universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Evaluation von Machine Learning Methoden zur Klassifikation von Wasserflächen kleiner Moorgebiete auf Basis von Satellitenbilddaten gezeigt am Beispiel des Leegmoors

verfasst von | submitted by

Julian Neisius BSc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 856
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Kartographie und Geoinformation
Betreut von | Supervisor:	Ass.-Prof. Mag. Dr. Andreas Riedl

Danksagung:

Ich möchte mich herzlichst bei meinem Betreuer Dr. Andreas Riedl für seine Geduld und seine Anregungen im Zuge meiner Arbeit bedanken. Des Weiteren möchte ich mich bei Herr Prof. Blankenburg bedanken, der mir das Thema zu dieser Arbeit vorgeschlagen hat und mich bei allen Fragen zur Projektfläche Leegmoor unterstützt hat und mir Daten zu dem Projekt zukommen ließ.

Außerdem möchte ich mich bei meinen Eltern bedanken, die mich während meines Studiums immer unterstützt haben.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Glossar

Machine Learning (ML)
Bezeichnet den Prozess eines selbst optimierenden Programmes nach einem Algorithmus

Support Vector Machine (SVM)
Ein Machine-Learning-Algorithmus

Random Forest (RF)
Ein Machine-Learning-Algorithmus

k- Nearest Neighbors (kNN)
Ein Machine-Learning-Algorithmus

Deep Learning (DL)
Ein spezieller Machine Learning Prozess, der dem Lernen des Gehirns mit seinen Neuronen nachempfunden ist

Accuracy Assessment
Aufstellung von Genauigkeitsgrößen für die Beurteilung eines Modelles

False positives
Wert der irrtümlicherweise als positiv ausgegeben wird

False negatives
Wert der irrtümlicherweise als negativ ausgegeben wird

Ground Truth Daten
Datensatz der durch beispielsweise Ortskenntnisse weiter verifiziert ist

Overfitting
Überanpassung eines Modelles an sein Trainingsdaten

Underfitting
Zu geringe Anpassung eines Modelles an seine Trainingsdaten

Top Of Atmosphere (TOP)
Produktklasse bei Satellitendaten

Bottom Of Atmosphere (BOA)
Produktklasse bei Satellitendaten

Schwarztorf
Stark zersetzte und komprimierte Torfmoose, die die unteren Schichten eines Moores aufbauen

Abstract

Diese Arbeit soll die beste Methode zur Klassifizierung von Wasserflächen in kleineren Moorgebieten in Deutschland herausfinden. Exemplarisch wird hierfür das Leegmoor als Testgebiet verwendet, da auf dieser Fläche seit den 80er Jahren Versuche zur Renaturierung von abgetorften Moorflächen durchgeführt werden und somit die Fläche und ihre Eigenschaften sehr gut bekannt sind. Die Arbeit untersucht dabei welcher Algorithmus, welche Satellitendaten und welche Software sich für die Umsetzung am besten eignen. Nach einer Vorauswahl von vier Algorithmen werden anhand dieser ein Accuracy Assessment vorgenommen und so die geeignetste Methode für die Klassifikation der Wasserflächen gefunden.

The aim of this work is to find the best method for classifying water areas in smaller wetlands in Germany. As an example, the Leegmoor is used as a test area, as experiments on the renaturalisation of drained wetlands have been carried out on this area since the 1980s. The area and its characteristics are therefore well known. The work analyses which algorithm, which satellite data and which software are best suited for implementation. After a pre-selection of four algorithms, an accuracy assessment is carried out on the basis of these and thus the most suitable method for the classification of the water areas is found.

1. Einleitung

1.1 Einführung in die Thematik

Feuchtgebiete und Moore sind im Zuge des Klimawandels von großer Bedeutung, denn ohne den Erhalt und Schutz dieser Flächen wird es kaum möglich sein, den Klimawandel zu verlangsamen. Obwohl die Moore nur einen flächenmäßig kleinen Teil der weltweiten Landbedeckung ausmachen, speichern sie mehr Kohlenstoff als alle anderen Ökosysteme (Bartz et al., 2015). Es ist daher von größter ökologischer Wichtigkeit, die Prozesse, die das Moorwachstum beeinflussen zu verstehen. Eine Möglichkeit hierfür ist die Fernerkundung, diese nutzt Satellitendaten oder Luftbildaufnahmen, um neue Daten für das Untersuchungsgebiet zu schaffen. Gerade Satellitendaten sind hierfür interessant, da keine extra Befliegung über dem Gebiet angeordnet werden muss und die Satelliten in der Regel wesentlich häufiger das Gebiet aufnehmen. In der Literatur finden sich hierfür viele Beispiele bei denen Satellitendaten, wie z.B. Landsat und Sentienel-2 genutzt werden um bei großflächigen Untersuchungsgebieten, wie die Westküste Australiens oder den Everglades, Dynamiken des Oberflächenwassers darzustellen (Tulbure & Broich, 2013), (Jones, 2015), (De Vries et al., 2017). In dieser Arbeit soll allerdings geprüft werden, ob sich diese Satellitendaten auch für kleinere Gebiete wie das Leegmoor in Norddeutschland eignen. Da die Feuchtgebiete in Europa deutlich kleiner ausfallen als beispielsweise die Everglades, könnte eine Methode geschaffen werden, mit der die Wasserflächen von Mooren und Feuchtgebieten in Europa effizient überwacht werden. Im Leegmoor in Surwold ist 1983 auf einer teilabgetorften Fläche, bei der nur noch stark zersetzte Torfe zurückbleiben, versucht worden, die ursprüngliche Moor-Flora und -Fauna wieder anzusiedeln. Ein wichtiger Schritt hierfür war die Wiedervernässung der teilabgetorften Fläche, da das Ökosystem eine dauerfeuchte Fläche benötigt. Um dies zu erreichen, wird im Winter die Fläche gestaut, so dass das Wasser nicht mehr aus der Fläche abfließen kann. So bleibt das Moor auch in den Sommermonaten feucht und trocknet nicht aus. Ziel der Arbeit ist es die optimalen Satellitendaten, sowie eine Methode zu finden mit der Wasserflächen im Leegmoor bestmöglich klassifiziert werden können.

Für die Mitarbeiter des Projektes können aus den klassifizierten Satellitendaten neue Erkenntnisse gewonnen werden. Dabei soll die Klassifizierung von Machine-Learning-Algorithmen übernommen werden, da so die Satellitendaten schnell und standardisiert ausgewertet werden können. Die entwickelte Methode könnte bei Erfolg auch für weitere Moorflächen umgesetzt werden. Viele Studien, die durch Fernerkundung die Bodenbedeckung analysieren zu verschiedenen Zeitpunkten, detektieren die Änderung der Bodenbedeckung umso eine Veränderung der Landschaft aufzuzeigen (Liu & Zhou, 2004), (Li & Zhou, 2009). Im Falle dieser Masterarbeit ist die Veränderung der Wasserflächen allerdings saisonal bedingt und ein zu erwartendes Ereignis. Der ML-Algorithmus für die Klassifizierung muss trotz der sich stark saisonal ändernden Landschaft im Moor, zuverlässige Ergebnisse liefern. Daher sollen unterschiedliche ML-Algorithmen auf ihre Eignung untersucht werden. Die Eignung der Methode soll über das Accuracy Assessment überprüft werden, da sich so die Genauigkeiten der verschiedenen Methoden, mit Hilfe einer Fehler Matrix einfach vergleichen lassen. Die Masterarbeit wird verschiedene Satellitendaten miteinander Vergleichen und die geeignetsten Satellitendaten für kleinere Untersuchungsgebiete wie das Leegmoor auswählen.

Die Entscheidung, welche Satellitendaten in dieser Arbeit verwendet werden, soll mit einer Entscheidungstabelle in einem eigenen Kapitel diskutiert werden. Dabei sollen die Unterschiede der Satellitendaten herausgearbeitet werden (zeitliche Auflösung vs räumliche Auflösung vs spektrale Auflösung). Die geeignete Software wird ebenfalls in einem eigenen Kapitel diskutiert und nach bestem Nutzen gewählt. In diesem Zuge soll auch auf die Eigenarten und Funktionen der verschiedenen Softwarepakete eingegangen werden. In einem Kapitel wird eine Vorauswahl an Machine Learning (ML) Methoden diskutiert und anhand des Ergebnisses vier verschiedene Methoden ausgewählt. Um

die beste Methode für die Klassifikation der Wasserflächen zu finden, soll ein Satellitenbild mit den ausgewählten ML-Methoden klassifiziert werden. Dabei sollen die Unterschiede herausgearbeitet werden und die am besten geeignete Methode für das Untersuchungsgebiet gefunden werden. Da in dieser Arbeit nur das Erkennen der Wasserflächen von Relevanz ist, also die Klassen bereits definiert sind durch die Aufgabenstellung werden beide Methoden als überwachte Klassifizierungen durchgeführt. Um die Methode zu validieren, wird das Satellitenbild, welches zur Findung der bestmöglichen Methode verwendet wurde, so ausgewählt, dass dieses zeitlich möglichst nah an einem Orthofoto des Untersuchungsgebietes liegt. Dabei soll eine Fehler Matrix bei der Erstellung des Accuracy Assessment zum Einsatz kommen (Congalton, 1991).

1.2 Umriss der Fragestellung

Diese Arbeit soll zeigen, ob sich Satellitendaten auch für die Klassifikation von kleinen Untersuchungsgebieten wie beispielsweise dem Leegmoor eignen oder, ob nur Klassifikationen von Drohnenaufnahmen ausreichende Genauigkeiten liefern können. Die Arbeiten von Adugna et al. (2022), Tulbure & Broich, (2013), Jones, (2015), De Vries et al., (2017) konnten bereits gute Ergebnisse für deutlich größere Untersuchungsgebiete liefern. Daher soll diese Arbeit die Frage beantworten, welcher ML-Algorithmus eignet sich am besten für die Klassifikation gerade von kleineren Moorgebieten. Dabei werden außerdem die Fragen beantwortet, welche Satellitendaten und welches Softwarepaket sich am besten eignet. Des Weiteren wird gezeigt, wie das Accuracy Assessment in einem hochdynamischen System, wie dem Moor, umgesetzt werden kann.

2. Stand des Wissens

2.1 Fernerkundung:

Die Fernerkundung ist ein indirektes Messverfahren, bei dem es zu keinem direkten Kontakt, mit dem zu dem beobachteten Objekt kommt und sich das Messgerät in einiger Entfernung zu diesem befindet. Bei den Daten, die dabei vom Messgerät erfasst werden, handelt es sich von der Erdoberfläche emittierte bzw. reflektierte elektromagnetische Strahlung oder aber auch Schallwellen. Die Sensoren sind dabei oftmals auf Flugzeugen, Satelliten, Drohnen oder auch Schiffen angebracht. Die Messverfahren können dabei auch noch einmal in zwei Kategorien unterteilt werden. Zum einen in eine **passive Messung** (Abb.1), bei der die von der Erdoberfläche ausgehende Strahlung gemessen wird und in eine **aktive Messung** (Abb.1), bei der eine Welle vom Messgerät auf die Erdoberfläche gesendet wird und die Reflektion dieser Welle gemessen wird (Albertz, 2009).

Ein Beispiel für die **passive Messung** (Abb. 1) wäre eine Luftbildaufnahme, bei der das Sonnenlicht auf der Erdoberfläche reflektiert wird und dann von einem Sensor erfasst wird. Als ein Beispiel für eine **aktive fernerkundliche Messung** (Abb.1) kann das Radar Verfahren genannt werden, bei der Microwellen Strahlung etwa von einem Flugzeug oder Satelliten ausgesendet wird und die reflektierte Mikrowellen Strahlung mit einer Antenne am Flugzeug oder Satelliten gemessen wird (Albertz, 2009).

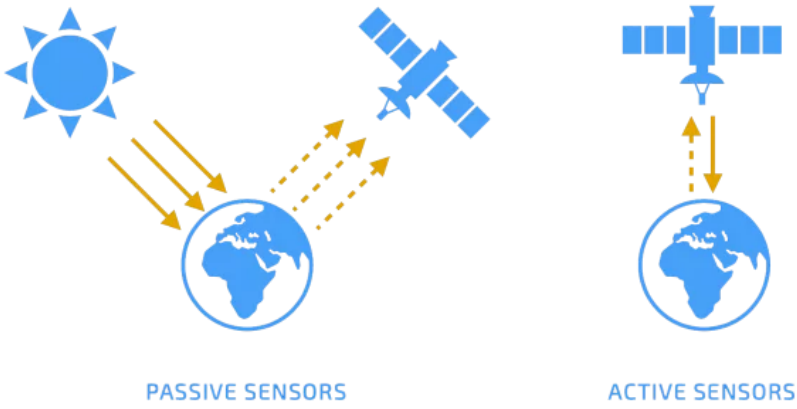

Abbildung 1 Passives Messverfahren und aktives Messverfahren am Beispiel eines Satelliten (https://www.skyrora.com/remote-sensing/ letzter Aufruf 18.12.23)

Die elektromagnetische Strahlung lässt sich anhand der Wellenlänge in mehrere Bereiche unterteilen (Abb. 2). Das elektromagnetische Spektrum beginnt mit der kosmischen Strahlung gefolgt von der Gammastrahlung und der Röntgenstrahlung (Abb.2). Da diese drei Kategorien der elektromagnetischen Strahlung nicht von der Atmosphäre durchgelassen werden, finden diese drei keine Anwendung in der Fernerkundung. An die Röntgenstrahlung grenzt die ultraviolette Strahlung, diese kann auch zum Großteil die Atmosphäre nicht durchdringen, allerdings wird die Atmosphäre ab einer Wellenlänge von 0,3 µm durchlässiger, so dass die nahe ultraviolette Strahlung für bestimmte Fragestellungen in der Fernerkundung genutzt werden kann (Abb.2). An die ultraviolette Strahlung grenzt das Spektrum des sichtbaren Lichtes mit einer Wellenlänge von 0,4 µm – 0,7 µm. Dies ist der Bereich in dem Menschen die elektromagnetische Strahlung sehen können. Es ist außerdem ein Bereich, der von der Atmosphäre nahezu vollständig durchlässig ist (Abb.2). Daher spielt die Wellenlänge des sichtbaren Lichtes eine enorm große Rolle in der Fernerkundung. An den Bereich des sichtbaren Lichtes grenzt die Infrarotstrahlung, diese deckt mit Wellenlängen von 0,78 µm – 1 mm ein deutlich größeres Spektrum ab, als es beim sichtbaren Licht der Fall ist. Die Infrarotstrahlung weist immer wieder Wellenlängen auf, die durchlässig für die Atmosphäre sind. Für die Fernerkundung werden dabei vor allem Strahlungen im nahen bis mittleren Infrarot verwendet, in einem Wellenlängenbereich von 0,78 µm – 8 µm. Auf die Infrarotstrahlung folgt die Microwellenstrahlung mit Wellenlängen von 1 mm – 30 cm. Microwellenstrahlung ist dabei vollständig durchlässig für die Atmosphäre und wird in der Fernerkundung mit dem Radarverfahren genutzt (Abb.2). An die Microwellenstrahlung grenzen die Radiowellen, diese haben in der Fernerkundung kaum eine Bedeutung, da die Wellenlängen zu groß sind, um Objekte noch ausreichend aufzulösen (Albertz, 2009).

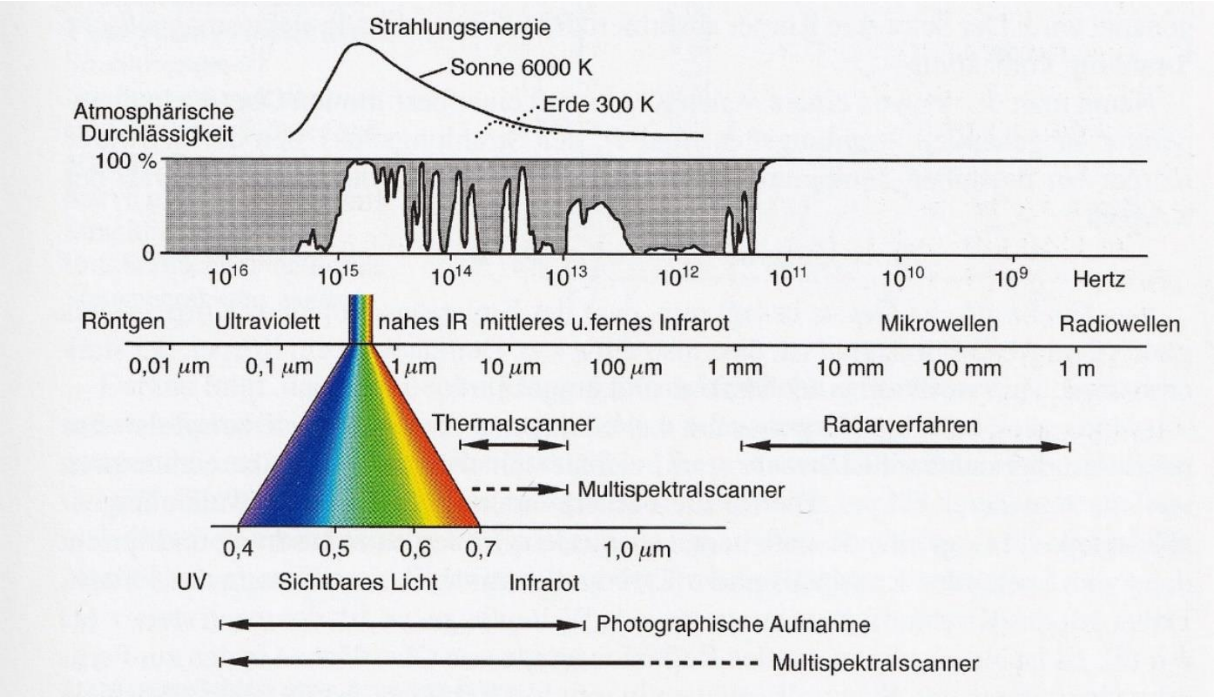

Abbildung 2 Elektromagnetisches Spektrum mit Durchlässigkeit der Atmosphäre eingezeichnetem Bereich des sichtbaren Lichts und Bereichen, in denen Fernerkundungsmethoden angewandt werden (Albertz, 2009)

Bei allen fernerkundlichen Aufnahmen durchläuft elektromagnetische Strahlung bzw. Schallwellen ein Medium. In den meisten Fällen ist dies die Atmosphäre der Erde, in selteneren Fällen kann dieses Medium allerdings auch Wasser sein z.B. bei der Nutzung eines Echolotes. Da für diese Arbeit nur die Aufnahmen durch die Atmosphäre eine Rolle spielen, werden nur ihre Einflüsse auf die Aufnahmen weiter beschrieben (Albertz, 2009).

Fällt elektromagnetische Strahlung durch die Atmosphäre auf ein strahlungsdurchlässiges Objekt kommt es zu drei verschiedenen Effekten: zum einen wird ein Teil der elektromagnetischen Strahlung absorbiert, ein anderer Teil wird reflektiert und wiederum ein anderer Teil der Strahlung durchdringt das Objekt. Dieser Prozess wird Transmission genannt. Bei einem strahlungsundurchlässigen Objekt ist, der Strahlungsfluss, der auf das Objekt trifft, gleich der Summe aus Absorption und Reflektion. Die Fernerkundung macht sich dabei stark zum Nutzen das verschiedene Objektkategorien, wie z.B. Wasser, Vegetation oder offener Fels sehr unterschiedliche Absorptions- und Reflektionseigenschaften besitzen. So absorbiert beispielsweise Wasser im nahen Infrarot fast die gesamte Strahlung, während das Chlorophyll im Pflanzengrün den großen Teil des nahen Infrarots reflektiert. In einer Satellitenaufnahme oder Luftbildaufnahme im nahen Infrarot, wären dann Wasserflächen als schwarze Flächen zu erkennen und Vegetationsflächen mit hohem Blattgrünanteil fast weiß, da der Sensor nur die reflektierte Strahlung aufnimmt. Es ist daher bei einer fernerkundlichen Untersuchung enorm wichtig, dass die spezifischen Reflektionseigenschaften, der im Untersuchungsgebiet vorkommenden Klassen bekannt sind, um so bestmöglich die Klassen durch einen Algorithmus trennen zu können (Albertz, 2009).

Um die Reflektionseigenschaften von Klassen noch genauer zu bestimmen, können beispielsweise sogenannter Feld-Goniometer zum Einsatz kommen. Dies ist ein Gerät, kann mit Hilfe eines Radiometers die Reflektionseigenschaften des Objekts ermittelt. Dabei ist es auch möglich das Radiometer aus unterschiedlichen Richtungen auf das Objekt zu richten. Die so erhaltenen Reflektionseigenschaften können gerade bei schwer unterscheidbaren Klassen noch einmal wichtige Aufschlüsse auf Unterschiede in den Reflektionseigenschaften liefern (Albertz, 2009).

2.1.1 Moorspezifische Klassifikationseigenschaften

Die zeitlich hochdynamischen Wechsel der Landbedeckung im Moor, ist ein Charakteristikum der Moorlandschaft, welches bei der Klassifizierung von Moorflächen berücksichtigt werden muss. So können die Wasserstände je nach Jahreszeit stark variieren. Im Winter und im Frühjahr sammelt sich das Wasser auf den Flächen, während es im Sommer verdunstet. So kann es Flächen geben, welche im Winter komplett mit Wasser bedeckt sind, im Sommer allerdings trockenfallen. Des Weiteren lassen sich die Klassen oftmals nicht klar voneinander trennen. So können auf Wasserflächen im Moor bestimmte Torfmoose an der Wasseroberfläche aufschwimmen. Dadurch vermischt sich das spezifische Spektralsignal der Klassen gerade in den Randbereichen zwischen zwei oder mehreren Klassen. Dies muss bei der Erstellung der Trainingsdaten berücksichtigt werden (Gallant, 2015; Mahdavi et al., 2018).

2.1.2 Satellitenbilddaten

Bei der Aufnahme von Satellitenbilddaten ist der Sensor auf einem Satelliten eingebaut, der dann die Erdoberfläche aufnimmt. Eine häufig gewählte Konfiguration der Bahnparameter, auch für andere Fernerkundungssatelliten, sind die der Landsat-Satelliten. So weisen die Landsat-Satelliten eine kreisförmige und polnahe Umlaufbahn auf, dadurch wird systematisch fast die gesamte Erdoberfläche überflogen. Einzig die Pole können mit diesen Bahnparametern von den Sensoren der Landsat-Satelliten nicht erfasst werden. Eine weitere Charakteristik der Bahnparametern der Landsat-Satelliten ist das diese sonnensynchron sind, dadurch überfliegen die Satelliten immer zur selben Uhrzeit den Äquator. Damit können durch dieselbe Aufnahmezeit gleichbleibende Aufnahmebedingungen geschaffen werden. Die Satellitenbahn hält dabei ihre Position im Raum und die Erdoberfläche bewegt sich unter dem Satelliten. Dadurch, dass sich die Lage der Satellitenbahn nicht verändert, wird durch die Rotation der Erde, die Bodenspur, die der Satellit aufnimmt, etwas zum vorherigen Umlauf verschoben (Albertz, 2009; ESA, 2012; U.S. Geological Survey, 2019b, 2019a).

In vielen Fernerkundungssatelliten wird als Aufnahmesystem eine Zeilenkamera genutzt, die das Untersuchungsgebiet Zeile für Zeile abtastet. Als Sensor dient bei einem solchen System ein CCD-Sensor (Charge-Coupled Device), der die Photonen des eingefangenen Lichtes von der Erdoberfläche in elektrische Ladungen umwandelt, welche wiederrum gemessen werden können und sich proportional zur ankommenden Lichtmenge verhalten. Ein solcher Sensor benötigt in der Regel keine mechanischen Teile beim Aufnahmevorgang und ist mit einer Videoaufnahme vergleichbar. Fernerkundungssatelliten die eine solche Zeilenkamera verwenden sind beispielsweise die Fernerkundungssatelliten, der SPOT-Reihe, die der Sentinell-2 Mission oder aber auch die letzten beiden Satelliten der Landsat Mission Landsat 8 und 9. In den früheren Satelliten der Landsat Reihe wird ein Aufnahmesystem genutzt, bei dem das ankommende Licht über einen rotierenden Spiegel dem Detektor zugeführt wird. Dadurch wird nicht die gesamte Zeile auf einmal aufgenommen, sondern nur ein Teil der Zeile. Um die gesamte Zeile aufnehmen zu können müssen dabei die Rotation des Spiegels und die Fluggeschwindigkeit genau aufeinander abgestimmt sein. Der Landsat 7 Satellit weist einen Defekt in diesem Spiegelsystem auf, so dass es in den Aufnahmen zu schwarzen Balken kommt (Albertz, 2009; ESA, 2012; U.S. Geological Survey, 2019b, 2019a).

2.1.3 Luftbildaufnahmen

Luftbildaufnahmen, werden anders als Satellitenbilddaten von einem Flugzeug oder einer Drohne aus durchgeführt. In den meisten Fällen dient dabei ein umgerüstetes Flugzeug, welches über einen Sensor lotrecht zur Flugrichtung über ein Loch im Rumpf Luftbildaufnahmen tätigen kann. Dabei gibt es zwei verschiedene Aufnahmesysteme. Zum einen sind dies Flächenkameras, die ähnlich wie bei einem normalen Foto eine gesamte Fläche aufnehmen und zum anderen werden auch Zeilensensoren genutzt. Anders als bei den Aufnahmen aus einem Satelliten unterliegt ein Flugzeug den Turbolenzen

der Atmosphäre. Die Position des Sensors ist somit nie komplett stabil im Raum. Dies ist vor allem ein Problem bei der Verwendung von Zeilensensoren, da sich nach jeder Aufnahme einer Zeile die Position des Sensors ändert. Es entstehen teilweise starke Verzerrungen des Luftbildes. Bei den Aufnahmen mit einer Flächenkamera muss allerdings genauso, wie es bei einem Zeilensensor der Fall ist, die Position des Sensors zum Zeitpunkt der Aufnahme bekannt sein. Um das Luftbild dann nutzen zu können, müssen diese erst einmal geometrisch entzerrt werden, dafür ist die genaue Position des Sensors im Raum zu allen Aufnahmezeitpunkten erforderlich. Da ein GNSS-Signal nur alle paar Sekunden empfangen wird und ein Flugzeug sich mit erheblicher Geschwindigkeit fortbewegt, reicht dies nicht aus, um die Lage des Sensors im Raum zu bestimmen. Um dieses Problem zu beheben, wird das GNSS-Signal mit einer INS (Inertial Navigation System) gekoppelt. Die INS misst jede Beschleunigung und Drehraten, so dass sich zwischen den Positionsdaten des GNSS-Signals und den Daten der INS die Position des Sensors genau bestimmen lässt (Albertz, 2009).

Luftbildaufnahmen sind immer zentralperspektivisch aufgenommen mit der Nähe zum Aufnahmerand erhöht sich die Verzerrung immer mehr (Abb. 3). Um eine geringere Verzerrung in den Luftbildaufnahmen zu den Bildrändern zu erhalten, werden die Luftbildaufnahmen überlappend durchgeführt. So lässt sich ein aufgenommenes Gebiet aus den zentralen Bereichen mehrerer überlappender Luftbildaufnahmen zusammensetzen. Die Verzerrungen zu den Rändern können so deutlich verringert werden. Damit die Luftbildaufnahmen für kartographische Produkte genutzt werden können, müssen die Aufnahmen in eine Parallelprojektion überführt werden (Abb.3). Dies wird mit Hilfe eines Geländemodells bewerkstelligt. Die Bildpunkte werden auf die äquivalente Position im Geländemodell gesetzt und so entzerrt. Die so entzerrten Luftbilder werden Orthofotos genannt und weisen durch ihre Parallelprojektion in jedem Punkt einen lotrechten Aufnahmepunkt auf, während die unbearbeiteten Luftbilder, die in der Zentralprojektion aufgenommen sind, nur einen lotrechten Punkt zum Aufnahmesensor aufweisen (Albertz, 2009).

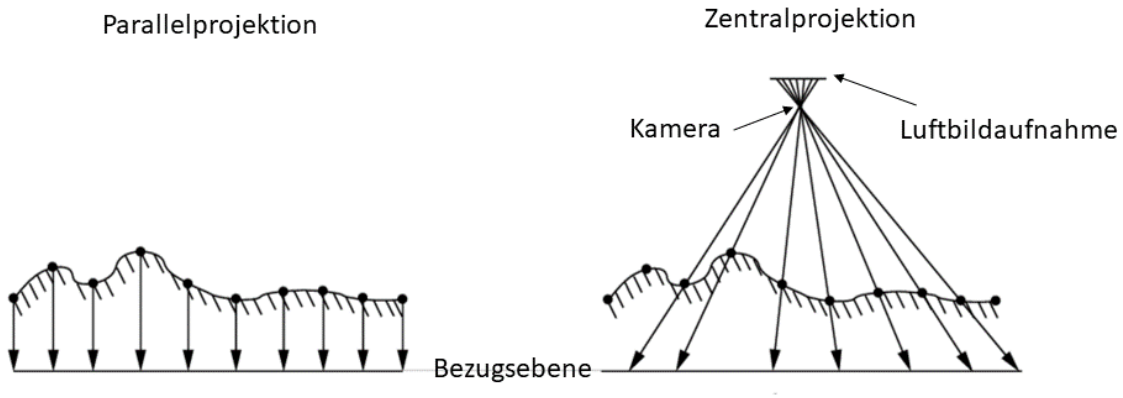

Abbildung 3 Unterschiede zwischen Zentralprojektion und Parallelprojektion. *nach* (Albertz, 2009)

2.2 Machine Learning

2.2.1 Arten der Klassifizierung

Im Folgenden werden die Arten der Klassifizierung: Unüberwachte-, überwachte-, semiüberwachte-, objektbasierte- und pixelbasierte Klassifizierung genauer erläutert. Es ist zu beachten, dass es sich entweder um eine unüberwachte Klassifizierung, semiüberwachte Klassifizierung oder um eine überwachte Klassifizierung handelt. Dabei können diese wiederum entweder mit einer objektbasierten Klassifizierung oder einer pixelbasierten Klassifizierung kombiniert werden. Bei komplexeren Arbeitsabläufen kann zunächst eine unüberwachte Klassifizierung verwendet werden. Dadurch kann

abgeschätzt werden welche Klassen existieren, um im Anschluss mit den erhaltenen Klassen noch einmal eine überwachte Klassifizierung durchzuführen. Dabei sind die Methoden aber immer noch voneinander getrennt und nur in einem Arbeitsablauf miteinander verbunden (Batta, 2020; Blaschke, 2010; Zerrouki & Bouchaffra, 2014).

