendes Schlickwatt & Sedimente nach dem Rückzug des Wassers; (f) bewässerte Felder; (g) Wasserstelle; (h) Wohnhaus; (i) Zaun; (j) Hochwasserentlastung/Damm mit Brücke am Lac Bam; (k) Motorpumpen; (l) Graben für Tiere. Fotoquellen: L. Moser und F. Betorz Martínez (Oktober 2013) und R. Ouedraogo (Oktober 2015).....	4
1.3 Niederschläge im Juni-Oktober in der Sahelzone 1920-2010. Gezeigt ist die Abweichung vom Mittel der Jahre 1950-2010 in mm pro Monat.	6
1.4 Grafische Darstellung der Passatzirkulation.....	7
1.5 Ergebnisse der multitemporalen Klassifizierung unter Verwendung von dual-polarimetrischen SAR -Intensität (links) sowie der einpolarimetrischen SAR-Intensität K0 (rechts). Folgende Klassen wurden dabei Klassifiziert: offenes Wasser (dunkelblau), Wasser zu Land (Hellblau), überschwemmte Vegetation an Land (hellgrün), überschwemmte Vegetation zum Feld (grün), bewässerte Felder (rot), Land mit permanenter Vegetation (dunkelgrün) und Land mit Erde, Felsen, städtische Region (beige)) (Quelle: (Linda Moser 2021))	10
2.1 Aufnahmemodi TerraSAR-X.....	14
2.2 a) Sentinel 2 Falschfarbenbild (band 5-3-2 / NIR-rot-grün) Referenzbild vom 10.04.2019 (@ESA (2019)) b-e stellen die vier Kennaugh-Elemente aus dem Dual-Pol-TSX-Bild vom 11.01.2021 (@DLR (2021) dar. b) K0 = die Gesamtintensität als Summe von HH-plus VV- Intensität; c) K3 (Differenz Double-Bounce minus Oberflächenstreuung); d) K4 (Differenz HH minus VV-Intensität); e) K7 (Phasenverschiebung zwischen Double-Bounce und Oberflächenstreuung)	17
2.3 Die orthogonale Transformation von Reflexionsgrad (blau) zu Kennaugh-Elementen (rot) im Zweidimensionalen Raum einer Messung (grün) gemäß Formel 2.2	18
2.4 First Guess Monthly Durchschnittlicher Niederschlagsinformation in mm von Juni 2020 in Afrika / Quelle: https://kunden.dwd.de/GPCC/Visualizer	21
5.1 Pixel Signatur Bewässerte Felder (x-Achse Anzahl Kanäle / y-Achse Digital Number)	33
5.2 Pixel Signatur schwimmende Vegetation (x-Achse Anzahl Kanäle / y-Achse Digital Number).....	33
5.3 Pixel Signatur offenes Wasser (x-Achse Anzahl Kanäle / y-Achse Digital Number)	33
5.4 Pixel Signatur Feuchtgebiete (x-Achse Anzahl Kanäle / y-Achse Digital Number)	33
5.5 Pixel Signatur Urbanes Gebiet (x-Achse Anzahl Kanäle / y-Achse Digital Number).....	34
5.6 Darstellung der Gesamtgenauigkeit der Unterschiedlichen Pixelstufen. Die Klassifikation erfolgte mit dem durch räumlich-hyperkomplexe Basen erweiterten Datensatz.....	34
5.7 Darstellung der Gesamtgenauigkeit der Unterschiedlichen Pixelstufen. Die Klassifikation erfolgte mit dem MutliLooking trainierten Datensatz.....	35
5.8 Darstellung der Gesamtgenauigkeit beider Datensätze.	36
5.9 F1-Score aller Klassen und Pixelstufen. Ergebnis des F1 Score derräumlichen-hyperkomplexen Basen.	37
5.10 Klassifizierte Aufnahmen a) 04.09.2019; b) 22.04.2020; c) 25.05.2020	39
5.11 Klassifizierte Aufnahmen a) 16.06.2020 b) 19.07.2020 c) 04.10.2020.....	40
5.12 Klassifizierte Aufnahmen a) 15.10.2020 b) 06.11.2020 c) 09.12.2020.....	40

5.13 Klassifizierte Aufnahmen a) 20.12.2020 b) 11.01.2021 c) 02.02.2021.....	41
5.14 Beispiel Abgeschlossene Klassifikation anhand der Aufnahme vom 11.01.2021 a) bewässerte Felder (braun), b) offenes Wasser (blau). c) schwimmende Vegetation (grün), d) feuchtgebiete (orange).....	42
5.15 Veränderung der Flächen bewässerte Felder, offenes Wasser, schwimmende Vegetation, Feuchtgebiete in Abhängigkeit des durchschnittlichen Niederschlages in mm.....	43
6.1 Fehlklassifizierung des urbanen Gebietes mit Feuchtgebieten.	46
6.2 SRTM-Höhenmodell für die Einordnung der Gebirgskette	47
6.3 Probleme bei der Klassifizierung von "offenem Wasser" am 29.07.2020	49

ABKÜRZUNGSVERZEICHNIS

DLR	Deutsches Luft- und Raumfahrtzentrum
DWD	Deutscher Wetterdienst
ESA	European Space Agency Europäische Weltraumorganisation
EPSG	European Petroleum Survey Group Geodesy
ITC	Innertropische Konvergenzzone
UTM	Universal Transverse Mercator
WGS	World Geodetic System
GIS	Geoinformationssystem
GPCC	Global Precipitation Climatology Centre Weltzentrum für Niederschlagsklimatologie
IPCC	Intergovernmental Panel on Climate Change Zwischenstaatlicher Ausschuss für Klimaänderungen
NIR	Nahes Infrarot
OSM	OpenStreetMap
RGB	Sichtbarer Farbraum Rot-Grün-Blau
SAGA	System für Automatisierte Geowissenschaftliche Analysen
SAR	Synthetic Aperture Radar
SRTM	Shuttle Radar Topography Mission
USGS	United States Geological Survey
WWF	World Wide Fund For Nature

1. EINLEITUNG

Feuchtgebiete und Flüsse sind wichtige Ressourcen in der halbtrockenen westafrikanischen Sahelzone. Überwacht werden diese Wasserquellen in regelmäßigen Abständen über punktuelle Wasserstandsmessungen. Wasser ist in dieser Region eine wertvolle Existenzgrundlage für Landwirte, Viehzüchter, die Bevölkerung, Fischer sowie die Tierwelt. Feuchtgebiete stellen außerdem wichtige Lebensräume für die biologische Vielfalt bereit. Zusätzlich tragen diese zum Erhalt des Ökosystems bei. (Linda Moser 2021) Die Feuchtgebiete werden in den folgenden Kapiteln noch genauer betrachtet.

Die Studie demonstriert die Anwendbarkeit von wetterunabhängigen dual-co-polarized HH-VV StripMap Daten des Terra-SAR-X Satelliten für die Analyse der räumlich-zeitlichen Dynamik von Feuchtgebieten anhand des Lac Bam Sees. Dabei soll das Verfahren der überwachten-Klassifizierung automatisiert werden. Zusätzlich wird untersucht, ob räumlich-hyperkomplexe Basen eine Genauigkeitssteigerung gegenüber dem standardisiertem MultiLooking Verfahren erzielen.

Aufnahmegeriet Lac Bam / Burkina Faso

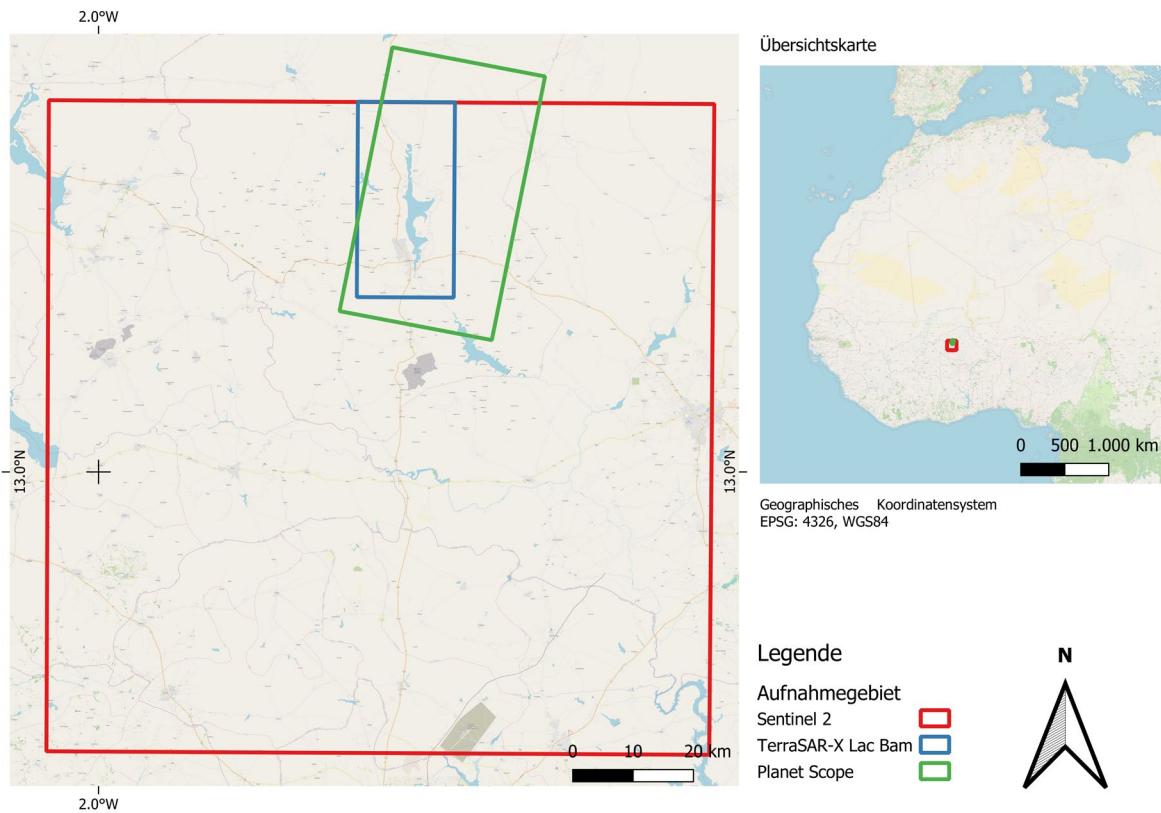


Abbildung 1.1 Untersuchungsgebiet Lac Bam. Zugeschnittene TerraSAR-X Aufnahme, Planet Scope, Sentinel 2

1.1 UNTERSUCHUNGSGEBIET

Der in dieser Arbeit zu untersuchende See Lac Bam befindet sich in Burkina Faso, das im Westen von Afrika gelegen ist. Damit liegt der See auch in der sogenannten Sahelzone. Die Zone ist bekannt für seine geringe Menge an Niederschlag. Der See liegt 300 m über dem Meeresspiegel und ist mit ca. 22-25km² der größte Süßwasserspeichersee in Burkina Faso. Dadurch ist er von großer Bedeutung für die Bewässerungslandwirtschaft, die Viehtränke, die Fischerei und die Entnahme von Wasser für Trinkwasser. Dafür werden vor allem für die Bewirtschaftung der umliegenden Felder Wasserpumpen errichtet. (Ambroise Ouédraogo)

Die örtlichen Behörden bestätigen, dass mindestens 1000 Motorpumpen Wasser aus dem See auch an bis zu zwei Kilometer entfernten Feldern gepumpt wird. Das Wasser wird dort für den Gemüseanbau verwendet. Es ist ebenfalls zu erkennen, dass sich die Felder im Laufe der Saison dynamisch verändern. Während der Trockenzeit (Mai-Oktober) gibt es bis auf kleine Wasserstellen oder Brunnen keine nennenswerten Süßwasserquellen. (Moser et al. 2016)

Durch die immer steigende Zunahme der Bewirtschaftung von Feldern im Gebiet Lac Bam, steigt gleichermaßen auch der Einsatz von Düngemitteln und Pestiziden. Diese werden vom See aufgenommen und tragen somit zur Verschmutzung bei. Der zunehmende Anbau führt auch dazu, dass bei der Kultivierung der Randgebiete die natürliche Vegetation entfernt wird. Durch die Entfernung der Vegetation entsteht ein zusätzlicher Sedimenteneintrag, der sich am Boden des Sees lagert. Im Zeitraum von 1963-2006 hat der See wegen der hohen Sedimentationsrate ein Drittel an Tiefe verloren. (Ouedraogo)

Die zunehmenden Probleme bedrohen die Existenz der für die in der Region lebenden Menschen und den Feuchtgebieten. Es ist wichtig den Süßwasserstand des Sees zu überwachen und diesen vor der Austrocknung zu schützen. Der Lac Bam See wurde im Jahre 2009 zu der weltweiten Ramsar-Liste der Feuchtgebiete von internationaler Bedeutung gemäß dem zwischenstaatlichen Ramsar-Vertrag aufgenommen. (Ramsar Convention on Wetlands 2022)

Die Ramsar-Konvention wurde 1971 von der Umweltschutzorganisation World Wide Fund For Nature (WWF) initiiert und setzt sich weltweit für den Erhalt von artenreichen Lebensräumen und Wassereinzugsgebiete für Millionen Menschen ein. Die in die Ramsar Konvention aufgenommenen Feuchtgebiete haben schätzungsweise ein Ausmaß von 12,8 Millionen Quadratkilometern. Zusätzlich setzt sich die WWF auch für die nachhaltige Fischerei ein, die in den Gebieten einer der Hauptlieferanten für tierisches Eiweiß darstellt. (Ramsar-Konvention: Schutz der Feuchtgebiete 2022)

1.1.1 FEUCHTGEBIETE

Feuchtgebiete im Allgemeinen besitzen keine einheitliche Definition. In dieser Arbeit wird die Definition des Ramsar-Übereinkommens über Feuchtgebiete gewählt. Zum Stand Juli 2021 umfasste die Datenbank mehr als 2422 Feuchtgebiete. Die Liste wird stets erweitert. In der Sahelzone bedecken die Feuchtgebiete bei Ihrer maximalen Ausdehnung mehr als 10% des gesamten Gebietes.

Feuchtgebiete bestehen aus mehreren verschiedenen Bodenbedeckungsarten. Die unterschiedlichen Typen besitzen verschiedene räumliche und zeitliche Merkmale. Durch die saisonale Überwachung oder das Erstellen einer Zeitreihenanalyse über viele Jahren können Trends wie z.B. von Dürren langfristig verursachten Schäden, die unter anderem Auswirkungen auf die Wasserversorgung haben, nachgewiesen werden.

In den westafrikanischen Ländern südlich der Sahara enthalten Feuchtgebiete in der Regel offenes Wasser, unterflutete / schwimmende Vegetation, bewässerte Felder und diverse Infrastrukturen, die Einflüsse auf die Feuchtgebiete haben. (Abbildung 1.2) Dazu zählen Dämme, Motorpumpen, Zäune und Tiergräber. Es ist auch möglich, dass städtische Gebiete innerhalb oder in der Umgebung der Feuchtgebiete liegen.



Abbildung 1.2 Typische Feuchtgebietslandschaft in Burkina Faso am Beispiel des Lac Bam: (a) trübes offenes Wasser; (b) schwimmende Seerosen; (c) auftauchende Gräser; (d) natürliche Küstenvegetation; (e) freiliegendes Schlickwatt & Sedimente nach dem Rückzug des Wassers; (f) bewässerte Felder; (g) Wasserstelle; (h) Wohnhaus; (i) Zaun; (j) Hochwasserentlastung/Damm mit Brücke am Lac Bam; (k) Motorpumpen; (l) Graben für Tiere. Fotoquellen: L. Moser und F. Betoriz Martínez (Oktober 2013) und R. Ouedraogo (Oktober 2015)(Quelle: (Linda Moser 2021))

Folgende Bodenabdeckungsklassen gilt es in dieser Arbeit zu klassifizieren:

Offene Wasserflächen können Flächen zwischen Quadratmetern und Kilometern erreichen. Die Wasserflächen verhalten sich im Gegensatz zur Umwelt homogen. Während es Teile von Seen gibt, die sich „stabil“ verhalten, werden andere Teile von Regen- und Trockenzeit beeinflusst.

Dabei können große Ausdehnungen während der Regenzeit und Rückzüge der Seen zur Trockenzeit festgestellt werden. Für die Überwachung der Wasserfläche eignen sich im Idealfall Fernerkundungsdaten, die in kurzen Intervallen aufgenommen wurden.

Zur **unterfluteten und schwimmenden Vegetation** in den Afrikanischen Feuchtgebieten zählen Bäume, Sträucher sowie Wasserpflanzen, die für Feuchtgebiete charakteristisch sind. Die Wasserpflanzen können aus Gräsern, Farne, Moosen und blühenden Pflanzen, die an oder im Wasser wachsen, bestehen. Die Flächenausdehnung der *unterfluteten / schwimmenden Vegetation* beträgt von kleinen Flächen bis hin zu solchen, die über ein paar hundert Metern groß werden können. Durch die jahreszeitlichen Veränderungen können diese Teile der Feuchtgebiete während der Trockenzeit vollständig austrocknen.

Nach Ramsar – Definition sind **bewässerte Felder** ein weiterer Teil von Feuchtgebieten. Sie werden mit dem Seewasser bewässert und dienen dem Gemüse- oder Reisanbau. Die Größen der Felder ändern sich dynamisch und sind von den Erntezeiten abhängig. Während der Regenzeit werden die Felder mit Regenwasser gespeist. Erst mit Beginn der Trockenzeit werden diese mit Seewasser bewirtschaftet. (Linda Moser 2021)

1.1.2 GESELLSCHAFT

Burkina Faso, auch genannt „Land der Aufrechten“, hat eine Einwohnerzahl von 21,5 Millionen Menschen. Die Bevölkerung ist ethnisch und religiös divers. Das Land hat, wie andere Sahel-Regionen auch, mit Terrorismus, hohem Bevölkerungswachstum, Klimawandel, Armut- und Ernährungssicherheit zu kämpfen.

So gab es allein 2022 (stand 01.11.22) schon den zweiten Militärputsch in einem Jahr, bei dem die gewählte Regierung abgesetzt wurde. Die Sicherheitslage trägt neben dem Klimawandel zu der größten humanitären Notlage des Landes bei. Die Folgen des Klimawandels führen zu Flucht und Migration. (Amt 2022)

Eine der größten Herausforderungen für Burkina ist das starke Wachstum der Bevölkerung. Es erschwert die politischen Bemühungen, die Armut zu verringern, genügend Arbeitsplätze zu schaffen, die Ernährung zu sichern und alle Einwohnerinnen und Einwohner mit Wasser und Energie, Gesundheitsdiensten und Bildungseinrichtungen zu versorgen. (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung 2022)

1.2 KLIMA

Das Klima in Burkina Faso ist im Allgemeinen heiß, sonnig und trocken. Es gibt verschiedene Klimazonen, die durch das Köppen-Geiger-Klimaklassifikationssystem definiert sind (Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F.) et al., 2006). Betrachtet man die Zone von Norden nach Süden, gilt folgende Einteilung: Die nördlichen Gebiete Burkina Fasos werden als Sahelzone bezeichnet, die mittleren Gebiete als sudanesische-sahel Savanne und die südlichen Gebiete als sudanesische Savanne. In der Sahelzone befindet sich der Lac Bam See. Die Niederschläge fallen fast ausschließlich während der Regenzeit (Juni-September). Während der Trockenzeit (Oktober-Mai) ist das Wasser in Feuchtgebieten für die Bevölkerung sehr wichtig und dient verschiedenen Lebensgrundlagen. (Linda Moser 2021)

Afrika ist häufig von Dürren betroffen. Hinzu kommt, dass das Klima abhängig von der innertropischen Konvergenzzone ist, die durch die Temperatur und die Druckverhältnisse der benachbarten Ozeane beeinflusst wird. Ausschlaggebend für einen hohen Niederschlag ist ein Temperaturunterschied zwischen Kontinent und Ozean. Wie man in der **Fehler! Verweisquelle konnte nicht gefunden werden.** erkennen kann, fand in den 70er und 80er Jahren eine große Dürre in der Region statt. Dies ist auf das Aufheizen des Ozeans zwischen Indonesien und Ostafrika zurückzuführen. Dadurch war der Temperaturunterschied zwischen Land und Wasser relativ gering und der Regen konnte durch den Temperaturdruck nicht bis ins Landesinnere getragen werden. Dieser Temperaturanstieg ist auf die Zunahme der Treibhausgase zurückzuführen. (Wasserprobleme im Sahel – Klimawandel 2022)

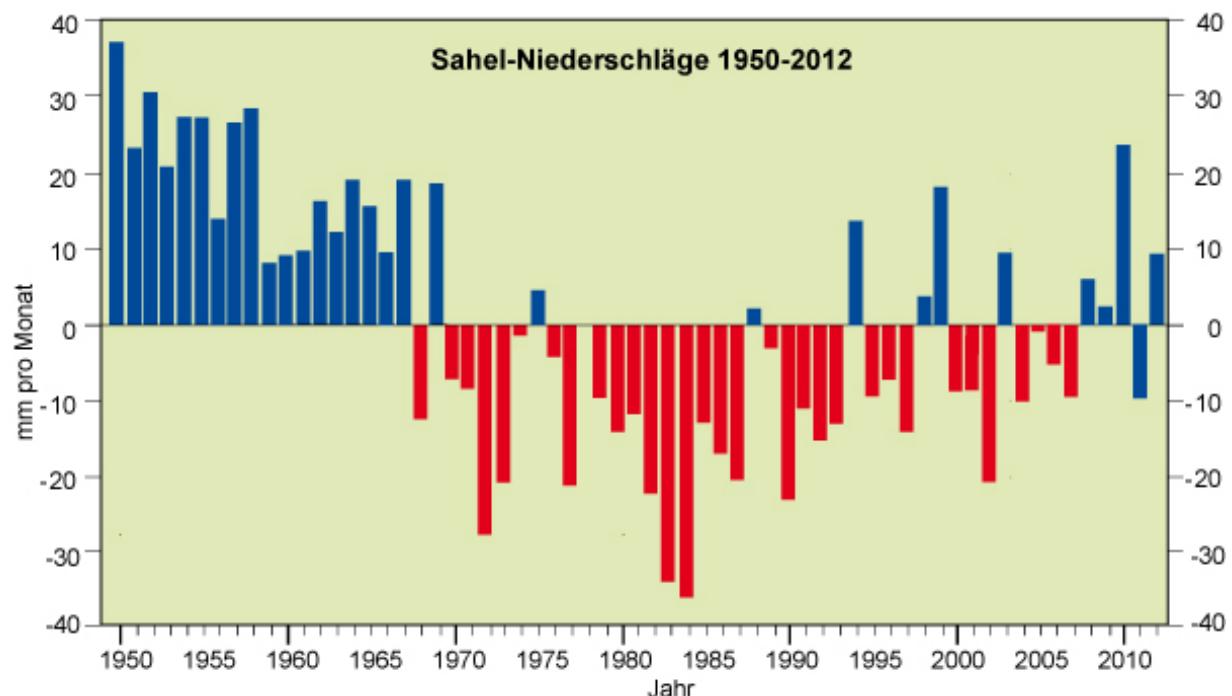


Abbildung 1.3 Niederschläge im Juni-Oktober in der Sahelzone 1920-2010. Gezeigt ist die Abweichung vom Mittel der Jahre 1950-2010 in mm pro Monat.

