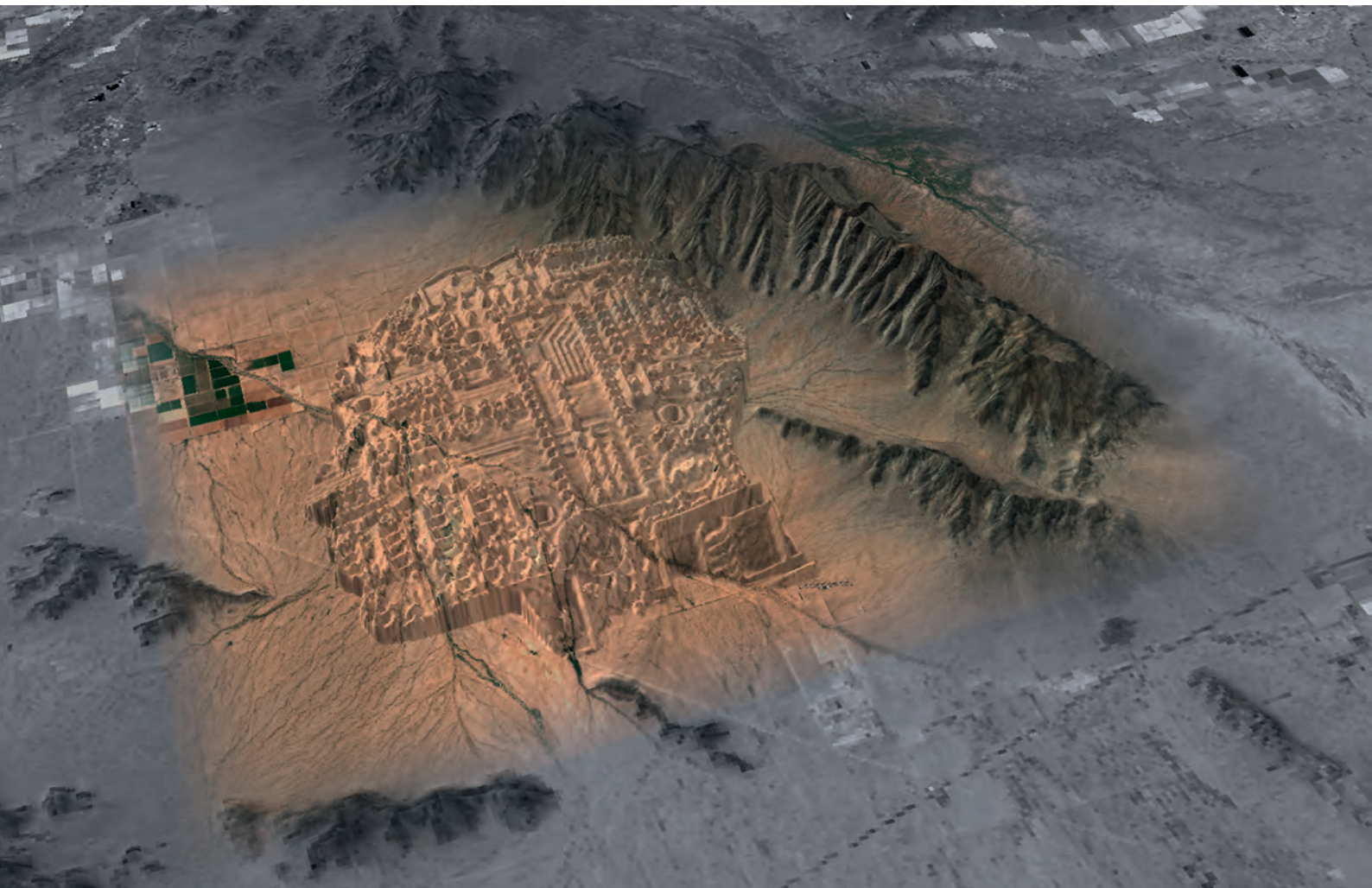


Maschinelles Lernen bei der Auswertung von Fernerkundungsdaten

*Karl Segl, Niklas Bohn, Sabine Chabrillat, Carsten Neumann, Sigrid Roessner, Kathrin Ward, Aleksandra Wolanin
Deutsches GeoForschungsZentrum GFZ, Potsdam*