2.2.1.1 Unüberwachte Klassifizierung

Bei der unüberwachten Klassifizierung werden vom Nutzer dem Algorithmus keine Klassen vorgegeben. Der Algorithmus soll anhand der vorgelegten Daten entscheiden welche Datenpunkte zu eigenen Klassen zusammengefasst werden können. Unüberwachte Klassifizierungen eignen sich besonders, wenn keine genaueren Daten für vorkommende Klassen vorliegen und eignen sich daher gut, um einen ersten Eindruck von einem Datensatz zu erhalten. Ein häufig hierfür genutztes Verfahren ist dabei die Cluster-Analyse, bei der Datenpunkte in Cluster zusammengefasst werden und eine Klasse bilden (Olaode et al., 2014).

2.2.1.2 Überwachte Klassifizierung

Anders als bei der unüberwachten Klassifizierung werden bei der überwachten Klassifizierung Klassen vom Nutzer vorgegeben. Der Klassifizierungsalgorithmus wird dabei mit Trainingsdaten einer Klasse spezifisch trainiert, um diese in einem unbekannten Datensatz zu identifizieren. Überwachte Klassifizierung eignet sich vor allem dann, wenn sich die Klassen im Untersuchungsgebiet gut abstecken lassen oder eine spezifische Klasse extrahiert werden soll. Für eine überwachte Klassifizierung stehen unterschiedliche Verfahren zur Wahl, wie z.B. k-Nearest Neighbors, Random Forest oder Support Vector Machine (Batta, 2020).

2.2.1.3 Semi-überwachte Klassifizierung

Die semi-überwachte Klassifizierung ist eine Mischform aus der unüberwachten- und der überwachten Klassifizierung. So enthalten nur Teile der Trainingsdaten eine zugewiesene Klasse und ein anderer Teil nicht. Das Verfahren eignet sich daher gut, wenn sich das Erhalten von Trainingsdaten schwierig gestaltet (Batta, 2020).

2.2.1.4 Objektbasierte Klassifizierung

Bei den objektbasierten Klassifizierungen wird noch vor dem Schritt der eigentlichen Klassifizierung ein Algorithmus zur Mustererkennung angewandt, dieser soll Objekte in einem Luftbild oder Satellitenbild erkennen und zusammenfassen. So könnte beispielsweise der Algorithmus eine Dachfläche als ein Objekt und einen Garten als ein anderes Objekt ausgeben. Diese Objekte werden dann vom eigentlichen Klassifizierungsalgorithmus anhand ihrer spektralen Eigenschaften einer Klasse zugeordnet. Ein Vorteil dieser Methode ist, dass die Klassen einem Objekt zugewiesen werden (Abb.4). Dies kann allerdings auch ein Nachteil sein, da es zu einer Generalisierung kommt, wenn nicht jeder Pixel für sich bewertet wird, sondern direkt ein ganzes Objekt einer Klasse zugewiesen wird. Deshalb ist die Methode der objektbasierten Klassifizierung mit hochauflösenden Luft- oder Satellitendaten eher zu empfehlen (Blaschke, 2010; Zerrouki & Bouchaffra, 2014).

2.2.1.5 Pixelbasierte Klassifizierung

Anders als bei der objektbasierten Klassifizierung wird bei der pixelbasierte Klassifizierung kein extra Algorithmus vor die eigentliche Klassifizierung geschaltet. Es wird stattdessen jeder Pixel von dem Klassifizierungsalgorithmus einzeln einer Klasse zugewiesen (Abb. 4). Das so erhaltene Ergebnis kann dabei etwas unruhiger wirken als das einer objektbasierten Klassifizierung, da jeder Pixel theoretisch jede Klasse annehmen kann. Allerdings ist die pixelbasierte Klassifizierung besser in der Lage auch kleinste Objekte richtig zuzuweisen, die eventuell in einer objektbasierten Klassifizierung

fälschlicherweise einem anderen größeren Objekt zugeordnet werden würden (Zerrouki & Bouchaffra, 2014).

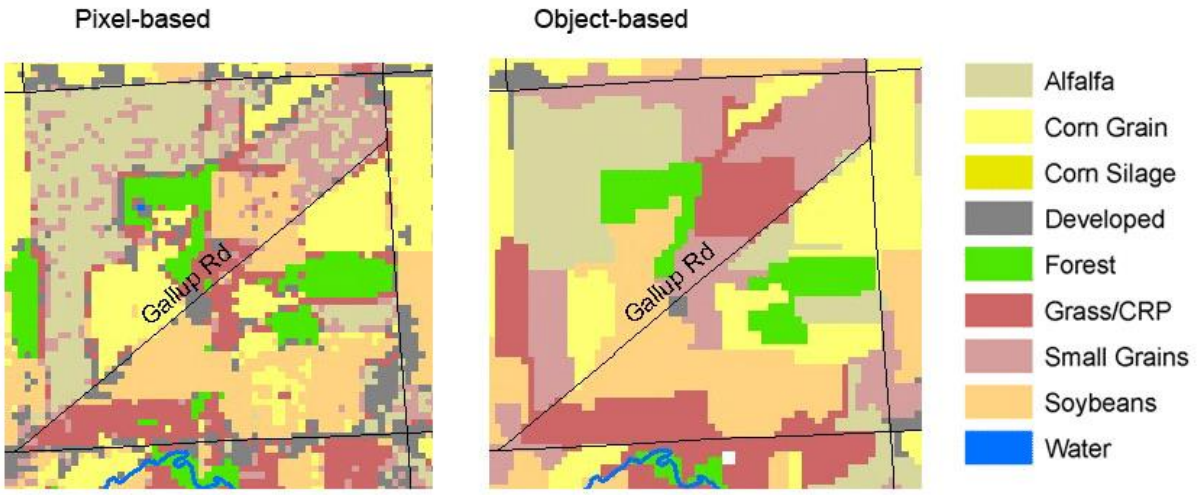

Abbildung 4 Unterschiede zwischen pixelbasierter Klassifizierung und objektbasierter Klassifizierung am Beispiel von unterschiedlicher Landbedeckung nach (Powell & Brooks, 2023).

2.2.2 Overfitting und Underfitting

Overfitting und Underfitting sind Zustände eines Machine Learning Modells, die auf Grund einer zu geringen Anpassung an die Trainingsdaten bzw. zu einer zu hohen Anpassung an die Trainingsdaten, auftreten. Bei einem Underfitting Modell kann es beispielsweise dazu kommen, dass das Modell einen linearen Zusammenhang in den Daten erkennt, obwohl diese optimalerweise als nicht linear zu beschreiben wären (Abb. 5). Das Modell ist dementsprechend nicht gut genug an die Trainingsdaten angepasst und sollte deswegen noch einmal überarbeitet werden, beispielsweise durch die Erhöhung der Lernzyklen. Bei Underfitting hat das Modell eine hohe Verzerrung. Es treten dabei starke Unterschiede zwischen den Werten des Modells und den des Trainingsdatensatzes auf. Ein Modelll bei dem Underfitting auftritt hat einen hohen systematischen Fehler oder auch Bias (Wilmott, 2020). Im Falle des Overfitting beschreibt das Modell zu genau den Trainingsdatensatz und ist daher kaum noch in der Lage zu generalisieren. Im extremen Fall von Overfitting beschreibt das Modell jeden einzelnen Trainingsdatenpunkt und gewichtet so den Noise in den Daten zu stark (Abb. 5). Ein Modell bei dem Overfitting auftritt hat eine hohe Varianz, dadurch wird das Modell bei einem anderen Datensatz, außer dem Trainingsdatensatz, eine hohe Streuung aufweisen. Um Overfitting zu vermeiden, könnten beispielsweise die Lernzyklen des Modells reduziert werden (Hawkins, 2004; Van Der Aalst et al., 2010).

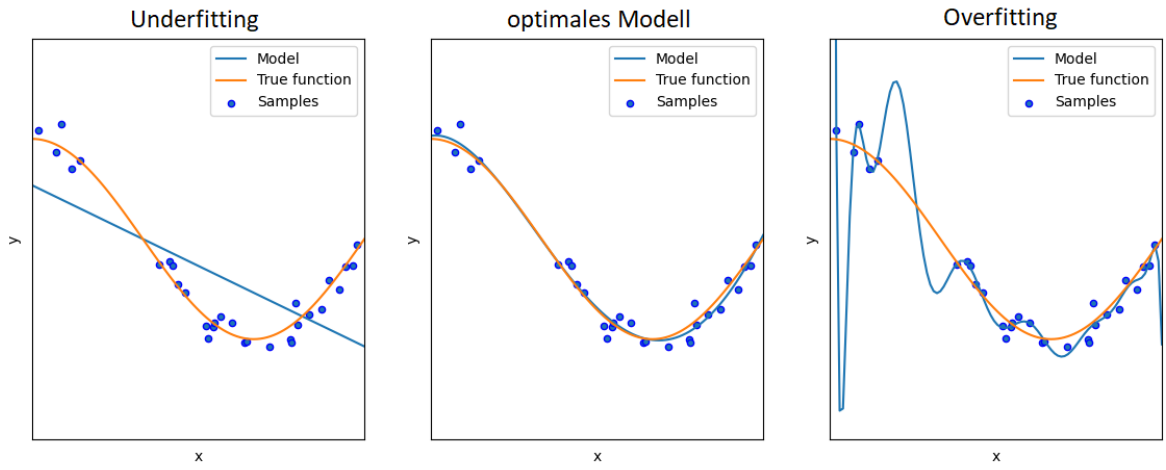

Abbildung 5 Unterschiedliches Verhalten des Modells für Underfitting, eine optimale Anpassung des Modells und bei Overfitting. nach (https://scikit-learn.org/stable/auto_examples/model_selection /plot_underfitting_overfitting.html). Letzer Aufruf: 20.04.2023

3. Daten und Methodik

3.1 Evaluation der Satellitendaten

Um die hochdynamischen Veränderungen in den Wasserständen im Moor abbilden zu können und auch diese zu analysieren, werden zeitlich gut aufgelöste Daten benötigt. Denn nur so können mehrere Satellitenbilder innerhalb eines Jahres erhalten werden. Gerade in Mitteleuropa kann eine wolkenfreie Satellitenaufnahme nur selten erreicht werden. Es ist daher unerlässlich, dass der Satellit häufig denselben Punkt der Erdoberfläche aufnehmen kann. Allerdings ist die räumliche Auflösung der Satellitendaten ebenfalls wichtig. In wieder vernässten Moorflächen gibt es oftmals kleinteilige Wechsel der Vegetation mit temporären Wasserflächen. Mit einer hohen räumlichen Auflösung lassen sich diese Details besser in der Klassifizierung wiederfinden. Die Daten müssen demnach sowohl eine hohe zeitliche Auflösung als auch eine möglichst hohe räumlich Auflösung aufweisen. Da beide Auflösungen sich allerdings zueinander umgekehrt proportional verhalten, ist ein Kompromiss zwischen beiden Auflösungen zu finden.

Eine Option für solche Satellitendaten ist, die seit 1972 aktive **Landsat-Mission**, diese können über die Website Earthexplorer des USGS heruntergeladen werden. Die Daten, reichen von 1972 bis zum jetzigen Tag. Die ersten 3 Satelliten der Landsat-Mission verfügen nur über eine räumliche Auflösung von 80 x 80 m pro Pixel, ab dem vierten Satelliten der Landsat-Mission weisen die Satellitenbilddaten eine räumliche Auflösung in ihren Spektralkanälen von 30 x 30 m pro Pixel auf. Seit dem siebten Landsat-Satelliten, der 1999 gestartet ist, gibt es zusätzlich einen panchromatischen Kanal mit einer Auflösung von 15 x 15 m. Mit Hilfe des panchromatischen Kanals könnte das sogenannte Pansharpening durchgeführt werden, bei dem die schlechter aufgelösten Spektralkanäle mit dem panchromatischen Kanal kombiniert werden, um ein höher aufgelöstes Farbbild zu erhalten. Die Satelliten ab dem vierten Landsat-Satelliten nehmen alle 16 Tage dasselbe Gebiet auf der Erdoberfläche auf (Albertz, 2009; U.S. Geological Survey, 2019a, 2019b).

Eine andere Satelliten-Mission die die Anforderungen dieser Arbeit erfüllen kann, sind die Satelliten der **Sentinel-2-Mission** deren erster Erdbeobachtungssatellit 2015 in den Erdorbit gebracht wurde. Etwa zwei Jahre später wurde der zweite Erdbeobachtungssatellit der Sentinel-2 Mission in den

Erdorbit gebracht. Die beiden Satelliten teilen sich dieselbe Umlaufbahn sind aber um 180° Grad zueinander versetzt. Dadurch wird alle fünf Tage derselbe Punkt auf der Erdoberfläche aufgenommen. Die Auflösung der Spektralkanäle reicht dabei von 10 x 10 m pro Pixel zu 60 x 60 m pro Pixel. Die Kanäle im Spektralbereich von Rot, Blau, Grün und einen im nahen Infrarot haben eine Auflösung von 10 m pro Pixel. Weitere sechs Kanäle im nahen Infrarot weisen eine Auflösung von 20 m pro Pixel auf. Anders als bei den Satelliten der Landsat-Mission wurde bei den Sentinel Satelliten auf einen panchromatischen Kanal verzichtet. Die Daten der Sentinel-2-Mission sind, wie die Daten der Landsat-Mission, frei verfügbar zum Download. Als Datenportal dient die Website der Erdbeobachtungsmission der ESA, Copernicus. Außerdem ist es möglich, mit einem Python-Skript die Daten über eine API direkt herunterzuladen (Albertz, 2009; ESA, 2012).

Als weitere potenzielle Quelle für Satellitendaten ist die **SPOT-Mission** der französischen Weltraumbehörde CNES zu nennen. Der erste SPOT-Satellit wurde 1986 in den Orbit gebracht, seitdem zeichnen die Satelliten der SPOT-Mission die Erdoberfläche auf. Die aktuell aktiven SPOT-Satelliten sind SPOT-Satellit 6 und 7, beide Satelliten sind baugleich und liefern hochauflösende Daten in den vier Spektralkanälen, blau, grün, rot und nahem Infrarot mit einer Auflösung von 6 x 6 m pro Pixel. Die Satelliten verfügen außerdem über einen panchromatischen Kanal mit einer Auflösung von 1,5 x 1,5 m pro Pixel. Die SPOT-Satelliten überfliegen alle 26 Tage denselben Punkt auf der Erdoberfläche. Eine Besonderheit der SPOT-Satelliten ist der schwenkbare Sensor und dadurch die Möglichkeit Gebiete mit einem besonderen Interesse deutlich häufiger aufzunehmen. Der schwenkbare Sensor wird dabei von Airbus Defense und Space kontrolliert. Anwendung kann so etwas beispielsweise bei einer Naturkatastrophe finden, nach der dringend immer wieder aktuelle Satellitendaten benötigt werden. Auch kann mit so einem System das Problem mit der Wolkenbedeckung umgangen werden, da der Sensor statt eines wolkenbedeckten Gebietes ein wolkenfreies aufnehmen kann. Daher ist die zeitliche Auflösung bei den SPOT-Satelliten deutlich schwerer konkret anzugeben, als es aus den reinen technischen Daten abzulesen ist. Die Daten sind anders als bei der Landsat- und der Sentinel-2-Mission nicht frei verfügbar. Eine Nutzung für wissenschaftliche Zwecke kann über das Copernicus Datenportal der ESA beantragt werden (Albertz, 2009; Airbus Defence and Space, n.d).

Über ein weiteres Portal von Copernicus lassen sich, die Metadaten aller Satellitenmission, die ihre Daten mit dem Copernicus Programm teilen, einsehen. So kann eine AoI(Area of Intrest) für das Untersuchungsgebiet eingezeichnet werden und das Portal gibt alle verfügbaren Datensetzte für das Untersuchungsgebiet an. Im Falle der SPOT-Mission sind 27 Datensätze für das Untersuchungsgebiet des Leegmoores zu verzeichnen, die sich über einen Zeitraum von 2015 bis 2019 erstrecken. Zusätzlich lassen sich zum Untersuchungsgebiet auch Parameter wie Wolkenbedeckung einstellen. Bei einer Wolkenbedeckung von maximal 15% weisen die gefundenen 27 Datensätze alle eine zu hohe Wolkenbedeckung auf, sie können daher für diese Arbeit nicht verwendet werden.

Für diese Arbeit kommen nur Daten der Sentinel-2-Mission in Frage, da diese sowohl eine gute räumliche Auflösung mit 10 x 10 m pro Pixel als auch eine gute zeitliche Auflösung bieten (Abb. 6). So können ab 2015 für das Untersuchungsgebiet im Leegmoor 158 Datensätze mit einer Wolkenbedeckung unter 15% über die Copernicus API heruntergeladen werden (Airbus Defence and Space, n.d.; Albertz, 2009; ESA, 2012; U.S. Geological Survey, 2019a, 2019b).das

Kriterien zur Wahl der Satellitendaten	Zur Auswahl stehende Satellitendaten		
	Landsat	Sentinel - 2	SPOT
Zeitliche Auflösung (Tage)	16	5	26
Räumliche Auflösung [m x m]	15 / 30	10 / 20	1,5 / 6
Spektrale Auflösung (Anzahl Kanäle)	11	13	5
Proprietär	✗	✗	✓

Abbildung 6 Tabellarische Evaluation der für diese Arbeit verwendeten Satellitendaten Nach (ESA, 2012; Albertz, 2009; Airbus Defence and Space, n.d)

3.2 Sentinel-2 Daten

Die Sentinel-2-Daten werden in unterschiedlichen Produkten über das Copernicus Datenportal angeboten. Zum einen werden Level-1C-Satellitendaten angeboten, bei denen es sich um Top Of Atmosphere (TOP) Daten handelt. Mit diesem Datenprodukt wird die gesamte elektromagnetische Strahlung einer Wellenlänge, die von der Erdoberfläche und Atmosphäre ausgeht, vom Sensor erfasst. Zusätzlich sind die Daten geometrisch über ein digitales Höhenmodell korrigiert (ESA, 2012).

Als zweites Datenprodukt wird über das Copernicus Datenprotal für die Sentinel-2-Mission Level-2A-Daten angeboten. Im Unterschied zu den Level-1C-Daten sind diese sogenannte Bottom Of Atmosphere (BOA) Daten. Bei diesen Daten wurde der Einfluss der atmosphärischen Strahlung auf die am Sensor detektierte Strahlung bereinigt, so dass nur die an der Erdoberfläche reflektierte Strahlung im Datenprodukt berücksichtigt wird. Die Level-2A-Datenprodukte basieren auf den Level-1C Daten, daher lassen sich auch alle Level-1C Daten in Level-2A-Daten umrechnen. Für eine Umrechnung stellt das Copernicus Datenportal eine eigene Sentinel-2-Toolbox bereit. Es lassen sich allerdings auch alle Sentinel-2 Datensätze direkt als Level-2A-Datenprodukte herunterladen (ESA, 2012).

Für eine Klassifikation der Erdoberfläche sind BOA-Datenprodukte zu verwenden, da anhand der Reflektionseigenschaften der Erdoberfläche die Trennung zwischen den Klassen erfolgt. Die Sentinel-2-Datenprodukte enthalten 13 Spektralkanäle (Abb. 7), von denen vier Bänder mit einer räumlichen Auflösung von 10 x 10 m pro Pixel, sechs Bänder mit einer räumlichen Auflösung von 20 x 20 m pro Pixel und drei Bänder mit einer räumlichen Auflösung von 60 x 60 m pro Pixel aufgenommen werden. Die vier Bänder mit der 10 m Auflösung befinden sich im sogenannten VNIR (Visible and Near-Infrared) also im Bereich des sichtbaren Lichtes und im nahen Infrarot. Die Kanäle im Bereich der 20-m-Auflösung liegen alle im infraroten Bereich des elektromagnetischen Spektrums (Abb. 7). Die drei Bänder mit einer räumlichen Auflösung von 60 m detektieren sowohl die Strahlung im elektromagnetischen Spektrum des sichtbaren Lichtes als auch im infraroten Spektrum (ESA, 2012; Louis et al., 2016).

Für diese Arbeit werden, um einen konsistenten Datensatz zwischen allen ML-Methoden zu behalten, die folgenden Bänder genutzt: Band 2, Band 3, Band 4, Band 8, Band 11 und Band 12 (Abb. 7). Diese Auswahl ist vor allem den Anforderungen des „DeepWaterMap“ Modells geschuldet, das Bänder mit diesen Spektraleigenschaften benötigt (F. Isikdogan et al., 2017).

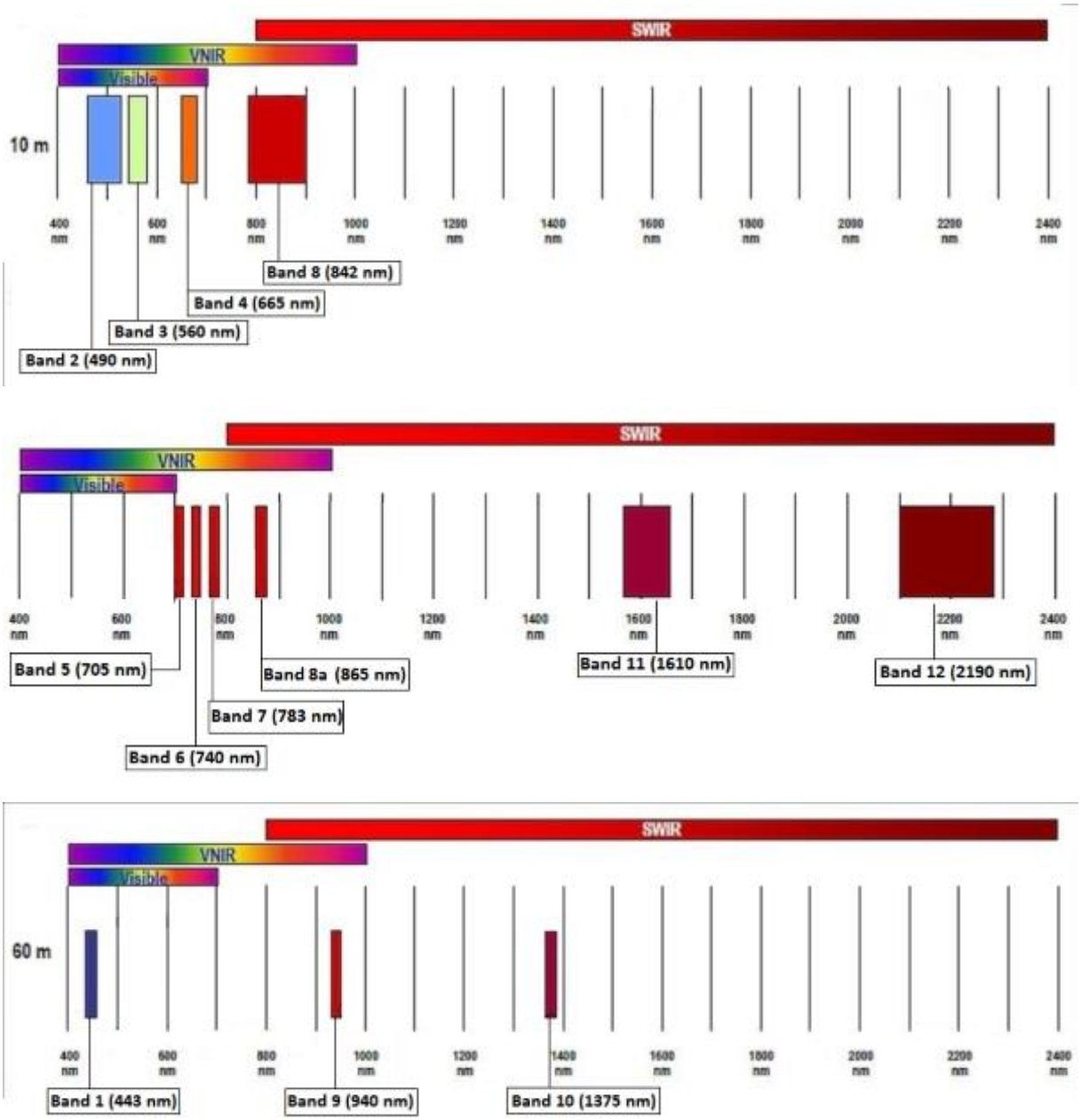

Abbildung 7 Alle Sentinel-2-Bänder verzeichnet im Spektralbereich zwischen 400 nm und 2400 nm. Oberste Abbildung enthält dabei alle Bänder mit den Auflösungen 10 x 10 m pro Pixel, die mittlere Grafik alle Bänder der Auflösung 20 x 20 m und die unterste Grafik enthält alle Bänder mit einer Auflösung von 60 x 60 m. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial; Letzter Aufruf: 20.04.2023

3.3 Drohnenorthofoto des Gebietes

Auf dem Gebiet des Leegmoors fand am 07.03.2020 eine Drohnenaufnahme statt mit einer Auflösung von 2 cm pro Pixel. Diese Aufnahme ist entzerrt und „downsampled“ auf eine Auflösung von 10 cm pro Pixel und dient als Grundlage für den Ground Truth Datensatz, mit welchem die Validierung der Klassifikationsergebnisse stattfindet.

3.4 Evaluation der verwendeten Machine-Learning-Algorithmen

Für diese Arbeit sollen mehrere Machine Learning (ML) Algorithmen auf ihre Eignung für die Klassifikation von Wasserflächen in Moorgebieten getestet werden. Es sollen für die Arbeit drei klassische Algorithmen und ein Deep-Learning-Algorithmus gewählt werden, um die Ergebnisse der verschiedenen Algorithmen Vergleichen zu können. Jeder dieser Algorithmen soll mit jeweils vier unterschiedlichen Sätzen an Parametern Ergebnisse liefern, um die besten Einstellungen für die Klassifikation der Wasserflächen im Moor zu finden. Einer dieser vier Parameter-Sätze soll dabei die Standardeinstellung des verwendeten Programms sein, um zeigen zu können, ob es signifikante Verbesserungen in den Ergebnissen gibt, wenn die Parameter abweichend vom Standard gewählt werden. Die ausgewählten Algorithmen sollen sich von ihrer Funktionsweise deutlich voneinander unterscheiden und weit verbreitet sein, damit sichergestellt werden kann, dass die Methoden auf möglichst vielen Plattformen umsetzbar sind. Als Grundlage für die Unterteilung der ML-Algorithmen wird die Arbeit von Batta (2020) herangezogen (Abb. 8).

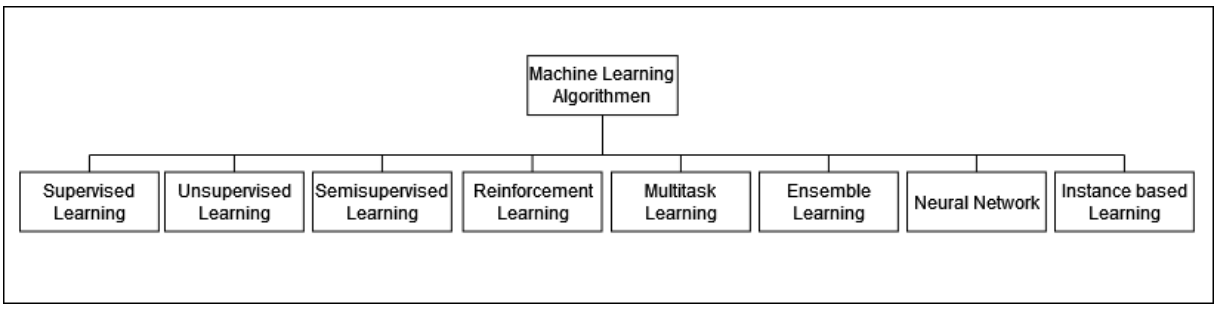

Abbildung 8 Schematische Unterteilung von Machine-Learning-Algorithmen nach (Batta, 2020).