(Quelle: https://wiki.bildungsserver.de/klimawandel/index.php/D%C3%BCrren_im_Sahel)

1.2.1 INNERTROPISCHE KONVERGENZZONE

Die innertropische Konvergenzzone (ITC) ist eines der wichtigsten Windsysteme innerhalb der atmosphärischen Zirkulation. Sie befindet sich zwischen den subtropischen Hochdruckgürtel der Nord- und Südhalbkugel (zwischen 30°N und 30°S). Die Sonne steht hier ganzjährig nahe dem Zenit und heizt somit die Luft auf.

Durch den niedrigen Luftdruck, der aufgrund der starken Erwärmung und des damit verbunden Aufstiegs der äquatornahen Luftmassen in der äquatorialen Tiefdruckrinne herrscht, wird ständig Luft von der nördlichen und südlichen Halbkugel angesaugt. Diese Winde heißen Passate, die auf das Klima und das Wetter in der Sahel Zone einen großen Einfluss haben.

Die Passatwinde verhindern einen Niederschlag, denn der sich in der Höhe befindende Wind wird auch Antipassat bezeichnet. Dieser kühlt ab und sinkt nach unten. Bei diesem Vorgang erwärmt sich die Luft zeitgleich. Somit wird ein Austausch der Luft verhindert. Die warme Luft überlagert die Kältere, es entstehen mehrere Luftsichten und somit ein Hochdruckgebiet. (Forkel 2015)

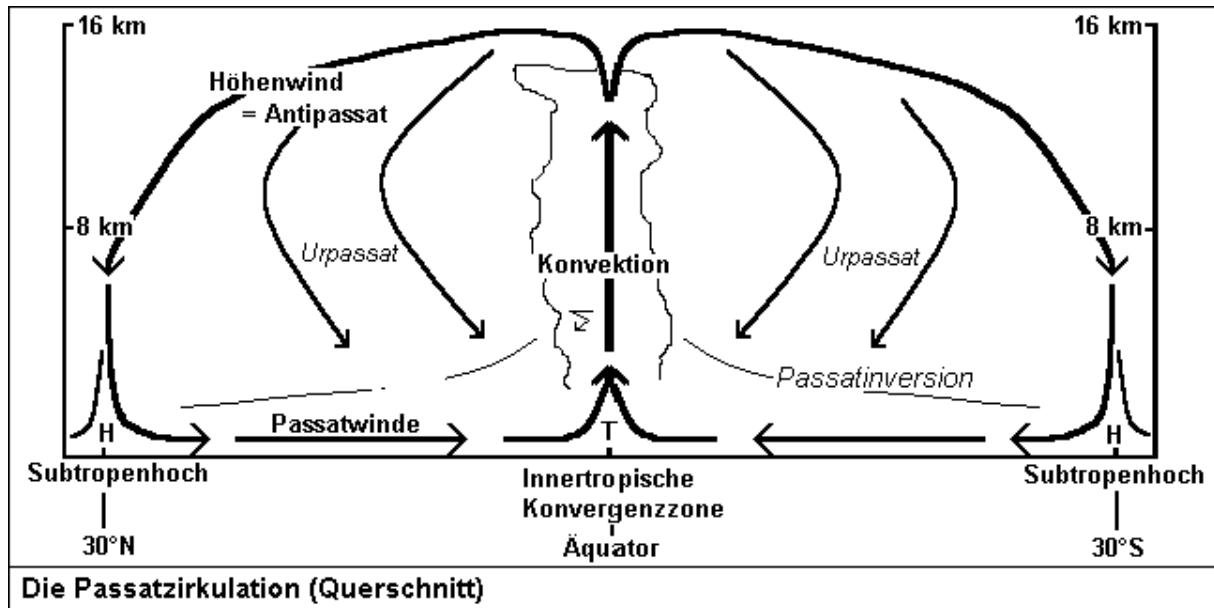


Abbildung 1.4 Grafische Darstellung der Passatzirkulation

(Quelle http://www.klima-der-erde.de/zirk_passat.html)

1.3 ROLLE DER FERNERKUNDUNG

Die Überwachung von Feuchtgebieten stellt durch die dynamisch zeitlich abhängigen Veränderungen eine große Herausforderung dar. Vor allem der Lac Bam See ist neben der Regen- und der Trockenzeit nicht nur ökologischen Einflüssen ausgesetzt, sondern auch der Mensch trägt zu dynamischen Veränderungen bei.

Dazu zählt, dass der See von der häuslichen Nutzung, wie Landwirtschaft, Viehzucht und Überfischung degradiert wird. Zusätzlich kommen die zunehmende Sedimentierung und Verschmutzung hinzu. Außerdem wird der Wasserstand durch vom Menschen angelegte Dämme versucht zu regulieren. (Linda Moser 2021)

Die Dynamik des Sees erfordert zeitliche Beobachtungsintervalle, die ein zentraler Bestandteil dieser Arbeit sind. Die Wasserstandsmessung erfolgt meist durch bodengestützte Verfahren. Dabei wird der Wasserstand an Dämmen oder Stauseen abgelesen. Solche Gebiete erfordern eine große räumliche Abdeckung, die durch die Wasserstandsmessung nicht gewährleistet werden kann. Die Fernerkundung kombiniert hierbei die große Flächenabdeckung sowie die erforderliche hohe räumliche Auflösung.

Die Fernerkundung stellt somit eine gute Möglichkeit bereit, den Wasserstand flächenabdeckend mit einer räumlichen Auflösung zu überwachen. Mit dieser Methode können bessere Messergebnisse erzielt werden als mit der sonst in Afrika üblichen bodengestützten Methode. Bei der bodengestützten Wasserstandsmessung kann der Wasserstand nur punktuell und nicht flächendeckend bestimmt werden. (Linda Moser 2021)

Neben der Berechnung der Wasserfläche können mithilfe der Fernerkundung auch die schwimmende Vegetation, die bewässerten Felder und die Feuchtgebiete überwacht werden.

In dieser Arbeit wird eine Zeitreihe mit räumlich-hyperkomplexen Basen erweiterten TerraSAR-X Zeitreihe auf eine dynamische Mustererkennung überwacht. Es sollen die Veränderungen des Sees und der umliegenden Felder während der Trocken- und Regenzeit mithilfe einer überwachten-Klassifikation automatisch erkannt werden. Die Klassifikationsergebnisse werden untereinander verglichen, um eine Genauigkeitssteigerung im Datensatz der räumlichen-hyperkomplexen Basen gegenüber dem MultiLooking nachzuweisen. Die TerraSAR-X Daten versprechen dabei gute Ergebnisse zu erzielen. Die hohe Wolkenbedeckung während der Regenzeit würde der Klassifikation mithilfe optischer Sensoren Probleme bereiten. Auch die Klassifikation der schwimmenden Vegetation wäre mit optischen Sensoren nur schwer zu erkennen. (Linda Moser 2021)

Zuvor gab es bereits Arbeiten, die mithilfe Fernerkundung gestützten Lösungen Feuchtgebiete und Gewässer in der Sahelzone überwacht haben. Das Ziel dieser Arbeit ist es, aus den gewonnenen Erkenntnissen, der Vorarbeiten von Moser.L (Linda Moser 2021) und Klingebiel C. (Klingebiel et al. 2019), aufzubauen. Dabei soll vor allem die von Klingebiel C. Klassifizierte 5- jährige Zeitreihe dafür genutzt werden, das Verfahren zu automatisieren. Im Nachfolgendem, Kapitel 1.7 Vorarbeiten werden diese noch genauer betrachtet.

Durch die beschriebenen Punkte, wie z.B. das ansteigende Klima und die daraus resultierenden Folgen, die vor allem in der Sahelzone spürbar sind, wurde das Augenmerk der Fernerkundung in den letzten Jahren auf diese Region gelenkt. So ist es wichtig, die Folgen durch Beobachtungen des Klimawandels mithilfe von Fernerkundung zu kontrollieren und an den Möglichkeiten dieser zu forschen.

Die wissenschaftliche Nutzung der TerraSAR-X Mission bezieht sich vor allem auf die Beobachtung von Vegetation, die für das menschliche Leben eine große Rolle spielt. Durch die Verteilung, Zusammensetzung und Änderung der verschiedenen Arten sind diese Erkenntnisse, die durch die Aufnahmen des Satelliten erzielt werden können, ein großer Bestandteil für viele Anwendungen.

1.4 VORARBEITEN

Die spürbaren Auswirkungen des Klimawandels und die Aufnahme des Bam-Sees in die *List of Wetlands of International Importance*, haben dazu geführt, dass die Bam Region einen höheren Stellenwert für Forschungsarbeiten erlangt hat.

So befasst sich die Dissertation von Moser L. (Linda Moser 2021) mit der Überwachung von Feuchtgebieten und Gewässern mithilfe der Fernerkundung. Hauptaugenmerk waren semiaride Gebiete in Burkina Faso, die von Regen- und Trockenzeit beeinflusst werden. Es wurden Fernerkundungszeitreihen von optischen Multispektraldaten und Synthetic Aperture Radar (SAR) Daten für die Analyse der räumlich-zeitlichen Entwicklung innerhalb saisonaler Zyklen und zwischen verschiedenen Jahren analysiert.

Des Weiteren wurden aus den Arbeiten von Klingebiel C. (Klingebiel et al. 2019) und Moser L. (Moser et al. 2016) die Erkenntnisse erlangt, dass die Multi-SAR Prozessierung sich für die Klassifizierung von Feuchtgebieten, eignet. Aus diesen Forschungsarbeiten wird verständlich, dass vor allem die schwimmende Vegetation, wie Seerosen durch die Kennaugh Elemente besser zu klassifizieren sind. Dies konnte man durch den Vergleich der unterschiedlichen Kennaugh Elementen (Multi-temporal) mit den herkömmlichen single-polarimetric SAR (KO) feststellen. Es kam zu Fehlklassifikationen der unterschiedlichen Klassen im Single Band. So ist es ohne die Verwendung der polarimetrischen Daten zu Verwechslungen der Klassen schwimmende Vegetation, bewässerter Anbau und Land gekommen.

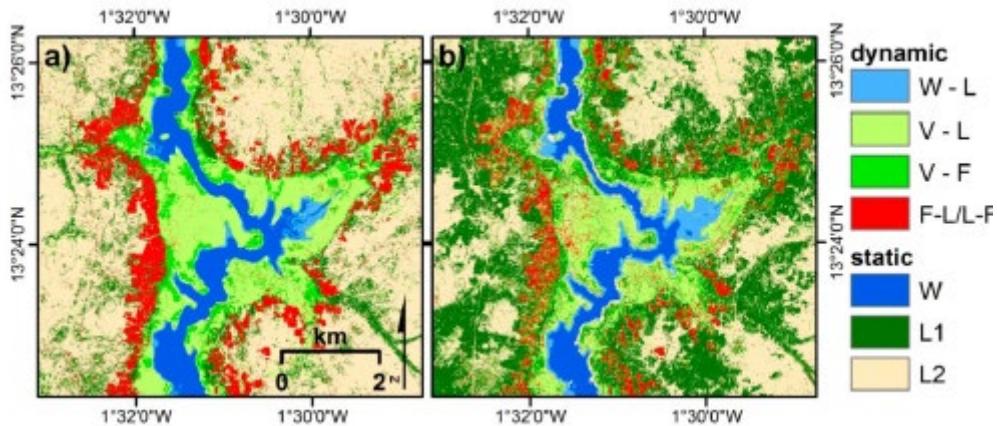


Abbildung 1.5 Ergebnisse der multitemporalen Klassifizierung unter Verwendung von dual-polarimetrischen SAR -Intensität (links) sowie der einpolarimetrischen SAR-Intensität K0 (rechts). Folgende Klassen wurden dabei Klassifiziert: offenes Wasser (dunkelblau), Wasser zu Land (Hellblau), überschwemmte Vegetation an Land (hellgrün), überschwemmte Vegetation zum Feld (grün), bewässerte Felder (rot), Land mit permanenter Vegetation (dunkelgrün) und Land mit Erde, Felsen, städtische Region (beige)) (Quelle: (Linda Moser 2021))

Die Klassifikation der Feuchtgebiete aus den hervorgegangenen Arbeiten erfolgte bei Klingebiel C. durch eine Kombination aus Maximum Likelihood und einer Schwellenwertberechnung. Die Klassen wurden in *offene Wasserfläche* (Schwellenwertberechnung), *Feuchtgebiete, bewässerte Landwirtschaft und schwimmende Vegetation* (Maximum Likelihood) eingeteilt. Dabei wurde für jede Szene und jede Klasse eine eigenständige Klassifikation durchgeführt. (Klingebiel et al. 2019) Die Klassifikationsergebnisse lieferte den in dieser Arbeit genutzten Referenzdatensatz. Anhand der Ergebnisse und der TSX-Datenzeitreihe konnten die angewendeten Algorithmen trainiert werden.

In der Arbeit von (Moser et al. 2016) wurde hauptsächlich eine überwachte Klassifizierung auf eine mit Kennaugh-Elementen verarbeitete TSX-Datenzeitreihe (elf tägliches intervall / 2013-2015) angewendet. Es wurden vier Klassen (*offene Gewässer, schwimmende Vegetation, bewässerte Felder und Land*) definiert und mithilfe des Maximum-Likelihood-Algorithmus klassifiziert. Für die Analyse der Klassen wurde für jede einzelne Aufnahme eine monosemrale Klassifizierung berechnet. So konnte die dynamische Veränderung des Sees zwischen jedem Aufnahmepunkt festgestellt werden. Um die Feuchtgebietsveränderungen über den gesamten Zeitraum der TSX-Daten nachzuweisen und zu veranschaulichen wurde zusätzlich noch eine multitemporale Klassifikation durchgeführt. Die in den Arbeiten von Moser L. (Moser et al. 2016) und Klingebiel C. (Klingebiel et al. 2019) verwendeten TSX-Datenzeiten hatten jeweils eine Genauigkeit von 5m-Pixel.

Auch der Artikel von Taylor C. (Taylor et al. 2017). beschäftigt sich mit der Auswirkung des Klimawandels in der Sahel-Zone. In dem Artikel wird auf die daraus resultierenden verstärkten Wetterextremen eingegangen. Vor allem die Erwärmung der Sahara und die erhöhten Windscherungen und Veränderungen in der Saharaluftschicht tragen dazu bei, dass es voraussichtlich im Laufe des 21. Jahrhunderts zu Anhäufungen von extremen Regenfällen in der Sahelzone kommen wird. Die resultierenden Erkenntnisse aus dieser Arbeit wurden durch eine Auswertung von Satellitendaten der letzten fünfunddreißig Jahre erforscht.

Der Artikel *Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases (Schmitt et al. 2020)* beschäftigt sich mit der Bildfusion einer orthogonalen Transformation der Reflexionskanäle. Dabei wurden die Kanäle der optischen Sensoren, die in Kennaugh Elemente vorprozessiert wurde, mit Hyperkomplexe Basen erweitert. Somit konnten SAR und die optischen Sensordaten zu einem Bilddatensatz fusioniert werden. Die Hyperkomplexe Basen werden in dieser Arbeit in räumliche-hyperkomplexe Basen erweitert. Diese werden im Kapitel 2.2.3 Die räumlich-hyperkomplexe Basis noch genauer betrachtet.

1.5 ZIELSETZUNG

Das Ziel dieser Arbeit ist es, den Informationsgewinn von räumlich-hyperkomplexen Basen mithilfe eines Random Forests oder K-Nearest-Neighbor nachzuweisen. Dabei werden die Algorithmen mit den vom DLR vorprozessierten Daten und die aus Lac Bam im Radarbild - jahreszeitliche und langfristige Entwicklung“ (Klingebiel et al. 2019) erhaltenen Klassifikationsergebnisse trainiert. Die Klassifikation soll für alle Aufnahmen und der jeweiligen Pixelstufen automatisch durchgeführt werden.

Durch die Erweiterung der dual-co-polarimetric HH-VV Daten des TerraSAR-x Satteliten mit Kennaugh Elemente und räumlich-hyperkomplexe Basen wurden jeweils fünf neue Pixelstufen berechnet. Nachdem die Klassifikation jeder Pixelstufe durchgeführt wurde, sollen die Ergebnisse validiert werden. Durch die Validierung der Klassifikationsergebnisse wird eine Genauigkeitsanalyse durchgeführt. Dadurch können beide Datensätze und deren Pixelstufen untereinander verglichen werden. Die Analyse soll Erkenntnisse liefern, ob eine Genauigkeitssteigerung der hyperkomplexen-Basen gegenüber dem Standardmultilooking vorliegen.

In dieser Arbeit sollen folgende Forschungsfragen beantwortet werden:

- Gibt es eine Genauigkeitssteigerung der räumlich-hyperkomplexen Basen gegenüber dem StandardmultiLooking.
- Ist eine vollautomatische Klassifikation möglich?

2. VERWENDETE DATEN

In dieser Arbeit wurde für die Klassifikation und der Validierung dual-co-polarimetric HH-VV Daten des polarimetrischen TerraSAR-X Satelliten verwendet. Als visuelle Referenz wurden Sentinel 2 sowie Planet Daten herbeigezogen. Zusätzlich wurden noch Daten des Deutschen Wetterdienst verwendet. Diese waren hilfreich bei der Einordnung von Regen- und Trockenzeit. Im Folgendem wird auf die verwendeten Daten näher eingegangen.

2.1 TERRASAR-X RADARDATEN

Für die Klassifizierung wurden die vorprozessierten polarimetrischen TerraSAR-X Daten des gleichnamigen deutschen Erdbeobachtungssatelliten genutzt. Die Daten wurden vom Deutschen Zentrum für Luft und Raumfahrt vorprozessiert und dank einer Kooperation mit der Hochschule München frei zur Verfügung gestellt.

Der Satellit ist mit einem Radarsensor ausgestattet, der in verschiedenen Modi benutzt werden kann. Diese ermöglichen Aufnahmen in unterschiedlichen Streifenbreiten, Auflösungen und Polarisationen. Die seit 2007 gestartete Mission stellt hochwertige Synthetic-Aperture-Radar (SAR) Daten im X-Band für verschiedenste Forschungszwecke bereit. Der Sensor bringt neben seiner hochauflösenden Abbildung der Erdoberfläche auch weitere Vorteile gegenüber herkömmlichen optischen Satelliten. SAR ist unabhängig von Sonneneinstrahlung und Wetterbedingungen. Vor allem Wolken bereiten bei optischen Aufnahmen große Probleme. Es können zu bestimmten Zeitpunkten Aufnahmen mit Radar gemacht werden, ohne sich durch diese Einflüsse einschränken zu lassen. (Albertz 2009)

Das SAR-System TerraSAR-X kann in verschiedenen Aufnahmemodi betrieben werden. Die hier bereitgestellten Daten wurden mit dem „StripMap-Modus“ aufgenommen. Das Gebiet wird von der Flugbahn des Satelliten nicht senkrecht, sondern schräg zur Erdoberfläche angestrahlt. In diesem Modus tastet der Satellit einen Streifen von 15 Kilometern Breite und einer maximalen Länge von 1500 Kilometern ab. Dabei wird eine Auflösung von drei Metern erreicht.