Remote sensing data analysis retrieves spatial-temporal information about the Earth's surface from remotely sensed optical and radar images. For this purpose accurate and efficient classification or parameter quantification techniques must be used. Consequently, there exists a long tradition in remote sensing to employ methods and techniques from the field of machine learning. They can be regarded as „universal function approximators“ that are able to link any data in order to derive connections, conclusions and predictions efficiently using different learning strategies. In the following, current research topics of the Remote Sensing section of the GFZ are presented, in which different forms of machine learning are used.



Maschinelles Lernen (ML) ist ein Teilgebiet der künstlichen Intelligenz, das in vielen Bereichen unseres Lebens immer häufiger anzutreffen ist. Intelligente Programme sind in vielen Berufsfeldern und auch im Alltag zu finden, hier vor allem bei der Nutzung von Diensten großer IT-Unternehmen wie z. B. Google, Apple, Amazon oder Facebook. Diese Programme lernen aus den verfügbaren Daten und erkennen Gesetzmäßigkeiten.

Im Bereich der Fernerkundung kommen ML-Verfahren seit etwa drei Jahrzehnten zum Einsatz. Hierbei werden aus der Fülle vorhandener Bilddaten von verschiedenen Fernerkundungssensoren problemspezifische Informationen gewonnen, die eine Beurteilung des Zustands und der Veränderung der Landoberfläche ermöglichen. ML-Verfahren können als „Universale Funktionsapproximatoren“ betrachtet werden, die in der Lage sind, beliebige Daten zu verknüpfen, um daraus rechentechnisch effizient Zusammenhänge, Rückschlüsse und Vorhersagen abzuleiten. Dabei kommen unterschiedliche Ansätze des Lernens zum Einsatz. Generell wird zwischen überwachtem und unüberwachtem Lernen unterschieden.

Sollen die Zusammenhänge zwischen vorgegebenen Eingabedaten (z. B. Fernerkundungsdaten) und den abzuleitenden Ausgabedaten (z. B. Klassifizierungsergebnisse) anhand von Beispieldaten (Trainingsinformationen) erlernt werden, spricht man von überwachtem Lernen. Dies beinhaltet die meisten Verfahren zur Klassifizierung und Zusammenhangsanalyse (Regression). Stellvertretend können hier Maximum-Likelihood-Klassifizierung, Support Vector Machine, Random Forest, Partielle-Kleinste-Quadrate-Regression oder künstliche neuronale Netze genannt werden. Deep Learning stellt die derzeit fortgeschrittenste Implementierung eines künstlichen neuronalen Netzes dar.

Das unüberwachte Lernen nutzt keine Lernvorgaben. Daher können die Algorithmen nicht trainiert werden. Hier werden die Verfahren des Data Mining genutzt, um bestimmten Ähnlichkeitskriterien genügende Strukturen aus den Fernerkundungsdaten zu extrahieren, die in einem zweiten Schritt mit thematischer Bedeutung verknüpft werden können. Beispiele für unüberwachtes ML sind alle Clusterverfahren (z. B. K-means) und Dimensionsreduktionsver-

fahren (z. B. Hauptkomponentenanalyse) sowie auch unüberwacht lernende künstliche neuronale Netze.

Im Folgenden werden aktuelle Forschungsthemen der Sektion „Fernerkundung“ am Deutschen GeoForschungsZentrum GFZ vorgestellt, bei denen verschiedene ML-Verfahren zum Einsatz kommen. Die dabei bearbeiteten Fragestellungen reichen von der Atmosphärenkorrektur über die Analyse von Bodeneigenschaften bis zu Vegetationsanwendungen.