Die Kategorien überwachtes Lernen, unüberwachtes Lernen und semiüberwachtes Lernen, können als eigene Unterkategorien für maschinelles Lernen gesehen werden. Alle fünf restlichen Kategorien wenden zum Teil auch Prinzipen des überwachten, unüberwachten und semiüberwachten Lernens an. Für eine bessere Übersichtlichkeit werden Algorithmen, die den fünf spezielleren Kategorien nicht zugeordnet werden können, unter den Kategorien des überwachten, unüberwachten und semiüberwachten Lernen zusammengefasst.

Die **überwachten Machine-Learning-Algorithmen** funktionieren über die Vorgabe der Klassen. Mit Hilfe der Trainingsdaten wird vom Algorithmus nach dem Schema der vordefinierten Klassen ein Modell erstellt. Für die Evaluation des so erstellten Modells wird der Gesamtdatensatz in ein Trainingsdatensatz und einen Testdatensatz gesplittet. Die Evaluation des Modells findet dann am Testdatensatz statt, den der Algorithmus zuvor für die Modellbildung noch nicht gesehen hat. Häufig verwendete Algorithmen in dieser Kategorie sind folgende Methoden:

Entscheidungsbaum

Entscheidungsbaum Algorithmus funktioniert ähnlich wie ein Flussdiagramm. Der Datensatz wird an jedem Knotenpunkt durch eine neue Bedingung aufgeteilt bis an den sogenannten Blättern das gewünschte Klassifikationsergebnis erreicht ist. Das Modell eines Entscheidungsbaums ist dardurch gut nachvollziehbar. Ein Nachteil dieser Methode ist allerdings, dass diese zu Overfitting neigt, das Modell sich also zu sehr an die Trainingsdaten anpasst (Batta, 2020; Wilmott, 2020).

Naiver Bayes-Klassifikator

Der Naive Bayes Algorithmus berechnet Wahrscheinlichkeiten mit der ein Datenpunkt einer bestimmten Klasse zugeordnet wird. Der Algorithmus findet vor allem in der Textanalyse Anwendung beispielsweise in einem Spamfilter (Batta, 2020; Wilmott, 2020).

Support Vector Machine (SVM)

Der Support Vector Machine (SVM) Algorithmus trennt die Klassen mit Hilfe einer Ebene, deren Abstand zu den Klassenrändern maximiert wird (Batta, 2020; Wilmott, 2020).

Die Algorithmen für eine **unüberwachte Klassifikation** eignen sich, um beispielsweise Klassen in unbekannten Daten zu finden, indem der Datensatz nach seinen Merkmalen unterteilt wird. So ist eine häufig verwendete Methode das Finden von Clustern mit ähnlichen Merkmalen. Weil in dieser Arbeit die Klassen bekannt sind, werden keine unüberwachten Algorithmen zum Einsatz kommen (Batta, 2020; Wilmott, 2020).

Ähnlich sieht es für Algorithmen der **semi-überwachten Klassifikation** aus, diese werden vor allem eingesetzt, wenn sich die Akquisition von Trainingsdaten schwierig gestaltet. Da dies in dieser Arbeit nicht der Fall ist, werden semi-überwachte Algorithmen ebenfalls nicht zum Einsatz kommen (Batta, 2020).

Als weitere Unterkategorie des maschinellen Lernens kann das verstärkte Lernen oder **Reinforcement Learning** genannt werden. Die Besonderheit des Reinforcement Learning Algorithmus ist das dieser über einen sogenannten Agenten verfügt, dieser ist an eine Umgebung gebunden, in der der Agent eine Aktion wählen kann, diese Aktionen verändert die Umgebung und es ergibt sich ein neuer Zustand, indem der Agent wieder eine Aktion wählen kann. Die Aktion, die der Agent wählt wird bewertet und mit einer Belohnung oder mit einer Bestrafung geahndet. Der Algorithmus versucht die Anzahl der Belohnungen zu optimieren. Verwendung findet der Algorithmus besonders dann, wenn Resultate im Voraus nicht bekannt sind. Beispiele hierfür sind Algorithmen, die Spiele lernen zu spielen oder Werbung personalisieren (Wilmott, 2020).

Eine weitere ML-Klasse ist das sogenannte **Multi Task Learning**. Der Ansatz des Multi Task Learning ist, dass an einem Inputdatensatz gleich mehrere unterschiedliche Operationen verrichtet werden. So kann beispielsweise neben einer Klassifikation auch noch eine Bild-Segmentation in einem Modell als Ergebnis erhalten werden, für die sonst mit anderen ML-Algorithmen zwei verschiedene Modelle benötigen werden würden. Dies kann besonderes dann interessant sein, wenn wenig Rechenleistung vorhanden ist und das Modell so effizient wie möglich laufen muss (Batta, 2020; Caruana et al., 1997).

Als eine weitere Kategorie für ML-Algorithmen werden die **Ensemble-Learning-Algorithmen** genannt. Bei diesem Verfahren erstellen mehrere verschiedene oder auch gleiche ML-Algorithmen jeweils ein Modell. Im Falle einer Klassifikation wird dann per Mehrheitsentscheid die Klasse zugewiesen. Ein Vorteil dieser Methode ist, dass es zu weniger Overfitting kommt. Ein Beispiel für einen Ensemble-Algorithmus ist der Random Forest Algorithmus, bei dem mehrere Entscheidungsbäume verwendet werden (Batta, 2020; Breiman, 2001; Wilmott, 2020).

Neuronale Netzwerke sind, wie es der Name schon andeutet, dem menschlichen Gehirn mit seinen Neuronen nachempfunden. Das technische Neuron besteht aus einem Input, einer Gewichtung und einem Schwellenwert. Ist dieser überschritten wird der Input der entsprechenden Klasse für den Schwellenwert zugewiesen. Ein Neuronales Netzwerk verfügt allerdings nicht nur über ein Neuron, sondern über mehrere Neuronen-Schichten. Eine Unterkategorie des Neuronalen Netzwerks ist ein sogenannter Deep-Learning-Algorithmus, der über eine hohe Anzahl an Neuronen Schichten verfügt. Der Vorteil eines Neuronalen Netzwerk Modells ist, dass es unterschiedliche Inputdaten verwenden

kann. So kann das Modell eispielsweise nicht nur auf mit einer Art von Satellitenbilddaten arbeiten kann, solange eine gewisse Ähnlichkeit in der Datenstruktur vorhanden ist. Ein Nachteil von Neuronalen Netzwerk Modellen ist die große Menge an Trainingsdaten, die benötigt werden, um ein gut angepasstes Modell zu erhalten. Hierfür müsste aus einer Großzahl an Satellitendaten manuell die Klassen ausgewiesen werden (F. Isikdogan et al., 2017; Steinwendner & Schwaiger, 2020; Rezaee et al., 2018).

Als letzte ML-Klasse wird für diese Arbeit die Methode des **Instance-Based Learning** vorgestellt. Anders als bei den anderen bisher besprochenen ML-Methoden wird kein Modell erstellt. Stattdessen werden die zu klassifizierenden Daten werden direkt mit den Trainingsdaten verglichen und ihnen in Abhängigkeit ihrer ähnlichen Eigenschaften, eine Klasse zugewiesen. Ein Beispiel hierfür ist der k-Nearest Neighbors Algorithmus, bei dem die k- nächsten Nachbarn von einem zu bestimmenden Punkt betrachtet werden und per Mehrheitsentscheid die Klasse zugewiesen wird (Batta, 2020; Wilmott, 2020; Zhang, 2016).

In der Kategorie des **überwachten Lernens** wurden drei Algorithmen vorgestellt: Entscheidungsbäume, Naive Bayes und Support Vektor Machine. Der Entscheidungsbaum ist anfällig für Overfitting und wird daher nicht für diese Arbeit verwendet. Der Naive Bayes Algorithmus ist vor allem für textanalytische Fragestellungen interessant, wenn auch für die Anwendung in der Fernerkundung eher weniger (Batta, 2020; Wilmott, 2020).

Der **SVM-Algorithmus** ist gerade bei kleineren Untersuchungsgebieten in der Fernerkundung interessant, da der Algorithmus mit kleinen Trainingsdatensätzen gute Erfolge erzielen kann (Wilmott, 2020).

Die Methoden des **Unüberwachten** sowie **Semiüberwachten Lernens** finden vor allem dann Anwendung, wenn wenig über die vorhandenen Klassen bekannt ist. Da diese Arbeit speziell Wasserflächen extrahieren möchte, werden diese Algorithmen in dieser Arbeit nicht verwendet. Aus einem ähnlichen Grund können auch Multitask Algorithmen ausgeschlossen werden, da diese darauf ausgelegt sind, mehr als nur ein Ergebnis mit einem Modell zu liefern (Batta, 2020).

Mit der Methode des **Reinforcement Learnings** kann ein Modell erzeugt werden, da beispielsweise besser als jeder Mensch ein komplexes Spiel wie Schach spielen kann. Da dies nicht dem Aufgabenprofil dieser Arbeit entspricht, wird die Methode für diese Arbeit nicht zum Einsatz kommen (Batta, 2020).

Die **Ensemble ML Algorithmen** verwenden mehrere andere ML-Methoden und erstellen mit diesen ein Modell. Dadurch fallen statistische Fehler einer einzelnen Methode weniger ins Gewicht, da diese sich herausmitteln. Eine Methode des Ensemble ML ist der Random Forest Algorithmus, welcher nur Entscheidungsbäume verwendet. So kann die Schwäche der Entscheidungsbäume, zu Overfitting zu neigen, deutlich verringert werden (Breiman, 2001; Wilmott, 2020).

Um einen **Deep-Learning-Algorithmus** zu trainieren, werden große Mengen an Trainingsdaten benötigt, die dieser Arbeit allerdings nicht zur Verfügung stehen (Rezaee et al., 2018). Um dennoch ein Deep-Learning-Algorithmus zu verwenden, kann ein bereits fertig trainiertes Modell eingesetzt werden, das auf die Erkennung von Oberflächengewässern trainiert wurde. Ein frei verfügbares Deep Learning (DL)-Modell für Oberflächengewässer ist „DeepWaterMap“ von F. Isikdogan et al., 2017. Das Modell wurde mit Landsat-Daten trainiert, ist aber auch in der Lage mit anderen Satellitendaten zu arbeiten, solange folgende Kanäle enthalten sind: B2: Blue, B3: Green, B4: Red, B5: Near Infrared (NIR), B6: Shortwave Infrared 1, (SWIR1) und B7: Shortwave Infrared 2 (SWIR2). Diese Flexibilität in den Datensätzen ist eine große Stärke der Deep Learning Algorithmen. Da diese neben den Spektraleneigenschaften auch Strukturen erkennen können (Steinwendner & Schwaiger, 2020).

Trotz seiner Einfachheit eignet sich der **k-Nearest Neighbors**, um Datenpunkte aufgrund seiner numerischen Werte einer Klasse zuzuordnen (Wilmott, 2020). Somit sind die drei Algorithmen, die in dieser Arbeit zum Einsatz kommen sollen: Support Vektor Machine (SVM), Random Forest (RF) und k-Nearest Neighbors (kNN) (Abb. 9). Zusätzlich soll das Deep Learning Modell „DeepWaterMap“ zur Extraktion von Wasserflächen getestet werden.

Kriterien zur Auswahl der verwendeten Klassifikationsalgorithmen

Unterschiedliche Klassifikationsalgorithmen	Überwacht	Unüberwacht	Häufiger für Anwendungsfall genutzt
Entscheidungsbaum	✓	×	✓
Naive Bayes Algorithmus	✓	×	×
Support Vector Machine	✓	×	✓
Cluster	×	✓	×
semi-überwachten Klassifikation	✓	✓	×
Reinforcement Learning	-	-	×
Multi Task Learning	✓	✓	✓
Ensemble-Learning-Algorithmen	✓	×	✓
Neuronale Netzwerke	✓	×	✓
Instance-Based Learning	✓	×	✓

Abbildung 9 Tabellarische Evaluation der für diese Arbeit verwendeten Klassifikationsalgorithmen mit den Kriterien überwacht, unüberwacht und häufiger für Anwendungsfall genutzt Nach (Batta, 2020; Breiman, 2001; Steinwendner & Schwaiger, 2020; Wilmott, 2020; Zhang, 2016)

3.4.1 k- Nearest Neighbors (kNN)

Der kNN-Algorithmus gehört zu den überwachten Machine-Learning-Algorithmen und kann dabei sowohl für die Regression und die Klassifikation eigesetzt werden. Der Algorithmus berechnet dabei die Distanzen zwischen dem zu bestimmenden Datenpunkt und den gelabelten Trainingsdaten (Wilmott, 2020). Für die Distanzberechnung zwischen den Trainingsdatenpunkten und den Testdatenpunkten gibt es verschiedene Verfahren, eines der Verfahren ist dabei der Euklidischer Abstand, welcher die direkte Distanz zwischen zwei Punkten angibt. Die Formel für den Euklidischen Abstand sieht dabei wie folgt aus:

$$D(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} \quad (1)$$

Dabei wird die Differenz zwischen den Koordinaten der Punkte *p* und *q* gebildet. Der aus dieser Differenz entstandene neue Punkt zieht mit den beiden vorhandenen Punkten *p* und *q* ein rechtwinkliges Dreieck auf, dessen Hypotenuse die direkte Distanz zwischen den beiden Punkten *p* und *q* ist. Mit Hilfe des Satzes von Pythagoras lässt sich aus beiden Seiten des rechtwinkligen Dreiecks die Hypotenuse berechnen. Die Distanz-Berechnung funktioniert dabei auch mit *n*-Dimensionen. Hat der Algorithmus alle Distanzen zu jedem Trainingsdatenpunkt berechnet, werden die *k* Punkte mit der geringsten Distanz zu dem zu klassifizierenden Punkt betrachtet und dem Punkt die Klasse zugewiesen, die am häufigsten bei den *k* nächsten Punkten vorkommen (Abb. 10). Der Hyperparameter *k* kann dabei vom Nutzer frei bestimmt werden, einzig ist zu beachten, dass *k* möglichst so gewählt werden sollte, dass es zu keinem Gleichstand im Mehrheitsentscheid kommen kann. Außerdem sollten auch nur Ganzzahlen verwendet werden, da es keine halben Datenpunkte gibt (Prasath et al., 2017; Wilmott, 2020; Zhang, 2016).

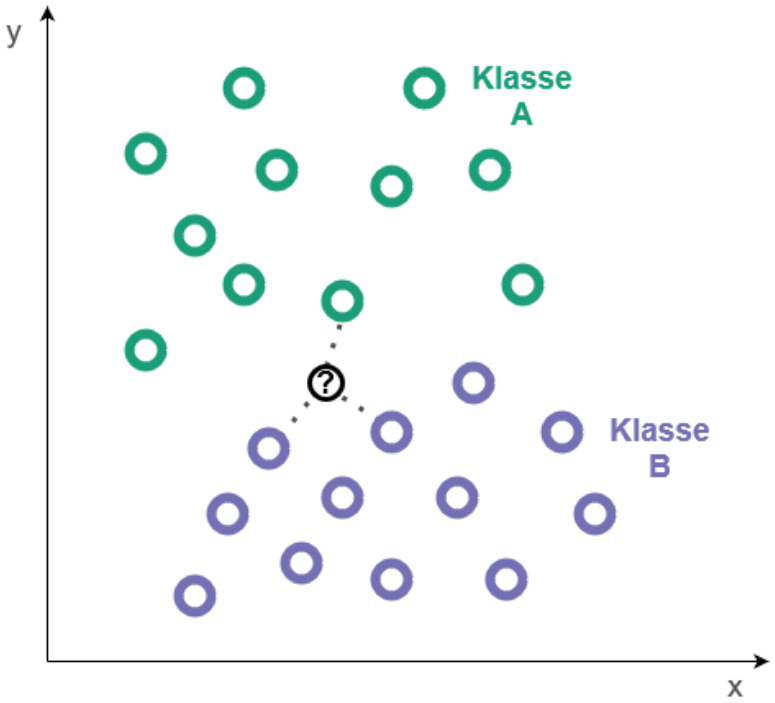

Abbildung 10 Funktionsweise des kNN-Algorithmus bei zwei Klassen (Klasse A und B). Die N nächsten Nachbarn werden betrachtet und per Mehrheitsentscheid die entsprechende Klasse vergeben. In diesem Fall werden die drei nächsten Nachbarn des Punktes "?" betrachtet. Nach https://www.ibm.com/de-de/topics/knn Letzer Aufruf 09.05.2023

Ein Vorteil des kNN-Algorithmus ist die einfache Funktionsweise, so kann der Nutzer immer nachvollziehen wie der Algorithmus arbeitet, da dieser einfach immer nur die *k* nächsten Nachbarn zu dem zu klassifizierenden Datenpunkt sucht. Dadurch entsteht für den Nutzer keine Blackbox, bei der dieser die Schritte nicht nachvollziehen kann, wie der Algorithmus aktuell arbeitet. Ein weiterer Vorteil ist, dass es nur einen Hyperparameter gibt. Es ist deutlich leichter das Optimum nur für einen Parameter zu finden als für zwei oder mehr Parameter. Ein Nachteil des kNN-Algorithmus ist gerade bei größeren Datensätzen der nicht unerhebliche Rechenaufwand, da für jeden zu klassifizierenden Punkt, die Distanzen zu allen Datenpunkten im gesamten Trainingsdatensatz bestimmt werden müssen. Als weiteren Nachteil kann noch genannt werden, dass das Verfahren zur Bestimmung der Distanz neben dem Hyperparameter *k* auch keinen unerheblichen Einfluss auf das Klassifikationsergebnis hat (Prasath et al., 2017; Zhang, 2016).

3.4.2 Support Vector Machine (SVM)

Der SVM-Algorithmus ist in der Lage sowohl Klassifikationen als auch Regressionen durchzuführen, dadurch gehört dieser zu den überwachten Machine-Learning-Algorithmen. Der Algorithmus versucht dabei eine Hyperebene zwischen die unterschiedlichen Klassen zu legen. Diese Ebene unterteilt dann für die Klassifikation den Merkmalsraum. Auf welcher Seite ein Datenpunkt liegt, entscheidet, welcher Klasse dieser zugeordnet wird (Abb. 11). Der Algorithmus betrachtet alle Datenpunkte als Vektoren und versucht den Abstand zwischen den Vektoren der unterschiedlichen Klassen und der Hyperebene zu maximieren. Die Vektoren, die dabei der Hyperebene am nächsten liegen, werden dabei als Support Vektoren bezeichnet. Alle anderen Vektoren werden für die Position der Hyperebene nicht benötigt und haben daher auch keinen Einfluss auf diese. Da eine Ebene nur lineare Datensätze voneinander trennen kann, wäre der SVM-Algorithmus für alle nicht linearen Klassengrenzen nutzlos. Daher wird der sogenannte Kernel-Trick eingesetzt, der den Vektorraum um Dimensionen erweitert. In einem Raum mit genügend hohen Dimensionen lassen sich die Klassen dann wieder durch eine Hyperebene linear trennen. Ist die Hyperebene im mehrdimensionalen Raum gefunden, wird diese wieder auf die

Ausgangsdimension zurück transformiert. Die so erhaltene Hyperebene verläuft nicht mehr linear und kann so auch nicht lineare Klassen trennen (Wilmott, 2020).

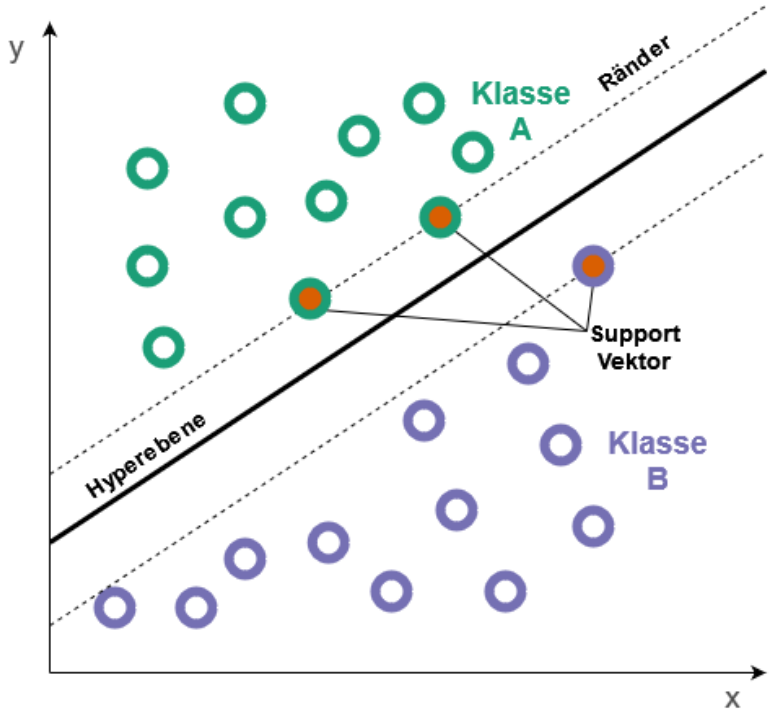

Abbildung 11 Prinzip des SVM-Algorithmus die Support Vektoren liegen auf den Rändern, welche den größtmöglichen Abstand zur Hyperebene aufweisen. Nach (Wilmott, 2020)

Einer der Vorteile, die der SVM-Algorithmus mit sich bringt, ist, dass dieser nur auf den Grenzen zwischen den Klassen beruht, dadurch kann der Algorithmus immer noch sehr gute Ergebnisse liefern, auch mit kleineren Trainingsdatensätzen (Foody & Mathur, 2004). Ein Nachteil des SVM-Algorithmus gerade bei größeren Datensätzen, ist die hohe Rechenleistung, die benötigt wird (Adugna et al., 2022). Ein weiterer Nachteil ist, dass der SVM-Algorithmus mehr als einen Hyperparameter nutzt. Somit kann das Finden der passenden Einstellung etwas komplexer ausfallen als es beispielsweise beim k-NN Algorithmus der Fall ist (Adugna et al., 2022).

3.4.3 Random Forest (RF)

Mit dem RF-Algorithmus lassen sich sowohl Klassifikationen als auch Regressionen durchführen. Der Algorithmus gehört zu den überwachten Machine-Learning-Algorithmen. Bei dem RF-Algorithmus handelt es sich um eine Weiterentwicklung des klassischen Entscheidungsbaums. Der Entscheidungsbaum Algorithmus ist vom Prinzip her ein Flussdiagramm. Der Datensatz wird an jedem Knotenpunkt durch eine neue Bedingung aufgeteilt bis zu den sogenannten Blättern das gewünschte Klassifikationsergebnis erreicht ist. Damit der Algorithmus Bedingungen wählt, die effizient zu dem gewünschten Klassifikationsergebnis führen, wird jede potenzielle Bedingung von einer Entropie Funktion bewertet. Diese Funktion bewertet, wie stark die Unordnung ist dabei ist 1 die höchste Unordnung und 0 die höchste Ordnung. Da in Klassifikationen mit einem Entscheidungsbaum die Blätter immer möglichst rein sein sollten, also in jedem Blatt möglichst nur eine Klasse enthalten sein sollte, hilft hierbei die Entropiefunktion enorm, um immer die Entscheidung für die nächste Bewertung zu treffen, die die Entropie am effektivsten verringert. Ein reines Blatt hat demnach eine Entropie von 0 und somit die höchste Ordnung erreicht. Bei komplexeren Datensätzen kann es dazu kommen, dass die Blätter nur noch sehr wenige Datenpunkte enthalten. Dadurch wird der Trainingsdatensatz perfekt klassifiziert allerdings wird das Klassifikationsergebnis mit neuen unbekannten Daten deutlich schlechter ausfallen, da das Modell überangepasst (Overfitting) ist (Breiman, 2001; Wilmott, 2020).

Dies ist der Nachteil des klassischen Entscheidungsbaums, da dieser sehr empfindlich auf die Trainingsdaten reagiert und es so zu hohen Varianzen in Zuge von Overfitting kommen kann (Breiman, 2001). Um diese starke Abhängigkeit von den Trainingsdaten zu reduzieren, werden beim Random Forest Algorithmus mehrere Entscheidungsbäume erstellt, an dessen Ende das Mehrheitsvotum der Entscheidungsbaum Modelle über die zugewiesene Klasse entscheidet (Abb. 12). Da die Entscheidungsbäume bei gleichen Datensätzen immer zum selben Ergebnis kommen würden, werden zufällig aus den Trainingsdaten Datenpunkte gewählt und so ein neuer Datensatz erstellt. Dabei können auch Duplikate der Datenpunkte auftreten. Dieser Prozess wird Bootstrap-Verfahren genannt und mit dem Parameter *n* für die Anzahl der so neu erstellten Datensätze beziehungsweise der Anzahl der Entscheidungsbäume beschrieben. An jedem der Knoten im Entscheidungsbaum werden zufällig Features gewählt, die das Kriterium für den nächsten Ast bilden. Die Anzahl dieser zufällig ausgewählten Features sollte geringer sein als die Gesamtanzahl der Features, die ein Datenpunkt aufzuweisen hat. Die maximale Anzahl an zufällig ausgewählten Features wird als Parameter *m* angegeben (Breiman, 2001; Wilmott, 2020).

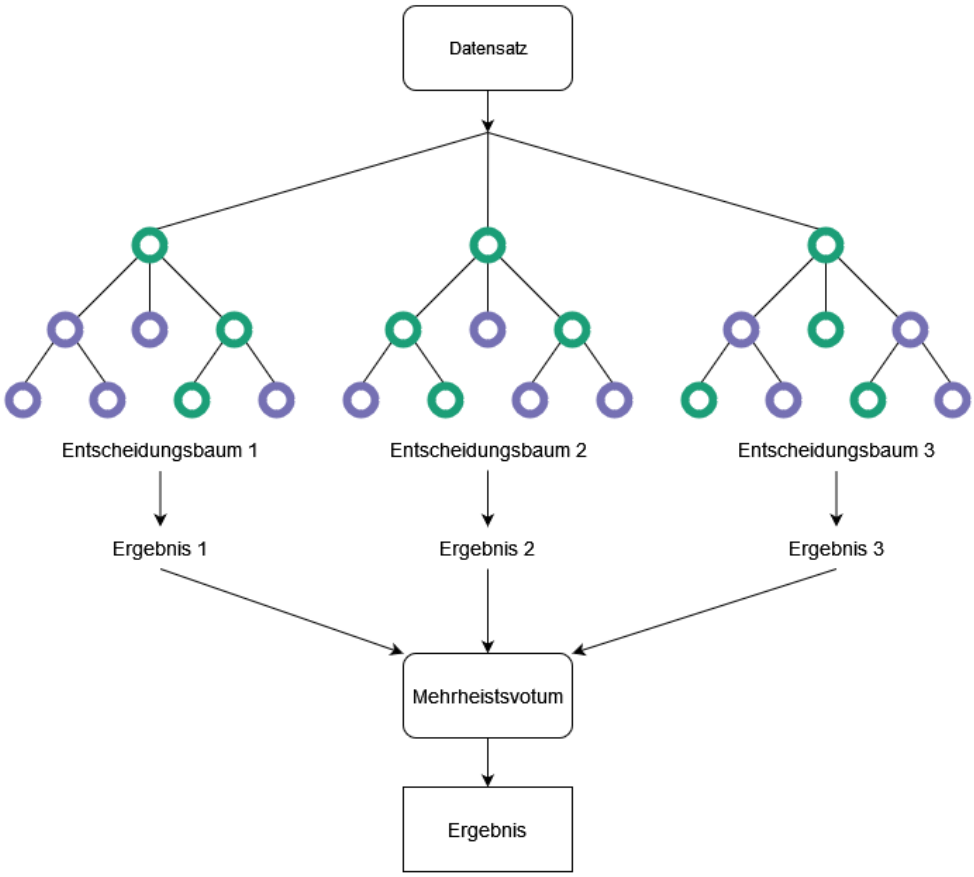

Abbildung 12 Schematische Funktionsweise eines Random Forest Modells. Die zu klassifizierenden Daten durchlaufen die unterschiedlichen Entscheidungsbäume. Das Ergebnis jedes Baumes geht dann als Stimme in ein Votum ein. Das Ergebnis mit dem höchsten Stimmenanteil gewinnt die Wahl und wird als Ergebnis ausgegeben. Nach (Breiman, 2001; Wilmott, 2020)

Ein Vorteil des RF-Algorithmus ist die Schnelligkeit, mit der die Modelle berechnet werden können, so steigt die Rechenleistung, die benötigt wird, linear mit der Anzahl der verwendeten Bäume an. Ein weiterer Vorteil, ist dass der Algorithmus nicht zum Overfitting neigt. Eine Erhöhung der Anzahl der Entscheidungsbäume erbringt ab einem gewissen Punkt keine nennenswerte Verbesserung aber auch keine Verschlechterung des Modells (Adugna et al., 2022). Durch diese Eigenschaft ist das Einstellen der Parameter deutlich einfacher als es sonst mit zwei oder mehr Parametern wäre, da bei einer hohen

Anzahl an Entscheidungsbäumen *n* sich das Modell weder verbessert noch verschlechtert. Ein Nachteil des RF-Algorithmus ist, dass dieser ab einer bestimmten Komplexität praktisch nicht mehr nachvollziehbar ist und daher der Entscheidungsprozess für den Nutzer oft eine Blackbox bleibt (Wilmott, 2020).