Der am 15.Juni 2007 gestartete Satellit wurde in einen 514m hohen Orbit gebracht. Die Erde wird dabei streifenweise abgetastet. Nach 11 Tagen befindet sich der Satellit auf seiner Ausgangsposition und ein neuer Zyklus wird gestartet. Die sogenannte Repeat Pass Time ist der geringste Abstand zwischen zwei Aufnahmen. (MISSION DLR Portal 2022)

Es standen für diese Arbeit insgesamt 119 Terra-SAR-X Aufnahmen aus der Arbeit von Klingebiel C. (Klingebiel et al. 2019) zum Training des Algorithmus zu Verfügung. Zusätzlich wurden zwölf weitere Aufnahmen aus den Jahren 2019,2020 und 2021 vom DLR vorprozessiert. Auf diese soll der trainierte Random Forest Algorithmus angewendet werden.

2.1.1 SYNTHETIC APERTURE RADAR (SAR)

SAR wird seit den 1980er in großem Umfang für die Fernerkundung von Land und Meer eingesetzt. SAR ist ein aktives Erfassungssystem, dessen Unterschied zu einem konventionellen Radar (Radio Detection and Ranging) darin besteht, dass es eine viel längere Antenne besitzt, die auf einer vorwärts bewegende Plattform montiert wurde. Durch die Bewegung des Satelliten ergeben sich aufeinanderfolgende Sende-/Empfangszeiten zu unterschiedlichen Positionen. Das aktive Sensorsystem beleuchtet das Gelände, indem es elektromagnetische Wellen im Mikrowellenbereich aussendet. Die zurückgestreuten Echos werden vom Empfänger in Form von Laufzeiten gemessen. Radar unterstützte Satelliten können in unterschiedlichen Frequenzbereichen des Radarspektrums beobachten. TerraSAR-X Satellit verwendet das X-Band, dass eine Wellenlänge von ca. 3 cm besitzt.

Der in dieser Arbeit verwendete „StripMap Mode“ ist einer von drei Aufnahmemodi des Satelliten. Bei den vorliegenden Aufnahmen handelt es sich um eine gleichpolarisierte (HH, VV) Kombination. Diese sind abhängig von den Impulsen, die von der Antenne ausgesandt werden. Dabei können diese horizontal(H) oder vertikal(V) polarisiert sein. (DLRARTICLE DLR Portal 2022)

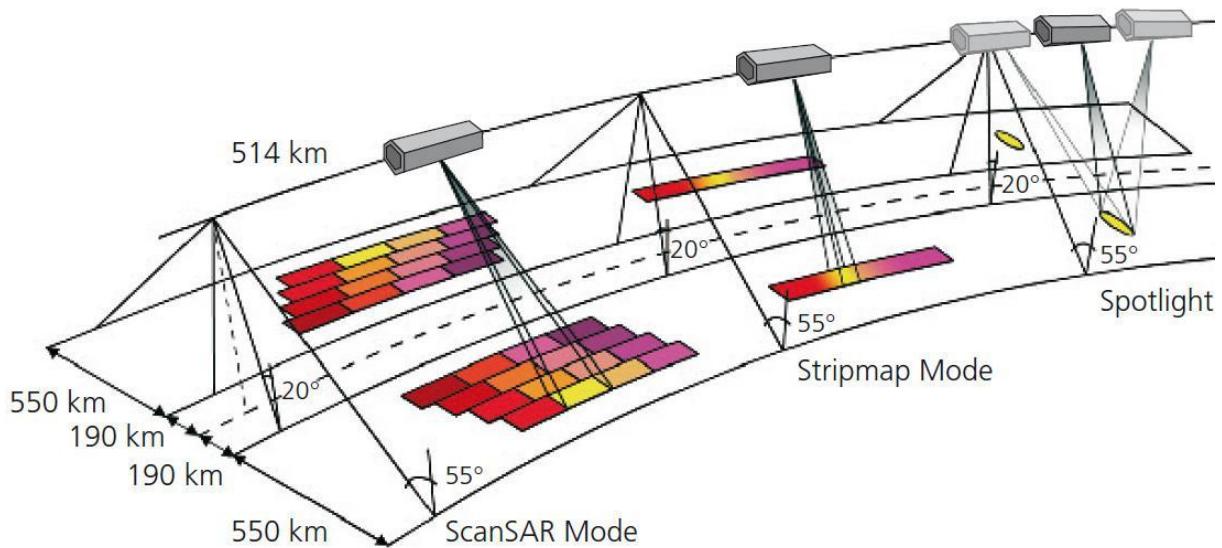


Abbildung 2.1 Aufnahmemodi TerraSAR-X

(Quelle:
<https://www.dlr.de/content/de/artikel/missionen-projekte/terrasar-x/synthetisches-apertur-radar.html>)

2.2 MULTI-SAR

Multi-SAR ist eine Vorverarbeitungssystem für SAR-Daten, zu diesen zählen auch die TerraSAR-X Aufnahmen. Die Daten können unabhängig von ihrer geometrischen, radiometrischen und polarimetrischen Auflösung prozessiert werden. Der große Vorteil von Multi-SAR ist, dass die prozessierten Daten in einem einheitlichen Format vorliegen. In diesem Format können sie unabhängig ihres Datenformates miteinander verglichen werden. (Schmitt et al. 2015) Die erhaltenen Kennaugh Elemente sind für die weitere Prozessierung mit den räumlich-hyperkomplexen Basen wichtig. (Schmitt et al. 2015)

2.2.1 DIE KENNAUGH ZERLEGUNG

Das Ziel der Kennaugh-Matrix ist die Beschreibung der polarimetrischen Informationen und die Interpretation physikalischer Mechanismen wie z.B. Doppelreflexion, Oberflächen- und Volumenstreuung. Jede Kennaugh-Matrix besteht aus zehn Elementen, die durch reelle Zahlen ausgedrückt werden.

In der Tabelle wird auf die Einteilung der Kennaugh-Elemente näher eingegangen. Diese kann man, abgesehen von K0, das die Gesamtintensität widerspiegelt, in drei Gruppen einteilen.

Tabelle 2.1 Einteilung Kennaugh Elemente

Kennaugh Element	Beschreibt	Bedeutung
K0	Gesamtintensität	Summe beider Intensitäten
K1, K2, K3	Absorption	Differenz zwischen Doppelreflexion und Oberflächenintensität
K4, K5, K6	Diattenuation	Änderung des Verhältnisses zwischen zwei Amplitudenwerten
K7, K8, K9	Verzögerung	Phasenwechsel zwischen Doppelreflexion und Oberflächenstreuung

Die Kennaugh Zerlegung ermöglicht somit unabhängig von Polarisation, Sensor oder Wellenlänge die Vergleichbarkeit und Kombination verschiedener SAR-Daten. Für die Zeitreihe der dual-polarisierten TerraSAR-X Daten wurde die von Schmitt A. (Schmitt et al. 2015) entwickelte normalisierte Kennaugh- Matrix K angewendet. (Formel 2.1)

$$[K] = \begin{bmatrix} K_0 & K_4 & K_5 & K_6 \\ K_4 & K_1 & K_9 & K_8 \\ K_5 & K_9 & K_2 & K_7 \\ K_6 & K_8 & K_7 & K_3 \end{bmatrix} = K_0 * \begin{bmatrix} 1 & k_4 & k_5 & k_6 \\ k_4 & k_1 & k_9 & k_8 \\ k_5 & k_9 & k_2 & k_7 \\ k_6 & k_8 & k_7 & k_3 \end{bmatrix} = K_0 * [k]$$

Formel 2.1 Kennaugh-Matrix K

Bei dieser wird die Gesamtintensität K0 als Skalierungsfaktor auf alle anderen Kennaugh- Elemente bezogen. Die normalisierte Kennaugh-Matrix ermöglicht die Trennung von Intensität und polarimetrischen Informationen. Außerdem wird ein geschlossener Wertebereich von]-1,+1[garantiert. Dieser ist für die Komprimierung und Archivierung notwendig.

Die jeweiligen Kennaugh-Elemente können abhängig von Ihrer Polarisationsrichtung berechnet werden. Die Dissertation von (Linda Moser 2021) hat ergeben, dass die HH-VV Polarisation eine gute Anwendung auf Feuchtgebiete verspricht. So sind u.a. beim Unterschied zwischen HH und VV- Rückstreuung aufgrund der geringen Rückstreuintensität gute Ergebnisse bei der Erkennung von schwimmender Vegetation zu erzielen. Die Berechnung von K0 und K4 führt zu intensiven Elementen. Die von K2 und K5 zu mäßigen. Außerdem wird noch K3 der den Realen und K7, der den imaginären Teil widerspiegelt, berechnet.

$$K_0 = \frac{1}{2} \{ |S_{HH}|^2 + |S_{VV}|^2 \}$$

Formel 2.2 K0 = Gesamtintensität von HH und VV

$$K_3 = -R_e \{ S_{HH} S_{VV} \}$$

Formel 2.3 Intensitätsverhältnis zwischen Doppelreflexion und Oberflächenstreuung

$$K_4 = \frac{1}{2} \{ |S_{HH}|^2 - |S_{VV}|^2 \}$$

Formel 2.4 Verhältnis zwischen HH- und VV-Intensität

$$K_7 = lm \{ S_{HH} S_{VV} \}$$

Formel 2.5 Phasenverschiebung zwischen Doppelreflexion und Oberflächenstreuung

Die vier normalisierten Kennaugh- Elemente zerlegen das Rückstreuungssignal in die Gesamtintensität von HH und VV (K0). Außerdem wird das Intensitätsverhältnis zwischen Doppelreflexion und Oberflächenintensität (K3), das Verhältnis zwischen HH- und VV-Intensität (K4) und die Phasenverschiebung zwischen Doppelreflexion und Oberflächenstreuung (K7) beschrieben. Für eine effiziente Archivierung oder einen effizienten Datenaustausch werden im Allgemeinen ganzzahlige Werte wegen der geringeren Dateigröße bevorzugt. Die vorprozessierten Daten wurden in 8-Bit-Bilder gespeichert. Dieses Format wird in der Praxis neben den 16-Bit-Bildern am häufigsten eingesetzt.(Schmitt et al. 2015)

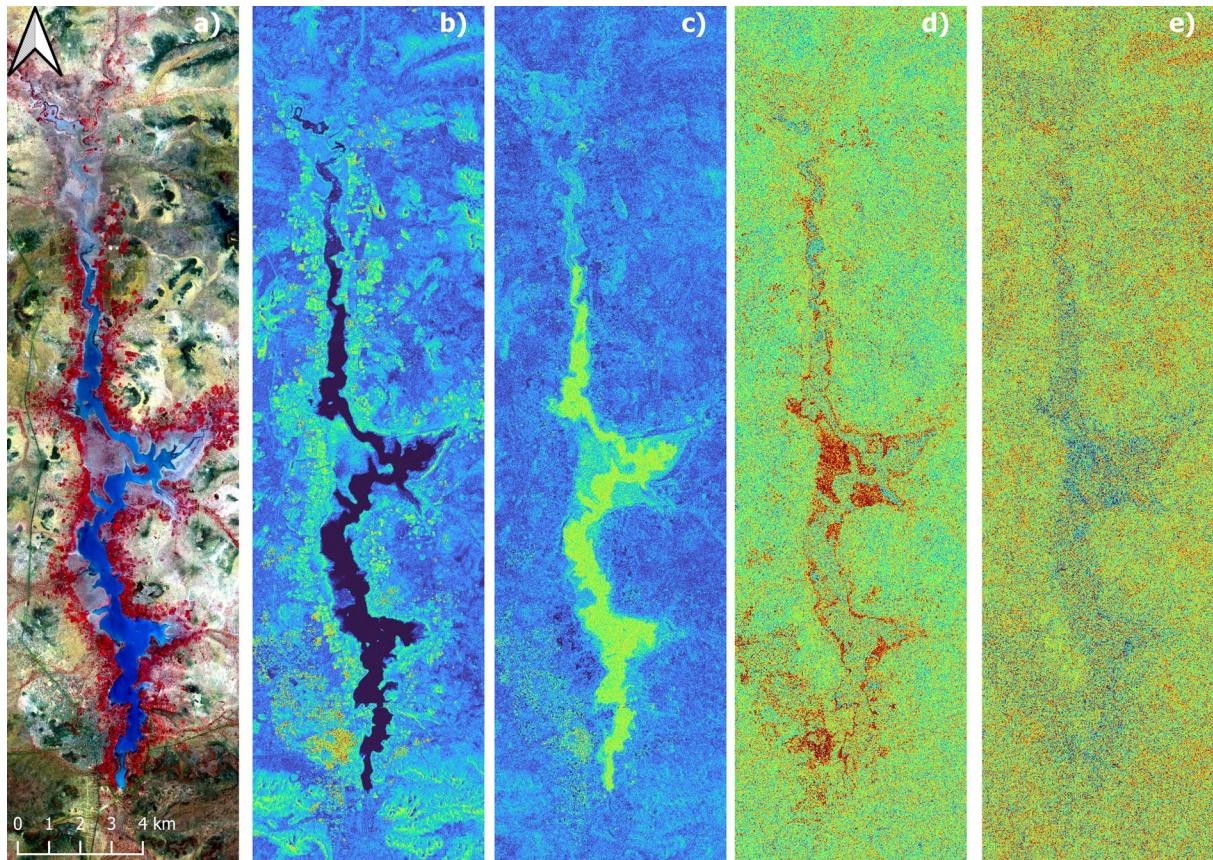


Abbildung 2.2 a) Sentinel 2 Falschfarbenbild (band 5-3-2 / NIR-rot-grün) Referenzbild vom 10.04.2019 (©ESA (2019)) b-e stellen die vier Kennaugh-Elemente aus dem Dual-Pol-TSX-Bild vom 11.01.2021 (©DLR (2021)) dar. b) K_0 = die Gesamtintensität als Summe von HH-plus VV-Intensität; c) K_3 (Differenz Doppelreflexion minus Oberflächenstreuung); d) K_4 (Differenz HH minus VV-Intensität); e) K_7 (Phasenverschiebung zwischen Doppelreflexion und Oberflächenstreuung)

2.2.2 DIE KOMPLEXE BASIS

Durch die Erweiterung der Komplexen Basis erhält man die auf die Kennaugh -Elemente angewandten Hyperkomplexen Basen. Im Folgenden wird auf die Berechnung der Basen höherer Ordnung eingegangen.

Die Transformation kann über eine lineare Transformation eines Vektors mithilfe zweier Elemente ausgedrückt werden. Dabei spannen beide Spalten der Transformationsmatrix eine Basis der komplexen Ebene C auf.

$$\vec{K_2} = C * R_2 * C = \frac{1}{2} * \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} * R_2 * \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Formel 2.6 Die Komplexe Basis von 2,5m-Pixel zu 5m-Pixel

Die Transformationsmatrix C, transformiert die linearen Vektoren der Reflexionsmaßen mit $r_i \geq 0$ in Kennaugh -Elemente. Wegen der Orthogonalität beschränkt sich die Transformation auf eine Rotation und Inversion der einfallenden Achsen.

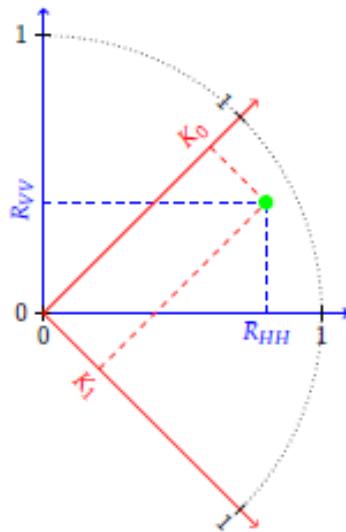


Abbildung 2.3 Die orthogonale Transformation von Reflexionsgrad (blau) zu Kennaugh-Elementen (rot) im Zweidimensionalen Raum einer Messung (grün) gemäß Formel 2.6

Um Korrelation zu vermeiden, bleiben die Skala wie auch die Winkel zwischen den Kanälen gleich. (Schmitt et al. 2020)

2.2.3 DIE RÄUMLICH-HYPERKOMPLEXE BASIS

Die hyperkomplexen Basen bezeichnen ein Set aus orthogonalen Transformationsmatrizen, die von Erweiterungen der Quaternionen-(4 Kanäle), Octonion (8 Kanäle) und Sedenion – Basis (16 Kanäle) abgeleitet werden.

$$\vec{K}_4 = Q * \vec{R}_4 = \frac{1}{4} \begin{bmatrix} C & C \\ C & -C \end{bmatrix} * \vec{R}_4 * \begin{bmatrix} C & C \\ C & -C \end{bmatrix} = \frac{1}{4} * \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} * \vec{R}_4 * \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Formel 2.7 **Quaternion Basis von 2,5m-Pixel zu 10m-Pixel**

$$\vec{K}_l = Q * \vec{P}_l * Q$$

Formel 2.8 **P als Patch (Bildausschnitt), die linke Multiplikation überführt die Zeilen, die rechte Multiplikation die Spalten auf räumlich-hyperkomplexe Basen.**

Die Erweiterten Basen wurden in der Arbeit “*Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases*” (Schmitt et al. 2020) bereits genauer untersucht. Die Überprüfung der polarimetrischen Informationsgehaltes nach der Bildfusion hat gezeigt, dass die polarimetrische Information der Eingangsbilder erhalten bleibt oder sogar ergänzt wird.

Aufgrund der konsistenten Überbrückung von Vektorräumen höherer Ordnung durch einfache Substitution der Eingabeelemente ist die Erweiterung theoretisch nicht eingeschränkt. Setzt man bei einer geringfügigen niedrigen Anzahl von unterschiedlichen Eingangskanälen den jeweiligen Kanal auf null, ist ein sehr flexibler, aber völlig konsistenter Merkmalsraum für die Fusion beliebiger Reflexions- und GIS-Layer gegeben. So können mehrschichtige Datensätze leicht in

Kennaugh-ähnliche Elemente transformiert werden, indem man die komplexe Ebene (Formel 2.6) mit sogenannten hyperkomplexen Basen erweitert. (Schmitt et al. 2020)

Formel 2.8 stellt die Berechnung der räumlich-hyperkomplexen Basen allgemein dar. I muss dabei als zweier Potenz gegeben sein. So lässt sich die Formel, wie in dieser Arbeit schrittweise erweitern, bis eine Pixelgröße von 80m erreicht wird.

Tabelle 2.2 Übersicht der mit räumlichen-hyperkomplexen Basen erweiterten TerraSAR-X Daten. TSX_0 zeigt die Originalaufnahme aus dem MultiSAR-Prozessor mit den vier verfügbaren Kennaugh-Elementen. Pro Stufe werden räumliche Kennaugh- Elemente hinzugefügt.

Pixelstufe	Pixelgröße	Kanäle
TSX_0	2,5m	4
TSX_1	5m	16
TSX_2	10m	64
TSX_3	20m	256
TSX_4	40m	1064
TSX_5	80m	4256

In der folgenden Tabelle 2.3 werden die Pixelstufen des MultiLooking vorgestellt. Die Auffälligkeit gegenüber den räumlichen-hyperkomplexen Basen ist, dass diese auf vier Kanäle beschränkt sind.

Tabelle 2.3 Pixelstufen MultiLooking

Pixelstufe	Pixelgröße	Kanäle
TSX_0	2,5m	4
TSX_1	5m	4
TSX_2	10m	4
TSX_3	20m	4
TSX_4	40m	4
TSX_5	80m	4

2.3 SPEKTRALDATEN

Als Sekundärquellen wurde unter anderem eine Aufnahme der Satelliten Sentinel 2 und Planet Scope verwendet. Hierbei handelt es sich um Satelliten, die mit optischen Sensoren die Erdoberfläche abtasten. Im folgenden Kapitel werden die Satelliten kurz beschrieben.

2.3.1 SENTINEL 2

Die beiden identischen Fernerkundungssatelliten Sentinel-2a und Sentinel 2b und deren dazugehörigen Instrumente wurden von der Europäischen Weltraumorganisation ESA für das Erdbeobachtungsprogramm Copernicus entwickelt. Das Programm soll eine moderne und leistungsfähige Infrastruktur für die Erdbeobachtung und Dienstleistungen der Geoinformation liefern. So können wertvolle Daten für Klimawandel, Sicherheit, Landüberwachung sowie Krisenmanagement gesammelt werden. Die Daten sind frei über das Portal „Copernicus Open Access Hub“ (<https://scihub.copernicus.eu/dhus/#/home> 2022) erhältlich.

2.3.2 PLANET

Planet bietet täglich mit 200 in der Umlaufbahn befindenden Satelliten einen hochauflösenden, kontinuierlichen und vollständigen Blick auf die Welt. So können Gebiete von geografischem Interesse beobachtet, Muster entdeckt, Infrastrukturen verfolgt und Veränderungen schnell erkannt werden.

Für diese Arbeit wurde eine Aufnahme eines PlanetScope-Satelliten als Referenz verwendet. Dabei handelt es sich um eine Satellitenkonstellation aus über 180 Satelliten, die Doves genannt werden. Diese kleinen CubeSats sind nur 10cm * 10 cm *30cm groß.