Klassifizierung von Wolken, Schnee, Schatten, Wasser und klarem Himmel

In der optischen Fernerkundung wird die von der Erdoberfläche reflektierte Sonnenstrahlung mit Hilfe von Sensoren aufgezeichnet. Diese Strahlung wird dabei nicht nur durch die Oberflächeneigenschaften beeinflusst, sondern auch durch den Sonnenstand und die aktuellen Streuungs- und Absorptionseigenschaften der Atmosphäre, die eine zeitliche Vergleichbarkeit von Fernerkundungsdaten stark beeinträchtigen können. Deshalb wird in der Regel eine Normierung der Fernerkundungsdaten in Form einer Atmosphärenkorrektur durchgeführt, in deren Ergebnis die Rückstreuung dem Reflexionsgrad an der Erdoberfläche entspricht. Da vor allem Wolken (insbesondere Zirren) einen starken Einfluss auf das Signal haben, kann eine Atmosphärenkorrektur deutlich verbessert werden, wenn zuvor eine möglichst gute Unterscheidung von klarem Himmel und Wolken sowie Schnee, Schatten und Wasser erreicht werden kann. Dazu eignen sich insbesondere ML-Verfahren, die auf überwachtem Lernen beruhen. *Hollstein et al. (2016)* entwickelten zwei effiziente Methoden zur Unterscheidung der oben genannten Phänomene in Sentinel-2-Satellitenbilddaten. Dabei wurde als Trainingsinformation für das ML ausschließlich die am oberen Rand der Atmosphäre gemessene Strahlung genutzt, um durch den Verzicht auf jegliche Art von Zusatzinformationen große Datenmengen besonders schnell prozessieren zu können. Zu diesem Zweck wurden die maschinellen Lerner mit zwei unabhängigen Datensätzen manuell klassifizierter Sentinel-2A-Spektren trainiert und validiert. Um die einzelnen Spektren klassifizieren zu können, wurden nicht nur die Reflexionswerte einzelner Sentinel-2-Kanäle, sondern auch mathematische Beziehungen zwischen diesen, wie z. B. die Differenz zwischen Kanal A und Kanal B oder ein Index in der Form $A-B/A+B$, zu einem sogenannten Merkmalsraum zusammengefasst. Im Rahmen eines solchen Merkmalsraums kann nun jedes Spektrum eines Sentinel-2-Datensatzes durch den Lerner einer der definierten Klassen zugeordnet werden.

Die erste von *Hollstein et al. (2016)* entwickelte Methode basiert auf sogenannten Entscheidungsbäumen, bei denen an jeder Verästelung eine Entscheidung getroffen werden muss. Diese

Links: Metaphorische Darstellung der Verbindung von Fernerkundung und maschinellem Lernen

Left: Metaphoric representation of the connection of remote sensing and machine learning



Kontakt: K. Segl
(segl@gfz-potsdam.de)