3.4.4 Deep Learning

Deep Learning Modelle bilden eine Unterkategorie der Neuronalen Netzwerke. Neuronale Netzwerke bestehen, wie der Name schon vermuten lässt, aus Neuronen, die dem menschlichen Gehirn nachempfunden sind. Ein sehr simples Neuronales Netz besteht dabei aus einem Input Wert, der wird vom Input Neuron aufgenommen. Dieser Wert des Neurons wird gewichtet und weiter an die Output Neuron Schicht gegeben, hier wird der gewichtete Wert mit einem Schwellenwert verglichen und als Output ausgegeben. Anhand des Output Wertes lässt sich dann eine Klasse zuweisen. Deep Learning Modelle weisen im Gegensatz zu einem einfachen Neuronalen Netzwerk eine große Anzahl an Gewichtslayer auf, diese werden auch hidden Layer genannt (Joachim Steinwendner & Roland Schwaiger, 2020). Um einen Deep-Learning-Algorithmus zu trainieren, werden große Mengen an Trainingsdaten benötigt, diese stehen in dieser Arbeit allerdings nicht zur Verfügung (Rezaee et al., 2018). Um dennoch ein Deep-Learning-Algorithmus zu verwenden, kann ein bereits fertig trainiertes Modell verwendet werden, welches auf die Erkennung von Oberflächengewässern trainiert wurde. Ein frei verfügbares DL-Modell für Oberflächengewässer ist „DeepWaterMap“ von (F. Isikdogan et al., 2017).

3.5 Evaluation der verwendeten Software

Für die Klassifizierung der Wasserflächen im Leegmoor werden verschiedene spezielle Softwarepakete benötigt. Zum einen wird eine GIS-Software (Geographic Information System) für das Darstellen der Ergebnisse benötigt. Dabei gibt es unterschiedliche GIS-Softwareanbieter mit unterschiedlichen Zugangsmodellen. So ist ArcGIS ein Produkt der Firma Esri der Marktführer im Bereich der GIS-Software und hat ein proprietäres Lizenzmodell, bei dem der Nutzer für eine aktuelle Version jährliche Gebühren entrichten muss. Ein weiteres sehr verbreitetes GIS-Softwarepaket ist QGIS, das anders als ArcGIS, ein Open Source Lizenzmodell ist. Anders als bei ArcGIS gibt es für QGIS kein offiziellen Kundensupport. Auch ist der Grundfunktionsumfang etwas geringer als bei ArcGIS. Allerdings gibt es gerade für spezialisiertere Funktionen oft Add-ons, die über ein Interface in QGIS direkt heruntergeladen werden können. Da beide Softwarepakete sich problemlos für das Darstellen der Klassifikationen eignen, wurde sich für diese Arbeit zu Gunsten des nicht proprietären Lizenzmodells für QGIS entschieden.

Zusätzlich wird für die Arbeit eine Software benötigt, die die Klassifikation, der Wasserflächen im Moor durchführen kann. Auch für diese Anwendungen kommen die beiden vorher genannten GIS-Software Pakete in Frage.

QGIS weist in seiner Grundfunktionalität allerdings keine Funktion zur Klassifizierung auf, allerdings kann das Add-on Semi-Automatic Classification Plugin (SCP) Rasterdaten klassifizieren. Dabei bietet das Add-on eine Klassifikation mit dem k-Nearest Neighbors Algorithmus und dem Random Forest Algorithmus an. Diese Arbeit soll allerdings neben dem kNN und dem RF-Algorithmus auch die Klassifikationen des SVM und die eines Deep-Learning-Algorithmus miteinander vergleichen. Damit wäre mit QGIS keine All-in-One Lösung möglich und es müsste ein weiteres Softwarepaket herangezogen werden (*QGIS.Org*, 2024).

ArcGIS bietet mit seinem Klassifikationstool alle drei für den Vergleich heranzuziehenden klassischen Machine-Learning-Algorithmen: kNN, RF und SVM an. Einzig die Möglichkeit zur Erstellung eines auf Deep Learning basierendem Modell, fehlt (*Esri*, 2024).

Da die Erstellung eines Modells mit Hilfe eines **Deep-Learning-Algorithmus** sehr komplex ist und große Mengen an Trainingsdaten erfordert, wird in dieser Arbeit ein bereits trainiertes Modell zur Erkennung von Oberflächengewässern verwendet. Die Klassifizierung durch ein solches Modell wird oft über Python durchgeführt. Als fertig trainiertes Modell kommt hierbei das „DeepWaterMap“ Modell von Isikdogan in Frage, da dieses freiverfügbar ist.

Als eine weitere Softwarelösung für Klassifikationen kann die von der ESA speziell für Sentinel-Daten entwickelte GIS **SNAP** angeführt werden. In diesem Softwarepaket lassen sich viele Korrekturen und Analyse Verfahren speziell für Satellitendaten durchführen. Die Software verfügt dabei auch über eine Funktion der Klassifikation, mit der sich mehrere Algorithmen ausführen lassen. Allerdings lassen sich auch hier genau wie bei den Klassifikationsoptionen bei QGIS nur RF und kNN durchführen. Da SNAP eine Vielzahl an Analysefunktionen, speziell für Satellitenbilder bietet, wird es für Grafiken in dieser Arbeit Anwendung finden (*SNAP*, 2024).

Als eine weitere Softwareoption bietet sich die für Klassifikationen eigens entwickelte Software **ENVI** der Firma Harris Geospatial Solutions an. Diese kann die von dieser Arbeit geforderten Klassifikationsalgorithmen durchführen und ist außerdem in der Lage Deep Learning Modelle zu erstellen. Allerdings ist das Lizenzmodell proprietär und erfordert für die Nutzung eine jährliche Gebühr. Da auf ein Schreiben an den österreichischen Distributor von ENVI, ob es möglich wäre eine kostenlose studentische Version zu erhalten, keine Antwort erfolgte, kann ENVI für diese Arbeit nicht verwendet werden (*ENVI*, 2024).

Als letzte Option bietet sich für die Arbeit an, alle benötigten Methoden, die für die Klassifikation von Oberflächengewässern in einem Moorgebiet benötigt werden, als **Python-Skript** umzusetzen. Eine Python Bibliothek, die die drei Methoden kNN, RF und SVM zur Klassifikation anbietet, ist scikit-learn. Scikit-learn läuft unter einem Open Source Lizenzmodell und ist für jeden frei zugänglich. Ein selbstgeschriebenes Python-Skript bietet den Vorteil, mehr Kontrolle über den gesamten Ablauf der Klassifikation zu haben, da jeder Schritt im Programm vom Nutzer geschrieben werden muss. Ein so erstelltes Python-Skript hat auch den Vorteil, dass es relativ einfach in ein vollautomatisiertes Python-Skript integriert werden kann, mit dem die Daten heruntergeladen, auf das Untersuchungsgebiet zugeschnitten und abschließend klassifiziert werden (Pedregosa et al., 2011).

Die Wahl für das optimale Softwarepaket fällt auf Grund des Umfangs und der einfachen Möglichkeit zur Automatisierung auf ein mit der Python Bibliothek scikit-learn geschriebenes Skript (Abb. 13).

Zur Auswahl stehende Software

Kriterien zur Wahl der Software

	ArcGIS	QGIS	ENVI	SNAP	Python
Random Forest (RF)	✓	✓	✓	✓	✓
k-nearest neighbors (k-NN)	✓	✓	✓	✓	✓
Support Vector Machine (SVM)	✓	×	✓	×	✓
Deep Learning	×	×	✓	×	✓
Proprietär	✓	×	✓	×	×

Abbildung 13 Evaluation der verwendeten Software in tabellarischer Form Nach (ENVI, 2024; Esri, 2024; QGIS.Org, 2024; Pedregosa et al., 2011; Isikdogan et al., 2017; SNAP, 2024)

3.6 Parameter der verwendeten Machine-Learning-Algorithmen

Für diese Arbeit werden nur die folgenden Parameter der drei verwendeten klassischen ML-Algorithmen verändert: für den SVM-Algorithmus wird C angepasst, für den kNN-Algorithmus wird k oder „n_neighbors“ angepasst und für den RF-Algorithmus wird „n_estimators“ oder E angepasst. Es werden im Folgenden noch weitere Parameter beschrieben, die von „scikit-learn“ genutzt werden, um ihren Einfluss auf das Ergebnis besser abschätzen zu können. Das Deeplearning Modell von Isikdogan (2017) ist bereits ein fertig trainiertes Modell und benötigt daher keine weiteren Parameter.

3.6.1 SVM

Im Folgenden werden die wichtigsten Parameter des SVC (C-Support Vector Classification) Algorithmus genauer beleuchtet, welche in der Python Bibliothek „scikit-learn“ eingestellt werden können. Für diese Arbeit werden folgende Werte für C genutzt: 0,5, 1, 20 und 50. Für den SVM-Algorithmus wurde sich für die Auswahl dieser Werte entschieden, um zum einen das Verhalten der Modelle bei einem Wert geringer als der Standartwert für C = 1 zu beobachten als auch für zwei größere Werte, um die Reaktion der Modelle bei deutlich größerem C zu untersuchen. Dabei wurden sowohl für den geringeren Wert als auch für die beiden größeren Werte verschiedene Werte getestet. Sich schlussendlich aber für die Auswahl dieser drei Werte entschieden, da diese immer noch eine gute Balance zwischen guten Klassifikationsergebnissen und dennoch Abweichungen in den Ergebnissen vom Standartwert C =1 liefern.

C

Der Standardwert welcher von „scikit-learn“ angenommen wird ist 1. C beeinflusst, wie sich die Grenzen zwischen den Klassen verhalten. C gibt dabei an wie stark Werte, die innerhalb des vom Algorithmus erstellten Trennberiechs, zum Gesamtfehler beitragen. Nimmt C den Wert 0 an werden diese Werte für den Gesamtfehler nicht berücksichtigt. Je größer C wird, desto stärker gehen Werte in den Trennbereich in den Gesamtfehler ein, folglich werden immer weniger Werte im Trennbereich für das Modell akzeptiert (Pedregosa et al., 2011).

Kernel

Ein weiterer wichtiger Parameter ist die gewählte Kernelfunktion, diese gibt an, welche Art von Hyperebene zwischen den Klassen gezogen wird. Über die „scikit-learn“ Bibliothek lassen sich folgende Kernelfunktionen einstellen: „linear“, „poly“, „rbf“, „sigmoid“, „precomputed“, dabei ist „precomputed“ eine Option, um eine vom Nutzer definierte Kernelfunktion zu verwenden. Die „linear“ Kernelfunktion trennt die Klassen linear. Die Kernelfunktion „poly“ verwendet eine Polynomfunktion, die eine nicht lineare Trennung der Klassen erlaubt. Die Kernelfunktion „rbf“ ist eine Radiale Basisfunktion und sorgt ebenfalls für eine nicht lineare Trennung der Klassen. Außerdem ist „rbf“ auch die Standardeinstellung für den SVC-Algorithmus in der „scikit-learn“ Bibliothek. Bei der Kernelfunktion „sigmoid“ handelt es sich um eine Sigmoidfunktion (Pedregosa et al., 2011).

Gamma

Gamma ist ein Parameter, der für nicht lineare Kernelfunktionen, wie z.B. „poly“, „rbf“ oder „sigmoid“ gewählt werden kann. Gamma gibt an, wie stark sich das Modell an die Trainingsdaten anpassen soll. Ein Modell mit einem hohen Gammawert versucht die Trainingsdaten perfekt abzubilden. „Scikit-learn“ bietet dabei zusätzlich zur manuellen Angabe eines Wertes für Gamma die beiden Modi „scale“ und “auto“ an. Bei der Option „scale“ wird das Produkt der Anzahl der Feature und Varianz von X durch 1 geteilt. Die „scale“ Option ist die Standardeinstellung in „scikit-learn“. Die „auto“ Option ist die Anzahl der Feature geteilt durch 1 (Pedregosa et al., 2011).

3.6.2 k-NN

Im Folgenden werden die wichtigsten Parameter des k-NN Algorithmus zusammengefasst, welche in der Python Bibliothek „scikit-learn“ zur Verfügung stehen. Für diese Arbeit werden folgende Werte für den Parameter k genutzt: 3, 5, 15 und 29. Für den k-NN-Algorithmus wurde sich für die Auswahl dieser Werte entschieden, um zum einen das Verhalten der Modelle bei einem Wert geringer als der Standartwert für k = 5 zu beobachten als auch für zwei größere Werte, um die Reaktion der Modelle bei deutlich größerem k zu untersuchen. Hierbei ist zu beachten, dass die Werte ungerade sein sollten, da die Klassenzugehörigkeit durch ein Mehrheitsvotum entschieden wird und so kein Unentschieden entstehen kann. Es wurden sowohl für den geringeren Wert als auch für die beiden größeren Werte, verschiedene Werte getestet, sich schlussendlich aber für die Auswahl dieser drei Werte entschieden, da diese gute Klassifikationsergebnisse liefern. Es gibt dabei weniger Abweichungen vom Standartwert, als es beim SVM-Algorithmus der Fall ist, allerdings verschlechtern sich die Ergebnisse deutlich bei zu extrem gewählten k Werten.

n_neighbors

Der „n_neighbors“ Parameter oder auch „k“, gibt an, wie viele nächstliegende Trainingsdatenpunkte für den Mehrheitsentscheid zur Klassifizierung des zu bestimmenden Datenpunktes herangezogen werden. Der Standardwert liegt hierbei von „scikit-learn“ bei 5(Pedregosa et al., 2011).

p

Gibt an welches Verfahren zur Distanzberechnung verwendet wird. Bei einem p-Wert von 1 wird „manhattan distance“ Verfahren angewandt. Bei einem p-Wert von 2 wird die euklidische Distanz zwischen den Datenpunkten berechnet. Der Standardwert der dabei von „scikit-learn“ angenommen wird ist 2(Pedregosa et al., 2011).

3.6.3 Random Forest

Im Folgenden werden die wichtigsten Parameter des RF (Random Forest) Algorithmus genauer beleuchtet, welche in der Python Bibliothek „scikit-learn“ eingestellt werden können. Die vier Werte für den Parameter E, welche von dieser Arbeit genutzt werden, sind: 100, 500, 1000 und 5000. Für den k-NN-Algorithmus wurde sich für die Auswahl diese Werte entschieden, um zum einen das Verhalten der Modelle bei größeren Werten als der Standartwert für E = 100 zu beobachten. In diesem Fall werden nur größere Werte untersucht, da ein E = 100 schon ein sehr niedriger Wert ist. Es wurden daher nur größere Werte als E = 100 getestet und sich auch hier für die mit dem besten Ergebnis entschieden.

n_estimators

In der Python Bibliothek „scikit-learn“ gibt „n_estimators“ an, wieviele Entscheidungsbäume erstellt werden. Der Standartwert ist dabei auf 100 gesetzt(Pedregosa et al., 2011). In dieser Arbeit wird statt „n_estimators“ auch E verwendet.

Criterion

Dieser Parameter gibt an mit welcher Funktion jede Entscheidung im Entscheidungsbaum bewertet wird. Die Standardeinstellung in der „scikit-learn“ Python Bibliothek ist „gini“. Weitere Funktionen, welche angegeben werden können, sind: „entropy“ und „log_loss“(Pedregosa et al., 2011).

max_depth

Der „max_depth“ Parameter gibt an, wieviele Knoten jeder Entscheidungsbaum maximal haben darf. Der Standardwert den „scikit-learn“ hierfür verwendet ist „None“. Der Entscheidungsbaum breitet sich

so lange aus, bis alle Blätter des Entscheidungsbaums rein sind und nur eine Klasse beinhalten oder die minimale Anzahl von Proben zum Aufsplitten eines Knotens erreicht ist. Dies wird mit dem Parameter „min_samples_split“ in „scikit-learn“ angegeben(Pedregosa et al., 2011).

min_samples_split

Dieser Parameter gibt an, wie hoch die minimale Anzahl an Proben sein soll, damit sich ein Knoten noch einmal aufsplitten kann. Der Standardwert für diesen Parameter ist in „scikit-learn“ mit 2 angegeben(Pedregosa et al., 2011).

min_samples_leaf

Der „min_samples_leaf“ Parameter wird verwendet, um die minimale Anzahl an Proben zu definieren, die benötigt werden, um ein Blatt am Entscheidungsbaum zu formen. Der Standardwert der dabei von „scikit-learn“ verwendet wird ist 1(Pedregosa et al., 2011).

max_features:

Die Random Forest Methode wählt neben den zufällig zusammen gesetzten Datensätzen auch zufällig einen Teil der Features aus, damit sich die Entscheidungsbäume stärker voneinander unterscheiden. Der Parameter „max_features“ gibt dabei an wieviele dieser zufällig gewählten Features für jeden Entscheidungsbaum verwendet werden sollen. Dabei sollte der Wert natürlich so gewählt sein, dass dieser niedriger ist, als die Gesamtanzahl von Features die der Datensatz besitzt. Der von „scikit-learn“ gewählte Standardwert ist „sqrt“. Dabei handelt es sich um die Wurzel aus der Gesamtanzahl der Feature des Datensatzes(Pedregosa et al., 2011).

3.7 Beschreibung des Projektgebietes

Die Wahl des Projektgebietes ist bei dieser Arbeit auf das Leegmoor gefallen, da auf dieser Fläche schon seit 1983 Renaturierungsmaßnahmen getestet werden. Die Eigenheiten des Gebietes sind somit genaustens bekannt, dadurch lassen sich weitere Genauigkeitsüberprüfungen, wie z.B. die Wasserkontrollzonen (siehe 3.8) etablieren. Gerade die Ortskenntnisse sind von großer Bedeutung, da auf einer Satelliten- oder auch Drohnenaufnahme Klassen kaum voneinander unterscheidbar sind. In dem Fall dieser Arbeit sind die nicht grüne Vegetation und die Bodenklasse nur schwer zu unterscheiden. Ein weiterer Grund für die Wahl des Leegmoores als Untersuchungsgebiet ist das Vorhandensein einer Drohnenaufnahme des Gebietes, die der Arbeit als Grundlage für den Ground Truth Datensatz dient (*OpenGeoData.NI*, 2024).

Das Leegmoor befindet sich im Nordosten Deutschlands im Landkreis Emsland in Niedersachsen und ist östlich des Dorfes Surwold gelegen (Abb. 14). Die Fläche ist ca. 450 ha groß und seit 1983 als Naturschutzgebiet mit der Kennzeichen NSG WE 136 ausgewiesen. Außerdem wird das Leegmoor unter den Natura 2000 Gebieten mit der Nummer 159 in den FFH-Richtlinien (Fauna-Flora-Habitat-Richtlinie) geführt. Das Leegmoor gehört zu dem Vogelschutzgebiet V 14 Esterweger Dose des Schutzgebietnetzwerkes Natura 2000. Beim Leegmoor handelt es sich um ein Hochmoor, das heißt, das Moor wird nur durch Regenwasser mit Wasser versorgt und hat keinen Zugang zum Grundwasser (Abb. 15, 16). Das Leegmoor wurde für den Torfabbau genutzt und daher trockengelegt. Die Fläche war bis zu den stark zersetzten Schwarztorf abgebaut. Im Jahr 1983 wurden dann auf der Fläche zahlreiche Maßnahmen vorgenommen, um das Leegmoor wieder zu vernässen, mit dem Ziel, die hochmoortypische Flora und Fauna wieder anzusiedeln. Dafür wurde die Fläche mit Gräben eingefasst, um so den Wasserstand in der Moorfläche zu kontrollieren. Außerdem wurden verschiedene hochmoortypische Pflanzen auf dem Gebiet ausgesät (Nick et al., 2001).

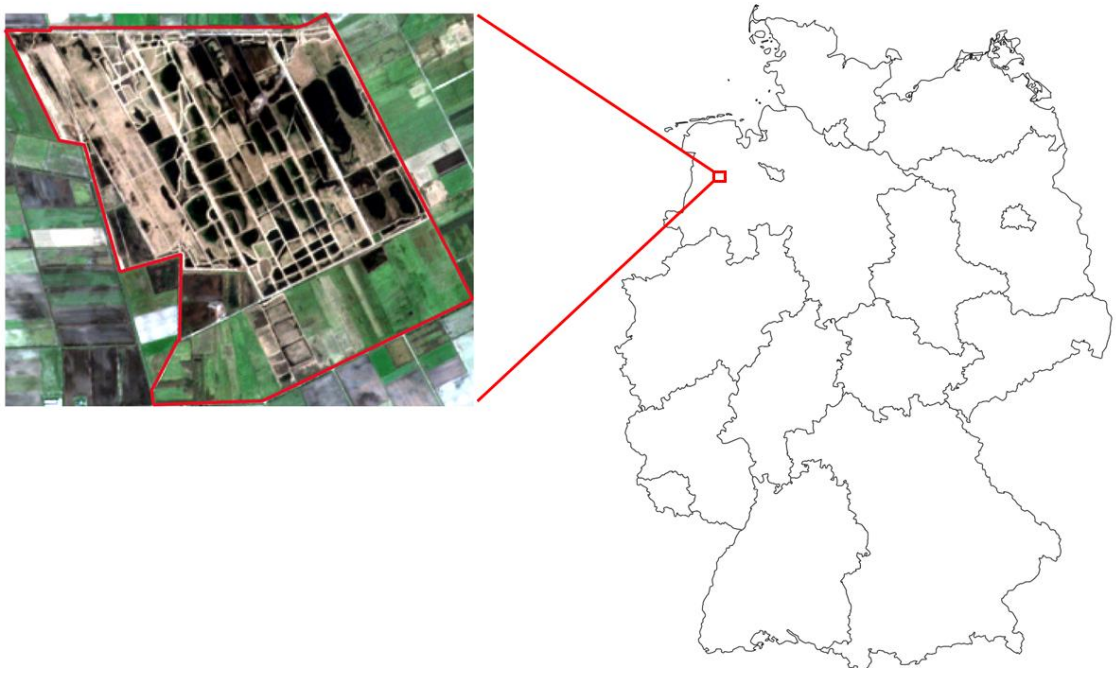

Abbildung 14 Lages des Leegmoores innerhalb von Deutschland (BKG)

Abbildung 15 Ausblick über das Leegmoor von einer Aussichtsplattform im Südosten des Gebietes. Aufgenommen am 20.03.2023

Abbildung 16 Aussicht auf Wasserfläche mit vorliegendem Bewuchs von Pfeifengras. Aufgenommen am 20.03.2023

3.8 Accuracy Assessment

Das Accuracy Assessment bildet die Grundlage zur Überprüfbarkeit eines Klassifikationsmodells. Es gibt an, mit welcher Genauigkeit eine bestimmte Klasse oder alle Klassen vom Modell bestimmt werden können. Das klassische Verfahren ist hierfür die **Konfusionsmatrix**, bei der in die Spalten die Anzahl an richtigen Bestimmungen für jede Klasse eingetragen werden (Abb. 17). In die Zeilen wird die Anzahl von Bestimmungen, welche vom Klassifikationsmodell bestimmt wurden, aufgetragen (Congalton et al., 1983; Congalton, 1991).

Aus dieser Matrix lassen sich zahlreiche Kennzahlen für die **Genauigkeit** des Modells ablesen. Der klassischste Wert wäre dabei die Angabe der Gesamtgenauigkeit (Abb. 17). Diese lässt sich aus der Summe der Diagonalen gegen die Gesamtanzahl der Bestimmungen berechnen. Die Diagonale bildet in der Matrix die Anzahl an Klassifizierungen, die vom Modell richtig zugeordnet wurden (Congalton et al., 1983; Congalton, 1991).

Eine weitere Kennzahl ist die sogenannte **Sensitivität**, die angibt, wie zuverlässig eine Klasse vom Klassifizierungsmodell erkannt werden kann (Abb. 17). Dafür wird die Anzahl an Bestimmungen, welche vom Klassifikationsmodell getroffen wurden, durch die richtige Anzahl an Bestimmungen einer Klasse geteilt (Powers, 2008).

Eine weitere klassenbezogene Kenngröße ist die sogenannte **Präzision** (Abb. 17)**.** Diese gibt an, zu welchem Anteil, die vom Modell bestimmte Klasse sich aus der wahren Klasse zusammensetzt. Um die Präzision zu berechnen, werden die richtig klassifizierten Bestimmungen vom Modell, durch die Anzahl der Gesamtbestimmungen vom Modell für diese Klasse, geteilt (Powers, 2008).

Eine weitere statistische Kenngröße zur Evaluierung des Klassifikationsmodells ist der **Kappa-Koeffizient**. Dieser gibt an, ob die Klassifikation statistisch signifikant ist oder auch einfach durch zufällig verteilte Werte hätte zustande kommen können. Der Kappa-Koeffizient kann dabei Werte

zwischen -1 und 1 annehmen. Dabei gibt ein negativer Wert an, dass zwischen Klassifikation und Referenzdaten eine schlechtere Übereinstimmung als zufällig verteilte Werte, besteht. Ein Kappa-Koeffizient von Null gibt an, dass sich das Klassifikationsergebnis gegenüber zufällig verteilten Klassen nicht unterscheidet. Ein Kappa-Koeffizient von eins zeigt, dass Referenzdaten und Klassifikation komplett miteinander übereinstimmen (Congalton et al., 1983; Congalton, 1991).

Als weitere Kenngröße kann noch der **f1-score** genannt werden (Abb. 17). Dieser wird aus dem Durchschnitt der Sensitivität und der Präzision berechnet und fasst somit beide Werte zusammen (Powers, 2008).

Um für diese Arbeit Referenzwerte zu erhalten, kann nicht auf eine schon vorhandene Klassifikation, wie etwa Daten zur Landbedeckung zurückgegriffen werden, da diese in der Regel nur alle paar Jahre aktualisiert werden. Da sich die Landbedeckung im Falle des Moores, durch Änderungen im Wasserstand, je nach Jahreszeit ändern, müssen die Referenzpunkte manuell bestimmt werden. Um eine Verfälschung der Ergebnisse zu vermeiden, müssen diese Referenzpunkte über eine andere Quelle bestimmt werden, als das Datenmaterial, welches für die Klassifikation genutzt wurde. In dieser Arbeit wird hierfür ein Drohnenbild des Untersuchungsgebietes genutzt, das zu einem sehr ähnlichen Zeitpunkt, wie eines der Sentinel-2 Satellitenbilder, aufgenommen wurde.

Als zusätzliche Genauigkeitskontrolle sollen Pixel markiert werden, die das gesamte Jahr über dauernass sind. Hierfür sind genaue Kenntnisse des Geländes erforderlich. Diese Ortskenntnisse sind in diesem Fall die Kreuzvalidierung der Methode. Im Falle des Leegmoores und im Zuge dieser Arbeit wurde sich für einen Kolk im Westen des Gebietes und eine der beiden großen Wasserflächen im Nordosten des Gebietes entschieden (Abb. 18). Ein Vorteil eines solchen Genauigkeitsparameter ist, dass dieser auf alle klassifizierten Szenen angewendet und so auch für Klassifikationsergebnisse genutzt werden kann, die über kein zusätzliches Luftbild – oder Satellitenbildaufnahme als Kreuzvalidierung verfügen. Allerdings ist ein solcher Genauigkeitsparameter deutlich weniger aussagekräftig, als die der **Konfusionsmatrix**, da sich nur auf sehr eindeutig zu klassifizierende Pixel konzentriert wird und keine zufällige Probenauswahl besteht. Aus diesem Grund sind Werte für den Wasserkotrollwert von an die 100% zu erwarten.