Die PlanetScope-Daten dienen, wie die der Sentinel-2 Aufnahmen, als hervorragende Quelle für die Vegetationsüberwachung. So können die Aufnahmen von den Sentinel-2-Daten durch die hohe räumliche Auflösung und die gute zeitliche Abdeckung ergänzt werden. Die schnelle zeitliche Abdeckung erhöht die Chancen, eine wolkenfreie Aufnahme in bewölkten Gebieten zu erhalten.

Die Daten sind gegen Entgelt zu erhalten und können über das Portal von Planeten heruntergeladen werden. Die Daten wurden für diese Arbeit kostenfrei zu Verfügung gestellt.(Nutzung für Lehre kostenlos) . (Satellite Imagery Analytics | Planet 2022)

2.4 WETTERDATEN

Als weitere Referenz für das zu untersuchende Gebiet dienten die Wetterdaten. Mithilfe der Niederschlagsdatensätze wurde die Eingrenzung der Regen- und Trockenzeit vereinfacht.

Der Deutsche Wetterdienst stellt die Daten des *Weltzentrum für Niederschlagsklimatologie* (WZN) für globale Niederschlagsanalysen und der Klimaüberwachung, sowie der Klimaforschung zur Verfügung. Das WZN ist der deutsche Beitrag zum Weltklimaforschungsprogramm und dem globalen Klimabeobachtungssystem. Durch den freien Zugang wurden diese Daten als Sekundärquelle hinzugezogen.

Verwendet wurden die First Guess Monthly Datensätze des *Global Precipitation Climatology Centre*. Die Daten geben eine erste Abschätzung der globalen Verteilung der Niederschlagssummen und -anomalien eines Monats an. Diese werden durch eine rein automatische Analyse mithilfe der über 7000 in Echtzeit verfügbaren Stationen erstellt. (*Wetter und Klima - Deutscher Wetterdienst - Our services - Global Precipitation Climatology Centre (GPCC)* 2022).

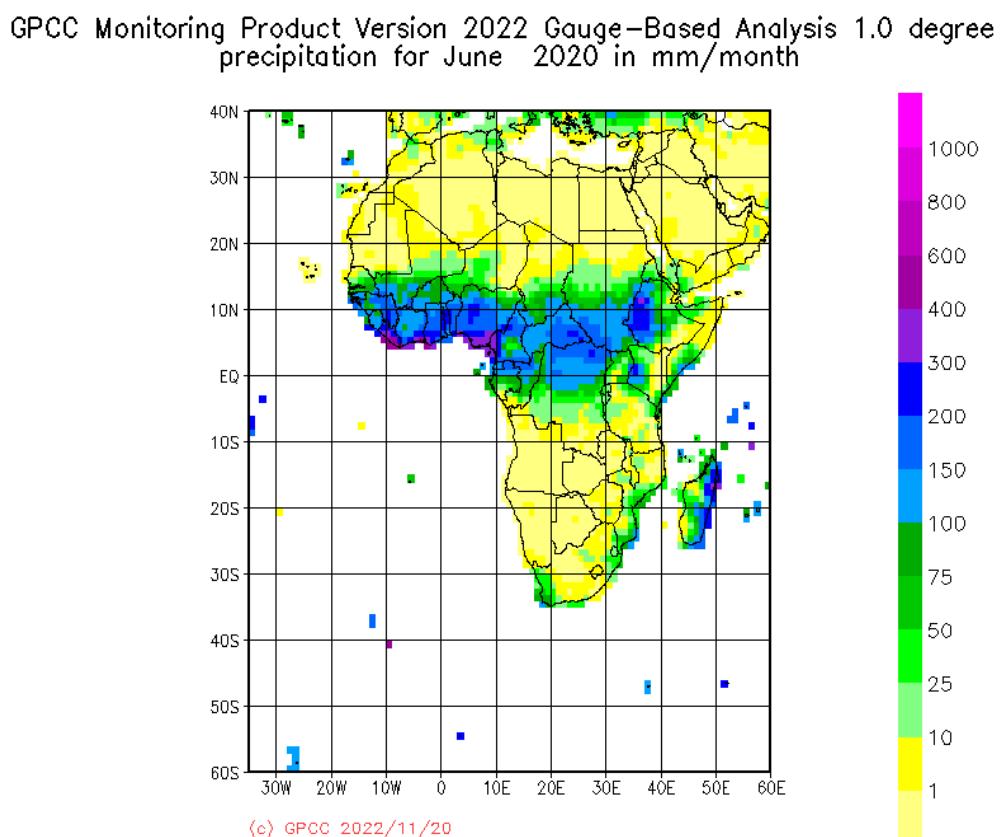


Abbildung 2.4 First Guess Monthly Durchschnittlicher Niederschlagsinformation in mm von Juni 2020 in Afrika (Quelle: <https://kunden.dwd.de/GPCC/Visualizer>)

3. VERWENDETE SOFTWARE

QGIS

QGIS ist ein frei zugängliches, benutzerfreundliches geographisches Informationssystem. Das Programm steht unter der General-Public-License und wurde für verschiedenste Betriebssysteme entwickelt. QGIS wird von mehreren freiwilligen Entwicklern bereitgestellt, die das System immer weiterentwickeln. Finanziert wird das Programm durch Spenden oder auch Förderungen. Das Programm ist vor allem für die Erstellung und Bearbeitung von Raster- und Vektordaten bekannt. Diese können analysiert und schlussendlich als visualisiertes Kartenmaterial ausgegeben werden. Eine weitere nützliche Funktion ist die Einbindung verschiedenster Plugins.

Für diese Arbeit wurde QGIS größtenteils für die Aufbereitung der Referenzdaten verwendet. Das Programm ermöglichte die Visualisierung der berechneten, mit räumlich-hyperkomplexen-Basen erweiterten, Kennaugh- Elemente in ihren verschiedensten Fusionsstufen, sowie die Bearbeitung der bereitgestellten Vektor Referenzdaten. Diese konnten dank QGIS zu jedem Aufnahmezeitpunkt in eine Vektordatei zusammengefasst werden.

Matlab

Matlab ist eine Plattform für Programmierung und numerische Berechnungen, die zur Analyse von Daten, Entwicklung von Algorithmen und das Erstellen von Modellen verwendet wird. Das Programm kombiniert iterative Analysen und Entwicklungsprozesse mit einer Programmiersprache, in der Matrizen-basierte Mathematik formuliert werden kann. Matlab kann durch verschiedene Toolboxes erweitert werden, die die Analysen, die Bildverarbeitung und das maschinelle Lernen vereinfachen. Außerdem können die Ergebnisse durch Plots bildlich veranschaulicht werden.

Für diese Arbeit wurde Matlab für die Erweiterung von Kennaugh Elemente zu räumlichen-hyperkomplexe Basen genutzt.

Excel

Excel ist ein Tabellenkalkulationsprogramm, das von Microsoft kostenpflichtig angeboten wird. Mithilfe Excel können Zahlen und Texte in tabellarischer Form dargestellt und ausgewertet werden. Excel wurde in dieser Arbeit hauptsächlich für das Erstellen von Diagrammen verwendet. So konnten die aus den Klassifizierungen berechneten Ergebnisse gut visualisiert und verglichen werden.

Python

Python ist eine universelle und objektorientierte Programmiersprache, die aufgrund ihrer einfachen Syntax anwendungsfreundlich ist. Die Anwendung kann durch sogenannte Bibliotheken erweitert werden. Mithilfe des Anaconda Navigator kann Anaconda auf einer grafischen Oberfläche auf den verschiedensten Betriebssystemen ausgeführt werden. Damit können Umgebungen und Bibliotheken in Anaconda einfach verwaltet und eingepflegt werden. Außerdem ist es möglich, die Webanwendung Jupyter Notebook über Anaconda zu starten. Jupyter Notebook ist eine open-Source Software die es ermöglicht, die für diese Arbeit notwendige Programmierung in einem Skript

auszuführen. Dabei wurden verschiedenste Bibliotheken verwendet, die im Folgenden erklärt werden.

Verwendete Bibliotheken:

- Geospatial Data Abstraction Library (GDAL): GDAL ist eine Übersetzungsbibliothek für Raster- und Vektor Geodatenformate. Die Bibliothek wird von der Open Source Geospatial Foundation zur Verfügung gestellt
- Scikit-Learn: Scikit-learn ist eine freie Software-Bibliothek für maschinelles Lernen. Die Bibliothek bietet verschiedene Klassifikations- Algorithmen an, wie z.B. in der Arbeit verwendete Random Forest oder k-Nearest-Neighbor.
- Pandas: Pandas dient für die Verarbeitung, Analyse und Darstellung von Daten. Pandas enthält Datenstrukturen und Operatoren für den Zugriff auf numerische Tabellen.
- Numpy: Mithilfe Numpy können Vektoren, Matrizen sowie große mehrdimensionale Arrays gehandhabt werden.

4. METHODIK

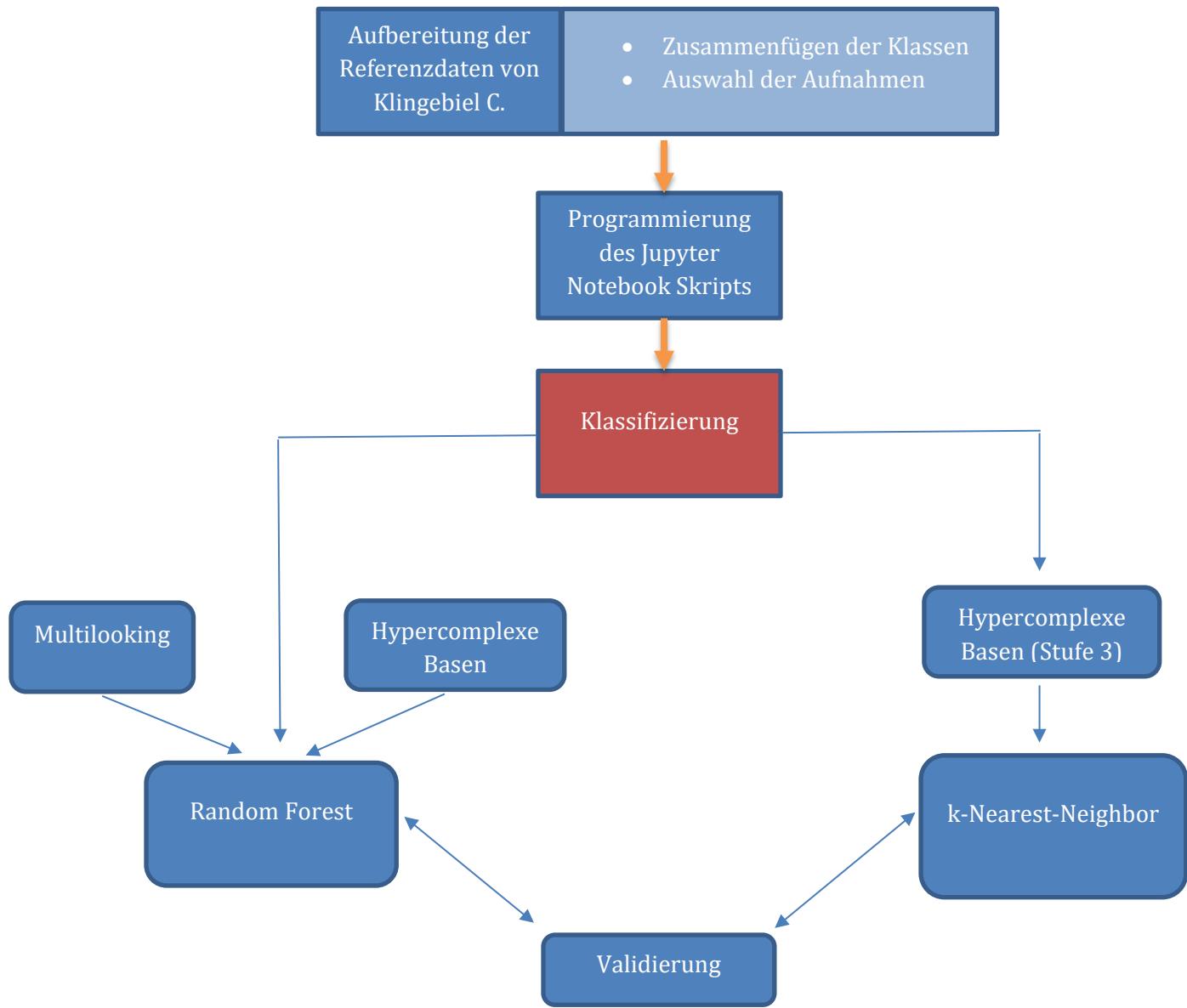


Abbildung 3.1 Workflow. Vorgehensweise von der Datenaufbereitung zur Validierung

4.1 AUFBEREITEN DER REFERENZDATEN

Bevor der Random Forest Algorithmus trainiert werden konnte, mussten die zur Verfügung gestellten Referenzdaten aufbereitet werden. Die Referenzdaten waren im Koordinatenbezugssystem EPSG:32630 - WGS 84 / UTM Zone 30N gegeben. Die Ergebnisse aus der Klassifizierung von Klingebiel C. (Klingebiel et al. 2019) waren nach Datum und den Klassen *schwimmende Vegetation, offenes Wasser, Feuchtgebiete und bewässerte Felder* sortiert. Die gegebenen Shapefiles, konnten in QGIS geöffnet und bearbeitet werden. In QGIS war es möglich, die Klassen jedes Aufnahmedatums zu einem Shapefile zusammenzufügen. Dabei wurde den Attributtabellen jeden Aufnahmedatums die Spalten „ID“ und „klasse“ hinzugefügt. Dies war für die automatische Klassifizierung des TSX-Datensatz notwendig. Durch die große Anzahl an Referenzdaten und dem Klassenungleichgewicht, wurden aus den 119 Aufnahmen eine Auswahl an Daten für die weitere Klassifikation gewählt. Hierbei wurde darauf geachtet, dass jede Klasse gleich oft vertreten ist.(siehe Anhang)

4.1.1 ERSTELLEN DER KLASSE LAND

Für die Klassifizierung war noch eine fünfte Klasse Land notwendig. Anhand dieser Klasse konnten bei der Klassifizierung die Flächen gefiltert werden, die nicht zu den Feuchtgebieten gezählt werden. Dafür wurden die Spektralen Referenzdaten und die Shapefiles in QGIS hinterlegt. So konnte die neue Klasse mit Shapefiles-Polygonen zu den ausgewählten Referenzdaten hinzugefügt werden. Auch hier wurde auf die Gesamtverteilung der Klassen geachtet.

4.1.2 ANPASSUNG DES TRAININGSGEBIETES

Die TerraSAR-X Aufnahmen wurden in QGIS geöffnet. Dabei wurden im Westen und im Osten „no-Data“-Gebiete festgestellt. Diese wurden bei der Betrachtung durch einen schwarzen Balken ersichtlich. Da es durch diese Gebiete bei der späteren Klassifikation zu Fehlklassifizierungen kommen kann, wurden diese mithilfe eines Clip Batch Prozess von allen bereit gestellten Aufnahmen in jeder Dimension entfernt.

Die Trainingsgebiete wurden mit Shapefiles nachträglich angepasst. Dadurch konnte ein „Mask-Layer“ über das Gebiet gelegt werden. Dieser Layer betrachtete nur das Gebiet um den See. Durch die Eingrenzung konnten Fehlklassifikationen vermieden werden. Auf dieser wird im Kapitel Ergebnisse noch genauer eingegangen.

4.2 POLARIMETRISCH-RÄUMLICHE SIGNATUREN

Auffälligkeiten der verschiedenen Klassen in Ihren Pixel Signaturen, wurde mithilfe des QGIS-Tool Temporal/Spectral Profile überprüft. Dabei wurden einzelne Pixel der Klassen *bewässerte Felder, offenes Wasser, schwimmende Vegetation und Feuchtgebiete* ausgewählt. Auch in dem Urbanen Gebiet von Kongoussi wurden die Pixel auf Ihre Digitale Signatur überprüft.

4.3 KLASIFIZIERUNG

Nachdem der Trainingsdatensatz ausgewählt wurde, wurden die dazugehörigen Aufnahmen in ihren jeweiligen Pixelstufen in unterschiedliche Verzeichnisse sortiert und abgespeichert. Für die Trainings-Vektordaten wurde ein separates Verzeichnis angelegt. Das war für den späteren Zugriff mithilfe Python wichtig. Hier wurden die Raster- mit den Vektordaten miteinander verschnitten. Python konnte durch den sortierten Datensatz jede Aufnahme einzeln abarbeiten. Dieser Vorgang musste für jede Pixelstufe analog durchgeführt werden. Die daraus erhaltenen Arrays wurden für die weitere Bearbeitung in mehrere CSV-Dateien gespeichert. Nun konnten die gespeicherten CSV-Daten in einem neuen Python Skript geöffnet und bearbeitet werden. Ziel ist es, eine vollautomatische Klassifizierung aller TSX-Daten in den unterschiedlichen Pixelstufen durchzuführen.

Der Datensatz wurde vor der eigentlichen Klassifikation in einem „Train und Test Datensatz“ geteilt. Hierbei erhielt der Trainingsdatensatz 70% und der Testdatensatz 30% des gesamten Datensatz. Die Aufteilung war für die späterer Validierung notwendig.

Um das restliche Klassenungleichgewicht zu beseitigen, wurde die Funktion „Oversampling“, die von „imbalanced-learn“ bereitgestellt wird, (2. Over-sampling — Version 0.10.0 2022) verwendet. Hierbei werden alle Klassen auf die Größe der größten Klasse angepasst. Durch dieses Verfahren gehen keine Informationen verloren. Die Funktion wurde nur auf den Trainingsdatensatz angewendet. Der Random Forest ist so konstruiert, dass er die Gesamtfehlerrate minimiert. Dadurch tendiert er dazu, sich mehr auf die Vorhersagegenauigkeit der Mehrheitsklasse zu konzentrieren. Dies führt oft zu einer schlechten Genauigkeit für die Minderheitsklasse.(Chao Chen 2004) „Oversampling“ ist eine Methode, die Klassenungleichgewicht ausgleicht.

Nach der Anpassung der Daten konnte der Random Forest Algorithmus trainiert werden. Ein Random Forest-Klassifikator ist ein Ensemble-Klassifikator, der mehrere Entscheidungsbäume erzeugt, indem er eine zufällig ausgewählte Teilmenge von Trainingsgebieten und Variablen benutzt. Er gehört zu den Verfahren der überwachten Klassifikation. In der Fernerkundung ist dieser Klassifikator aufgrund seiner Genauigkeit besonders beliebt. Auch der Umgang mit hoher Daten Dimensionalität und Multikollinearität ist ein Vorteil des Random Forests. (Breiman 2001)

Bei Random Forests Algorithmus wird jeder Baum des Ensembles aus einer mit Ersatz gezogenen Stichprobe (d. h. einer Bootstrap-Stichprobe) aus der Trainingsmenge erstellt. Darüber hinaus wird bei der Aufteilung jedes Knotens während der Erstellung eines Baumes die beste Aufteilung entweder aus allen eingegebenen Merkmalen oder einer zufälligen Teilmenge der Parametergröße „max_features“ gefunden.

Der Zweck dieser beiden Zufallsquellen (Bootstrap und „max_features“) besteht darin, die Varianz des Forest-Schätzers zu verringern. Einzelne Entscheidungsbäume weisen in der Regel eine hohe Varianz auf und neigen zu einer Überanpassung. Die in die Forests eingespeiste Zufälligkeit führt zu Entscheidungsbäumen mit leicht entkoppelten Vorhersagefehlern. Durch die Bildung eines Durchschnitts aus diesen Vorhersagen können sich einige Fehler ausgleichen. Zufallswälder erreichen eine geringere Varianz, indem sie verschiedene Bäume kombinieren, manchmal um den Preis einer geringfügigen Zunahme der Verzerrung. In der Praxis ist die Varianzreduzierung oft signifikant, was zu einem insgesamt besseren Modell führt. (Breiman 2001) Die Baumtiefe wurde für alle Pixelstufen (MultiLooking & räumliche-hyperkomplexe Basen) auf 30 gesetzt.

Die Klassifikation wurde analog mit dem Datensatz der erweiterten Hyperkomplexen Basen und dem Datensatz des MultiLooking durchgeführt. Bevor dem Algorithmus den Datensatz des MultiLooking übergeben werden konnte, wurden nur die ersten vier Kanäle der räumlichen-hyperkomplexen Basen betrachtet. Dies erfolgte mit einer Nachbearbeitung der CSV-Datei. Darauf musste auch bei der Klassifikation der neuen Aufnahmen geachtet werden.

Um die Modellauswahl dieser Arbeit zu begründen, wurde ein zweiter Algorithmus, der K-Nearest-Neighbour über scikit-learn implementiert. Die nachbarschaftsbasierte Klassifizierung ist eine Art der Instanz basierten Lernens oder nicht verallgemeinernden Lernens: Es wird nicht versucht, ein allgemeines internes Modell zu erstellen, sondern es werden einfach Instanzen der Trainingsdaten gespeichert. Die Klassifizierung wird anhand einer einfachen Mehrheitsabstimmung der nächsten Nachbarn jedes Punktes berechnet. Einem Abfragepunkt wird die Datenklasse zugewiesen, die die meisten Vertreter innerhalb der nächsten Nachbarn des Punktes hat. (Cover und Hart 1967) Die Anzahl der Nachbarn wurde in dieser Arbeit auf 10 gesetzt. Der K-Nearest Neighbor wurde ausschließlich mit der Pixelstufe 20m der erweiterten-räumlichen Basen trainiert.