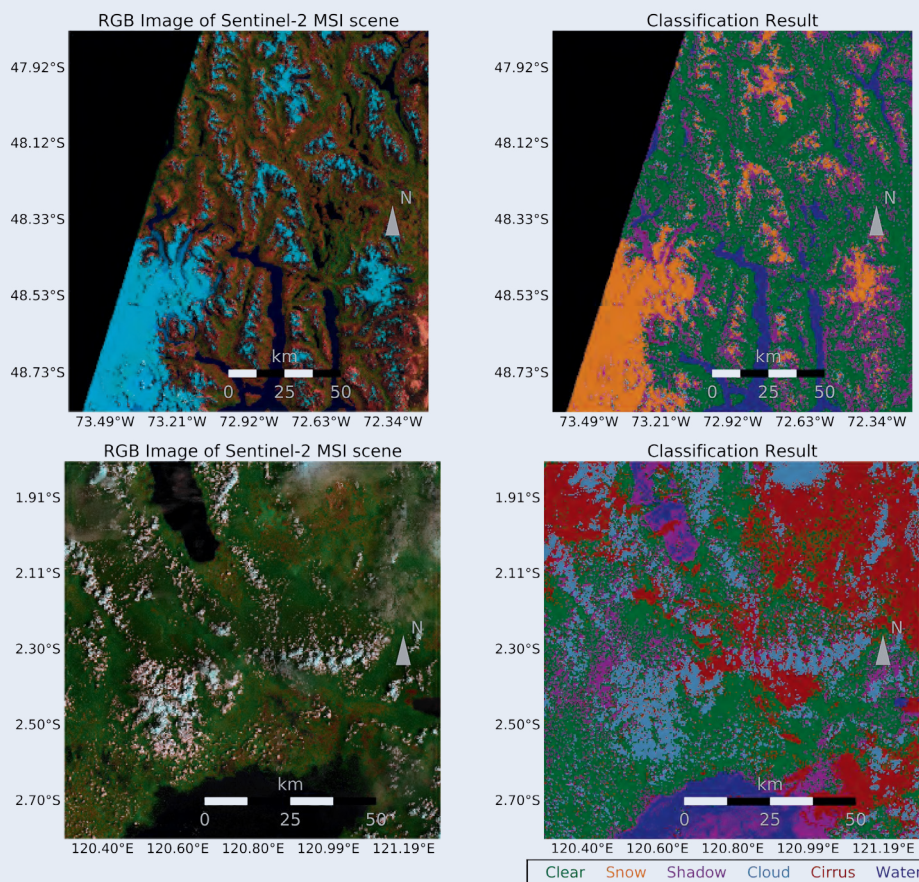


Abb. 1: Klassifizierungsergebnisse basierend auf der Bayes-Methode für Sentinel-2-Szenen. Links: RGB-Bilder (Band 11, 8 und 3); rechts: Klassifizierungsergebnisse

Fig.1: Classification results based on the Bayesian algorithm applied on Sentinel-2 images. The left panel shows RGB images (Bands 11, 8, and 3) and the right panel the corresponding classification results.

erfolgt anhand von festgelegten Schwellwerten, die mit Hilfe der Trainingsdaten zuvor geschätzt wurden. Am Ende der Baumstruktur stehen dann die Klassen, in die das Spektrum eingeordnet wird. Die Analyse von Validierungsdaten hat gezeigt, dass mit Hilfe der auf Entscheidungsbäumen basierenden maschinellen Lerner Klassifizierungsgenauigkeiten von 87 bis 91 % erzielt werden konnten.

Die zweite in der Studie von *Hollstein et al. (2016)* entwickelte Klassifizierungsmethode basiert auf dem Satz von Bayes, der die Berechnung bedingter Wahrscheinlichkeiten beschreibt (*Hollstein et al., 2015*). Sie betrachtet die Eintrittswahrscheinlichkeit der einzelnen Klassen unter der Bedingung, dass die jeweiligen mathematischen Beziehungen des Merkmalsraums bestimmte Werte aufweisen. Die Trainingsspektren werden für den Merkmalsraum zu einem Histogramm zusammengefasst, so dass für jedes Testspektrum die Eintrittswahrscheinlichkeit der Klassen abgeleitet werden kann. Abschließend wird für jedes Spektrum diejenige Klasse ausgewählt, welche die höchste Wahrscheinlichkeit aufweist.

Die Bayes-Methode führte mit 98 % zu einer deutlich höheren Klassifizierungsgenauigkeit. Der Vergleich mit den originalen

Bilddaten hat jedoch gezeigt, dass trotz der erreichten hohen Genauigkeiten beide Techniken zum Teil noch Probleme in der Unterscheidung von klarem Himmel und Schatten sowie auch von dunklen Wasserflächen und Schatten haben. Klassifizierungsergebnisse für die Bayes-Methode sind exemplarisch in Abb. 1 dargestellt. Vor allem in Gebirgen lässt sich eine gute Unterscheidung zwischen Eis und Schatten erkennen. Die Bayes-Methode weist eine besonders schnelle Rechengeschwindigkeit auf, was von wesentlicher Bedeutung für den operationellen Einsatz eines solchen Verfahrens ist.