		Vorhergesagte Klassen: Negativ (VN)	Vorhergesagte Klassen: Positiv (VP)	
Tatsächliche Klassen	Negativ (N)	True negative (TN)	False positive (FP)	
	Positiv (P)	False negative (FN)	True positive (TP)	Sensitivität (Sen) = $\frac{TP}{P}$
			Präzision (Prä) = $\frac{TP}{VP}$	Gesamtgenauigkeit = $\frac{TP + TN}{P + N}$
				F1-Score = $\frac{Sen + Prä}{2}$

Abbildung 17 Schematische Error Matrix mit Herleitung zur Berechnung von Kenngrößen des Accuracy Assessments Nach (Powers, 2008)

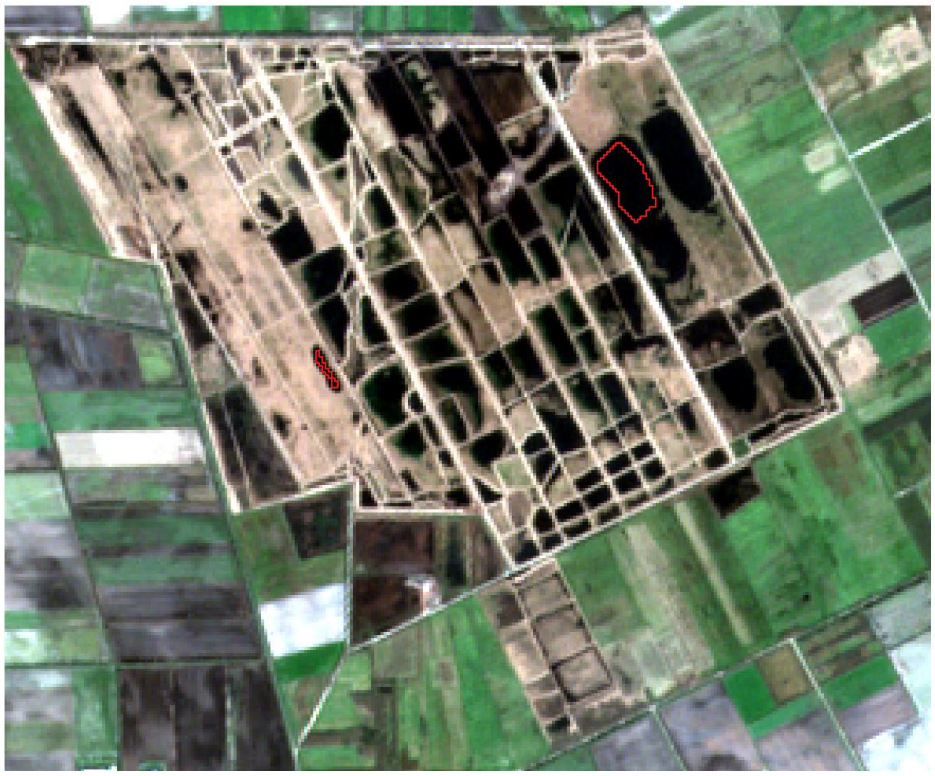

Abbildung 18 Satellitenaufnahme des Leegmoores vom 23.03.2020 mit in Rotumrahmten Wasserkontrollzonen

3.9 Trainingsdaten

Um ein ML-Modell zu erstellen, werden Trainingsdaten benötigt anhand deren der Algorithmus lernen kann, wie die Eigenschaften einer Klasse sind. Üblicherweise wird dabei ein Datensatz in einen Test und einen Train Split aufgeteilt. Häufig wird dem Train Split ein höherer Prozentsatz der Daten zur Verfügung gestellt. Diese Aufteilung der Daten ist notwendig, um zu verhindern, dass das Modell nicht auch die Daten klassifiziert, an denen das Modell trainiert wurde (Adugna et al., 2022). Dies würde zu einem verfälschten Ergebnis führen, da das Modell an die Trainingsdaten angepasst ist. Daher wird für diese Arbeit eine Sentinel-2 Aufnahme vom 28.03.2020 verwendet, um die Trainingsdaten zu erstellen. Die Sentinel-2 Aufnahme, zur Klassifizierung der Daten und für die Erstellung des Accuracy Assessment ist vom 23.03.2020 und demnach nur 5 Tage vor der Aufnahme für die Trainingsdaten erstellt worden. Die Trainingsgebiete wurden als Shapefile in QGIS als Polygone eingezeichnet und eine der vier Klassen, Wasser, Vegetation, Boden oder nicht grüne Vegetation zugewiesen (Abb. 20). Für den Umfang der Trainingsgebiete gibt es unterschiedliche Richtwerte, die in der Arbeit von Adugna, 2022 wie folgt zusammengefasst werden:

Es sollen pro Klasse **mindestens 100 Pixel** als Trainingsdaten ausgewiesen sein. Als ein anderer Richtwert wird angeführt, dass jede Klasse das **zehnfache der Anzahl an Klassen** als Trainingsdaten enthalten sein sollen, was in dem Fall dieser Arbeit 40 Pixel pro Klasse wären. Als letzten Richtwert für Trainingsdaten wird ausgewiesen, dass es sich um **mindestens 0,25% der Fläche** handeln soll. In dieser Arbeit werden alle drei Richtwerte erfüllt (Tab. 1).

Tabelle 1 Auflistung der Anzahl verwendeter Pixel pro Klasse für Trainingsdaten, sowie deren prozentualer Anteil an der Gesamtanzahl an Pixeln im Satellitenbild

Klasse	Anzahl der Pixel	Anzahl der Pixel in Prozent zur Gesamtfläche
Wasser	1551	2,0%
Vegetation	3892	5,1 %

Boden	2121	2,8%
Nicht grüne Vegetation	1923	2,5%

Die drei verwendeten Klassifizierungsalgorithmen, SVM, kNN und RF benötigen die Trainingsdaten (Abb. 19) in Form einer ROI-Maske als Rasterdatei, daher wird das in QGIS erstellte Shapefile mit Hilfe eines Python Skriptes in eine Raster Maske umgewandelt.

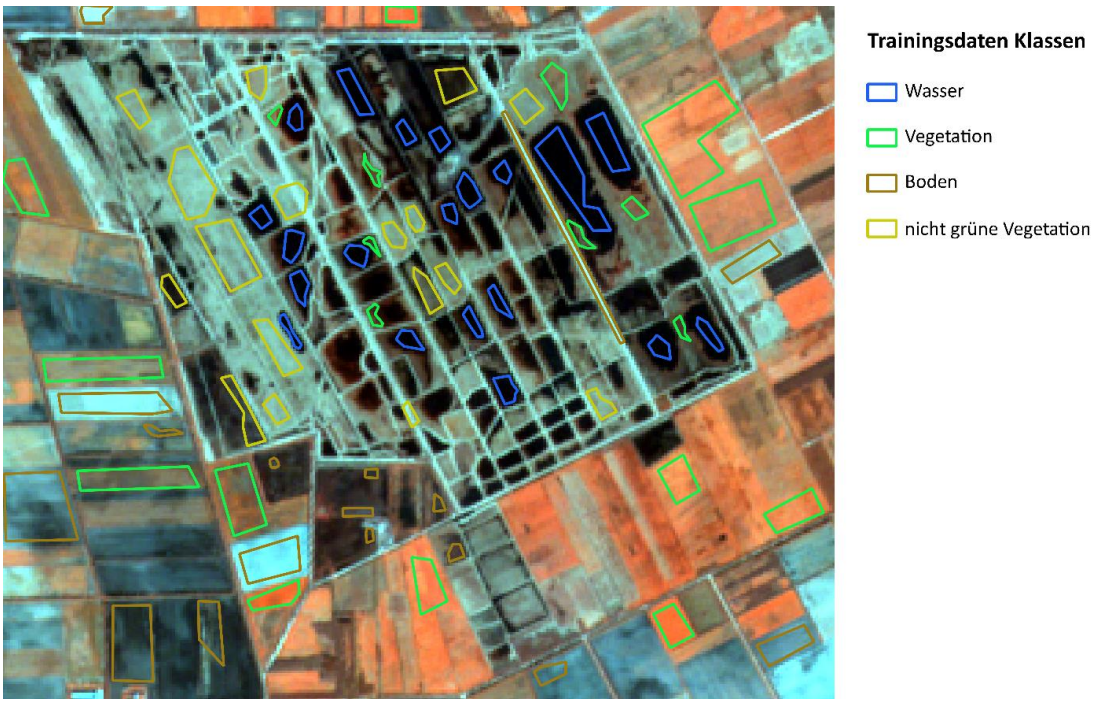

Abbildung 19 Alle verwendeten Trainingsdatenflächen mit der jeweiligen Klassesind farblich umrandet. Mit falschfarben Sentinel- 2 Aufnahme (NIR ersetzt den Rot Kanal)

Abbildung 20 Geländeaufnahmen, vom 20.03.2023, der vier verwendeten Klassen von links oben nach rechts unten: Wasser, Vegetation, Boden und nicht grüne Vegetation

3.10 Ground Truth Daten

Um die Qualität eines Modells im Accuracy Assessment zu bestimmen, wird ein im besten Falle vom Modell unabhängiger Datensatz zur Kontrolle verwendet. In dieser Arbeit wurde auf Grundlage einer

Drohnenaufnahme des Gebietes vom 07.03.2020 mit einer Auflösung von 10 cm ein Ground Truth Datensatz erstellt (*OpenGeoData.NI*, 2024). Dabei wurde händisch der gesamte Datensatz in eine der vier Klassen, Wasser, Vegetation, Boden oder nicht grüne Vegetation unterteilt (Abb. 21). Die Drohnenaufnahme ist dabei 16 Tage vor der Satellitenaufnahme vom 23.03.2020 aufgenommen, an dessen das Modell validiert werden soll. Der Zeitabstand zwischen den beiden Aufnahmen sollte gering genug sein, dass es noch keine nennenswerten Unterschiede in der Landbedeckung gibt.

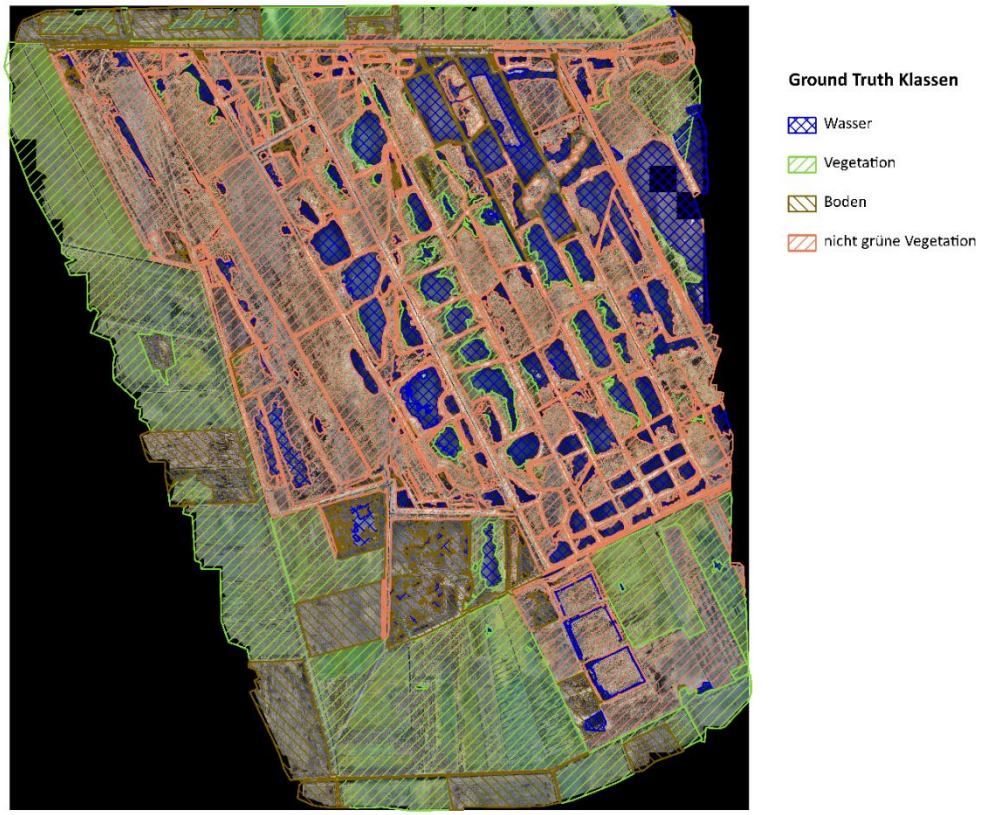

*Abbildung 21 Ground Truth Daten mit Drohnenaufnahme des Gebietes in 10 cm Auflösung (*OpenGeoData.NI*, 2024)*

3.11 Daten Prozessierung

Im ersten Schritt werden die Sentinel-2 Daten über die Copernicus API mit einem Python-Script für das Untersuchungsgebiet bei weniger als 15% Wolkenbedeckung heruntergeladen. Im Anschluss werden die Daten mit einem weiteren Python-Script geclipped. Dadurch das alle Sentinel 2 Bänder als einzelne Raster vorliegen werden diese im Schritt des Layer Stacks zu einer Raster Datei zusammengeführt. Um eine Klassifizierung der Daten durchführen zu können werden Trainingsdaten benötigt, diese werden in einem GIS-Programm eingezeichnet und mit einem Python-Script rasterisiert. Nun kann mit Hilfe der Trainingsdaten ein Modell trainiert werden, welches die Satellitendaten klassifiziert. Das klassifizierte Satellitenbild kann dann mit einem Ground Truth Datensatz verglichen werden, um so das Accuraccy Assessment zu erhalten (Abb. 22).

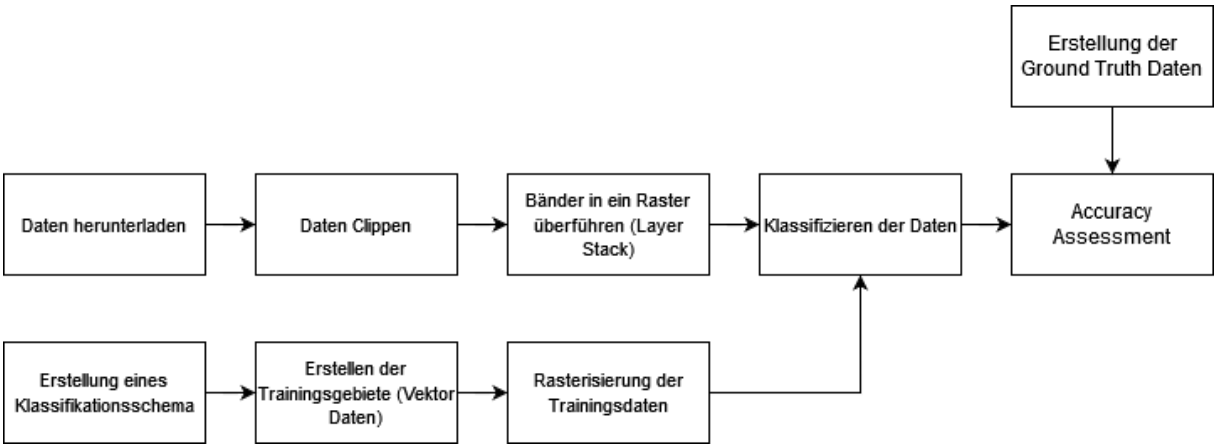

Abbildung 22 Workflow für die Arbeitsschritte dieser Arbeit

4. Ergebnisse

Die Modelle der SVM-, kNN- und RF-Algorithmen geben ein GeoTiff aus, bei dem jedes Pixel einen Wert entsprechend seiner Klasse, zugewiesen bekommt. In dieser Arbeit sind es die vier Klassen: Wasser, Vegetation, Boden und nicht grüne Vegetation, welchen in dieser Reihenfolge die Nummern 1-4 zugeordnet werden.

4.1 Ergebnisse Deep Learning

Die Ergebnisse des DeepWaterMap Modells von Isikdogan (2017) werden mit Hilfe von Graustufen dargestellt, welche die Wahrscheinlichkeit für Oberflächengewässer angeben. Je heller dabei ein Pixel, desto höher die Wahrscheinlichkeit, dass das Modell ein Oberflächengewässer erkannt haben will. Das Deeplearning-Modell gibt dabei die Ergebnisse als einfaches PNG aus. Die Ergebnisse müssten demnach wieder georeferenziert werden, damit diese wieder ihren räumlichen Bezug erhalten. Außerdem müsste noch eine Interpretation der Grauwerte durchgeführt werden, um die Pixel in Wasser und nicht Wasser zu unterteilen. Dies würde entweder einen festen Schwellenwert erfordern oder wiederrum einen ML-Algorithmus, der ein Modell erstellt, um die Grauwertergebnisse des Deeplearning-Modells zu interpretieren.

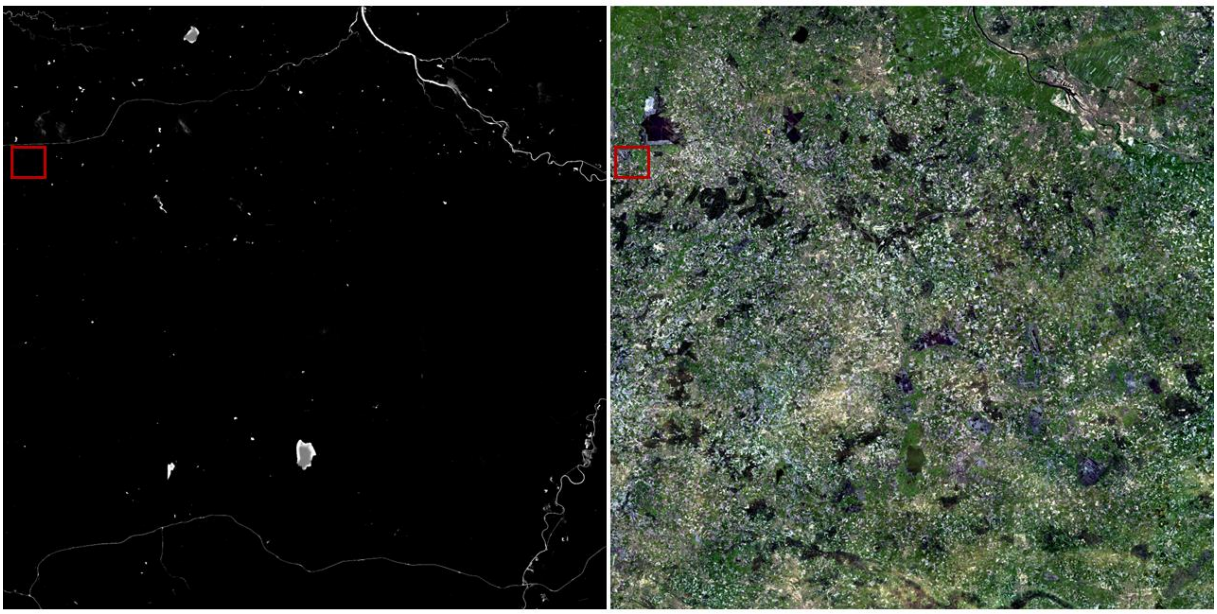

Abbildung 23 Ergbeniss des DeepWaterMap Modells für eine komplette Sentinel Kachel von 28.03.2020. Das rote Rechteck markiert das Leegmoor

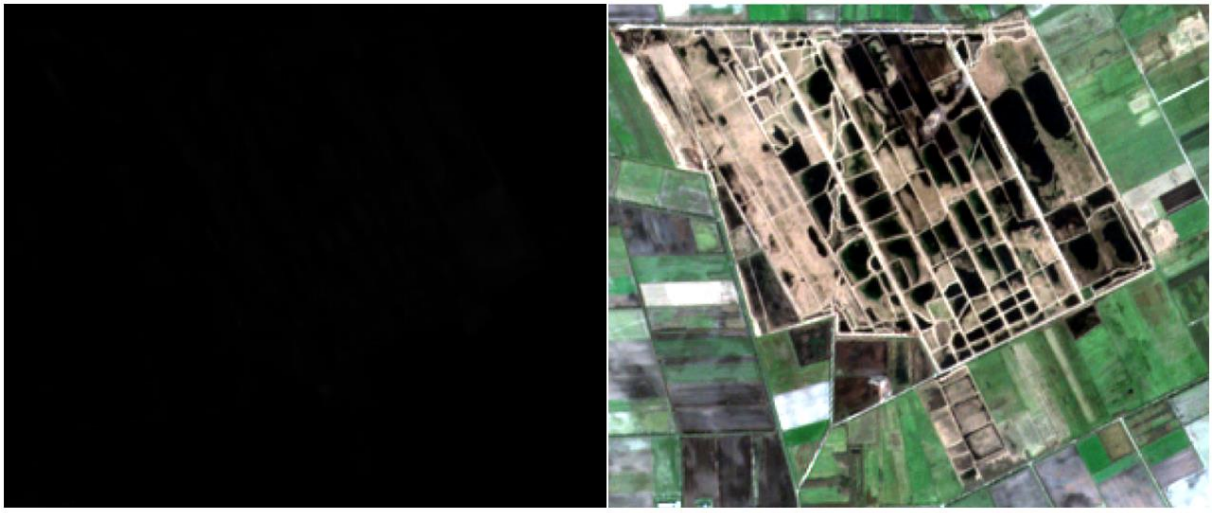

Abbildung 24 Links Ergebnis aus dem DeepWaterMap Modell für das Leegmoor und rechts die Sentinel 2 Datengrundlage vom 28.03.2020

Die Ergebnisse des DeepWaterMap Modells zeigen, dass Flussstrukturen sehr gut erkannt werden können. So ist der Küstenkanal, der sich oberhalb des Leegmoors befindet und Richtung Nordosten weiter verläuft, im Sentinel-2-Satellitenbild kaum zu erkennen (Abb. 23). Im Ergebnis des DeepWaterMap Modells ist dieser wiederum sehr gut zu erkennen. Das Modell scheint allerdings Probleme bei der Erkennung und Abgrenzung von Feuchtflächen mit niedrigem Wasserstand, wie es bei Mooren der Fall ist, zu haben. So erkennt das Modell für den Ausschnitt des Leegmoores keine Wasserfläche (Abb. 24). Das ausgegebene PNG ist komplett schwarz. Es wurde somit vom Modell kein Oberflächengewässer erkannt. Daher können die Ergebnisse aus dem DeepWaterMap Modell auch nicht weiter für ein Accuracy Assessment genutzt werden.

4.2 Ergebnisse der ML-Algorithmen kNN, SVM, RF

Im Folgenden werden die Klassifikationsergebnisse der drei ML-Algorithmen kNN, SVM und RF vorgestellt.

4.2.1 Klassifikationsergebnisse kNN

Die Unterschiede zwischen den vier gewählten Parametern sind optisch im Klassifikationsergebnis kaum zu erkennen. Alle vier Einstellungen weisen rundherum um das Gebiet des Leegmoors großflächige Vegetationsflächen auf, die teilweise von Bodenflächen unterbrochen werden. Es finden sich auch einige Flächen der nicht grüne Vegetationsklasse außerhalb des Gebietes des Leegmoors wieder. Dennoch dominiert hier deutlich die Vegetationsklasse. Auf dem Gebiet des Leegmoors ändert sich die Klassenzusammensetzung deutlich (Abb. 25). Hier dominiert vor allem die nicht grüne Vegetationsklasse, welche von den häufig vorkommenden Wasserflächen unterbrochen wird. Um die Wasserflächen herum finden sich in der Regel hierbei auch immer Säume von grüner Vegetation (Abb. 25). Die genaueren Unterschiede in den verschiedenen Parameter Einstellungen werden sich dann im Accuracy Assessment widerspiegeln.

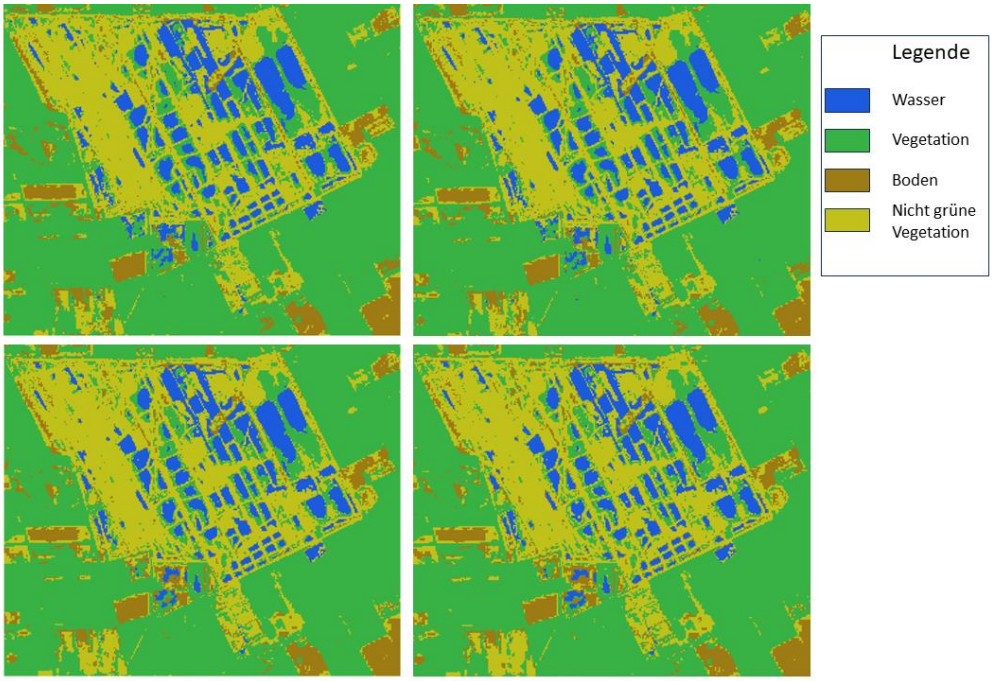

Abbildung 25 Klassifikationsergebnisse des kNN - Algorithmus Parameter k von links oben nach rechts unten sind wie folgt: k=3, k=5, k=15 und k=29

4.2.2 Klassifikationsergebnisse SVM

Die Unterschiede zwischen den Parametern sind für den SVM – Algorithmus gerade, wenn die Parameter C = 0,5 und C= 1 mit den Ergebnissen der Parameter C = 20 und C = 50 verglichen werden deutlich besser sichtbar. So gibt es deutliche Unterschiede in der Klassifikation der Wasserflächen am Rand des Leegmoors im Südwesten. Die Parameter C = 0,5 und C = 1 weisen hier deutlich zusammenhängender und größer ausgeprägte Wasserflächen auf als es bei den anderen beiden Parametern der Fall ist (Abb. 26). Auch gibt es deutliche Unterschiede bei der Klassifikation einer außerhalb des Leegmoores liegenden Fläche im Südwesten. Die Parameter C = 0,5 und C = 1 erkennen dort deutlich mehr Boden als es die anderen beiden Parameter tun. Von diesem Unterschied abgesehen, ähneln sich allerdings die Klassifikationsergebnisse der Flächen außerhalb des Leegmoors sehr. Auch hier dominiert die Vegetationsklasse, welche von Bodenflächen immer wieder

unterbrochen werden. Es finden sich außerdem einige wenige Flächen nicht grüne Vegetation außerhalb der Moorfläche. Auf der Fläche des Moores dominiert die nicht grüne Vegetationsklasse, die durch häufig vorkommende Wasserflächen unterbrochen wird. Entlang der Ränder der Wasserflächen finden sich häufig streifen von Vegetation wieder. Die Bodenklasse kommt beinahe gar nicht auf der Fläche des Moores vor (Abb. 26).

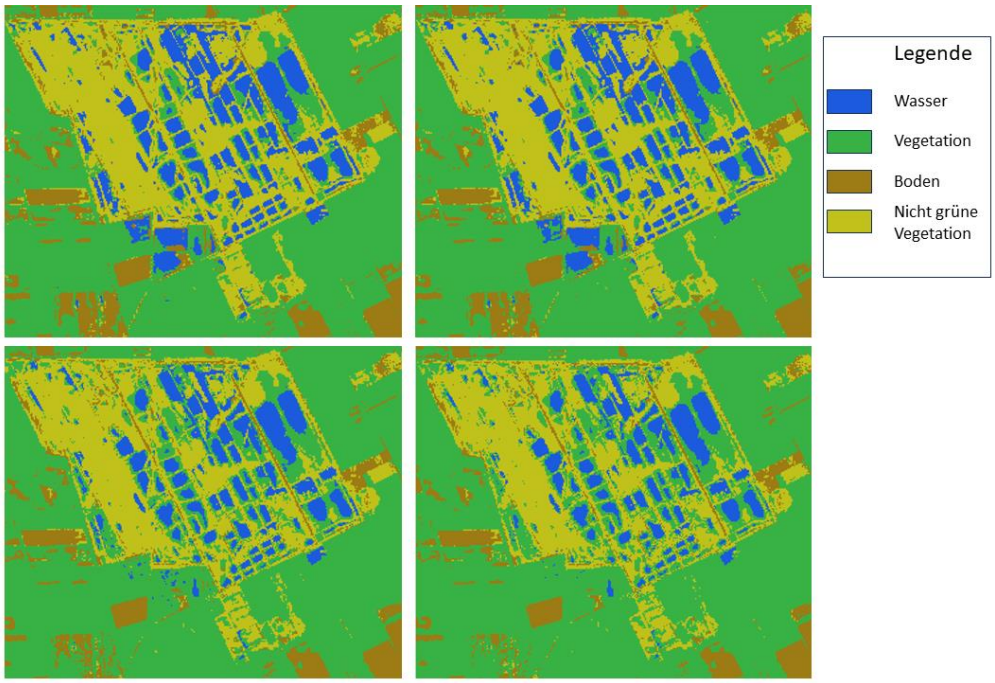

Abbildung 26 Klassifikationsergebnisse des SVM - Algorithmus Parameter C von links oben nach rechts unten sind wie folgt: C=0,5, C=1, C=20 und C=50

4.2.3 Klassifikationsergebnisse RF

Im Klassifikationsergebnis des RF-Algorithmus sind zwischen den Parametern auch optisch kaum Unterschiede zu erkennen. So ist das Klassifikationsergebnis außerhalb des Leegmoors von der Vegetationsklasse geprägt, welche vereinzelt durch die Bodenklasse und die nicht grüne Vegetationsklasse unterbrochen wird (Abb. 27). Innerhalb der Moorfläche ist die nicht grüne Vegetationsklasse stark vertreten und wird dabei häufig von Wasserflächen unterbrochen. Um die Wasserflächen herum finden sich dabei häufig Bereiche von der Vegetationsklasse. Die Bodenklasse ist auf der Moorfläche fast nicht vorhanden. Die Unterschiede zwischen den vier Parametern werden sich für den RF-Algorithmus eher im Accuracy Assessment zeigen.