4.4 VALIDIERUNG

Nachdem der Algorithmus den Datensatz trainiert hat, konnte die Validierung mithilfe des Trainingsdatensatzes durchgeführt werden. Hier wurde der Datensatz mithilfe zweier Bewertungsmaßzahlen validiert. Durch die erlangten Ergebnisse der Validation, kann überprüft werden, ob HCP gegenüber dem Standardmultilooking eine Genauigkeitssteigerung erbracht hat.

Die erste Metrik, die für den Vergleich verwendet wurde, war die Berechnung des F1-Scores. Der F1-Score ist eine der wichtigsten Bewertungsmaßzahlen im Maschinellen Lernen. Er fasst die Vorhersageleistung eines Modells zusammen, indem er die Metriken, Precision und Recall kombiniert. Die Präzision misst das Ausmaß der durch Falsch-Positiv verursachten Fehler, während Recall das Ausmaß der durch Falsch-Negativ verursachten Fehler misst.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Formel 4.1 Precision

$$\text{Recall} = \frac{TP}{TP + FN}$$

Formel 4.2 Recall

Aus Precision und Recall lässt sich der F1-Score berechnen.

$$F1 - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Formel 4.3 F1-Score

Die Range des F1-Scores geht von 0 bis 1. Je näher er an 1 ist, desto besser ist das Modell. (Sokolova et al. 2006)(siehe Abbildung 5.9)

Mithilfe der berechneten Genauigkeit des Klassifizierungsalgorithmus ist es möglich, eine Aussage über die Güte des verwendeten Modells zu geben. Die Genauigkeit ist die Anzahl der korrekt vorhergesagten Datenpunkten unter allen Datenpunkten.

$$\text{Accuracy} = \frac{\text{correct classification}}{\text{all classifications}}$$

Formel 4.4 Accuracy

In dieser Arbeit werden alle fünf Klassen (*bewässerte Felder, offenes Wasser, schwimmende Vegetation, Feuchtgebiete und Land*) in die Genauigkeit mit einberechnet, um so die Gesamtgenauigkeit zu erhalten.

4.5 VISUALISIEREN DER KLASSEKTIIONSERGEBNISSE

Bevor die Klassifikationsergebnisse als Karten ausgegeben werden konnten, wurden die Raster-Dateien mithilfe des Verarbeitungswerkzeugs *SAGA Majority Filter in QGIS* bereinigt. Der Filter ersetzt Zellen in einem Raster auf Basis der Mehrheit der zusammenhängenden Nachbarzellen. Dadurch können einzelne Fehlklassifikationen ausgeschlossen werden.

Die bereinigten Daten konnten danach zu Vektordateien umgewandelt (Vektorisieren von Raster zu Vektor) und als Shapefiles abgespeichert werden. Anschließend wurden die Kanten geglättet. Dieser

Algorithmus glättet die Geometrien in dem Polygonlayer. Ein neuer Layer mit den gleichen Objekten wie der Eingabelayer wird erzeugt, jedoch mit geglätteten Geometrien mit höherer Anzahl von Stützpunkten und Ecken in den Geometrien.

Für jeder der zwölf klassifizierten Aufnahmen wurde eine Karte erstellt. Hierbei wurden jeweils drei Aufnahmen nebeneinander dargestellt. Bei der Erstellung der Karten wurde jeder der vier Klassen eine unterschiedliche Farbe zugewiesen. Es wurde auch darauf geachtet, dass jede Karte den gleichen Kartenausschnitt und denselben Maßstab erhält. Dadurch konnten die Ergebnisse nebeneinander gut veranschaulicht werden. (siehe Abbildung 5.10 - Abbildung 5.13)

4.6 BERECHNUNG DER FLÄCHE

Um die Dynamische Veränderung des Sees über den Zeitraum der klassifizierten Aufnahmen besser veranschaulichen zu können, wurde eine Flächenberechnung der Klassen *bewässerte Felder, offenes Wasser, schwimmende Vegetation und Feuchtgebiete* erstellt. Die Flächenberechnung erfolgte in QGIS. Dort konnte die Fläche der Ausgewählten Polygone mithilfe des für die Verschmelzung genutzten Werkzeug aus einer Tabelle abgelesen werden.

Anhand der Flächenberechnung konnten die Ergebnisse der zwölf klassifizierten Aufnahmen in Abhängigkeit des monatlichen Niederschlages in einem Diagramm dargestellt werden. Es wurde für jede Aufnahme die Gesamtfläche der vier klassifizierten Klassen zum Aufnahmepunkt abgebildet. Zusätzlich wurde die Niederschlagsmenge in mm pro Monat, als Sekundärachse in dem Diagramm hinzugefügt.

5. ERGEBNISSE

In diesem Kapitel werden die Pixel Signaturen, Klassifikationsergebnisse, die aus dem Random Forest Algorithmus und dem des K-Nearest-Neighbor präsentiert. Dabei werden die Validierungsergebnisse der beiden Algorithmen gezeigt. Vor allem der Vergleich zwischen den Ergebnissen der Trainingsdatensätze mit räumliche-hyperkomplexen Basen und MultiLooking werden mithilfe von Diagrammen veranschaulicht.

Außerdem werden die neu klassifizierten Aufnahmen aus den Jahren 2019, 2020 und 2021 mit den klassifizierten Klassen schwimmende Vegetation, offenes Wasser, bewässerte Felder und Feuchtgebiete als Karten dargestellt.

5.1 POLARIMETRISCH-RÄUMLICHE SIGNATUREN

Vor Beginn der Klassifizierung wurden die verschiedenen Aufnahmen in Ihren Pixel Signaturen untersucht. Die gezeigten Diagramme wurden Anhand der Pixelstufe 20m erstellt. Zu sehen sind auf der X-Achse die enthaltenen 256 Kanäle. Jeder Kanal enthält einen Digital Number (DN) Wert. Der Wert einer 8-Bit Aufnahme wird durch die Zahlen von 1– 255 beschrieben. Durch die DN eines Satellitenbildes wird der Pixel Wert in jedem Kanal ausgedrückt. (Schmitt et al. 2015) Die Digital Number wird in der Y-Achse dargestellt. Wegen seiner Auffälligkeiten wurde die Pixel Signatur, des Urbanen Gebietes Kongoussi einzeln betrachtet.

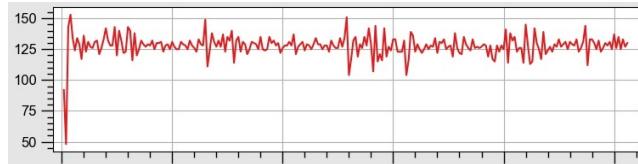


Abbildung 5.1 Pixel Signatur Bewässerte Felder (x-Achse Anzahl Kanäle / y-Achse Digital Number)

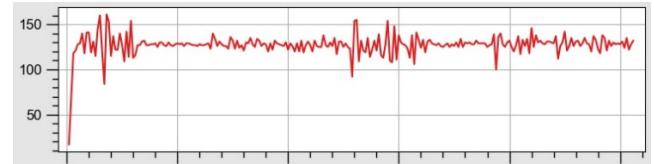


Abbildung 5.2 Pixel Signatur schwimmende Vegetation (x-Achse Anzahl Kanäle / y-Achse Digital Number)

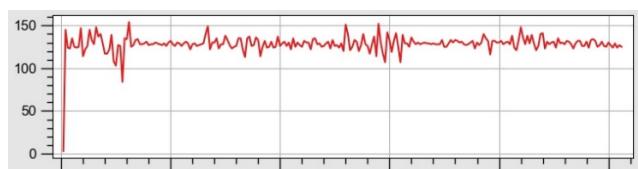


Abbildung 5.3 Pixel Signatur offenes Wasser (x-Achse Anzahl Kanäle / y-Achse Digital Number)



Abbildung 5.4 Pixel Signatur Feuchtgebiete (x-Achse Anzahl Kanäle / y-Achse Digital Number)

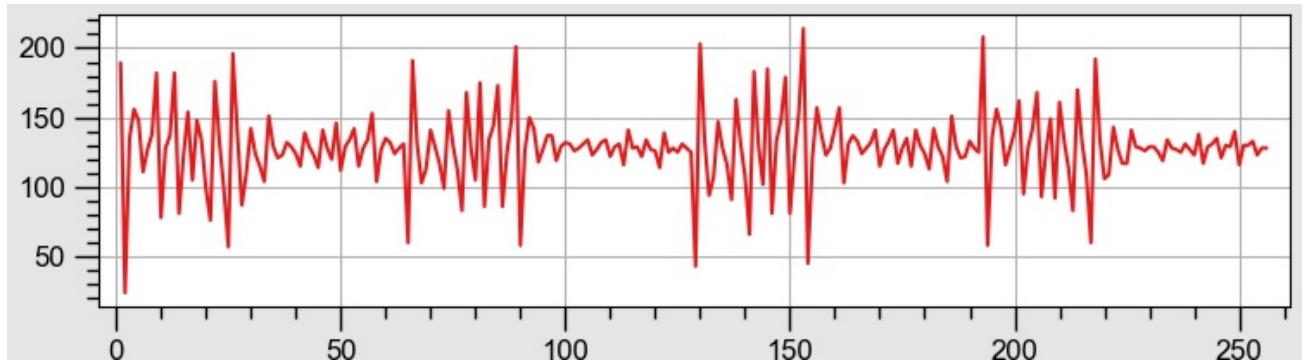


Abbildung 5.5 Pixel Signatur Urbanes Gebiet (x-Achse Anzahl Kanäle / y-Achse Digital Number)

5.2 ERGEBNISSE VALIDIERUNG

In den Unterkapiteln werden die Ergebnisse der Validierung vorgestellt.

5.2.1 ERGEBNISSE GESAMTGENAUIGKEIT

Im Folgenden werden die durch die Validierung berechneten Genauigkeiten der Hyperkomplexen Basen in ihren unterschiedlichen Pixelgrößen dargestellt. Das Diagramm zeigt, dass die Genauigkeit ab der Pixelstufe 2,5m von 61% bis zur Pixelgröße 20m ansteigt. Hier erzielte die Gesamtgenauigkeit den Besten Wert von 73%.

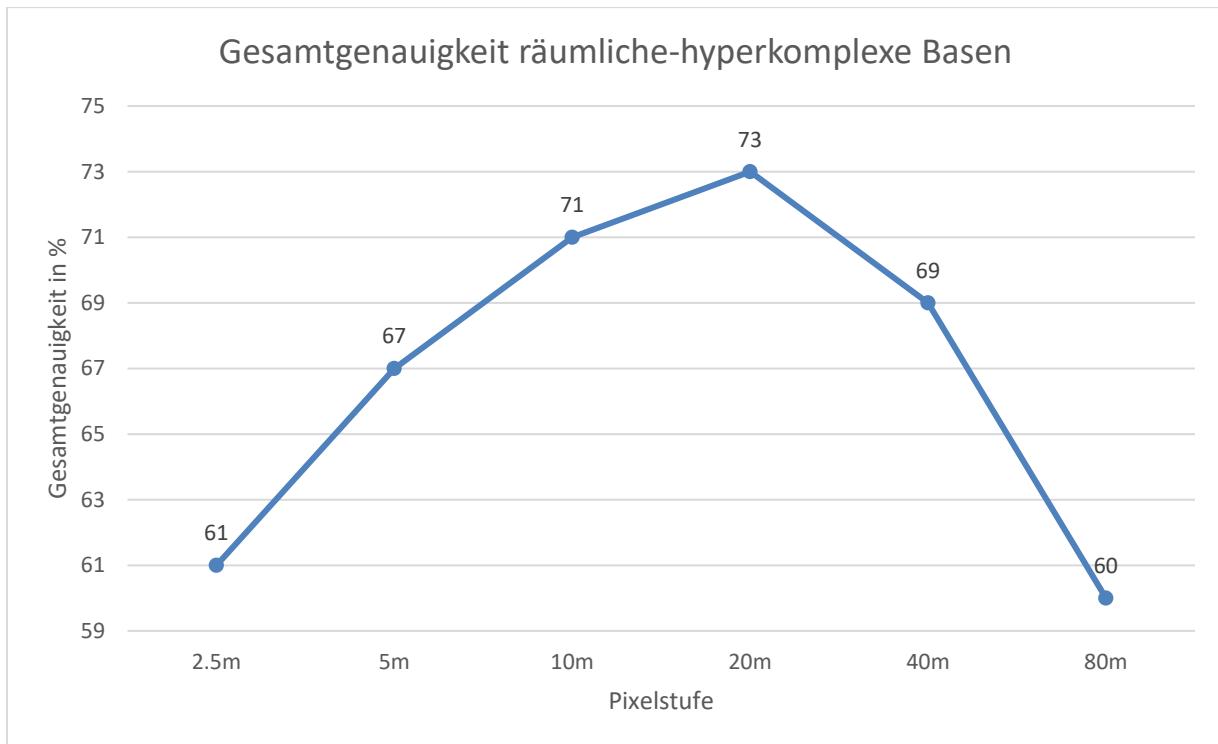


Abbildung 5.6 Darstellung der Gesamtgenauigkeit der Unterschiedlichen Pixelstufen. Die Klassifikation erfolgte mit dem durch räumlich-hyperkomplexe Basen erweiterten Datensatz.

Das zweite Diagramm (Abbildung 5.7) beschreibt die berechnete Gesamtgenauigkeit des mit den MultiLooking-Daten trainierten Random Forests. Hierbei wird ersichtlich, dass dieser die Genauigkeit mit dem höchsten Wert in der Pixelstufe 20m erzielt.

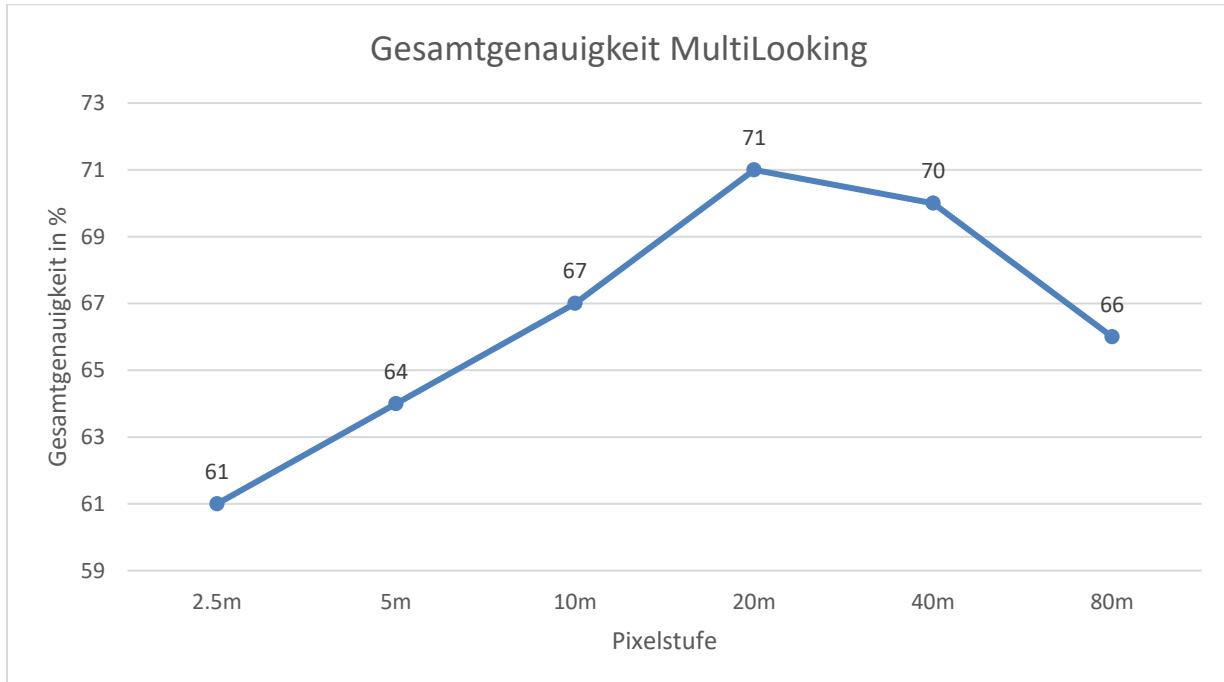


Abbildung 5.7 Darstellung der Gesamtgenauigkeit der Unterschiedlichen Pixelstufen. Die Klassifikation erfolgte mit dem MutliLooking trainierten Datensatz.

Um die trainierten Datensätze in einem direkten Vergleich gegenüberzustellen, wurden die Ergebnisse in ein gemeinsames Diagramm geplottet. Daraus wird erkenntlich, dass sich beide Datensätze nur minimal unterscheiden. Der mit Hyperkomplexen Basen erweiterte Datensatz hat im direkten Vergleich in den Stufen 5m, 10m, und 20m einen höheren Wert. Beide Datensätze haben in der Stufe 2,5m den gleichen Wert von 61%. Dies liegt daran, dass die erste Stufe in beiden Datensätzen vier Kanäle aufzeigt und somit noch nicht mit räumlichen-hyperkomplexen Basen erweitert wurde.

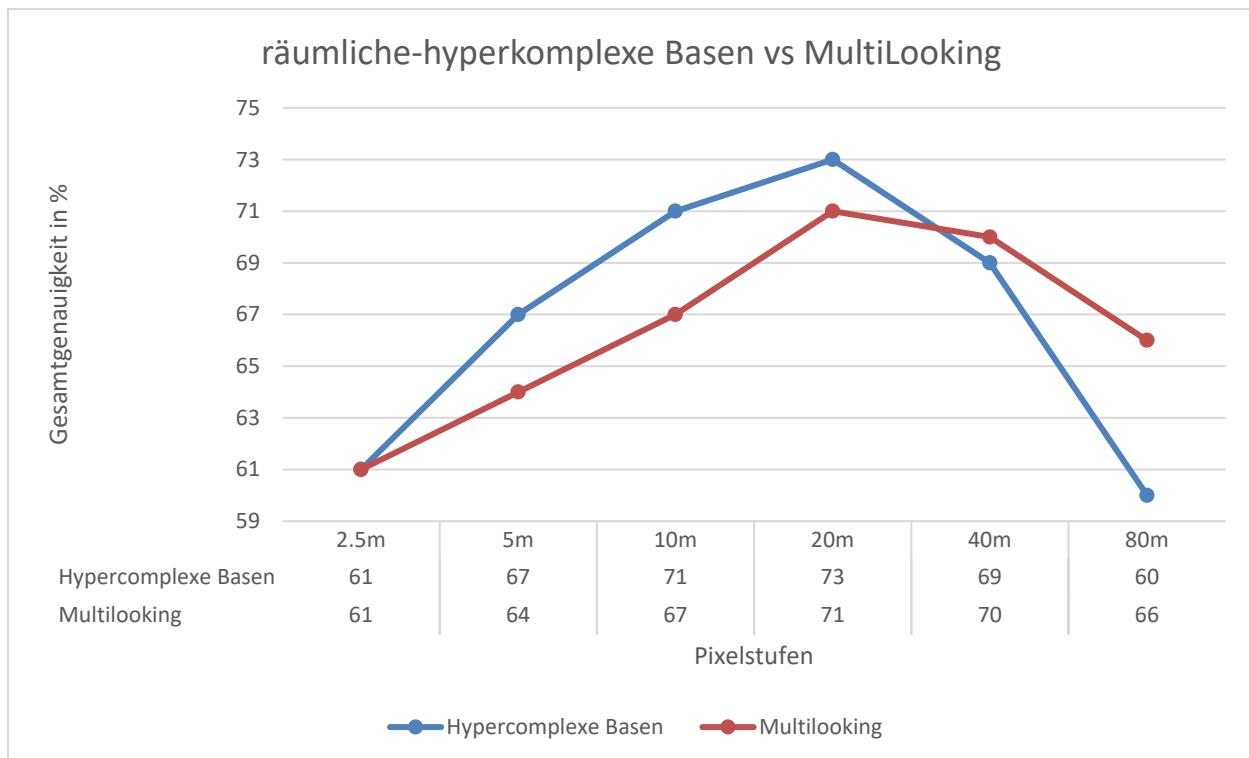


Abbildung 5.8 Darstellung der Gesamtgenauigkeit räumlich-hyperkomplexe Basen und MutliLooking.

5.2.2 ERGEBNIS F1-SCORE

Hier werden die Berechnungen des F1-Scores in einem Balkendiagramm gezeigt. Durch farbliche Kennung aller Pixelstufen und Trennung der *Klassen offenes Wasser, schwimmende Vegetation, bewässerte Felder und Feuchtgebiete*. Aus dem Diagramm wird ersichtlich, dass der Wert von *offenem Wasser* von 72%-91% erzielt. Der berechnete F1-Score der *bewässerten Felder* wurde zwischen 64% und 84% berechnet. Die *Feuchtgebiete* weisen einen Wert von 64%-68% auf. Bei der *schwimmenden Vegetation* wurden Werte im Bereich von 32%-43% erzielt. Bei den Klassen *schwimmende Vegetation, bewässerte Felder, Feuchtgebiete* wurde jeweils in der Pixelstufe 20m der höchste Wert erzielt. Die Werte bei der Klasse *offenes Wasser* zeigen in den ersten vier Stufen keine großen Abweichungen untereinander auf.