Erfassung und Quantifizierung von Bodeneigenschaften

Der Boden erfüllt eine Reihe wichtiger Umweltfunktionen, die für den menschlichen Lebensunterhalt unerlässlich sind, wie z. B. die Nahrungsmittel-, Faser- und Holzproduktion, Wasserspeicherung und -umverteilung sowie Schadstofffilterung und die Kohlenstoffspeicherung. Inzwischen entwickelt sich ein neues Bewusstsein für die Endlichkeit der weltweiten Bodenressourcen sowie eine wachsende Besorgnis darüber, ob die Böden dem wachsenden Nutzungsdruck standhalten (*Hartemink und McBratney, 2008; FAO und ITPS, 2015*). Die Bodenwissenschaften stehen damit vor der Herausforderung, genaue und aktuelle Bodeninformationen

von regionaler bis globaler Ebene bereitzustellen (Grunwald *et al.*, 2011; Arrouays *et al.*, 2017).

Vor diesem Hintergrund stellt die Fernerkundung mittels flugzeug- und satellitengestützter Systeme eine aussichtsreiche Methodik dar. Die hyperspektrale Fernerkundung, die auch als bildgebende Spektroskopie bezeichnet wird, besitzt ein besonderes Potenzial zur quantitativen Bestimmung und Modellierung verschiedener Bodeneigenschaften auf Grundlage ihrer spezifischen spektralen Reflexionseigenschaften. Dazu gehören die mineralogische Zusammensetzung des Oberbodens, der Gehalt an organischem Kohlenstoff an der Bodenoberfläche sowie Textur- und Strukturinformationen. Zukünftige hyperspektrale Satellitenmissionen, wie z. B. EnMAP (Guanter *et al.*, 2015), eröffnen dabei neue Möglichkeiten für eine globale, quantitative Bodenkartierung und -überwachung.

Da die Bodenreflexionseigenschaften multivariat sind und viele Wechselwirkungen aufweisen (Schwartz *et al.*, 2011), eignen sich ML-Verfahren insbesondere zur quantitativen Ableitung von Bodeneigenschaften auf der Basis von spektralen Eingangsdaten. Dabei kommen häufig Kombinationen von Verfahren aus dem Bereich des überwachten und unüberwachten Lernens zum Einsatz, wie z. B. multiple lineare Regressionsanalyse, Hauptkomponentenanalyse, die Partielle-Kleinste-Quadrate-Regression (PLSR), künstliche neuronale Netze sowie Random Forest.

In einer am GFZ durchgeführten Studie wurde das Potenzial der zukünftigen EnMAP-Satellitenmission zur fernerkundlichen Erfassung und Quantifizierung von Bodenparametern untersucht (Steinberg *et al.*, 2016). Die Arbeit basiert auf Bodenproben verschiedener

Untersuchungsgebiete in Nord- und Südeuropa. Mit Hilfe der PLSR-Methode konnten quantitative Zusammenhänge zwischen aus simulierten EnMAP-Satellitendaten abgeleiteten spektralen Reflexionseigenschaften und aus Bodenproben bestimmten Bodeneigenschaften hergestellt werden (Abb. 2).

Ein weiteres Anwendungsgebiet für ML-Verfahren am GFZ liegt in der Analyse von großen Bodenspektraldatenbanken, wie der europaweiten spektralen Bodendatenbank LUCAS, die vom Joint Research Centre der Europäischen Kommission (JRC) bereitgestellt wird. Momentan beinhaltet LUCAS für 20 000 Probenpunkte Informationen über zwölf verschiedene Bodenparameter, inklusive dem Schlüsselparameter SOC (organischer Kohlenstoffgehalt). Aufgrund der großen Datenmenge kamen im Rahmen der EnMAP-Nutzungsvorbereitung verschiedene ML-Methoden zum Einsatz. Dabei wurden verschiedene Clustering-Verfahren getestet, die mit PLSR-Modellen kombiniert wurden. Zum einen wurde eine unüberwachte Klassifizierung mit dem K-means-Verfahren durchgeführt und zum anderen eine angepasste Version einer lokalen PLSR angewendet.