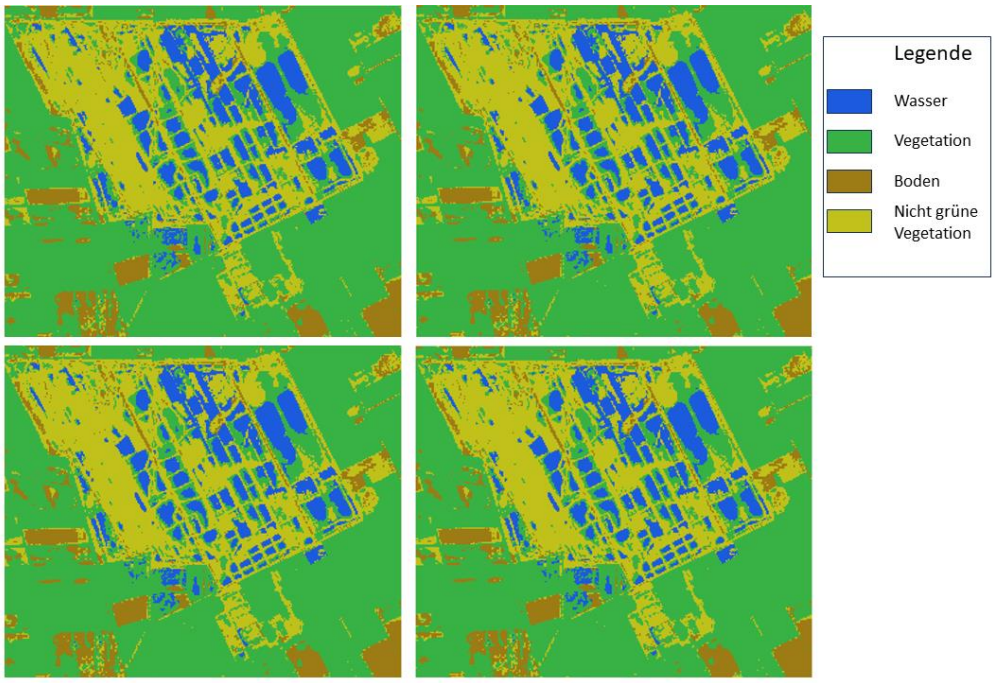

Abbildung 27 Klassifikationsergebnisse des RF - Algorithmus Parameter E von links oben nach rechts unten sind wie folgt: E=100, E=500, E=1000 und E=5000

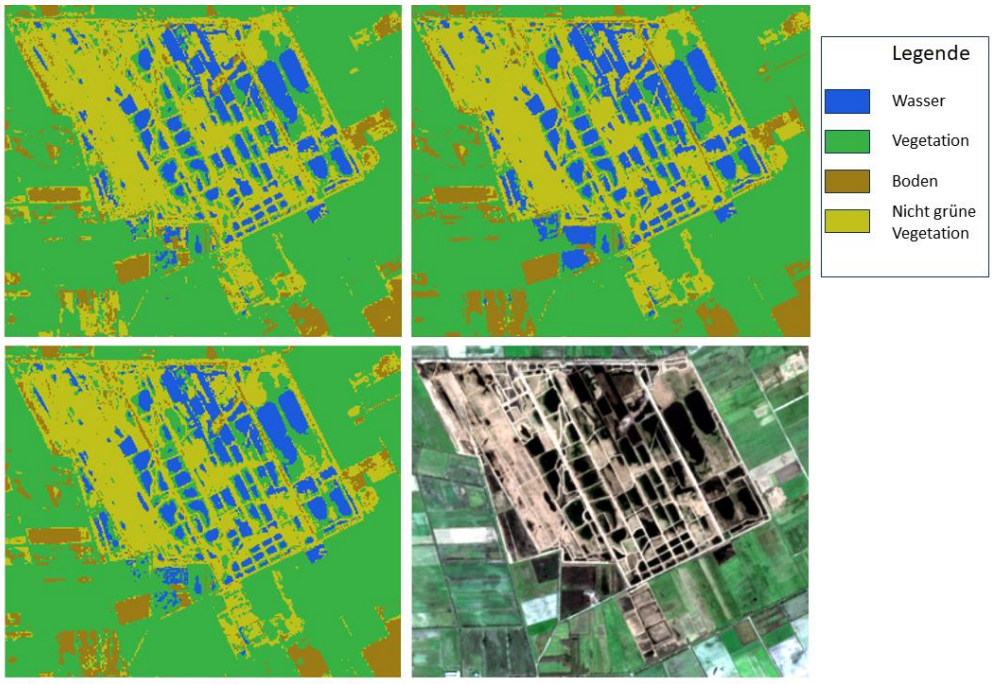

Abbildung 28 Klassifikation der drei verwendeten ML-Algorithmen mit entsprechender Sentinel 2 Aufnahme. Von links oben nach links unten die folgenden Algorithmen mit verwendetem Parameter: kNN bei k 5, SVM bei C 1 und RF bei E 100

4.3 Ergebnisse des Accuracy Assessment

Im folgendem werden die Ergebnisse des Accuracy Assessment für die drei ML-Algorithmen kNN, Rf und SVM beschrieben. Dabei wird die Error Matrix für jeden gewählten Parameter beschrieben, sowie zusätzlich Kennwerte des Accuracy Assessment, die sich in den Tabellen finden.

4.3.1 Ergebnisse des Accuracy Assessment für den kNN – Algorithmus:

Beginnend mit dem ersten gewählten Algorithmus kNN wurden vier Parameter im Accuracy Assessment untersucht, so wurde für k die folgenden Werte genutzt: 3, 5, 15 und 29.

Für den Parameter k = 3 wird für die Wasserklasse eine Sensitivität von 0,79 erreicht (Tab. 2 & Abb. 29). Diese setzt sich aus dem Verhältnis der richtig bestimmen Klassen (True positive) und der als fälschlicherweise als andere Klassen (False negative), klassifizierten, zusammen. Die Werte der False negative oder auch FN für die Wasserklasse liegen zwischen 0,064 und 0,075. Der höchste Wert mit 0,075 für fälschlicherweise nicht als Wasser klassifizierten Klassen, der Klasse nicht grüne Vegetation zugeordnet wird. Der niedrigste Wert mit 0,064 für nicht erkannte Wasserfläche findet sich in der Klasse Vegetation. Neben der Sensitivität ist die Präzision ein weiterer wichtiger Parameter im Accuracy Assessment. Dieser gibt das Verhältnis zwischen den richtig klassifizierten Klassen und den fälschlicherweise als positiv ausgegebenen Klassen wieder. Der Parameter gibt also an mit welcher Wahrscheinlichkeit es sich bei einer positiven Klassifikation einer Klasse auch tatsächlich um die Klasse handelt. Die Präzision für k = 3 liegt für die Klasse Wasser bei 0,87. Die Werte für fälschlicherweise als positiv (False positive) angenommene, Wasserklassen liegen für k = 3 zwischen 0,025 und 0,069 und haben ihr Maximum in der Klasse Boden und ihr Minimum in der Vegetationsklasse. Der f1-score, der sich aus Präzision und Sensitivität zusammensetzt, erreicht für die Klasse Wasser einen Wert von 0,83.

Die Klasse Vegetation bei einem Parameter von k = 3 erreicht eine Sensitivität von 0,8. Die Werte, der nicht erkannten Klassen reichen von 0,025 bis 0,12 (Tab. 2 & Abb. 29). Dabei ist das Minimum der Wasserklasse zuzuordnen und das Maximum der Bodenklasse. Die Vegetation Klasse erreicht eine Präzision von 0,64. Die Werte der fälschlicherweise als Vegetation ausgegebenen Klassen reichen dabei von 0,064 bis 0,15. Der geringste Wert ist in der Wasserklasse zu finden und der höchste Wert wird von der nicht grünen Vegetationsklasse ausgegeben. Der f1-score ist hier für die Vegetationsklasse 0,71.

Für die Bodenklasse ist die Sensitivität 0,71 bei einem k = 3 (Tab. 2 & Abb. 29). Die Werte für die fälschlicherweise negativen Klassifikationen liegen zwischen 0,069 und 0,12, dabei ist das Minimum in der Wasserklasse zu verorten und das Maximum in der Vegetationsklasse. Die Bodenklasse weist bei k = 3 eine Präzision von 0,55 auf. Die Werte der als fälschlicherweise als positiv ausgegebenen Klassifikationen in Bezug auf die Bodenklasse reichen von 0,071 bis 0,21. Das Minimum ist hier in der Wasserklasse zu finden und das Maximum in der nicht grünen Vegetationsklasse. Der f1-score liegt für die Bodenklasse bei 0,62.

Die Klasse nicht grüne Vegetation erreicht seine Sensitivität von 0,6 (Tab. 2 & Abb. 29). Die Werte für fälschlicherweise negative Klassifizierungen liegen dabei zwischen 0,044 und 0,21. Hier ist das Minimum in der Klasse Wasser zu finden und das Maximum in der Klasse Boden. Der Wert für die Präzision liegt für die Klasse nicht grüne Vegetation bei 0,79. Die Werte für fälschlicherweise positive Annahmen der Klasse nicht grüne Vegetation liegen zwischen 0,055 für die Vegetationsklasse und 0,11 in der Bodenklasse. Der f1-score liegt für die nicht grüne Vegetationsklasse bei 0,68.

Als weitere allgemeine und nicht klassenbezogene Größen werden für jeden Parameter die allgemeine Genauigkeit, die der Durchschnitt aller Sensitivitätswerte ist und der Kappa Koeffizient angegeben. Dieser gibt mit einem Wert zwischen -1 und 1 an, wie weit sich das Ergebnis vom Zufall unterscheidet. Ein Wert von -1 gibt dabei an das sich das Ergebnis nicht von zufällig gewählten Werten unterscheidet. Im Falle des Parameters k = 3 für den kNN – Algorithmus ergibt sich eine Gesamtgenauigkeit von 71,23 % und ein Kappa Wert von 0,62, sowie einen durchschnittlichen f1-score über alle Klasse von 0, 71 (Tab. 2 & Abb. 29).

Die Sensitivität für die Wasserklasse bei einem k von 5 beträgt 0,8 (Tab. 2 & Abb. 29). Die Werte für die False negatives liegen zwischen 0,053 und 0,083 und haben den niedrigsten Wert in den Klasse

Vegetation und den höchsten Wert in der Klasse Boden. Die Präzision für die Wasserklasse beträgt 0,87, während die Werte der False positives zwischen 0,03 und 0,061 liegen. Das Minimum ist dabei in der Vegetationsklasse mit 0,03 zu finden und das Maximum in der Bodenklasse mit 0,061. Der f1-score liegt bei einem Wert von 0,83.

Die Sensitivität für die Vegetationsklasse lieg bei 0,82 (Tab. 2 & Abb. 29). Der Anteil der fälschlicherweise nicht erkannten und anderen Klassen zugeordneten, liegt zwischen 0,03 und 0,11 Dabei ist der kleinste Wert der Wasserklasse zuzuordnen und der höchste Wert der Bodenklasse. Die Präzision der Vegetationsklasse beträgt 0,71, und die False positives liegen zwischen 0,053 und 0,12. Das Minimum befindet sich dabei in der Wasserklasse mit einem Wert von 0,053 und das Maximum findet sich in der nicht grüne Vegetationsklasse mit 0,12. Der f1-score liegt bei 0,76.

Die Klasse Boden weist eine Sensitivität von 0,71 bei einem k = 5 auf (Tab. 2 & Abb. 29). Der Anteil der False negatives liegt zwischen 0,061 und 0,13, dabei werden die niedrigsten Werte in der Wasserklasse erreicht und die höchsten Werte in der nicht grüne Vegetationsklasse. Die Präzision erreicht für die Bodenklasse einen Wert von 0,54 während die Werte der False positives zwischen 0,083 und 0,21 liegen. Der geringste Wert ist dabei innerhalb der Wasserklasse zu finden, während der höchste innerhalb der nicht grünen Vegetationsklasse zu beobachten ist. Der f1-score für die Bodenklasse bei einem k von 5 liegt bei 0,61.

Die Sensitivität der nicht grünen Vegetationsklasse liegt bei einem Wert von 0,62 bei einem k von 5 (Tab. 2 & Abb. 29). Die Werte für fälschlicherweise als negativ angenommen Klassen liegen dabei zwischen 0,047 und 0,21. Das Minimum liegt mit 0,047 in der Wasserklasse, während das Maximum bei 0,21 innerhalb der Bodenklasse liegt. Die Präzision für die nicht grüne Vegetationsklasse liegt bei einem Wert von 0,8. Die Anzahl der False positives innerhalb der Klassen liegt zwischen 0,046 und 0,13, dabei ist der niedrigste Wert innerhalb der Vegetationsklasse zu finden und der höchste Wert innerhalb der Bodenklasse zu finden. Der f1-score beträgt 0,70 für die nicht grüne Vegetationsklasse bei einem k von 5.

Die allgemeine Genauigkeit bei einem k von 5 liegt bei 72,93 % und einem Kappa Score von 0,64 sowie einen durchschnittlichen f1-score über alle Klasse von 0, 725 (Tab. 2 & Abb. 29).

Bei einem k von 15 ist der Wert für die Sensitivität der Wasserklasse 0,76 (Tab. 2 & Abb. 29). Die Werte für die False negatives liegen hierbei zwischen 0,052 und 0,14, während der niedrigste Wert von der Vegetationsklasse erreicht wird und der höchste von der Bodenklasse. Die Präzision beträgt für die Wasserklasse 0,90, dabei liegen die Werte der False positives zwischen 0,017 und 0,048. Der niedrigste Wert wird dabei mit 0,017 von der Vegetationsklasse erreicht, während der höchste Wert innerhalb der Bodenklasse bei 0,048 liegt. Der f1-score liegt für die Wasserklasse bei einem k von 15 bei 0,82.

Die Vegetationsklasse weist eine Sensitivität von 0,81 auf, bei der der höchste Wert für False negatives in der Bodenklasse mit 0,12 erreicht wird und der niedrigste mit 0,025 in der Wasserklasse zu finden ist (Tab. 2 & Abb. 29). Die Präzision der Vegetationsklasse liegt bei einem Wert von 0,67, während die Werte der False positives zwischen 0,052 und 0,14 liegen. Der niedrigste Wert ist dabei innerhalb der Wasserklasse anzutreffen und der höchste findet sich in der nicht grünen Vegetationsklasse wieder. Der f1-score liegt für die Vegetationsklasse bei einem k von 15 bei 0,74.

Die Bodenklasse weist eine Sensitivität von 0,7 auf (Tab. 2 & Abb. 29). Hier liegen die Werte der False negatives zwischen 0,048 und 0,13. Der geringste Wert liegt in der Wasserklasse und der höchste in der Vegetationsklasse. Die Präzision erreicht einen Wert von 0,47 für die Bodenklasse. Die Werte der False positives liegen dabei zwischen 0,12 in der Vegetationsklasse und 0,2 in der nicht grünen Vegetationsklasse. Der f1-score liegt bei 0,57 für die Bodenklasse bei einem k von 15.

Die nicht grüne Vegetationsklasse weist eine Sensitivität von 0,62 auf (Tab. 2 & Abb. 29). Dabei liegen die Werte der False negatives zwischen 0,038 innerhalb der Wasserklasse und 0,2 in der Bodenklasse. Die Präzision der nicht grünen Vegetationsklasse liegt bei 0,81. Die Werte der False positives liegen hierbei zwischen 0,053 und 0,12 und erreichen ihr Minimum in der Vegetationsklasse und ihr Maximum in der Bodenklasse. Der f1-score für die nicht grüne Vegetation bei einem k von 15 liegt bei 0,70.

Die allgemeine Genauigkeit bei einem k von 15 liegt bei 71,41 % und der Kappa Score erreicht einen Wert von 0,64, sowie einen durchschnittlichen f1-score über alle Klasse von 0, 708 (Tab. 2 & Abb. 29).

Die Sensitivität der Wasserklasse liegt bei einem k von 29 bei 0,74 (Tab. 2 & Abb. 29). Die False negatives reichen. Dabei von 0,044 innerhalb der Vegetationsklasse und 0,17 in der Bodenklasse. Die Präzision beträgt 0,91 für die Wasserklasse bei einem k von 29, dabei liegen die False positives zwischen 0,02 innerhalb der Vegetationsklasse und 0,042 in der Bodenklasse. Der f1-score beträgt 0,82 für die Wasserklasse bei einem k von 29.

Die Sensitivität der Vegetationsklasse liegt bei 0,83 für die Vegetationsklasse bei einem k von 29, dabei reichen die False negatives von 0,03 in der Wasserklasse bis 0,11 in der Bodenklasse (Tab. 2 & Abb. 29). Die Präzision erreicht einen Wert von 0,68 für die Vegetationsklasse bei einem k von 29. Die False positives liegen in einem Bereich von 0,044 innerhalb der Wasserklasse und 0,14 in der nicht grüne Vegetationsklasse. Der f1-score beträgt für die Vegetationsklasse 0,74.

Die Sensitivität der Bodenklasse beträgt 0,65 bei einem k von 29 (Tab. 2 & Abb. 29). Die False negatives erreichen dabei Werte von 0,042 innerhalb der Wasserklasse bis 0,17 in der nicht grüne Vegetationsklasse. Die Präzision liegt bei einem Wert von 0,43 bei einem k von 29. Die False positives erreichen dabei Werte von 0,088 innerhalb der Vegetationsklasse bis 0,23 in der nicht grüne Vegetationsklasse. Der f1-score der Bodenklasse liegt bei 0,52 bei einem k von 29.

Die Sensitivität, der nicht grünen Vegetationsklasse beträgt, 0,6 bei einem k von 29 (Tab. 2 & Abb. 29). Hier liegen die False negatives in einem Bereich von 0,037 für die Wasserklasse bis zu einem Wert von 0,23 in der Bodenklasse. Die nicht grüne Vegetationsklasse erreicht eine Präzision von 0,77 bei einem k von 29. Die Werte der False positives liegen dabei zwischen 0,047 innerhalb der Wasserklasse und 0,17 in der Bodenklasse. Der f1-score liegt dabei für die nicht grüne Vegetationsklasse bei 0,688 bei einem k von 29.

Die Gesamtgenauigkeit für ein k von 29 liegt bei 69,76%. Dabei erreicht der Kappa Score einen Wert von 0,60 und einen durchschnittlichen f1-score über alle Klasse von 0, 688 (Tab. 2 & Abb. 29).

Alle Modelle des kNN-Algorithmus weisen einen Wasserkontrollwert von 100% auf (Tab. 2).

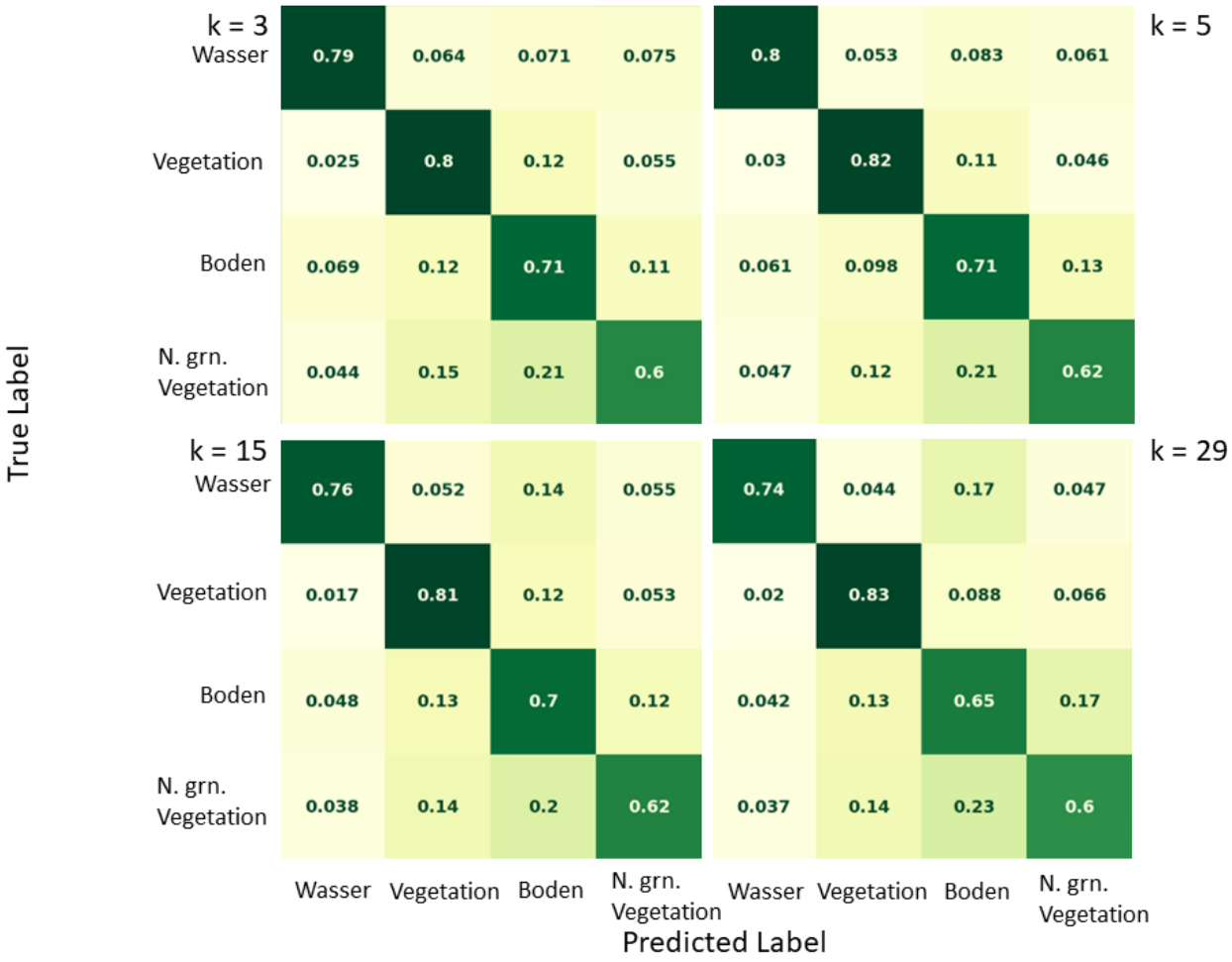

Abbildung 29 Ergebnisse des Accuracy Assessment für den kNN-Algorithmus bei variierendem Parameter k. Der Standardwert in der Scikit Learn Python Bibliothek ist dabei für k = 5.

Tabelle 2 Ergebnistabelle für verschiedene Parameter des kNN-Algorithmus. Der Standardwert für den Parameter k ist dabei in Scikit Learn 5. Zum Vergleich stehen die statistischen Kenngrößen Präzision, Sensitivität, f1-Score, Gesamtgenauigkeit, Wasserkontrollwert und Kappa

Klasse	Parameter	Präzision	Sensitivität	f1-score	Gesamt-genauigkeit	Wasserkontroll-wert [%]	Kappa
k	3						
Wasser		0,87	0,79	0,83			
Vegetation		0,64	0,80	0,71			
Boden		0,55	0,71	0,62			
Nicht grüne Vegetation		0,79	0,60	0,68			
Durchschnitt f1-score				0,71	71,23	100	0,62
k	5						
Wasser		0,87	0,80	0,83			
Vegetation		0,71	0,82	0,76			
Boden		0,54	0,71	0,61			
Nicht grüne Vegetation		0,80	0,62	0,70			
Durchschnitt f1-score				0,725	72,93	100	0,64

k	15						
Wasser		0,90	0,76	0,82			
Vegetation		0,67	0,81	0,74			
Boden		0,47	0,70	0,57			
Nicht grüne Vegetation		0,81	0,62	0,70			
Durchschnitt f1-score				0,708	71,41	100	0,62
k	29						
Wasser		0,91	0,74	0,82			
Vegetation		0,68	0,83	0,74			
Boden		0,43	0,65	0,52			
Nicht grüne Vegetation		0,77	0,60	0,67			
Durchschnitt f1-score				0,688	69,76	100	0,60

4.3.2 Ergebnisse des Accuracy Assessment für den SVM – Algorithmus

Für den SVM – Algorithmus wurden folgende Werte für den Parameter C gewählt: 0,5, 1, 20 und 50.

Bei einem C von 0,5 wird in der Wasserklasse eine Sensitivität von 0,8 erreicht (Tab. 3 & Abb. 30). Die False negatives liegen dabei zwischen 0,046 und 0,082, während der niedrigste Wert innerhalb der Vegetationsklasse zu finden ist und der höchste Wert in der Bodenklasse erreicht wird. Die Präzision für die Wasserklasse hat einen Wert von 0,79, während die False positives zwischen 0,034 und 1,4 liegen. Das Minimum ist in der Vegetationsklasse anzutreffen, während das Maximum in der Bodenklasse liegt. Der f1-score ist bei einem C von 0,5 in der Wasserklasse bei 0,79.

Der Wert der Sensitivität erreicht für die Vegetationsklasse 0,81 (Tab. 3 & Abb. 30). Die Werte für die False negatives reichen von 0,034 bis 0,12. Der niedrigste Wert wird hier sowohl von der Wasser - als auch von der nicht grünen Vegetationsklasse mit 0,034 erreicht und der höchste Wert liegt innerhalb der Bodenklasse mit 0,12. Die Präzision liegt für die Vegetationsklasse bei 0,76. Die Werte der False positives liegen dabei zwischen 0,046 und 0,098. Der geringste Wert ist in der Wasserklasse zu finden und der höchste in der nicht grünen Vegetationsklasse. Der f1-score liegt für die Vegetationsklasse bei einem C von 0,5 bei 0,79.

Die Sensitivität für die Bodenklasse bei einem C von 0,5 liegt bei 0,73 (Tab. 3 & Abb. 30). Die Werte der False negatives reichen von 0,14 in der Wasserklasse bis zu 0,051 innerhalb der nicht grünen Vegetationsklasse. Die Präzision der Bodenklasse liegt bei 0,54. Die False positives liegen dabei zwischen 0,082 innerhalb der Wasserklasse und 0,2 in der nicht grüne Vegetationsklasse. Der f1-score für die Bodenklasse liegt bei 0,62, während der Parameter C einen Wert von 0,5 hat.

Die Sensitivität liegt für die nicht grüne Vegetationsklasse bei 0,64 bei einem C von 0,5 (Tab. 3 & Abb. 30). Die Werte für die False negatives liegen zwischen 0,059 innerhalb der Wasserklasse und 0,2 in der Bodenklasse. Die Präzision liegt dabei bei einem Wert von 0,86 für die nicht grüne Vegetationsklasse. Die Werte der False positives erreichen dabei Werte von 0,034 in der Vegetationsklasse und 0,074 in der Wasserklasse. Der f1-score hat einen Wert für die nicht grüne Vegetationsklasse von 0,73 bei einem C von 0,5.

Die Allgemeine Genauigkeit für einen C von 0,5 liegt bei 73,78% und einem Kappa Score von 0,65 sowie einen durchschnittlichen f1-score über alle Klasse von 0, 733 (Tab. 3 & Abb. 30).

Bei einem C von 1 liegt die Sensitivität für die Wasserklasse bei einem Wert von 0,83, während die Werte der False negatives zwischen 0,53 innerhalb der Bodenklasse und 0,58 innerhalb der nicht grünen Vegetationsklasse liegen (Tab. 3 & Abb. 30). Die Präzision liegt für die Wasserklasse bei einem C von 1 bei 0,78. Die False positives reichen dabei von 0,034 in der Vegetationsklasse und 0,13 in der Bodenklasse. Die Wasserklasse erreicht einen f1-score von 0,81 bei einem C von 1.

Für die Vegetationsklasse liegt die Sensitivität bei einem C von 1 bei einem Wert von 0,8. Die Werte der False negatives reichen dabei von 0,034 bis 0,12 (Tab. 3 & Abb. 30). Der niedrigste Wert wird innerhalb der Wasserklasse erreicht, während der höchste innerhalb der Bodenklasse liegt. Die Präzision hat für die Vegetationsklasse bei einem C von 1 den Wert von 0,74. Der f1-score für die Vegetationsklasse bei einem C von 1 liegt bei 0,77.

Die Sensitivität erreicht innerhalb der Bodenklasse bei einem C von 1 einen Wert von 0,72 (Tab. 3 & Abb. 30). Die False negatives reichen hier von 0,066 bis 0,13. Das Minimum wird dabei von der nicht grüne Vegetationsklasse erreicht, während das Maximum von der Wasserklasse erreicht wird. Die Präzision für die Bodenklasse liegt bei einem C von 1 bei 0,57. Die False positives nehmen dabei Werte zwischen 0,053 innerhalb der Wasserklasse und 0,2 innerhalb der nicht grünen Vegetationsklasse an. Der f1-score beträgt 0,64 für die Bodenklasse bei einem C von 1.

Die Sensitivität für die nicht grüne Vegetationsklasse liegt bei einem C von 1 bei 0,63 (Tab. 3 & Abb. 30). Die False negatives nehmen hier Werte zwischen 0,062 in der Wasserklasse und 0,2 in der Bodenklasse an. Die Präzision hat einen Wert von 0,86 für die nicht grüne Vegetationsklasse bei einem C von 1. Dabei liegen die False positives zwischen 0,039 innerhalb der Vegetationsklasse und 0,066 innerhalb der Bodenklasse. Der f1-score liegt für die nicht grüne Vegetationsklasse bei 0,73 für ein C von 1.

Die Gesamtgenauigkeit für ein C von 1 liegt bei 73,69% und weist einen Kappa Score von 0,65 auf sowie einen durchschnittlichen f1-score über alle Klasse von 0, 738 (Tab. 3 & Abb. 30).

Die Sensitivität für die Wasserklasse bei einem C von 20 liegt bei 0,83 (Tab. 3 & Abb. 30). Die Werte der False negatives liegen zwischen 0,034 innerhalb der Bodenklasse und 0,093 innerhalb der Vegetationsklasse. Die Präzision der Wasserklasse liegt bei einem C von 20 bei 0,87. Die False positives liegen dabei zwischen 0,039 in der Bodenklasse und 0,048 in der Vegetationsklasse. Der f1-score beträgt 0,85 für die Wasserklasse bei einem C von 20.