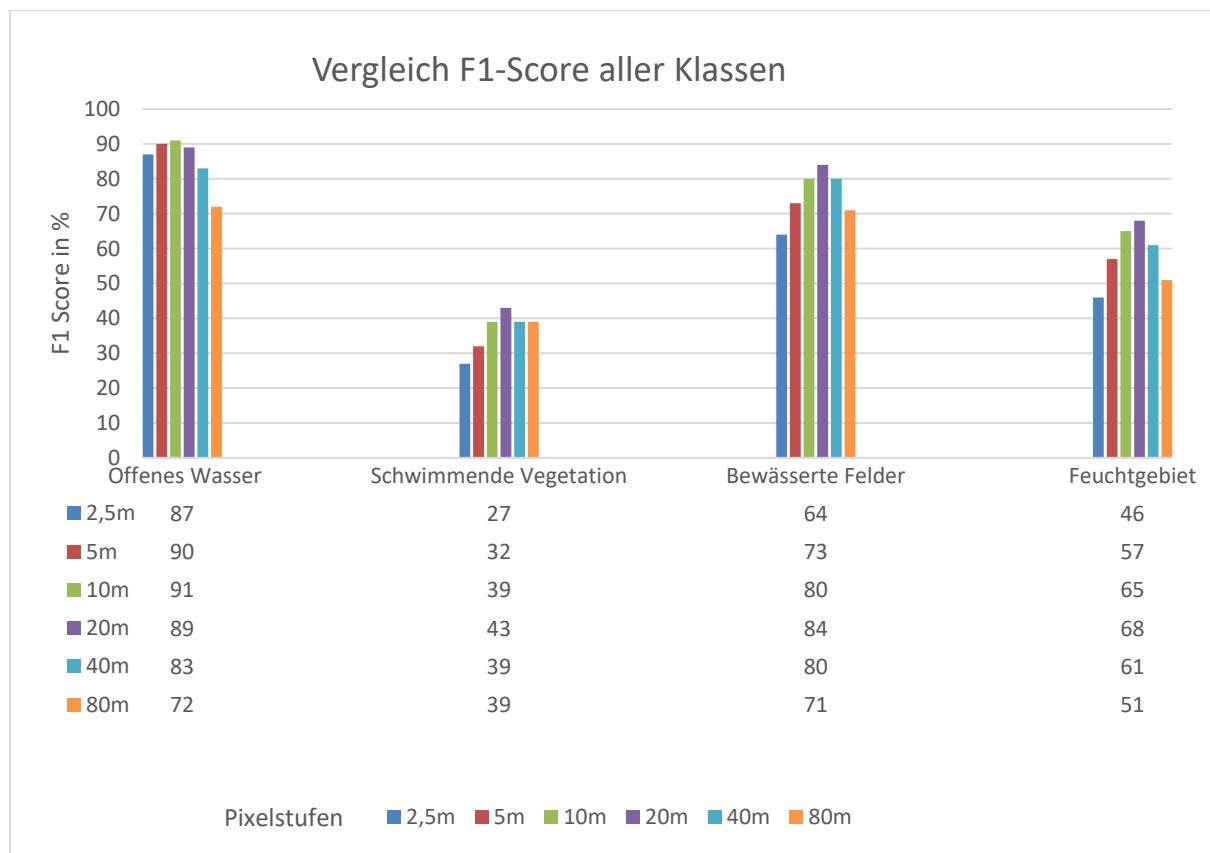


Abbildung 5.9 F1-Score aller Klassen und Pixelstufen. Ergebnis des F1 Score der räumlichen-hyperkomplexen Basen.

5.2.3 ERGEBNIS K-NEAREST-NEIGHBOR

In dieser Arbeit wurde das Hauptaugenmerk auf die Klassifikation mithilfe eines Random Forest Algorithmus gesetzt. Da es sich bei der Klassifikation mithilfe des K-Nearest-Neighbour um eine Modellüberprüfung handelt, wurde dieser Algorithmus mit einem Datensatz trainiert und klassifiziert. Der Algorithmus wurde mit dem Datensatz der Pixelstufe 20m (erweitert mit räumlichen-hyperkomplexen Basen) trainiert. Es wurde eine Gesamtgenauigkeit von 66% erzielt. In derselben Pixelstufe erzielte der Random Forest 73%.

Tabelle 5.1 F1-Score aller Klassen und die Gesamtgenauigkeit in der Pixelstufe 20m des K-Nearest-Neighbor Algorithmus

bewässerte Felder	80%
offenes Wasser	80%
schwimmende Vegetation	25%
Feuchtgebiete	60%
Gesamtgenauigkeit	66%

5.3 VISUALISIEREN DER KLASIFIZIERUNGSERGEBNIS

Um die dynamischen Änderungen des Lac Bam Sees bildlich zu veranschaulichen, wurde von jeder Aufnahme eine Karte erstellt. Die Karten wurden nach Aufnahmzeitraum sortiert. Der Zyklus fängt mit der Aufnahme vom 09.04.2019 an und endet am 02.02.2021.

Um die dynamische Entwicklung des Sees über den zu untersuchenden Zeitraum zu visualisieren, wurde von jeder Aufnahme eine Karte, die das Klassifizierungsergebnis zeigt, erstellt. Die Klassen wurden in unterschiedlichen Farben dargestellt. Diese sind der Legende zu entnehmen.

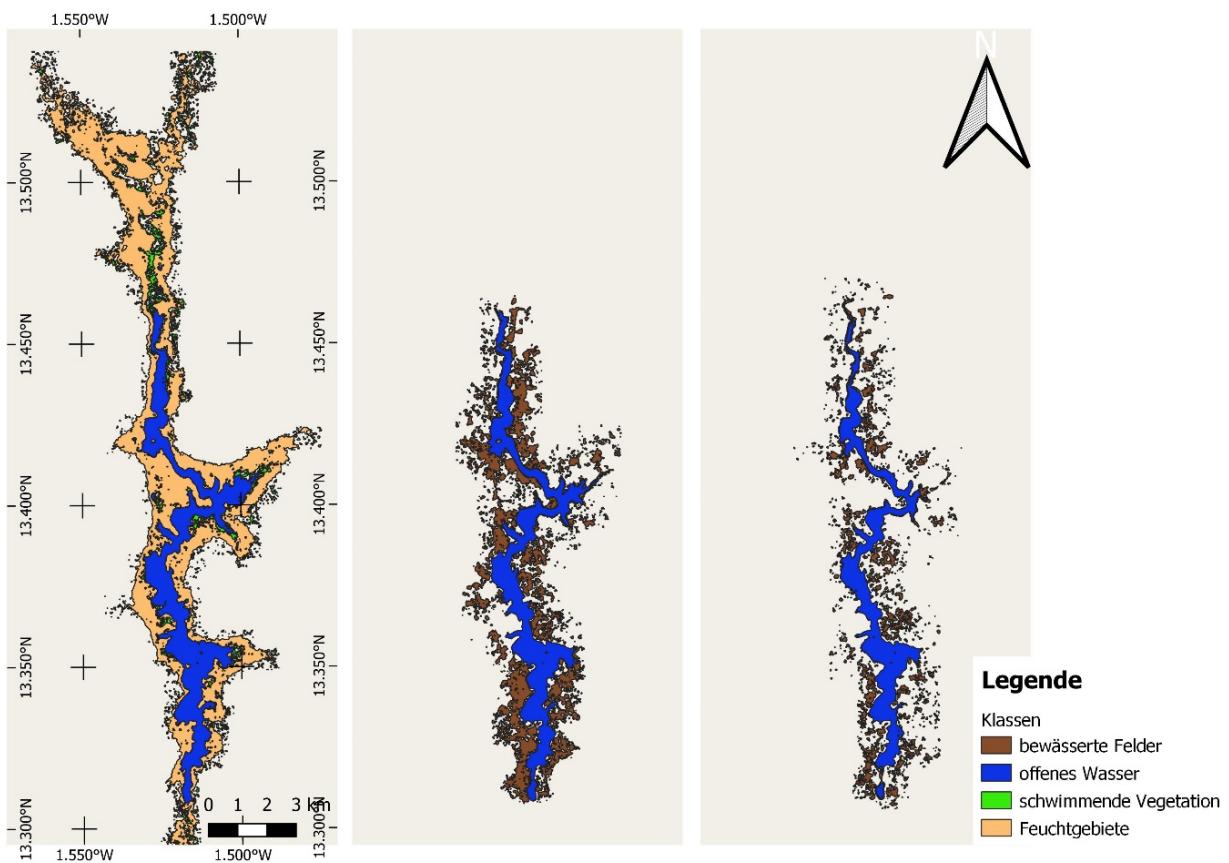


Abbildung 5.10 klassifizierte Aufnahmen von links nach rechts 04.09.2019; 22.04.2020; 25.05.2020

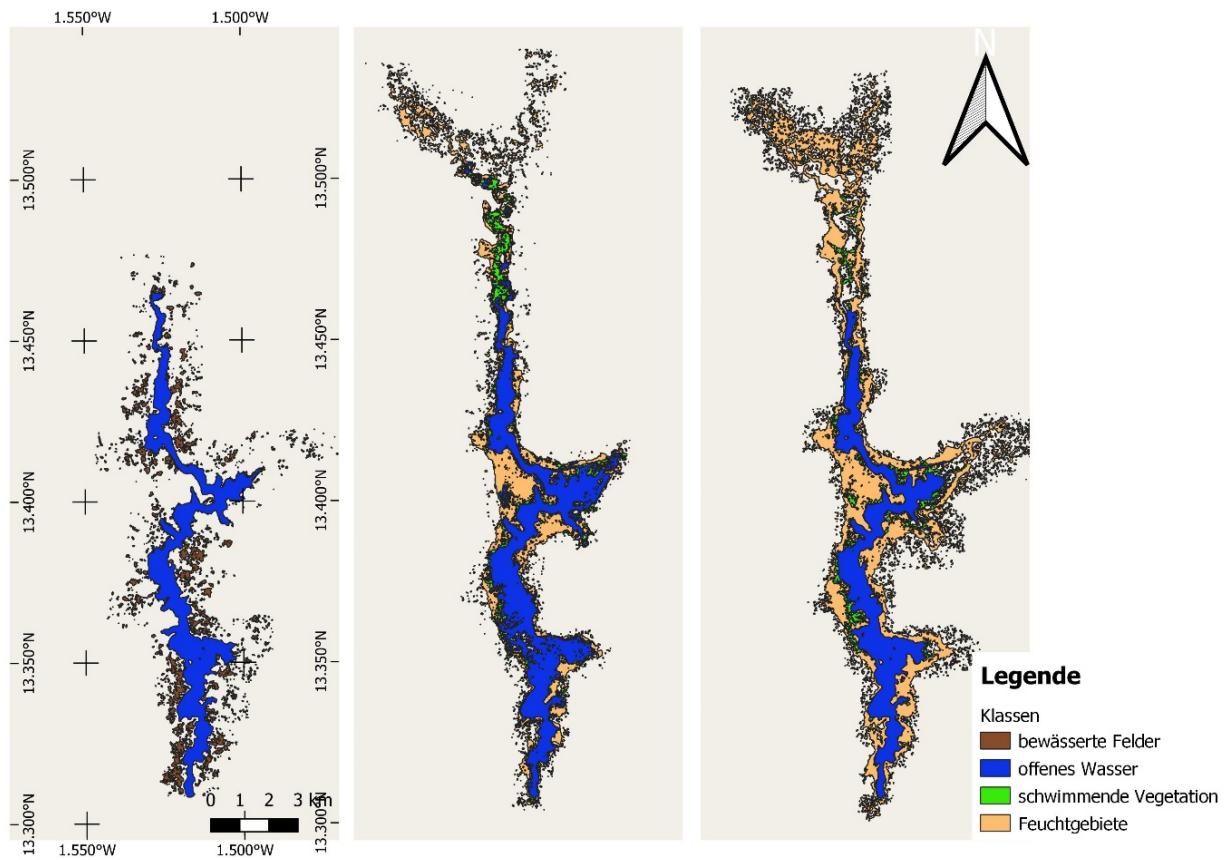


Abbildung 5.11 klassifizierte Aufnahmen von links nach rechts 16.06.2020; 19.07.2020; 04.10.2020

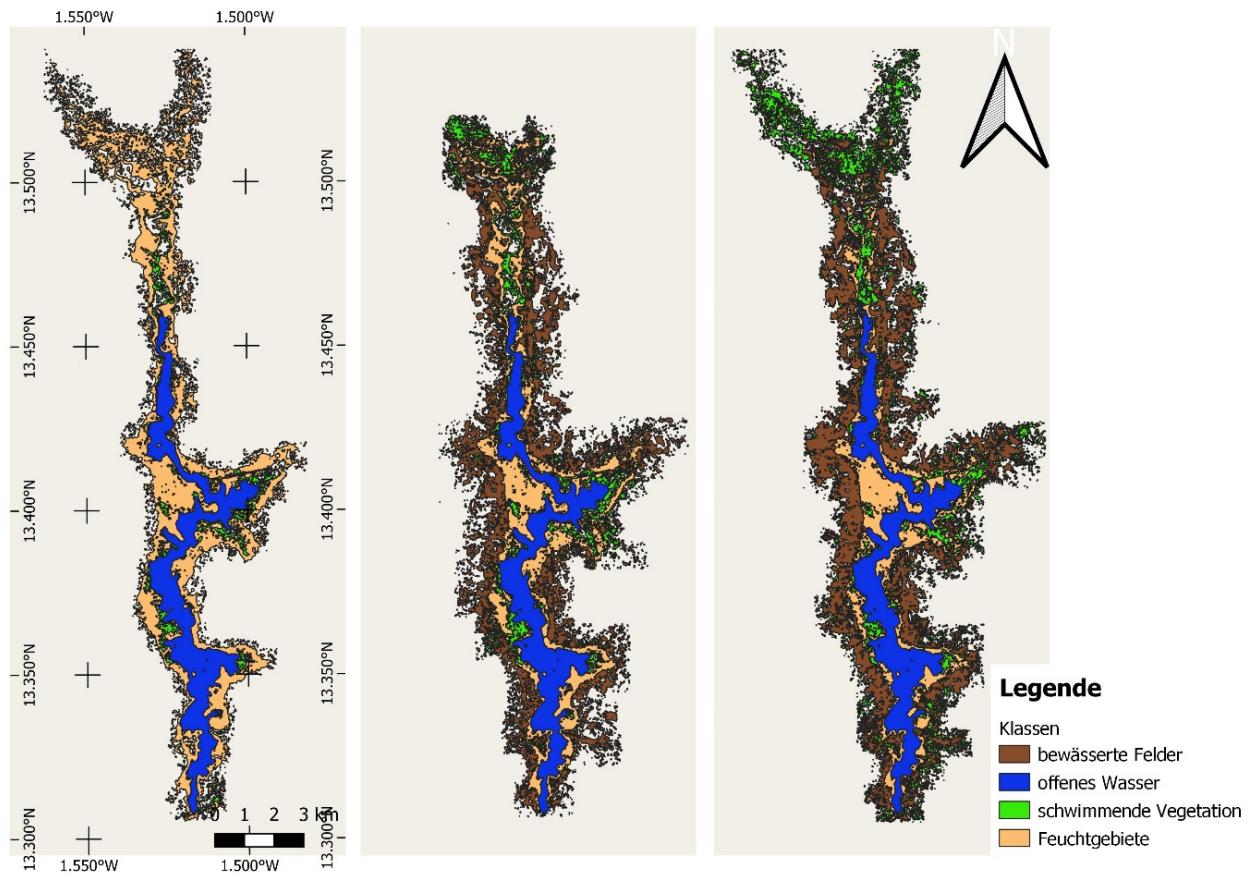


Abbildung 5.12 klassifizierte Aufnahmen von links nach rechts 15.10.2020; 06.11.2020; 09.12.2020

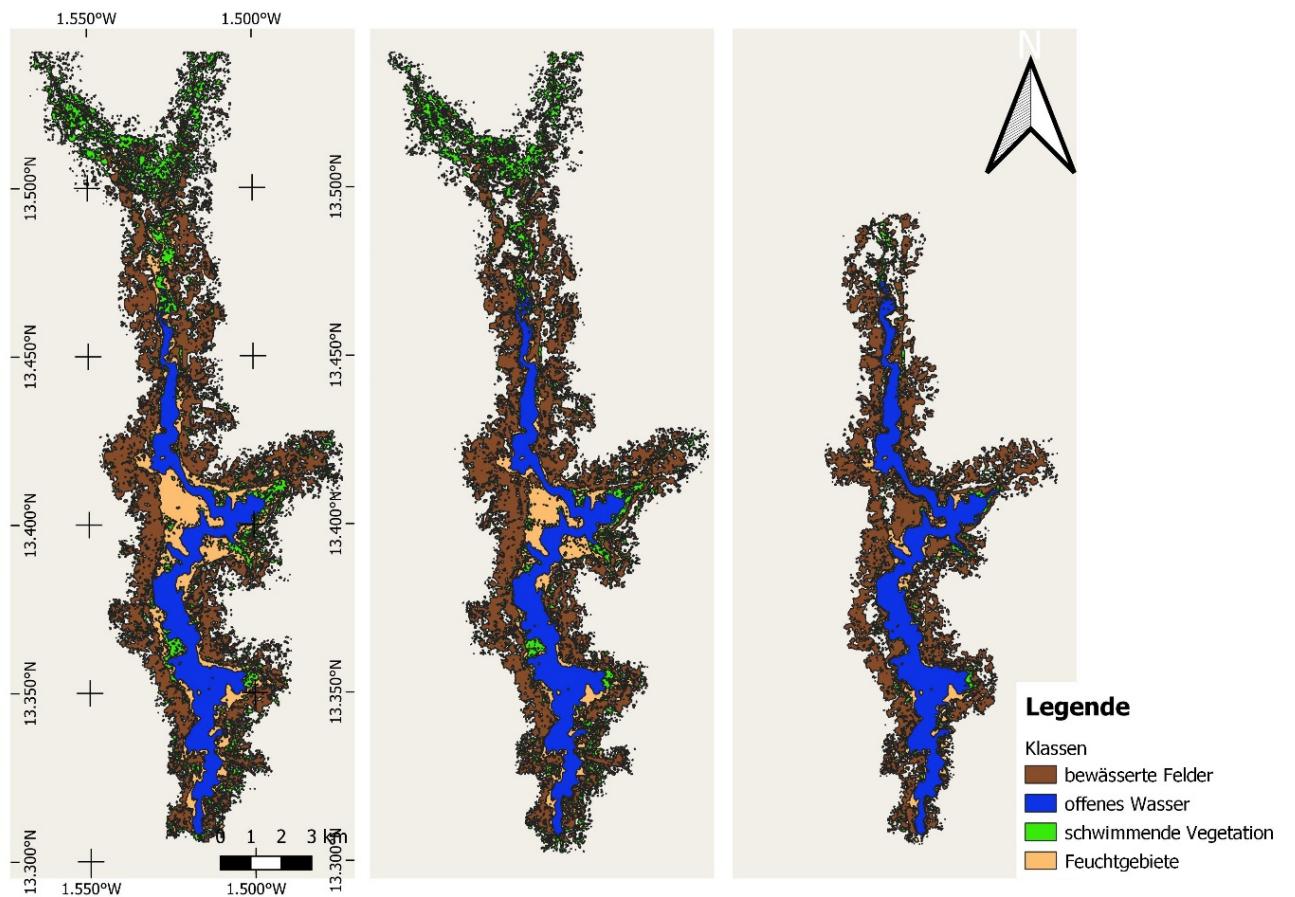


Abbildung 5.13 klassifizierte Aufnahmen von links nach rechts 20.12.2020; 11.01.2021; 02.02.2021

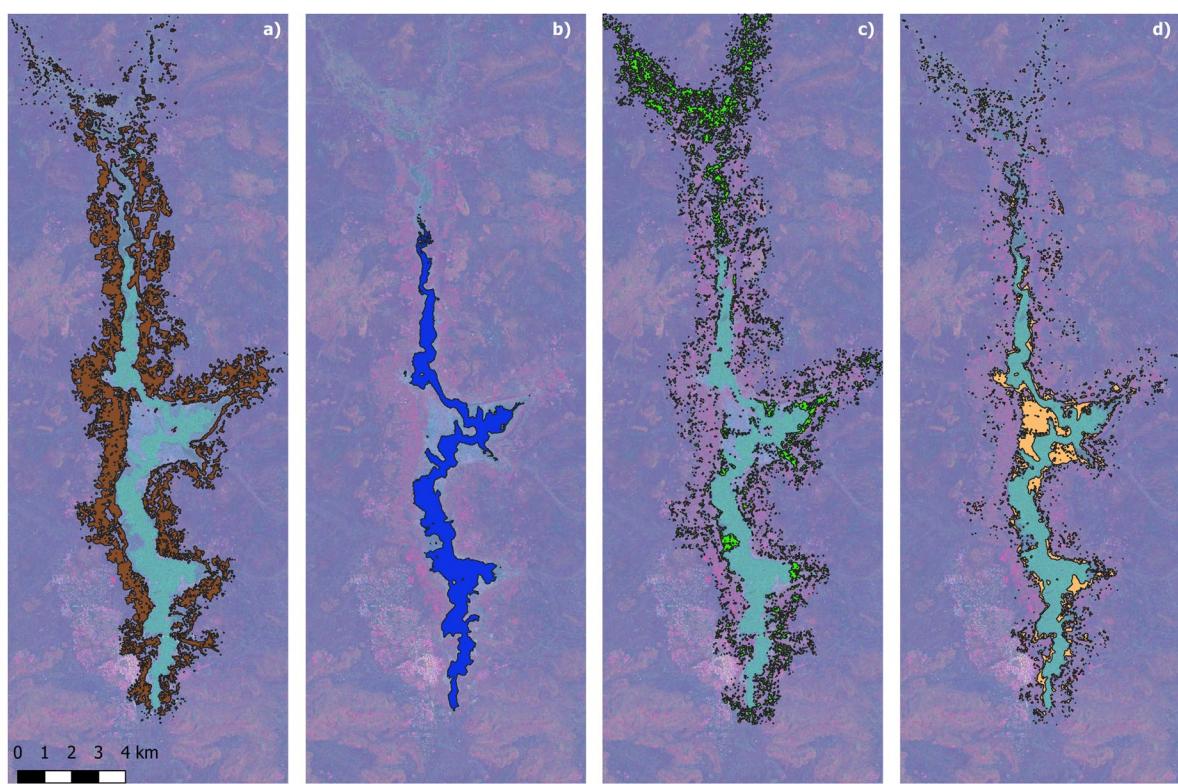


Abbildung 5.14 Beispiel abgeschlossene Klassifikation anhand der Aufnahme vom 11.01.2021 a) bewässerte Felder (braun), b) offenes Wasser (blau). c) schwimmende Vegetation (grün), d) Feuchtgebiete (orange)

5.4 ERGEBNIS FLÄCHENBERECHNUNG

Im folgenden Diagramm werden alle Klassen der Aufnahmen von 22.04.2020 – 02.02.2021 in Abhängigkeit des monatlichen Niederschlags dargestellt. Die Aufnahme von 19.04.2019 wurde dabei nicht berücksichtigt, da diese aufgrund des hohen zeitlichen Abstands zur Aufnahme vom 22.04.2020 nicht zu diesem Zyklus hinzugezählt werden konnte.