Die Sensitivität für die Wasserklasse ergibt bei einem C von 20 einen Wert von 0,74 (Tab. 3 & Abb. 30). Dabei liegen die False negatives zwischen 0,048 innerhalb der Wasserklasse und 0,16 innerhalb der Bodenklasse. Die Vegetationsklasse erreicht eine Präzision von 0,58 bei einem C von 20. Die Werte der False positives liegen dabei zwischen 0,093 in der Wasserklasse und 0,15 sowohl bei der Bodenklasse als auch bei der nicht grünen Vegetationsklasse. Der f1-score beträgt für die Vegetationsklasse 0,65 bei einem C von 1.

Die Sensitivität für die Bodenklasse hat einen Wert von 0,74 bei einem C von 1 (Tab. 3 & Abb. 30). Hier liegen die Werte der False negatives zwischen 0,039 in der Wasserklasse und 0,15 in der Vegetationsklasse. Der Wert der Präzision ist im Falle der Bodenklasse 0,50 für ein C von 20. Die Werte der False positives liegen dabei zwischen 0,034 für die Wasserklasse und 0,23 für die nicht grüne Vegetationsklasse. Der f1-score beträgt für die Bodenklasse 0,60 bei einem C von 20.

Der Wert der Sensitivität für die nicht grüne Vegetationsklasse liegt bei 0,58 bei einem C von 20 (Tab. 3 & Abb. 30). Die False negatives zeigen Werte zwischen 0,047 innerhalb der Wasserklasse und 0,23 für die Bodenklasse. Die nicht grüne Vegetationsklasse erreicht eine Präzision von 0,86 bei einem C

von 20. Dabei liegen die False positives Werte zwischen 0,044 für die Wasserklasse und 0,66 innerhalb der Bodenklasse. Der f1-score ist für die nicht grüne Vegetationsklasse mit 0,68 anzugeben.

Die Gesamtgenauigkeit bei einem C von 20 beträgt 70,19%. Dabei wird ein Kappa Score von 0,60 erreicht sowie ein durchschnittlicher f1-score über alle Klasse von 0,698 (Tab. 3 & Abb. 30).

Die Sensitivität für die Wasserklasse beträgt 0,84 bei einem C von 50 (Tab. 3 & Abb. 30). Die Werte der False negatives liegen dabei zwischen 0,016 innerhalb der Bodenklasse und 0,11 innerhalb der Vegetationsklasse. Die Präzision liegt für die Wasserklasse bei 0,92, während die False positives zwischen 0,017 innerhalb der Bodenklasse und 0,044 innerhalb der Vegetationsklasse liegen. Der f1-score beträgt für die Wasserklasse 0,88 bei einem C von 50.

Der Wert für die Sensitivität liegt für die Vegetationsklasse bei 0,72 bei einem C von 50 (Tab. 3 & Abb. 30). Die False negative Werte erreichen dabei Werte zwischen 0,044 und 0,16. Das Minimum liegt innerhalb der Wasserklasse und das Maximum innerhalb der Bodenklasse. Die Präzision der Vegetationsklasse liegt 0,53. Hier reichen die Werte der False positives zwischen 0,11 für die Wasserklasse und 0,17 innerhalb der Bodenklasse. Der f1-score beträgt 0,61 für die Vegetationsklasse bei einem C von 50.

Die Sensitivität der Bodenklasse beträgt 0,71 bei einem C von 50 (Tab. 3 & Abb. 30). Die Flase negatives reichen dabei von 0,017 innerhalb der Wasserklasse und 0,17 innerhalb der Vegetationsklasse. Die Bodenklasse erreicht dabei eine Präzision von 0,50 bei einem C von 50. Die False positives liegen dabei zwischen 0,016 in der Wasserklasse und 0,25 innerhalb der nicht grünen Vegetationsklasse. Der f1-score liegt bei 0,59 für die Bodenklasse bei einem C von 50.

Die Sensitivität für die nicht grüne Vegetationsklasse beträgt 0,57 bei einem C von 50 (Tab. 3 & Abb. 30). Die False negatives reichen dabei von 0,026 bis 0,25. Der geringste Wert wird von der Wasserklasse und der höchste Wert von der Bodenklasse erreicht. Die Präzision beträgt 0,84 für die nicht grüne Vegetationsklasse bei einem C von 50. Die False positives liegen zwischen 0,038 in der Wasserklasse und 0,097 innerhalb der nicht grünen Vegetationsklasse. Der f1-score liegt bei 0,68 für die nicht grüne Vegetationsklasse.

Die Gesamtgenauigkeit beträgt 69,72% und ein Kappa Score von 0,59 sowie einem durchschnittlichen f1-score über alle Klasse von 0,69 bei einem C von 50 (Tab. 3 & Abb. 30).

Alle Modelle des SVM-Algorithmus weisen einen Wasserkontrollwert von 100% auf (Tab. 3).

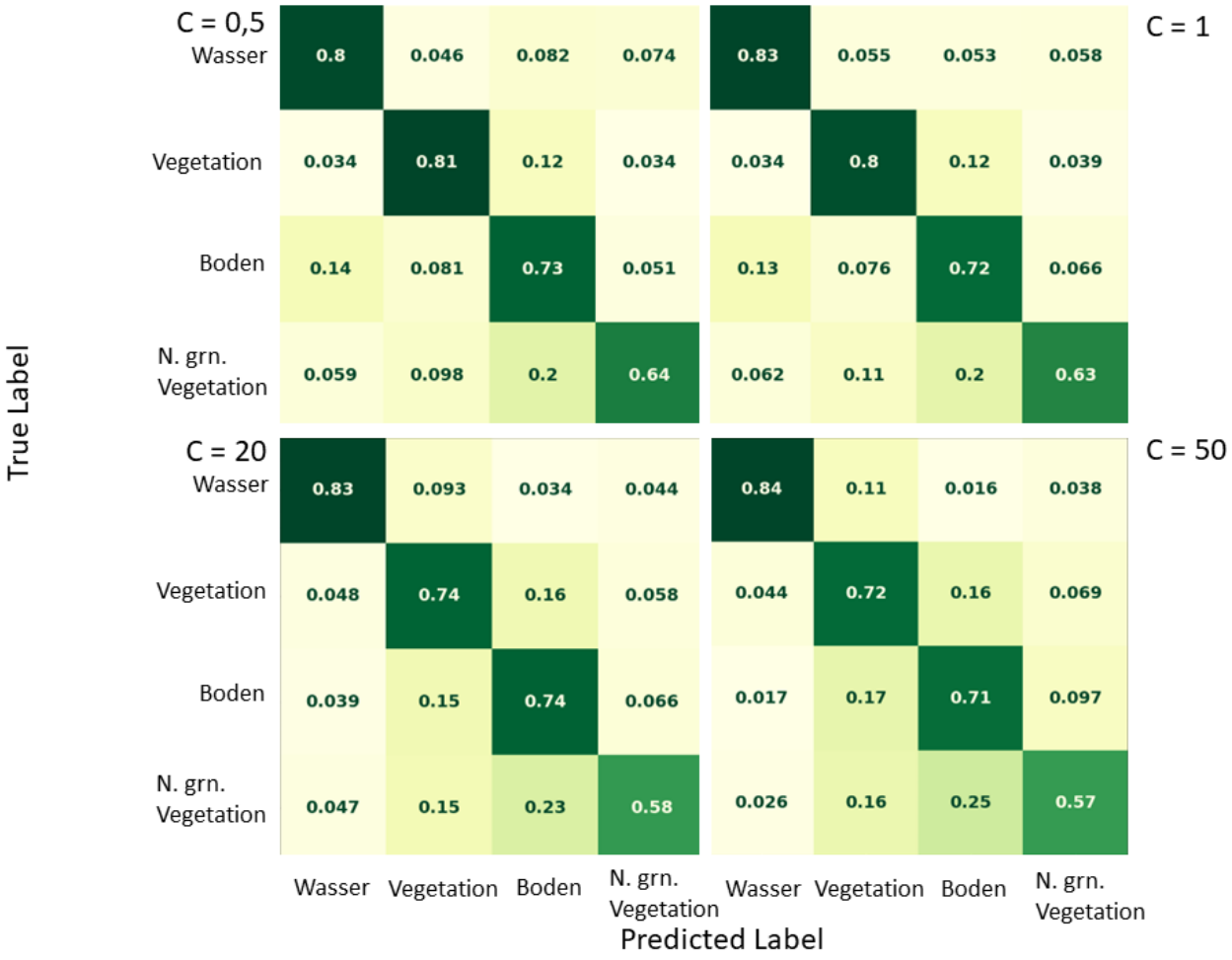

Abbildung 30 Ergebnisse des Accuracy Assessment für den SVM-Algorithmus bei variierendem Parameter C. Der Standardwert in der Scikit Learn Python Bibliothek ist C = 1.

Tabelle 3 Ergebnistabelle für verschiedene Parameter des SVM-Algorithmus. Der Standardwert für den Parameter C ist dabei in Scikit Learn 1. Zum Vergleich stehen die statistischen Kenngrößen Präzision, Sensitivität, f1-Score, Gesamtgenauigkeit, Wasserkontrollwert und Kappa

Klasse	Parameter	Präzision	Sensitivität	f1-score	Gesamt-genauigkeit	Wasserkontroll-wert [%]	Kappa
C	0,5						
Wasser		0,79	0,80	0,79			
Vegetation		0,76	0,81	0,79			
Boden		0,54	0,73	0,62			
Nicht grüne Vegetation		0,86	0,64	0,73			
Durchschnitt f1-score				0,733	73,78	100	0,65
C	1						
Wasser		0,78	0,83	0,81			
Vegetation		0,74	0,80	0,77			
Boden		0,57	0,72	0,64			
Nicht grüne Vegetation		0,86	0,63	0,73			

Durchschnitt f1-score				0,738	73,69	100	0,65
C	20						
Wasser		0,87	0,83	0,85			
Vegetation		0,58	0,74	0,65			
Boden		0,50	0,74	0,60			
Nicht grüne Vegetation		0,86	0,58	0,69			
Durchschnitt f1-score				0,698	70,19	100	0,60
C	50						
Wasser		0,92	0,84	0,88			
Vegetation		0,53	0,72	0,61			
Boden		0,50	0,71	0,59			
Nicht grüne Vegetation		0,84	0,57	0,68			
Durchschnitt f1-score				0,69	69,72	100	0,59

4.3.3 Ergebnisse des Accuracy Assessment für den RF – Algorithmus:

Als letztes werden die Ergebnisse für den Rf – Algorithmus beschrieben. Hierbei ändert sich der Parameter E, welcher die Anzahl der erstellten Entscheidungsbäume innerhalb des Algorithmus angibt. Die Python Bibliothek Scikit Learn gibt dabei einen Standartwert für E gleich 100 an. Es werden im Folgenden die Ergebnisse des Accuracy Assessment für die Werte für E beschrieben: 100, 500, 1000 und 5000.

Die Sensitivität der Wasserklasse bei einem E von 100 ist gleich 0,81 (Tab. 4 & Abb. 31). Die False negatives erreichen dabei Werte zwischen 0,052 in der Bodenklasse und 0,074 in der Vegetationsklasse. Die Präzision liegt für die Wasserklasse bei 0,84 bei einem E von 100. Die False positives liegen zwischen 0,05 innerhalb der Vegetationsklasse und 0,083 in der Bodenklasse. Der f1-score beträgt 0,83 für die Wasserklasse bei einem E von 100.

Die Sensitivität der Vegetationsklasse liegt bei 0,78 bei einem E von 100 (Tab. 4 & Abb. 31). Die False negatives erreichen dabei Werte, die zwischen 0,045 innerhalb der Wasserklasse und 0,12 in der Bodenklasse liegen. Die Präzision liegt bei einem E von 100 für die Vegetationsklasse bei 0,66. Die False positives zeigen Werte im Bereich zwischen 0,074 innerhalb der Wasserklasse und 0,13 in der nicht grünen Vegetationsklasse. Der f1-score ist für die Vegetationsklasse 0,71 bei einem E von 100.

Die Sensitivität der Bodenklasse liegt bei 0,73 für ein E von 100 (Tab. 4 & Abb. 31). Die False negatives erreichen hier Werte zwischen 0,073 in der Wasserklasse und 0,12 in der Vegetationsklasse. Die Präzision beträgt für die Bodenklasse 0,58 bei einem E von 100. Die False positives liegen im Bereich von 0,052 innerhalb der Wasserklasse und 0,2 innerhalb der nicht grünen Vegetationsklasse. Der f1-score beträgt für die Bodenklasse 0,65 bei einem E von 100.

Die Sensitivität für die nicht grüne Vegetationsklasse liegt bei 0,62 (Tab. 4 & Abb. 31). Dabei erreichen die False negatives Werte von 0,049 in der Wasserklasse und 0,2 in der Bodenklasse. Die nicht grüne Vegetationsklasse hat eine Präzision von 0,83 bei einem E von 100. Die False positives liegen in einem Bereich von 0,055 innerhalb der Wasserklasse und 0,078 in der Bodenklasse. Der f1-score für die nicht grüne Vegetationsklasse liegt bei 0,71.

Die Gesamtgenauigkeit bei einem E von 100 liegt bei 72,63%, während ein Kappa Score von 0,64 erreicht wird sowie einem durchschnittlichen f1-score über alle Klasse von 0,725 (Tab. 4 & Abb. 31).

Die Sensitivität der Wasserklasse beträgt 0,81 bei einem E von 500 (Tab. 4 & Abb. 31). Die False negatives liegen dabei im Bereich von 0,039 in der Bodenklasse und 0,076 innerhalb der Vegetationsklasse und der nicht grüne Vegetationsklasse. Die Präzision beträgt für die Wasserklasse 0,82 bei einem E von 500, dabei erreichen die False positives Werte zwischen 0,052 in der Vegetationsklasse und 0,074 in der Bodenklasse. Der f1-score beträgt 0,82 bei einem E von 500.

Die Sensitivität der Vegetationsklasse beläuft sich auf 0,76 bei einem E von 500 (Tab. 4 & Abb. 31). Die False negatives liegen dabei im Bereich zwischen 0,042 in der nicht grünen Vegetationsklasse und 0,14 für die Bodenklasse. Die Präzision für die Vegetationsklasse ist mit einem Wert von 0,61 zu beziffern bei einem E von 500. Die False positives reichen von 0,076 innerhalb der Wasserklasse bis hin zu 0,16 in der nicht grünen Vegetationsklasse. Der f1-score beträgt für die Vegetationsklasse 0,68 bei einem E von 500.

Die Sensitivität der Bodenklasse liegt bei 0,7 (Tab. 4 & Abb. 31). Dabei erreichen die False negatives Werte von 0,074 für die Wasserklasse und 0,11 sowohl für die Vegetationsklasse als auch für die nicht grüne Vegetationsklasse. Die Präzision der Bodenklasse beträgt 0,53 bei einem E von 500. Die Werte der Flase positives liegen zwischen 0,039 innerhalb der Wasserklasse und 0,22 in der nicht grüne Vegetationsklasse. Der f1-score beträgt für die Bodenklasse 0,60.

Die Sensitivität liegt für die nicht grüne Vegetationsklasse bei 0,57 (Tab. 4 & Abb. 31). Dabei liegen die False negatives im Bereich von 0,055 innerhalb der Wasserklasse und 0,22 in der Bodenklasse. Die Präzision beträgt für die nicht grüne Vegetationsklasse 0,80 bei einem E von 500. Die False positives erreichen dabei Werte von 0,042 innerhalb der Vegetationsklasse und 0,11 für die Bodenklasse. Der f1-score liegt bei 0,66 für die nicht grüne Vegetationsklasse.

Die Gesamtgenauigkeit für ein E von 500 beträgt 69,30% bei einem Kappa Score von 0,59 und einem durchschnittlichen f1-score über alle Klasse von 0,69 (Tab. 4 & Abb. 31).

Die Sensitivität der Wasserklasse bei einem E von 1000 beträgt 0,82 (Tab. 4 & Abb. 31). Die False negatives liegen zwischen 0,055 für die nicht grüne Vegetationsklasse und 0,067 innerhalb der Vegetationsklasse. Die Präzision beträgt 0,83 bei einem E von 1000 für die Wasserklasse. Hier erreichen die False positives Werte von 0,05 für die Vegetationsklasse und 0,083 innerhalb der Bodenklasse. Der f1-score für die Wasserklasse beträgt 0,82 bei einem E von 1000.

Die Sensitivität der Vegetationsklasse beträgt 0,73 bei einem E von 1000 (Tab. 4 & Abb. 31). Die Werte der False negatives sind dabei in einem Bereich zwischen 0,045 in der Wasserklasse und 0,14 innerhalb der Bodenklasse. Die Präzision der Vegetationsklasse beträgt 0,61. Die False positives liegen zwischen 0,067 innerhalb der Wasserklasse und 0,17 in der nicht grünen Vegetationsklasse. Der f1-score liegt für die Vegetationsklasse bei 0,67 bei einem E von 1000.

Die Sensitivität der Bodenklasse liegt bei einem E von 1000 bei 0,73, dabei erreichen die False negatives Werte von 0,078 für die nicht grüne Vegetationsklasse und 0,11 innerhalb der Vegetationsklasse (Tab. 4 & Abb. 31). Die Präzision für die Bodenklasse liegt bei 0,56 bei einem E von 1000. Die False positives liegen hierbei zwischen 0,061 innerhalb der Wasserklasse und 0,18 in der nicht grünen Vegetationsklasse. Der f1-score beträgt für die Bodenklasse 0,64 bei einem E von 1000.

Die Sensitivität der nicht grünen Vegetationsklasse beträgt 0,6 bei einem E von 1000 (Tab. 4 & Abb. 31). Die False negatives erreichen dabei Werte von 0,051 innerhalb der Wasserklasse bis hin zu 0,18 für die Bodenklasse. Die Präzision der nicht grünen Vegetationsklasse liegt dabei bei 0,82. Die Werte der False positives liegen dabei in einem Bereich zwischen 0,055 in der Wasserklasse und 0,08

innerhalb der Bodenklasse. Der f1-score für die nicht grüne Vegetationsklasse beträgt 0,69 bei einem E von 1000.

Die Gesamtgenauigkeit bei einem E von 1000 beträgt 70,72% und erreicht einen Kappa Score von 0,61 mit einem durchschnittlichen f1-score über alle Klasse von 0,705 (Tab. 4 & Abb. 31).

Die Sensitivität der Wasserklasse liegt bei 0,82 bei einem E von 5000 (Tab. 4 & Abb. 31). Die False negatives reichen dabei von 0,048 innerhalb der nicht grünen Vegetationsklasse und 0,076 in der Vegetationsklasse. Die Präzision liegt bei 0,84 für die Wasserklasse bei einem E von 5000. Die False positives erreichen dabei Werte von 0,04 für die Vegetationsklasse bis 0,063 für die nicht grünen Vegetationsklasse. Der f1-score beträgt 0,83 für die Wasserklasse bei einem E von 5000.

Die Sensitivität der Vegetationsklasse liegt bei 0,76 für ein E von 5000 (Tab. 4 & Abb. 31). Dabei liegen die False negatives zwischen 0,04 innerhalb der Wasserklasse und 0,15 für die Bodenklasse. Die Präzision für die Vegetationsklasse beträgt 0,65 bei einem E von 5000. Die False positives liegen in einem Bereich zwischen 0,075 innerhalb der Wasserklasse und 0,14 in der nicht grüne Vegetationsklasse. Der f1-score der Vegetationsklasse beträgt 0,70 bei einem E von 5000.

Die Sensitivität der Bodenklasse liegt bei 0,72 bei einem E von 5000 (Tab. 4 & Abb. 31). Die False negatives reichen von 0,055 in der Wasserklasse bis 0,12 innerhalb der nicht grünen Vegetationsklasse. Die Präzision erreicht einen Wert von 0,52 für die Bodenklasse bei einem E von 5000, dabei erreichen die False positives Werte von 0,06 für die Wasserklasse bis zu 0,21 in der nicht grünen Vegetationsklasse. Der f1-score beträgt 0,61.

Die Sensitivität der nicht grünen Vegetationsklasse erreicht einen Wert von 0,59 bei einem E von 5000 (Tab. 4 & Abb. 31). Hier liegen die Werte der False negatives zwischen 0,063 für die Wasserklasse und 0,21 innerhalb der Bodenklasse. Die Präzision erreicht einen Wert von 0,82 für die nicht grüne Vegetationsklasse bei einem E von 5000. Die False positives liegen dabei zwischen 0,048 innerhalb der Wasserklasse und 0,12 in der Bodenklasse. Der f1-score für die nicht grüne Vegetationsklasse liegt bei 0,69 bei einem E von 5000.

Die Gesamtgenauigkeit bei einem E von 5000 liegt bei 70,84% bei einem Kappa Score von 0,61 und einem durchschnittlichen f1-score über alle Klasse von 0,708 (Tab. 4 & Abb. 31).

Alle Modelle des RF-Algorithmus weisen einen Wasserkontrollwert von 100% auf (Tab. 4).

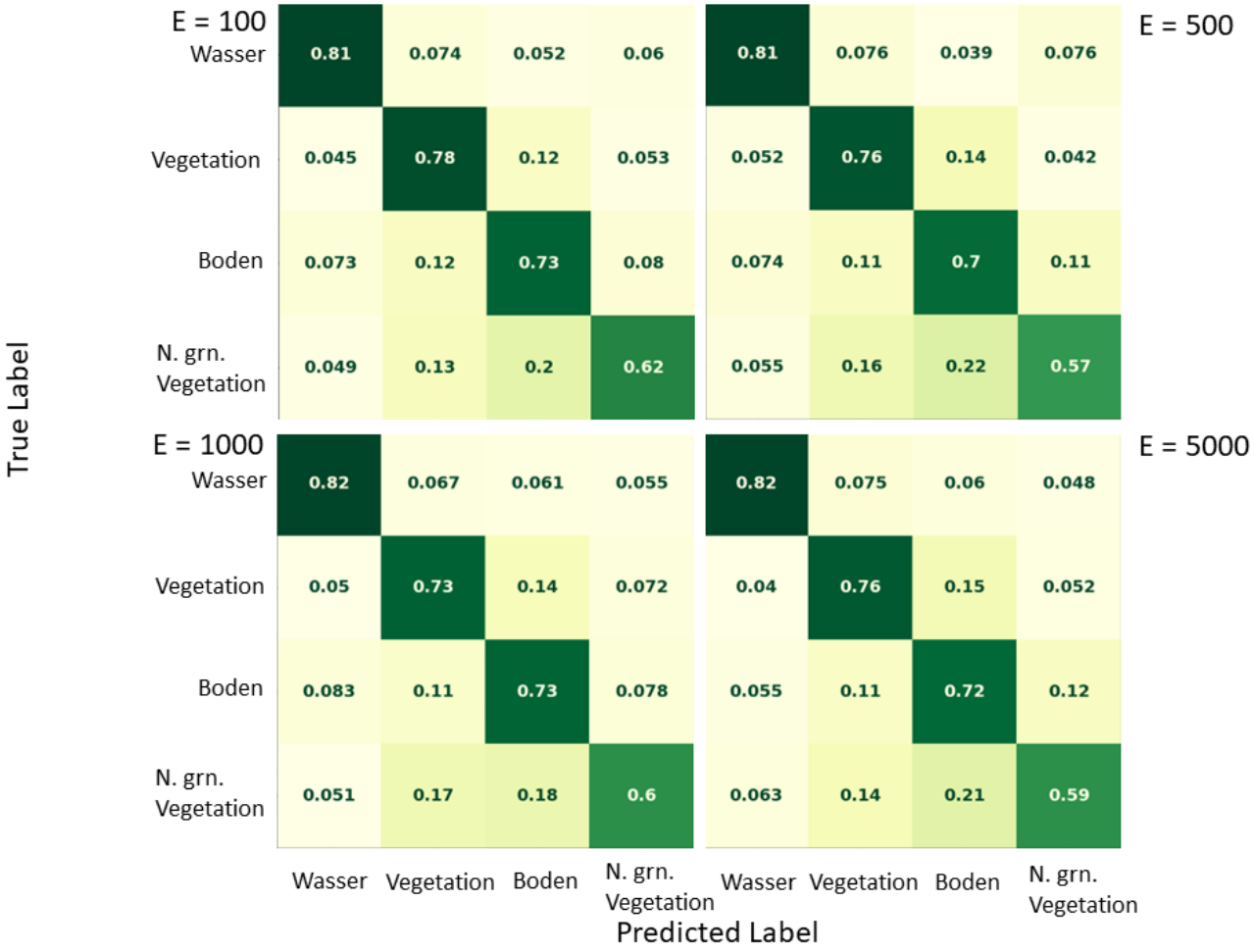

Abbildung 31 Ergebnisse des Accuracy Assessment für den RF-Algorithmus bei variierendem Parameter E. Der Standardwert in der Scikit Learn Python Bibliothek ist dabei für E = 100.

Tabelle 4 Ergebnistabelle für verschiedene Parameter des RF-Algorithmus. Der Standardwert für den Parameter Estimator ist dabei in Scikit Learn 100. Zum Vergleich stehen die statistischen Kenngrößen Präzision, Sensitivität, f1-Score, Gesamtgenauigkeit, der Wasserkontrollwert und Kappa

Klasse	Parameter	Präzision	Sensitivität	f1-score	Gesamt-genauigkeit	Wasserkontroll-wert [%]	Kappa
E	100						
Wasser		0,84	0,81	0,83			
Vegetation		0,66	0,78	0,71			
Boden		0,58	0,73	0,65			
Nicht grüne Vegetation		0,83	0,62	0,71			
Durchschnitt f1-score				0,725	72,63	100	0,64
E	500						
Wasser		0,82	0,81	0,82			
Vegetation		0,61	0,76	0,68			
Boden		0,53	0,70	0,60			

Nicht grüne Vegetation		0,80	0,57	0,66			
Durchschnitt f1-score				0,69	69,30	100	0,59
E	1000						
Wasser		0,83	0,82	0,82			
Vegetation		0,61	0,73	0,67			
Boden		0,56	0,73	0,64			
Nicht grüne Vegetation		0,82	0,60	0,69			
Durchschnitt f1-score				0,705	70,72	100	0,61
E	5000						
Wasser		0,84	0,82	0,83			
Vegetation		0,65	0,76	0,70			
Boden		0,52	0,72	0,61			
Nicht grüne Vegetation		0,82	0,59	0,69			
Durchschnitt f1-score				0,708	70,84	100	0,61

5. Diskussion und Interpretation

Die Ergebnisse des DeepWaterMap Modells ließen keine Auswertung durch das Accuracy Assessment zu, da das Modell im Untersuchungsgebiet keinerlei Oberflächengewässer erkannt hat. Das Modell hat allerdings vielversprechende Ergebnisse im Bereich der Fließgewässer gezeigt und diese sehr gut detektieren können. Um das DeepWaterMap Modell für die Klassifikation von Wasserflächen in Moorgebieten zu nutzen, wäre daher eine Anpassung des Modells mit neuen Trainingsdaten von Nöten oder die Erstellung eines Deep Learning Modells spezifisch für diese Aufgabe.

In den klassifizierten Ergebnissen des kNN-Algorithmus sind optisch über alle Parameter kaum Unterschiede auszumachen. In den Ergebnissen der vier Parameter, sind die gut abgegrenzten Wasserflächen im Zentrum und im Osten des Gebietes gut erkannt (Abb. 29). Im Süden des Leegmoors befindet sich eine Fläche von trockenem Schwarztorf. Dieser wirkt optisch in den Satellitendaten sehr ähnlich den Wasserflächen, da beide sehr dunkel sind. Der Algorithmus erkennt über alle Parameter hier eine Mischung aus Wasser, Boden, Vegetation und nicht grüne Vegetation. Auf der Schwarztorffläche kann in der Drohnenaufnahme erkannt werde, dass sich in bestimmten Bereichen etwas Wasser angesammelt hat, dies erkennt der Algorithmus gut. Allerdings erkennen die Modelle für die Parameter k = 3 und k = 5 auf dieser Fläche sehr viel Vegetation, welche so gut wie gar nicht vorhanden ist. Die Modelle k = 15 und k = 29 hingegen klassifizieren die Fläche vor allem als Boden und zu geringen Anteilen als nicht grüne Vegetation. Dieser Unterschied kann darauf zurück zuführen sein, dass mit k = 3 und 5 eine zu geringe Menge an Nachbarn betrachtet werden und eine etwas größere Generalisierung des Modells, bessere Ergebnisse für diese Schwarztorfflächen erzielen würde, wie es in den Ergebnissen mit den Parametern k = 15 und k = 29 auch zu sehen ist. Dem Modell gelingt es demnach erst bei höheren k Werten besser, diese Schwarztorfflächen der Bodenklasse zu zuordnen.

Eine weitere Fläche, auf der Unterschiede zwischen den Modellen, der vier Parameter zu sehen sind, sind zwei größere Sandwege, die sich vom Norden in den Süden des Leegmoors erstrecken (Abb. 29). Dabei ordnen die Modelle der Parameter k = 3 und k = 5 diese beiden Wege besser als Bodenklasse ein. Die Modelle der Parameter k = 15 und k = 29 erkennen auf diesen Flächen zum großen Teil nur

nicht grüne Vegetation. Eine Ursache hierfür könnte die Größe der Struktur sein. So ist der Weg etwa 3- 4m breit und füllt somit kein ganzes Pixel bei der Sentinel 2 Auflösung von maximal 10 m pro Pixel aus. Jedes Pixel in dem einer der beiden Wege enthalten ist, wird daher immer noch andere Signale, anderer Klassen enthalten. Die Modelle k = 15 und k = 29 generalisieren allerdings mehr als die Modelle der Parameter k = 3 und k =5, es kann daher sein, dass solche Subpixel Strukturen besser von den Modellen mit geringem k erkannt werden können.

Um die Genauigkeit und die Güte eines Modells zu bewerten, reicht die alleinige optische Betrachtung der Ergebnisse nicht aus. Die Ergebnisse des Accuracy Assessment geben hierfür spezielle Parameter aus, welche die Qualität des Modells beschreiben. Die Kennzahlen für die Genauigkeit für die Klassifizierung von Wasserflächen ist bei den Modellen k = 3 und k =5 fast identisch, beide Modelle weisen eine Präzision von 0,87 auf und weichen auch bei der Sensitivität nur um 0,01 voneinander ab (Tab. 2). Die Präzision von 0,87 ist ein guter Wert für die Modelle, sie gibt an, dass in 87% der Fälle, eine vom Modell erkannte Wasserfläche, auch tatsächlich eine ist. Die Sensitivität für die beiden Modelle ist etwas geringer, diese liegt bei 0,79 für k = 3 und bei 0,8 bei einem k = 5. Die Werte sind als ausreichend bis gut zu bewerten. Die Modelle erkennen also 79% bzw. 80% der Wasserflächen im Untersuchungsgebiet. Für die beiden Modelle der Parameter k = 3 und k = 5 ergibt sich daher ein f1-score von 0,83, dieser bildet den Durchschnitt von Präzision und Sensitivität (Tab. 2).

Die Modelle für die Parameter k = 15 und k = 29, weisen noch einmal einen Zuwachs für die Präzision auf 0,9 für k = 15 und 0,91 bei k = 29, allerdings sinkt die Sensitivität beider k Parameter im Vergleich zu den Modellen von k = 3 und k = 5. Die Sensitivität beträgt von k = 15 nur noch 0,76 und für k = 29 sinkt sie noch einmal auf 0,74. Dies spiegelt sich auch im f1-score wider, der leicht sinkt und bei beiden Modellen bei 0,82 liegt (Tab. 2).

Für die Klassifizierung der Wasserflächen erreichen alle kNN-Modelle gute Klassifikationsergebnisse. Allerdings sind die Klassifizierungsergebnisse der anderen Klassen deutlich weniger zufriedenstellend. So ist der höchst erreichte f1-score außerhalb der Wasserklasse über alle Modelle, bei einem k = 5 in der Vegetationsklasse mit 0,76 zu finden (Tab. 2). Der niedrigste f1-score außerhalb der Wasserklasse lässt sich für die Bodenklasse bei einem k = 29 mit 0,52 finden. Dieser niedrige Wert ist nach der optischen Begutachtung der Ergebnisse überraschend, da die Modelle k = 15 und k = 29 zumindest für die Schwarztorffläche im Südwesten des Gebietes, bessere Ergebnisse liefern als die der Modelle k = 3 und k = 5. Dies zeigt umso mehr, wie wichtig ein Accuracy Assessment mit einer Fehlermatrix ist, da nur so die Qualität eines Modells objektiv bewertet werden kann.

Bei der Betrachtung der Bodenklasse fällt auf, dass die False positives über alle Modelle bei einem Wert von ca. 0,2 in der nicht grünen Vegetationsklasse liegen. Ein Grund für diese hohen False positive Werte in der nicht grünen Vegetationsklasse für die Bodenklasse, könnte der Ground Truth Datensatz sein, da dieser auch mit einem gewissen Grad an Generalisierung aufgenommen ist. Des Weiteren ist die Trennung zwischen der Bodenklasse und der nicht grünen Vegetationsklasse optisch relativ schwierig, da in beiden Klassen Brauntöne dominieren. Dies ist entscheidend, da die Drohnenaufnahme, welche die Grundlage für den Ground Truth Datensatz bildet, nur über einen rot, grün und blauen Kanal verfügt und deswegen nicht auf Spektralkanäle im nahen Infrarot zurückgegriffen kann, die eventuell mehr Informationen zur Trennung der Klassen hätten liefern können. Durch die Spektrale Ähnlichkeit der Klassen Boden und nicht grüne Vegetation kann es auch zu nicht optimal gewählten Trainingsgebieten kommen, gerade bei der Auflösung der Sentinel 2 Daten von 10 x 10 m.

Außerdem ist zu erkennen, dass mit steigendem k auch der False positive Wert der Wasserklasse steigt. So liegt dieser bei einem k = 3 bei 0,071 steigt dann bei einem k = 29 aber auf einen Wert von 0,17. Das Modell generalisiert hier deutlich zu stark und klassifiziert mehr Wasserflächen als Boden als es

noch bei den Modellen k = 3 und k = 5 der Fall ist. Es ist daher nicht zu empfehlen eines der Modelle k = 15 oder k = 29 für die Klassifizierungen von Wasserflächen im Moor zu verwenden, da die Modelle zu häufig statt Wasserflächen Boden erkennen.

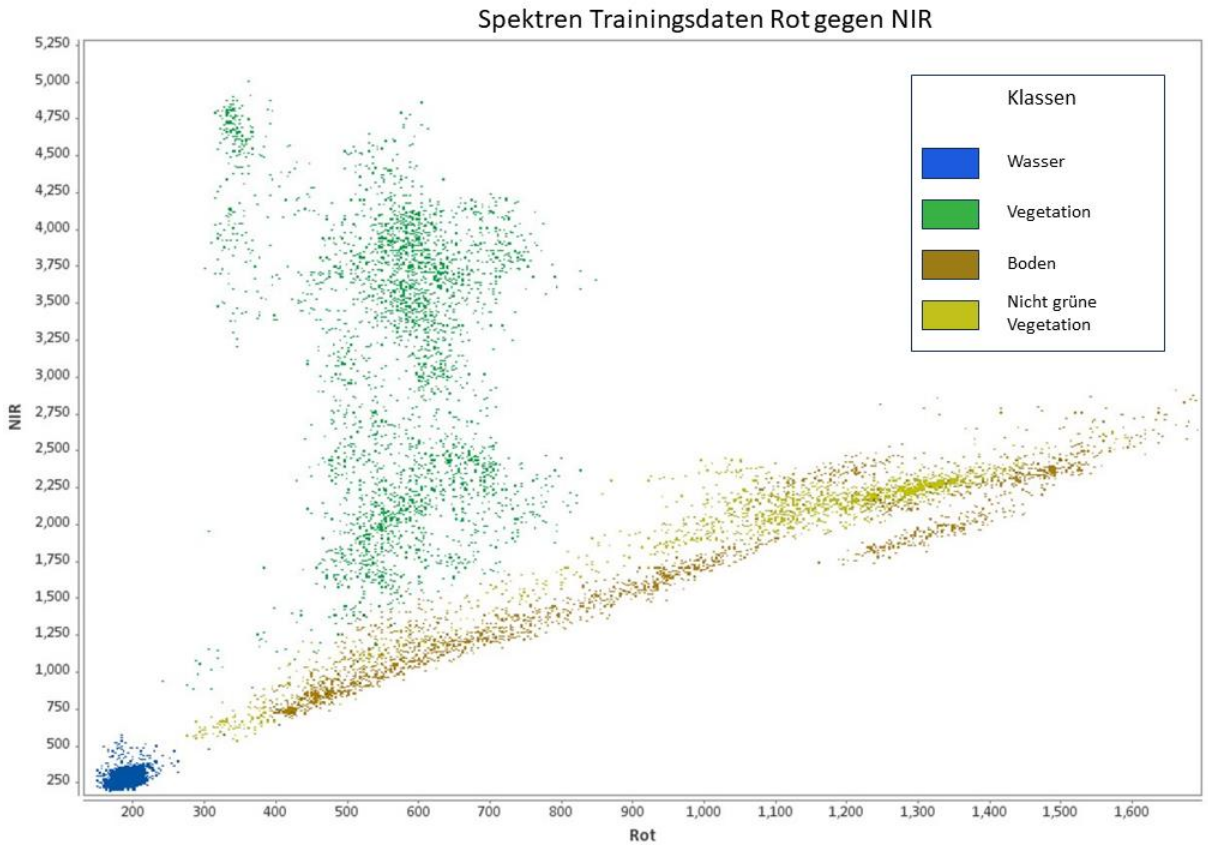

Abbildung 32 Spektralbereiche der Klassen innerhalb der Trainingsdaten in der Ansicht Rot gegen NIR

Bei der optischen Betrachtung der SVM Modell Ergebnisse, lassen sich direkt Unterschiede in den Klassifikationsergebnissen erkennen (Abb. 30). So klassifizieren die Modelle C = 0,5 und C = 1 für fast die gesamte Schwarztorffläche im Südwesten des Gebietes Wasser mit kleineren Boden- und Vegetationsflächen. Dies ändert sich sehr deutlich mit den Modellen C = 20 und C = 50, so verschwinden hier fast vollständig die Wasserflächen und werden vor allem durch die Vegetationsklasse ersetzt. Im Modell C = 20 sind mehr Wasserflächen auf der Schwarztorffläche zu finden, als es für das Modell C = 50 der Fall ist. Das Modell interpretiert diese Schwarztorfflächen mit zunehmenden C demnach immer mehr als Vegetation.

Dies ist auch am südwestlichen Rand des Leegmoors gut zu erkennen (Abb. 30). Hier finden sich Wasserflächen umschlossen von nicht grüner Vegetation. Die Modelle C = 0,5 und C = 1 erkennen für diese Flächen vor allem Wasser, aber auch Vegetation und nicht grüne Vegetation. Die Modelle C = 20 und C = 50 erkennen in diesen Flächen nur noch sehr wenig Wasser dafür deutlich mehr Vegetation.

Dies ist auch im Norden des Leegmoores zu beobachten, die Wasserflächen in den Modellen C = 0,5 und C = 1 werden in den Modellen C = 20 und C = 50 als Vegetation erkannt (Tab. 2). Dies könnte an den Trainingsdaten liegen. Um eine bessere Vergleichbarkeit zu liefern, werden für alle drei klassischen ML-Methoden die gleichen Trainingsdaten verwendet. Der SVM-Algorithmus betrachtet allerdings nur die Trainingspunkte im Grenzbereich der Klassen, dadurch kann es sein, dass die Trainingsdaten nicht optimal an die Methode angepasst sind. Dies ist wahrscheinlich, da höhere C Werte eine stärkere Diskriminierung der Werte im Randbereich bedeuten (Harte Ränder) und sich somit stärker an den Trainingsdaten orientieren als die Modelle mit geringerem C.

Die Modelle mit einem C = 0,5 und C = 1 erreichen bei der Klassifikation von Wasserflächen einen recht ähnlichen f1-score, so liegt dieser bei einem C = 0,5 bei einem Wert von 0,79 und bei einem C = 1 bei 0,81. Diese leichte Verbesserung des f1-scores vom Modell C = 1 im Vergleich zum Modell C = 0,5 ist vor allem auf eine Verbesserung der Sensitivität von 0,8 auf 0,83 zurückzuführen (Tab. 3). Die Modelle C = 20 und C = 50 sehen vielversprechender aus. So erreicht das Modell C = 20 einen f1-score von 0,85 und das Modell C = 50 einen Wert von 0,88. Dieser Sprung im f1-score verglichen, mit den anderen beiden SVM-Modellen, ist vor allem auf einen starken Anstieg in der Präzision zurückzuführen. So erreicht das Modell C = 20 hier einen Wert von 0,87 und das Modell C = 50 einen von 0,92.

Diese Verbesserung ist auf die genauere Unterscheidung zwischen Wasserflächen und Boden zurückzuführen, so hat die Bodenklasse bei den Modellen C = 0,5 und C = 1 bei der Wasserklasse für den größten Anteil an False positives gesorgt (Tab. 3). Die Sensitivität der beiden Modelle bleibt dabei etwa auf dem Niveau des Modells C = 1.

Die Modelle mit einem höheren C sind für die Wasserklasse wesentlich präziser. Dies kann damit begründet werden, dass die Wasserklasse gerade im nahen Infrarot eine sehr spezifische und gut eingrenzbare Spektralsignatur hat (Abb. 32) und je strikter (harte Ränder) der Algorithmus sich an die Trainingsdaten hält, umso bessere Ergebnisse liefert dieser.

Bei der Betrachtung der anderen Klassen fällt auf, dass auch ähnlich dem kNN-Algorithmus, die Präzision der Bodenklasse sehr gering ausfällt mit dem höchsten Wert 0,57 bei einem C = 1 und den niedrigsten Werten bei den Modellen C = 20 und C = 50 mit 0,5. Die Sensitivität liegt dabei zwischen 0,74 und 0,71 über alle SVM-Modelle für die Bodenklasse. Die Bodenklasse hat dabei in der nicht grünen Vegetationsklasse den höchsten Anteil an False positives, gefolgt von der Vegetationsklasse (Tab. 3).

Ein Grund hierfür könnte die Bandbreite in der Spektralsignatur der Bodenklasse sein. Sie lässt sich nicht so gut abgrenzen wie die der Wasserklasse und hat daher Überschneidungen mit der nicht grünen Vegetationsklasse und der Vegetationsklasse (Abb. 32). Für eine Verbesserung wäre eventuell eine weitere Untergliederung der Bodenklasse vonnöten, um eindeutigere Trainingsdaten zu erhalten.

Die Modelle mit einem höheren C weisen außerdem eine deutliche Verschlechterung der Präzision in der Vegetationsklasse auf, diese fällt von ihrem höchsten Wert 0,76 im Modell C = 0,5 auf ihren niedrigsten Wert 0,53 im Modell C = 50 (Tab. 3). Dies konnte auch schon gut bei der optischen Betrachtung der Ergebnisse festgestellt werden, dass mit höherem C immer mehr Flächen vom Algorithmus als Vegetation erkannt wurden (Abb. 30). Dieser Anstieg in den False positives ist für die Vegetationsklasse bei den Modellen C = 20 und C = 50 über alle anderen Klassen zu erkennen. Auch in diesem Fall könnte die Verschlechterung der Ergebnisse an nicht speziell für den Algorithmus gewählten Trainingsdaten liegen, da sich das Modell mit steigendem C immer stärker an die Klassengrenzen hält. Außerdem wäre es eventuell sinnvoll die Vegetationsklasse weiter zu unterteilen, um klarere Signaturen der daraus entstehenden Subklassen zu erhalten. Dennoch sind die Parameter des Accuraccy Assessments für die Wasserklasse gerade bei den SVM-Modellen mit höherem C sehr viel versprechend. Sowohl C = 20 als auch C = 50 erreichen dabei sehr gute Werte und sie eignen sich daher sehr gut für die Extraktion von Wasserflächen im Leegmoor.

Bei der optischen Betrachtung der RF-Modell-Ergebnisse lassen sich keine nennenswerten Unterschiede feststellen (Abb. 31). So sieht selbst die Schwarztorfstelle, im Südwesten des Gebietes, in allen RF-Modellen identisch aus. Die Modelle erkennen auf dieser Fläche zum Großteil Vegetation mit einzelnen Wasser- und Bodenflächen. Die Vegetation auf dieser Fläche ist hierbei deutlich überinterpretiert vom Modell. In den Ground Truth Daten finden sich auf dieser Fläche fast ausschließlich Wasser- und Bodenflächen. Gerade die Bodenflächen werden vom Modell, auf der Schwarztorfläche, fälschlicherweise als Vegetation erkannt. Ein Grund hierfür könnte die geringe

Anzahl an Schwarztorfflächen im Untersuchungsgebiet sein, so dass das Modell kaum Trainingsflächen für Schwarztorfflächen hat. Außerdem könnte eine Unterteilung der Boden- sowie der Vegetationsklasse helfen, die Spektralsignaturen beider Klassen zu spezifizieren.

Da optisch keine Unterschiede auszumachen waren, kann eventuell die Betrachtung der Ergebnisse des Accuracy Assessment Unterschiede hervorbringen. Der f1-score für die Wasserklasse schwankt über alle Modelle zwischen 0,82 und 0,83 (Tab. 4). Der Trend der sehr ähnlichen optischen Ergebnisse, zeichnet sich demnach auch in der Betrachtung des Accuracy Assessments ab.

In den anderen Klassen sind ebenfalls sehr ähnliche Ergebnisse über alle Modelle zu beobachten. Der durchschnittlich höchste f1-score ist dabei im Modell E = 100 mit 0,725 zu verzeichnen, während der niedrigste im Modell E = 500 mit 0,69 zu finden ist. Dennoch liegen auch hier beide Modelle sehr nah beieinander mit einer Differenz von nur 0,035 (Tab. 4). Auffällig ist, auch bei den Modellen des RF-Algorithmus, dass wie bei den vorangegangenen beiden Methoden, die Präzision der Bodenklasse sehr niedrig ist und gerade in den Klassen Vegetation und nicht grüne Vegetation hohe False positive Werte zu finden sind.

Ähnliches lässt sich über die Sensitivität, der nicht grünen Vegetationsklasse sagen. Diese ist ebenfalls über alle Methoden sehr gering und hat dabei ihre größten False negative Werte in der Bodenklasse und Vegetationsklasse (Tab. 4). Gerade an der Präzision der Bodenklasse und den hohen Wert für die False positives in der nicht grünen Vegetationsklasse bzw. den hohen False negative Werten für die nicht grüne Vegetationsklasse in der Bodenklasse, lassen darauf schließen, dass die drei verwendeten Algorithmen, Probleme haben, diese Klassen sauber zu trennen.

Ein Grund dafür könnte der Ground Truth Datensatz sein (Abb. 21), da dieser händisch anhand einer Drohnenaufnahme, mit einer Auflösung von 10 cm x 10 cm, erstellt wurde. Dieser ist durch den Prozess der Erstellung auch zu einem gewissen Grad generalisiert und kann daher nicht bis ins kleinste Detail richtig sein. Auch ist zu beachten, dass die Satellitendaten eine Auflösung von 10 m x 10 m bzw. 20 m x 20 m haben und viele Details, welche noch in der sehr viel besser aufgelösten Drohnenaufnahme zusehen sind, im Satellitenbild nicht mehr zu erkennen sind. Gerade auf einem Gebiet von 10 m x 10 m bzw. 20 m x 20 m können mehrere Klassen auftreten, die dann als Mischsignal aufgenommen werden. Im Falle eines Mischsignals würden sich die Klasse Vegetation bzw. Wasser eher den anderen beiden Klassen gegenüber durchsetzen, da die beiden Klassen gerade in den Infrarot Kanälen sehr extreme Signale aufweisen, während die Signale der Bodenklasse und der nicht grünen Vegetationsklasse wesentlich unspezifischer sind.

Falls ein besseres Klassifikationsergebnis in diesen Klassen benötigt wird, müssten daher die Klassen Boden und nicht grüne Vegetation weiter unterteilt werden, um spezifischere Spektralcharakteristiken herauszuarbeiten, damit der Algorithmus diese ähnlich den Wasserflächen oder auch Vegetationsflächen eindeutiger zuordnen kann.

Der Kappa Wert schwankt über alle Methoden und ihrer Modelle zwischen 0,59 und 0,65 (Tab. 2,3,4), um ein besseres Ergebnis hier zu erreichen müssten die Klassifikationsergebnisse in der Boden- und grünen Vegetationsklasse sich verbessern. Dennoch sind die Werte von Kappa deutlich größer als 0 und unterscheiden sich somit deutlich von einem zufällig verteiltem Klassifikationsergebnis (Congalton, 1991).

Alle Modelle erkennen 100% der Wasserkontrollzonen, dementsprechend können alle getesteten Modelle sehr gut eindeutige Wasserpixel klassifizieren (Tab. 2,3,4).

6. Zusammenfassung

Die Wahl der besten Methode für die Klassifikation von Wasserflächen im Leegmoor, ist davon abhängig, wie wichtig die Qualität der Klassifizierung der anderen Klassen ist. Wenn nur die Qualität der Klassifizierung für die Wasserklasse entscheidend ist, sollte sich für den SVM - Algorithmus mit einem C von 50 entschieden werden (Tab. 3). Mit dieser Einstellung werden die besten Klassifikationsergebnisse für die Wasserklasse erreicht mit einem f1-score von 0,88. Insgesamt wirkt der SVM – Algorithmus sehr geeignet für die Klassifikation von kleineren Moorgebieten, da der Algorithmus keine große Menge an Trainingsdaten benötigt. Denn auch bei der Betrachtung der Gesamtgenauigkeit und des durchschnittlichen f1-scores fällt auf, dass bei einem C = 1 die besten Ergebnisse über alle Methoden erreicht werden. Dies geht entgegen den Ergebnissen von Adugna et al., bei denen der RF – Algorithmus besser abgeschnitten hat als der SVM, allerdings wurden in der Arbeit auch schlechter aufgelöste Satellitenbilddaten verwendet und für wesentlich größere Flächen. Um das bestmögliche Modell für die Klassifikation von Wasserflächen in kleinen Moorflächen zu erreichen, wäre eine Arbeit, welche sich primär auf das Hyperparametertuning des SVM – Algorithmus konzentriert von Nöten. So könnte ein Trainingsdatensatz speziell angepasst an die Methode sowie, das Einbeziehen weiterer Parameter außer C, zu einer Verbesserung des Klassifikationsergebnis führen. Dennoch erreichen die Modelle gute Werte bei dem Erkennen von Wasserflächen. Die Satellitendaten sollten demnach als Datengrundlage der Klassifikationen mehr als ausreichend sein und es kann darauf verzichtet werden Drohnenaufnahmen oder höher räumlich aufgelöste Satellitendaten als Datengrundlage für die Klassifikation von Wasserflächen auf kleinen Moorgebieten zu verwenden.

7.Quellen

Adugna, T., Xu, W., & Fan, J. (2022). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sensing*, *14*(3), 1–22. https://doi.org/10.3390/rs14030574

Airbus Defence and Space. (n.d.). *SPOT Imagery User Guide DEFENCE AND SPACE Intelligence*.

Albertz, J. (2009). *Einführung in die Fernerkundung*. WBG.

Batta, M. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJ*, *9*(1), 381–386. https://doi.org/10.21275/ART20203995

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Caruana, R., Pratt, L., & Thrun, S. (1997). *Multitask Learning ** (Vol. 28). Kluwer Academic Publishers.

Congalton, Oderwald, & Mead. (1983). Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. In *Photogrammetric Engineering and Remote Sensing: Vol. v. 49*.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

ENVI. (2024). Exelis Visual Information Solutions. 2010. Boulder, Colorado: Exelis Visual Information Solutions. Https://Www.Nv5geospatialsoftware.Com/Products/ENVI.

ESA. (2012). *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services (ESA SP-1322/2)*. ESA Communications.

Esri. (2024). ArcGIS Pro Dokumentation Https://Www.Esri.Com/En-Us/Arcgis/Products/Arcgis-pro/Resources.

Foody, G. M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(6), 1335–1343. https://doi.org/10.1109/TGRS.2004.827257

Gallant, A. L. (2015). The challenges of remote monitoring of wetlands. In *Remote Sensing* (Vol. 7, Issue 8, pp. 10938–10950). MDPI AG. https://doi.org/10.3390/rs70810938

Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Isikdogan, F., Bovik, A. C., & Passalacqua, P. (2017). Surface water mapping by deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(11), 4909–4918. https://doi.org/10.1109/JSTARS.2017.2735443

Isikdogan, L. F., Bovik, A., & Passalacqua, P. (2020). Seeing through the Clouds with DeepWaterMap. *IEEE Geoscience and Remote Sensing Letters*, *17*(10), 1662–1666. https://doi.org/10.1109/LGRS.2019.2953261

Joachim Steinwendner, & Roland Schwaiger. (2020). *Neuronale Netze programmieren mit Python* (2nd ed.). Rheinwerk Computing.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., & Gascon, F. (2016). Sentinel-2 SEN2COR: L2A processor for users. *European Space Agency, (Special Publication) ESA SP*, *SP-740*(August), 9–13.

Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., & Huang, W. (2018). Remote sensing for wetland classification: a comprehensive review. In *GIScience and Remote Sensing* (Vol. 55, Issue 5, pp. 623–658). Taylor and Francis Inc. https://doi.org/10.1080/15481603.2017.1419602

Nick, K.-J., Löpmeier, F.-J., Schiff, H., Blankenburg, J., Gebhardt, H., Knabke, C., Weber, H. E., Främbs, H., & Mossakowski, D. (2001). *NICK, K.-J. & WEBER, H. E. (2001): Entwicklung der Vegetation auf dem wiedervernäßten Leegmoor (Landkreis Emsland) in den Jahren 1989 bis 1996. In: BUNDESAMT FÜR NATURSCHUTZ (Ed.): Moorregeneration im Leegmoor/Emsland nach Schwarztorfabbau und Wiedervernä*.

Olaode, A., Naghdy, G. A., Todd, C., & Naghdy, G. (2014). Unsupervised classification of images: A review. *International Journal of Image Processing (IJIP)*, *8*(5), 325–342. https://www.researchgate.net/publication/265729668

OpenGeoData.NI. (2024). Datenlizenz Deutschland – Namensnennung – Version 2.0. Https://Opengeodata.Lgln.Niedersachsen.de/#dop.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Powell, R., & Brooks, C. (2023). *Land Use Land Cover Mapping in the Tiffin River Watershed*.

Powers, D. (2008). Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Mach. Learn. Technol.*, *2*.

Prasath, V. B. S., Alfeilat, H. A. A., Hassanat, A. B. A., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., & Salman, H. S. E. (2017). *Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review*. 1–39. https://doi.org/10.1089/big.2018.0175

QGIS.org. (2024). QGIS Geographic Information System. QGIS Association. Http://Www.Qgis.Org.

Rezaee, M., Mahdianpari, M., Zhang, Y., & Salehi, B. (2018). Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(9), 3030–3039. https://doi.org/10.1109/JSTARS.2018.2846178

SNAP. (2024). SNAP - ESA Sentinel Application Platform v2.0.2, Http://Step.Esa.Int.

U.S. Geological Survey. (2019a). Landsat 7 (L7) Data Users Handbook. In V. Ihlen & K. Zanter (Eds.), *Department of the Interior U.S. Geological Survey* (Vol. 7, Issue November).

U.S. Geological Survey. (2019b). Landsat 8 Data Users Handbook. In V. Ihlen & K. Zanter (Eds.), *Department of the Interior U.S. Geological Survey* (Vol. 8, Issue November). https://landsat.usgs.gov/documents/Landsat8DataUsersHandbook.pdf

Van Der Aalst, W. M. P., Rubin, V., Verbeek, H. M. W., Van Dongen, B. F., Kindler, E., & Günther, C. W. (2010). Process mining: A two-step approach to balance between underfitting and overfitting. *Software and Systems Modeling*, *9*(1), 87–111. https://doi.org/10.1007/s10270-008-0106-z

Wilmott, P. (2020). *Grundkurs Machine Learning* (1st ed.). Rheinwerk Computing.

Zerrouki, N., & Bouchaffra, D. (2014a). Pixel-based or object-based: Which approach is more appropriate for remote sensing image classification? *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *2014-Janua*(January), 864–869. https://doi.org/10.1109/SMC.2014.6974020

Zerrouki, N., & Bouchaffra, D. (2014b). Pixel-based or Object-based: Which approach is more appropriate for remote sensing image classification? *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 864–869. https://doi.org/10.1109/SMC.2014.6974020

Zhang, Z. (2016). Introduction to machine learning: K-nearest neighbors. *Annals of Translational Medicine*, *4*(11), 1–7. https://doi.org/10.21037/atm.2016.03.37

URL-Quellen:

Bundesamt für Kartographie und Geodäsie (BKG) Deutschland Shapefile http://www.geodatenzentrum.de/geodaten/gdz_rahmen.gdz_div?gdz_spr=deu&gdz_akt_zeile=5&gdz_anz_zeile=4&gdz_user_id=0 Letzer Aufruf 20.05.2024

ENVI (2024). Exelis Visual Information Solutions. 2010. Boulder, Colorado: Exelis Visual Information Solutions. Https://Www.Nv5geospatialsoftware.Com/Products/ENVI. Letzer Aufruf 20.05.2024

ESA (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services (ESA SP-1322/2). ESA Communications. Letzer Aufruf 20.05.2024

Esri (2024). ArcGIS Pro Dokumentation Https://Www.Esri.Com/En-Us/Arcgis/Products/Arcgis-pro/Resources.

IBM ://www.ibm.com/de-de/topics/knn Letzer Aufruf 09.05.2023

OpenGeoData.NI. (2024). Datenlizenz Deutschland – Namensnennung – Version 2.0. Https://Opengeodata.Lgln.Niedersachsen.de/#dop. Letzer Aufruf 20.05.2024

QGIS.org. (2024). QGIS Geographic Information System. QGIS Association. Http://Www.Qgis.Org. Letzer Aufruf 20.05.2024

Scikit Learn Overfitting Underfitting https://scikit-learn.org/stable/auto_examples/model_selection

/plot_underfitting_overfitting.html). Letzer Aufruf: 20.04.2023

Sentinel https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial; Letzter Aufruf: 20.04.2023

Skyrora https://www.skyrora.com/remote-sensing/ letzter Aufruf 18.12.23

SNAP (2024). SNAP - ESA Sentinel Application Platform v2.0.2, Http://Step.Esa.Int. Letzer Aufruf 20.05.2024