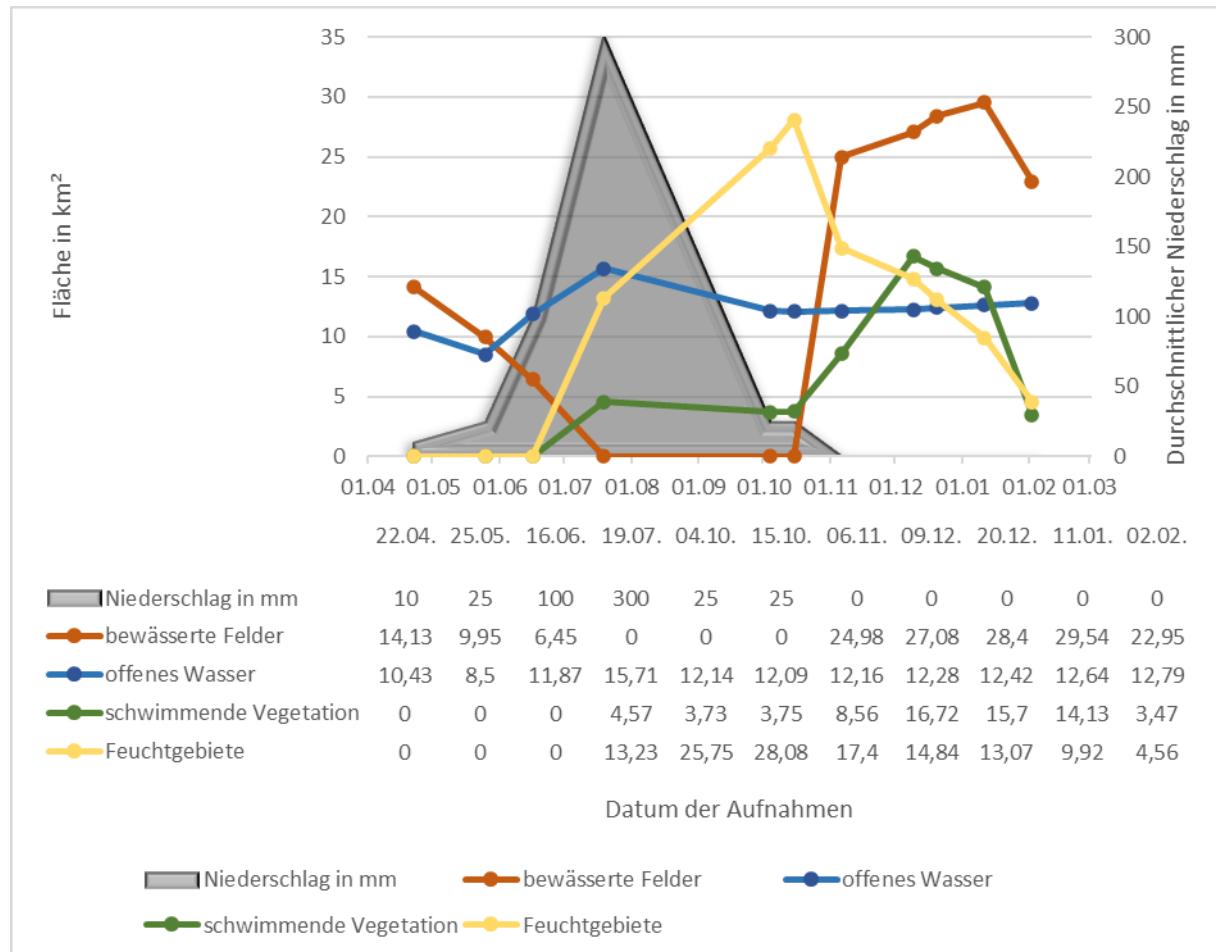


Abbildung 5.15 Veränderung der Flächen bewässerte Felder, offenes Wasser, schwimmende Vegetation, Feuchtgebiete in Abhängigkeit des durchschnittlichen Niederschlages in mm

6. DISKUSSION

6.1 BEWERTUNG DER REFERENZDATEN

Die Referenzdaten, die aus der Arbeit von Klingebiel C. (Klingebiel et al. 2019) resultierten, stellten nach der Aufbereitung die Grundlage für diese Arbeit bereit. Die Referenzdaten lieferten ausreichende Daten, um die Algorithmen zu trainieren.

Die TerraSAR-X Mission liefert seit dem Jahre 2013 durchgängige Aufnahmen von dem Untersuchungsgebiet Lac Bam. So wurden bis zu 119 Aufnahmen aus den Jahren 2013-2018 in (Klingebiel et al. 2019) klassifiziert. Auf diese Daten konnte in dieser Arbeit zurückgegriffen werden. Bei der genaueren Betrachtung der Datensätze mit den Klassen *bewässerte Landwirtschaft, offenes Wasser, schwimmende Vegetation und Feuchtgebiet*, wurden einige Aufnahmen aus dem Jahre 2018 nicht weiter berücksichtigt. Hier überlagerten sich einige Polygone und es konnte nicht nachvollzogen werden, welche Klasse, die richtig klassifizierte ist.

Um den Random Forest korrekt auszuführen, war es wichtig, dass alle Klassen gleich oft vertreten sind. Bei einem zu großen Ungleichgewicht der Klassen können bei der Klassifizierung Probleme auftreten. (Chao Chen 2004) Für die Klasse schwimmende Vegetation, standen über die Zeiträume am wenigsten Daten zur Verfügung. Dies lag auch daran, dass die Klasse nicht über das ganze Jahr verfügbar war.

Das Ungleichgewicht der Daten, konnte durch die Auswahl eines bestimmten Datensatz (siehe Anhang) aus den 119 Aufnahmen größtenteils beseitigt werden. Auch die „scikit-learn“ Bibliothek imbalanced learn (imbalanced learn, 2022), hat geholfen, den Datensatz durch die „Oversampling“ Methode auszugleichen.

Durch die Eingrenzung der Referenzdaten wurde die Größe des Datensatzes minimiert. Es stellte sich heraus, dass der Random Forest mit allen zur Verfügung gestellten Daten bis zu einem Tag Berechnungszeit brauchte. Durch die kleinere Speicherkapazität, die durch das Selektieren der Trainingsdaten erhalten wurde, konnte die Performance und die Berechnungszeit verbessert werden.

Bei der Selektierung der Referenzdaten wurde neben dem Klassen-Gleichgewicht auch darauf geachtet, dass die Klassen über das Jahr gleich oft, falls vorhanden, vertreten waren.

6.2 SATELLITENDATEN, HÖHENMODELL UND OPEN STREET MAP

Wie bereits erwähnt, wurden Aufnahmen des PlanetScope sowie des Sentinel 2 Satelliten als Referenzdaten genutzt. Durch diese konnte das zu klassifizierende Gebiet eingegrenzt und somit Fehlklassifikationen vermieden werden. Vor allem die hochauflösende PlanetScope Aufnahme war für die Nachbearbeitung von Vorteil.

Anhand der Referenzdaten konnten die fehklassifizierten Gebiete, wie z.B. die Stadt Kongoussi identifiziert werden und so das Trainingsgebiet neu eingegrenzt werden. (Abbildung 6.1)

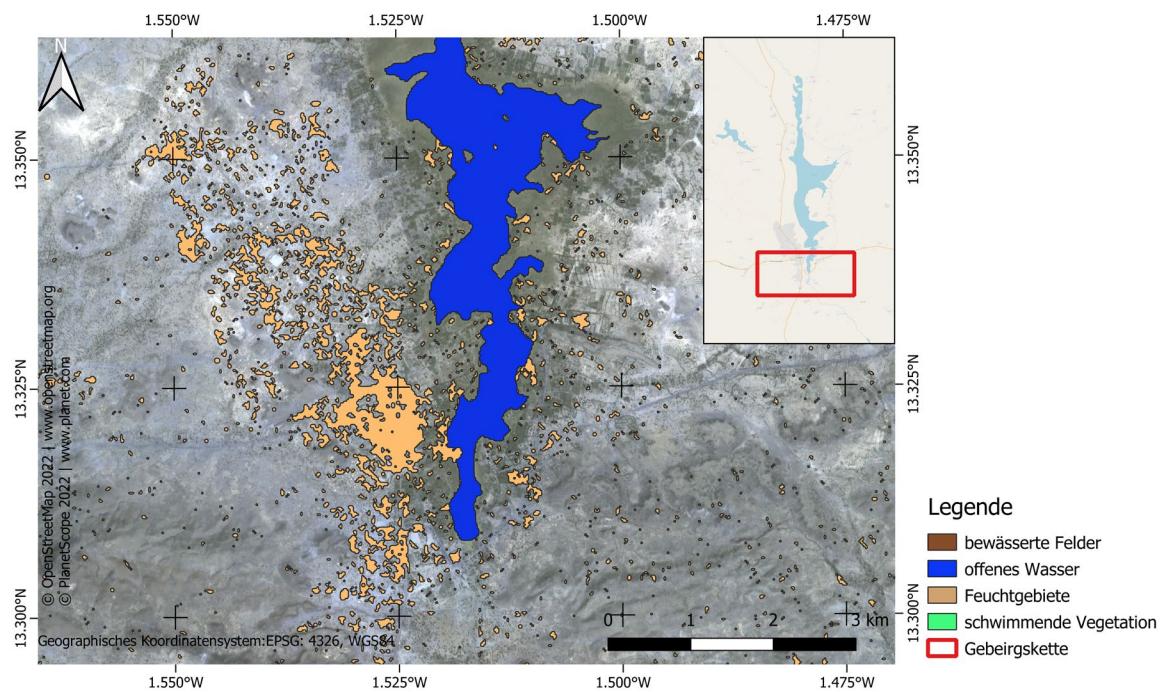


Abbildung 6.1 Fehlklassifizierung des urbanen Gebietes mit Feuchtgebieten.

Des Weiteren diente das Höhenmodell der SRTM-Mission als Referenz. Mithilfe der im Jahre 2000 gestarteten Weltraummission konnten Fehlklassifikationen der sich um den See befindenden Hügelketten ausgeschlossen werden. Die Gebirgsketten wurden bei der Klassifikation als bewässerte Landwirtschaft erkannt. Dieser Effekt ist darauf zurückzuführen, dass das zum Sensor gerichtete Gestein stark zurückstrahlt und somit die gleichen Eigenschaften wie die Klasse bewässerte Landwirtschaft besitzt.

Die Daten von Open Street Map waren für die Veranschaulichung als Hintergrund für die Darstellung von Karten oder deren Übersicht hilfreich.

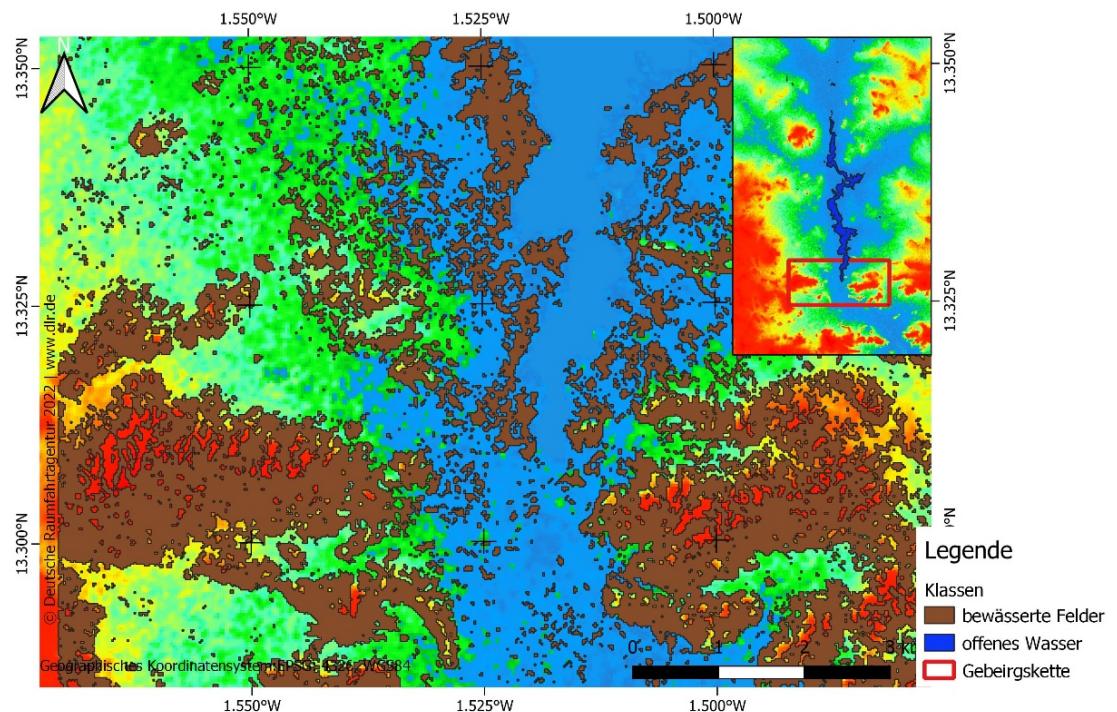


Abbildung 6.2 SRTM-Höhenmodell für die Einordnung der Gebirgskette

6.3 BEWERTUNG DER WETTERDATEN

Die vom Deutschen Wetterdienst GPCC First Guess Monthly Daten dienten zur Einordnung von Regen und Trockenzeit. Hier war es wegen der geringen Auflösung der zur Verfügung gestellten Daten schwer, den genauen Wert der monatlichen Niederschlagsmenge in Burkina Faso und dem Untersuchungsgebiet zu bestimmen. Trotz der geringen Auflösung waren diese Daten für die Bestimmung der Jahreszeiten für diese Arbeit ausreichend.

6.4 BEWERTUNG DER VALIDIERUNGSERGEBNISSE

Durch die berechnete Validierung lässt sich die erste Forschungsfrage dieser Arbeit beantworten. Es wurde eine Genauigkeitssteigerung mithilfe eines Random Forests und den räumlichen-hyperkomplexen Basen erreicht.

Die Validierungsergebnisse zeigen, dass die räumlich-hyperkomplexen Basen eine minimale Genauigkeitssteigerung gegenüber dem MultiLooking erzielen konnten. So kann aus dem Diagramm (Abbildung 5.8) entnommen werden, dass die Genauigkeit, ausgehend von der Pixelstufe 2,5m bis zur Stufe 20m, einen schnelleren Anstieg erzielten. Das beste Ergebnis der Genauigkeitsanalyse lieferte die Pixelstufe 20m mit 73%. Die Gesamtgenauigkeit des MultiLooking erzielte dort 71%. Nach der Pixelstufe 20m nimmt die Genauigkeit beider Datensätze ab. Da die ursprüngliche Klassifikation auf 5m-Pixel angefertigt wurde, kann die Abnahme der 40m- und 80m-Pixel auf Mischpixel zurückgeführt werden.

Die K-Nearest-Neighbor Klassifikation ergab eine Gesamtgenauigkeit der Stufe 20m mit räumlich-hyperkomplexen Basen von 66%. Die des Random Forests erzielte dort einen Wert von 71%. Durch die hohe Berechnungszeit, des K-Nearest-Neighbor wurde in dieser Arbeit nur die Stufe 20m berechnet werden. Hinsichtlich zukünftiger Forschungsarbeit wäre es interessant, die räumlich-hyperkomplexe Basen in Ihren unterschiedlichen Pixelstufen weiterführend mit einem K-Nearest-Neighbor zu erforschen.

6.4.1 FEHLKLASSIFIKATIONEN

Neben der schon erwähnten Fehlklassifikationen der Gebirgsketten und dem Urbanen Gebiet in Kongoussi traten weitere saisonal bedingte Probleme bei der Klassifikation der zwölf TerraSAR-X Daten auf.

Während der Regenzeit wurde das umliegende Land als bewässerte Felder klassifiziert. Aufgrund des starken Niederschlags in der Regenzeit wurden die bewässerten Felder in Regenbewässerten Landwirtschaft kategorisiert. Somit wurde die Klasse bewässerte Felder in den Monaten September und Oktober nicht berücksichtigt. Dieselben Probleme traten auch bei der Arbeit von Moser.L (Moser et al. 2016) auf. Im Hinblick auf zukünftige Studien könnte das Problem mit einem Klassen-Wahrscheinlichkeitslayer eingegrenzt werden. Der Layer gibt an, zu wie viel Prozent das klassifizierte Pixel z.B. als „offenes Wasser“ klassifiziert wurde. Eine Random Forest Klassifizierung stuft das Pixel entweder zu 0% (nicht Wasser) oder 100% (Wasser) ein.

Ein weiteres Problem ist während der Trockenzeit aufgetreten. Dort wurden die Flächen der Schwimmende Vegetation mit der Vegetation der trockenen Flächen außerhalb der Feuchtgebiete verwechselt. Das Problem konnte mit der Eingrenzung des Trainingsgebietes behoben werden. In den hervorgegangenen Arbeiten von Moser L. (Moser et al. 2016) Klingebiel C. (Klingebiel et al. 2019) wurde dieses Problem auch schon beschrieben.

Die Aufnahme vom 29.07.2020 zeigt im südwestlichen Teil des Sees in der Fläche offenes Wasser einige als Land klassifizierte Lücken auf. Diese sind nach Abgleich mit den Referenzdaten als klare Fehlklassifikation zu identifizieren. Eine Ursache dafür können zum Zeitpunkt der Aufnahme Stürme

oder Unwetter gewesen sein. So lassen sich aus den GPCC-Daten ableiten, dass im Monat Juli die Niederschlagsmenge mit 300ml am höchsten war.

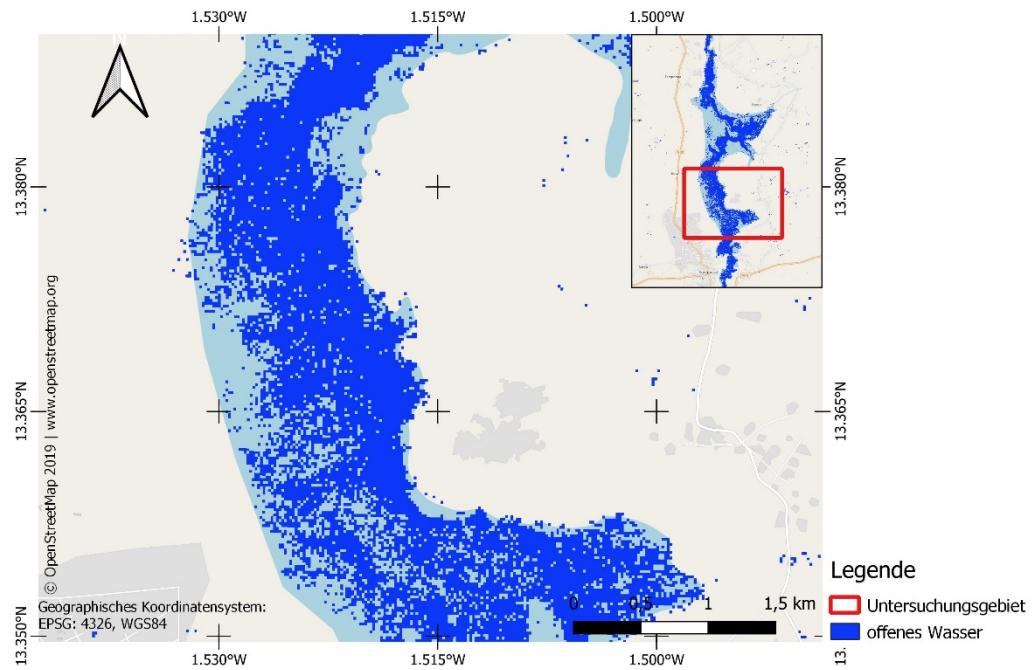


Abbildung 6.3 Probleme bei der Klassifizierung von "offenem Wasser" am 29.07.2020

Aufgrund der hohen Anzahl an Fehlklassifikationen und die dadurch benötigte Nachbearbeitung war es nicht möglich, eine vollautomatische Klassifikation durchzuführen. Durch die in diesem Kapitel (6.4.1 Fehlklassifikation) und in Abbildung 6.1/Abbildung 6.2 SRTM-Höhenmodell für die Einordnung der Gebirgskette Abbildung 6.2 gezeigten Fehlklassifikationen, die eine demensprechende Anpassung des Trainingsgebiet erforderten, kann die zweite Forschungsfrage mit nein beantwortet werden.

6.5 BEWERTUNG DER KLASSEN

In diesem Abschnitt wird auf den berechneten F1-Score aller Klassen eingegangen. Die Klasse offenes Wasser haben in allen Pixelstufe eine gute bis sehr gute Aussagekraft. Während der Klassifikation gab es außer der in Abbildung 6.3 gezeigten Aufnahme keine Probleme. Ein weiterer Grund für die gute Klassifikation war, dass diese Klasse in jeder klassifizierten Aufnahme verfügbar war. Die Werte der Stufen 0-4 liegen sehr nahe im Bereich von 83%-91% F1-Score zusammen. Bei den letzten zwei Pixelstufe nimmt die Genauigkeit etwas ab. So liegt der Wert bei der Stufe 80m bei 72 %.

Die Klasse bewässerte Felder besitzt in der Stufe 20m ihren besten F1-Score mit 84%. Generell besitzt die Klasse einen guten Wert in allen Pixelstufen. Die Stufe mit dem niedrigsten Wert ist die Stufe 2,5m mit 64%.

In der Klasse Feuchtgebiete erzielte die Stufe 20m den höchsten F1-Score. Dort hatte dieser einen Wert von 68% und ist somit noch mit gut zu bewerten.

Als letztes werden die Werte des F1-Scores der Klasse schwimmende Vegetation beschrieben. Diese hatte im Vergleich zu den anderen Klassen den niedrigsten Wert. Hier wurde ein maximaler Score von 43% in der Stufe 20m erreicht. Der geringe erzielte F1-Score kann auf die Klassenverteilung über das Jahr zurückzuführen sein.

Zusammenfassend wurde in jeder Klasse bis auf „offenes Wasser“ in der Stufe 20m der Höchste Wert erzielt.

6.6 BEWERTUNG DER KARTENDARSTELLUNG

Die Darstellung der zwölf Aufnahmen und deren Klassifizierungsergebnisse konnte mit QGIS gut umgesetzt werden. Durch diese konnten die Dynamik des Lac Bam Sees gut dargestellt werden. Durch die geringe Anzahl an Aufnahmen und das nicht erreichten Aufnahmeintervall von elf Tagen, wurden diese nebeneinander abgebildet und nicht wie in „Lac Bam im Radarbild - jahreszeitliche und langfristige Entwicklung beobachtet mit TerraSAR-X (Klingebiel et al. 2019) als Animation dargestellt. Eine weitere Einschätzung der Aufnahmen erfolgt im nächsten Kapitel (6.7 Bewertung der neuen Aufnahmen)

6.7 BEWERTUNG DER NEUEN AUFNAHMEN

Es wurden für die Arbeit zwölf neue Aufnahmen des Lac Bams Sees für die Klassifikation vom DLR bereitgestellt. Dabei gab es einige Auffälligkeiten. So startete die erste Aufnahme am 19.04.2019. Diese knüpft ein halbes Jahr später an den von Christoph Klingebiel (Klingebiel, 2019) klassifizierten Datensatz an.

Im EOWEB GeoPortal des DLR (EOWEB GeoPortal, 2022) wurde die Verfügbarkeit des TerraSAR-X Satelliten in dem zu untersuchendem Gebiet überprüft. Das Portal gab einen Aufschluss darüber, dass im Jahr 2019 vom TerraSAR-X Satelliten nur eine Aufnahme getätigten wurde. Die Nächste verfügbare Aufnahme wurde ein Jahr später am 22.04.2020 aufgenommen.

Generell war ein Aufnahmeintervall von 11 Tagen zu erwarten. Bei der Überwachung des Untersuchungsgebiete mit dem TerraSAR_X handelt es sich um eine „background mission“. Vermutlich wurde ein anderes Gebiet in der Zeit, in der keine Aufnahmen vom Lac Bam See zu Verfügung standen, vom DLR priorisiert.

Wegen der fehlenden Aufnahmen konnte erst wieder mit der Aufnahme vom 22.04.2020 zum Ende der Trockenzeit ein Klassifizierungsintervall gestartet werden. Auch im Jahr 2020 war kein 11-tägige Aufnahmeintervall des Satelliten zu erkennen. Eine Background Mission ist für eine kontinuierliche Überwachung nicht ausreichend. Es werden elftägige Aufnahmen benötigt. In dem Zeitraum von 2013- 2018 waren elftägige Aufnahmen meistens verfügbar. (Klingebiel et al. 2019) Dadurch konnten die TSX-Daten in die Zyklen (Regen- Trockenzeit) eingeteilt werden. Auch die dynamische Veränderung konnte in einer Zeitreihenanalyse gezeigt werden.

6.8 BEWERTUNG DER FLÄCHENBERECHNUNGEN

Die in QGIS berechneten Flächen sollen den dynamischen Verlauf des Sees tabellarisch darstellen. Auffälligkeiten zeigen die Klassen schwimmende Vegetation und Feuchtgebiete. Diese Klassen sind in den Aufnahmen vom 22.04-, 25.05- und 26.06.2020 komplett verschwunden (Trockenzeit). Anhand der Wetterdaten konnte nachgewiesen werden, dass die Regenzeit im Juni eingesetzt hat. In der Aufnahme vom 19.07.2020 ist ein Anstieg beider Klassen zu erkennen.

Die bewässerten Felder konnten in den Aufnahmen vom 19.07, 04.10, 15.10- nicht klassifiziert werden. Durch den starken Regen in der Regenzeit war es nicht möglich, die Klassen zu dem Zeitpunkt nachzuweisen. In dieser Zeit werden die Felder ohnehin von dem Regen bewässert und benötigen kein aus dem See abgepumptes Wasser. Die Ausdehnung der bewässerten Felder erreichte ihren Höhepunkt am 11.01.2021 mit 29,54 km². Wie auch auf der klassifizierten Abbildung 5.13 b) zu erkennen ist, wurde hier die komplette Fläche um den See bewirtschaftet.

Die Klasse offenes Wasser hatte über den betrachteten Zyklus vom 22.04.2020 bis 02.02.2021 keinen großen Schwankungen in der Flächenausdehnung. Den niedrigsten Stand hatte der See zum Ende der Trockenzeit am 22.05.2020 mit einer Fläche von 8,5 km². Am 19.07.2020 wurde die höchste Fläche mit 15,71km² berechnet. Dies kann entweder an der hohen Niederschlagsmenge von 300ml pro Quadratmeter oder an den durch das stürmische Wetter beeinflusste Ungenauigkeiten liegen. Durch die vom Sturm beeinflusste Aufnahme stieg die Fläche des offenen Wassers auf ca. 3km² an. Da Unwetter nur einzelne Aufnahmen betreffen und nicht kontinuierlich verlaufen, könnte man diese durch eine zeitliche Glättung ausgleichen. Die Fläche war das restliche Jahr konstant bei ca.12,5km². Hierbei sollte beachtet werden, dass zu der Gesamtwaterfläche die schwimmende Vegetation dazugezählt werden muss. Diese hatte eine weitaus größere jahreszeitliche Ausdehnung.

Um die Flächenausdehnung anhand der Niederschlagsmenge zu veranschaulichen, wurde diese über eine Sekundärachse in Abhängigkeit der zeitlichen Aufnahmen in einem Diagramm dargestellt. Durch die fehlenden Aufnahmen konnte dies nicht für die erwarteten 11-tägigen Aufnahmen geplottet werden. Trotzdem liefert das Diagramm eine gute Übersicht über die Veränderung der Flächen in der Trocken- und Regenzeit. Die Aufnahme aus 2019 wurde dabei nicht berücksichtigt.

6.9 EMPFEHLUNGEN FÜR WEITERFÜHRENDE FORSCHUNG

Räumliche-hyperkomplexe Basen sind ein noch unerforschtes Gebiet der Fernerkundung. Bei der spektralen Analyse der Pixel in den unterschiedlichen Klassen wurde vor allem im Bereich des urbanen Gebietes (Abbildung 5.5) ein hoher Informationsgehalt in den weiteren Kanälen festgestellt.

Durch die erlangten Erkenntnisse dieser Arbeit, empfiehlt es sich, die räumlichen-hyperkomplexen Basen vor allem in urbanen Gebieten weiterführend zu untersuchen. Dort könnte ein Vergleich mittels einem Feature Ranking neue Erkenntnisse ergeben.

In dieser Arbeit wurden die Pixel kachelweise untersucht. Die Daten sind aufgrund ihrer Orthogonalität nicht redundant. Dies könnte durch ein “moving Window” Verfahren, in dem jedes Pixel kontinuierlich untersucht wird, behoben werden. Durch so ein Verfahren würde die Speicherkapazität erhöht werden. Gleichermaßen wäre es aber auch möglich, die Genauigkeit zu steigern. Dies gilt es in zukünftigen Arbeiten zu untersuchen.

Für weiterführende Forschungen könnten ebenfalls von Interesse sein, weitere Algorithmen sowie neuronale Netzwerke bei der Klassifikation mithilfe räumlichen-hyperkomplexen Basen erweiterte TerraSAR-X Aufnahmen zu nutzen. Neuronale Netzwerke untersuchten bis jetzt (stand 01.12.2020) nur Aufnahmen mit maximal vier Kanälen. Methoden, wie Sequence-to-Sequence Classification Using 1-D Convolutions (MathWorks 2022), UMAP for Clustering (UMAP 2022) oder satellite-image-deep-learning (Cole 2022) könnten viel versprechende Ansätze sein, um in naher Zukunft die räumlichen-hyperkomplexen Basen fortführend zu erforschen.

Die Spektrale Analyse, lässt auch weitere Fragen für zukünftige Arbeiten offen. Es konnte festgestellt werden, dass der Intensitätsunterschied im Betrag bei den Klassen „*bewässerte Felder, offenes Wasser, schwimmende Vegetation und Feuchtgebiete*“ in den weiteren Kanälen abnimmt. Die Analyse, des *urbanen Gebietes*, hatte weiterhin einen hohen Ausschlag des Intensitätsunterschied. Hier wäre ein *Feature Ranking* von Interesse.

7. ZUSAMMENFASSUNG

In dieser Arbeit wurde erfolgreich an den Forschungsstand der Arbeit „*Lac Bam im Radarbild - jahreszeitliche und langfristige Entwicklung beobachtet mit TerraSAR-X*“ (Klingebiel et al. 2019) angeknüpft. Die Klassifikationsergebnisse, die aus dieser Arbeit erlangt wurden, konnten als Trainingsdatensatz für die überwachte Klassifizierungen aufbereitet und genutzt werden.

Anhand der Validierung der Klassifikationen, konnte eine Genauigkeitssteigerung der räumlich hyperkomplexen Basen gegenüber dem Standard-MultiLooking erreicht werden. Eine weitere Forschungsfrage dieser Arbeit war die Automatisierung der Klassifikation neuer Aufnahmen. Durch die in Kapitel Diskussion aufgezählten Probleme und die daraus resultierende Nachbearbeitung der Klassifizierung war dies nicht vollständig möglich. Die Probleme sind bereits in den hervorgegangenen Arbeiten((Klingebiel et al. 2019); (Moser et al. 2016)) aufgetreten.

Es gilt, in weiterführenden Arbeiten die räumlichen hyperkomplexen Basen weiter zu erforschen. Durch weitere Verfahren kann es möglich sein, dass eine höhere Genauigkeit Steigerung erzielt, sowie die Klassifikation vollautomatisch durchgeführt werden kann.

8. LITERATURVERZEICHNIS

2. Over-sampling — Version 0.10.0 (2022). Online verfügbar unter https://imbalanced-learn.org/stable/over_sampling.html, zuletzt aktualisiert am 09.12.2022, zuletzt geprüft am 09.12.2022.

Albertz, Jörg (2009): Einführung in die Fernerkundung. Grundlagen der Interpretation von Luft- und Satellitenbildern. 4., aktualisierte Aufl. Darmstadt: WBG (Wissenschaftliche Buchgesellschaft).

Ambroise Ouédraogo, directeur du projet de restauration du Lac Bam : « le Lac a perdu actuellement les ¾ de ses ressources en eau » - leFaso.net (2022). Online verfügbar unter <http://lefaso.net/spip.php?article60373>, zuletzt aktualisiert am 14.12.2022, zuletzt geprüft am 14.12.2022.

Amt, Auswärtiges (2022): Burkina Faso: Politisches Porträt. Online verfügbar unter <https://www.auswaertiges-amt.de/de/aussenpolitik/laender/burkinafaso-node/politisches-portrait/212366>, zuletzt aktualisiert am 20.12.2022, zuletzt geprüft am 20.12.2022.

Breiman, Leo (2001): Random Forests. In: *Machine Learning* 45 (1), S. 5–32. DOI: 10.1023/A:1010933404324.

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (2022): Burkina Faso. Online verfügbar unter <https://www.bmz.de/de/laender/burkina-faso/>, zuletzt aktualisiert am 20.12.2022, zuletzt geprüft am 20.12.2022.

Chao Chen (2004): Using Random Forest to Learn Imbalanced Data. In: *undefined*. Online verfügbar unter <https://www.semanticscholar.org/paper/Using-Random-Forest-to-Learn-Imbalanced-Data-Chain/2138b37bfced70599d26dfccbf93a8e7a4b7ad85#citing-papers>.

Cole, Robin (2022): satellite-image-deep-learning | Robin Cole | Substack. Online verfügbar unter <https://www.satellite-image-deep-learning.com/>, zuletzt aktualisiert am 09.12.2022, zuletzt geprüft am 09.12.2022.

Cover, T.; Hart, P. (1967): Nearest neighbor pattern classification. In: *IEEE Trans. Inform. Theory* 13 (1), S. 21–27. DOI: 10.1109/TIT.1967.1053964.

DLRARTICLE DLR Portal (2022): Synthetic Aperture Radar (SAR). Online verfügbar unter <https://www.dlr.de/content/de/artikel/missionen-projekte/terrasar-x/synthetisches-aperture-radar.html>, zuletzt aktualisiert am 12.12.2022, zuletzt geprüft am 12.12.2022.

Forkel, Matthias (2015): Das Klima der Erde | Passatzirkulation. Online verfügbar unter http://www.klima-der-erde.de/zirk_passat.html, zuletzt aktualisiert am 04.11.2015, zuletzt geprüft am 20.12.2022.

<https://scihub.copernicus.eu/dhus/#/home> (2022). Online verfügbar unter <https://scihub.copernicus.eu/dhus/#/home>, zuletzt aktualisiert am 16.12.2022, zuletzt geprüft am 20.12.2022.

Klingebiel, Christoph; Schmitt, Andreas; Wendleder, Anna; Moser, Linda (2019): Lac Bam Shown by TerraSAR-X - The Visualisation of Seasonal and Long Term Changes. In: TerraSAR-X/TanDEM-X Science Meeting. Oberpfaffenhofen, Germany, 21.-24.Okt 2019. Online verfügbar unter <https://elib.dlr.de/130011/>.

Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F.: World map of the Köppen-Geiger climate classification updated. In: *Meteorologische Zeitschrift*, 15(3), 259-263 2006.

Linda Moser (2021): Monitoring Wetlands and Water Bodies in Semi-arid Sub-Saharan Regions. Friedrich-Schiller-Universität Jena, Jena.

MathWorks (2022): Sequence-to-Sequence Classification Using 1-D Convolutions - MATLAB & Simulink. Online verfügbar unter

<https://www.mathworks.com/help/deeplearning/ug/sequence-to-sequence-classification-using-1-d-convolutions.html>, zuletzt aktualisiert am 10.12.2022, zuletzt geprüft am 10.12.2022.

MISSION DLR Portal (2022): TerraSAR-X – Deutschlands Radar-Auge im All. Online verfügbar unter <https://www.dlr.de/content/de/missionen/terrasar-x.html>, zuletzt aktualisiert am 08.12.2022, zuletzt geprüft am 08.12.2022.

Moser, Linda; Schmitt, Andreas; Wendleder, Anna; Roth, Achim (2016): Monitoring of the Lac Bam Wetland Extent Using Dual-Polarized X-Band SAR Data. In: *Remote Sensing* 8 (4), S. 302. DOI: 10.3390/rs8040302.

Ouedraogo, Raymond: Fish and fisheries prospective in arid inland waters of Burkina Faso, West Africa / composed and submitted by Raymond Ouedraogo. Wien. Online verfügbar unter <https://epub.boku.ac.at/obvbokhs/content/titleinfo/1931637>.

Ramsar Convention on Wetlands (2022). Online verfügbar unter <https://www.ramsar.org/>, zuletzt aktualisiert am 14.12.2022, zuletzt geprüft am 14.12.2022.

Ramsar-Konvention: Schutz der Feuchtgebiete (2022). Online verfügbar unter <https://www.wwf.de/themen-projekte/fluesse-seen/wasser-politik-maerkte/ramsar-konvention>, zuletzt aktualisiert am 20.12.2022, zuletzt geprüft am 20.12.2022.

Satellite Imagery Analytics | Planet (2022). Online verfügbar unter <https://www.planet.com/products/planet-imagery/>, zuletzt aktualisiert am 17.12.2022, zuletzt geprüft am 20.12.2022.

Schmitt, Andreas; Wendleder, Anna; Hinz, Stefan (2015): The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 102, S. 122–139. DOI: 10.1016/j.isprsjprs.2015.01.007.

Schmitt, Andreas; Wendleder, Anna; Kleynmans, Rüdiger; Hell, Maximilian; Roth, Achim; Hinz, Stefan (2020): Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases. In: *Remote Sensing* 12 (6), S. 943. DOI: 10.3390/rs12060943.

Sokolova, Marina; Japkowicz, Nathalie; Szpakowicz, Stan (2006): Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In: Abdul Sattar und Byeong-ho Kang (Hg.): AI 2006: advances in artificial intelligence. 19th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, December 4-8, 2006 ; proceedings, Bd. 4304. Berlin, Heidelberg: Springer (Lecture notes in computer science Lecture notes in artificial intelligence, 4304), S. 1015–1021.

Taylor, Christopher M.; Belušić, Danijel; Guichard, Françoise; Parker, Douglas J.; Vischel, Théo; Bock, Olivier et al. (2017): Frequency of extreme Sahelian storms tripled since 1982 in satellite observations. In: *Nature* 544 (7651), S. 475–478. DOI: 10.1038/nature22069.

UMAP (2022): Using UMAP for Clustering — umap 0.5 documentation. Online verfügbar unter <https://umap-learn.readthedocs.io/en/latest/clustering.html>, zuletzt aktualisiert am 10.12.2022, zuletzt geprüft am 10.12.2022.

Wasserprobleme im Sahel – Klimawandel (2022). Online verfügbar unter https://wiki.bildungsserver.de/klimawandel/index.php/Wasserprobleme_im_Sahel#Wassernutzung, zuletzt aktualisiert am 20.12.2022, zuletzt geprüft am 20.12.2022.

Wetter und Klima - Deutscher Wetterdienst - Our services - Global Precipitation Climatology Centre (GPCC) (2022). Online verfügbar unter <https://www.dwd.de/EN/ourservices/gpcc/gpcc.html>, zuletzt aktualisiert am 20.12.2022, zuletzt geprüft am 20.12.2022.

ANHANG

Verwendete TSX-Aufnahmen	fields	open water	fv	wetlands
26.08.2013				
06.09.2013				
09.10.2013				
14.12.2013				
16.01.2014				
27.01.2014				
07.02.2014				
28.05.2014				
15.09.2014				
07.10.2014				
18.10.2014				
29.10.2014				
12.12.2014				
14.01.2015				
10.03.2015				
01.04.2015				
05.10.2015				
27.10.2015				
07.11.2015				
10.12.2015				
21.12.2015				
03.06.2016				
28.07.2016				
08.08.2016				
30.08.2016				
02.10.2016				
04.11.2016				
07.12.2016				
09.01.2017				
11.02.2017				
21.05.2017				
04.07.2017				
26.07.2017				
28.08.2017				
05.12.2017				
18.01.2018				

Neue TSX Aufnahmen	
04.09.2019	
22.04.2020	
25.05.2020	
16.06.2020	
19.07.2020	
04.10.2020	
15.10.2020	
06.11.2020	
09.12.2020	
20.12.2020	
11.01.2021	
02.02.2021	
Optische Aufnahmen	
Sentinel 2	10.04.2019
Planet	25.12.2020