Die europaweiten Ergebnisse der Clusteranalyse basierend auf dem K-means-Ansatz sind in Abb. 3 dargestellt. Die Clusteranalyse wird es ermöglichen, die anhand von Laborspektren entwickelten Modelle zur Quantifizierung des Gehalts an organischem Kohlenstoff später auf hyperspektrale Satellitendaten für die globale Kartierung von Bodeneigenschaften in vegetationsfreien Böden zu übertragen. Die in dieser Studie entwickelten Methoden können die Überwachung von Bodenzustandsveränderungen deutlich verbessern.

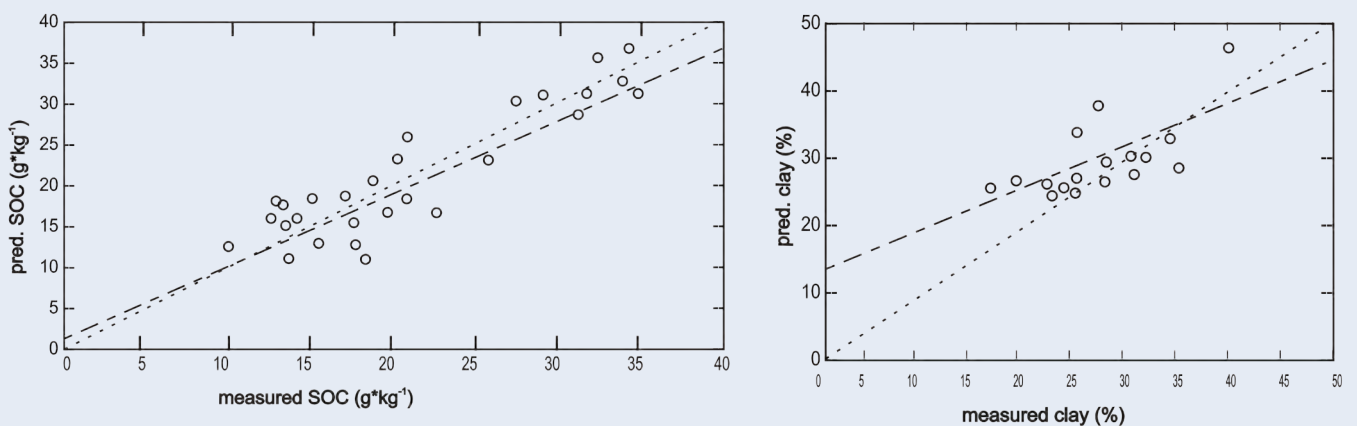


Abb. 2: Quantitative Bestimmung des organischen Kohlenstoffgehalts (SOC) und des Tongehalts (clay) in Böden auf der Basis von simulierten hyperspektralen EnMAP-Satellitendaten mit Hilfe von ML-Verfahren

Fig. 2: Quantitative estimation of soil organic carbon (SOC) and clay content in soils based on simulated hyperspectral EnMAP satellite data using ML techniques

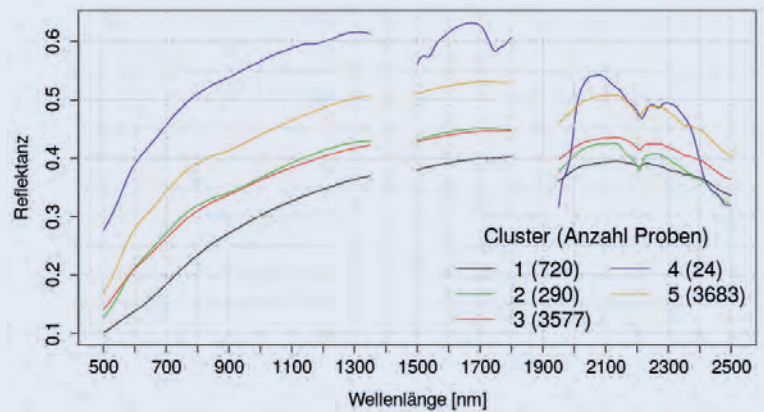
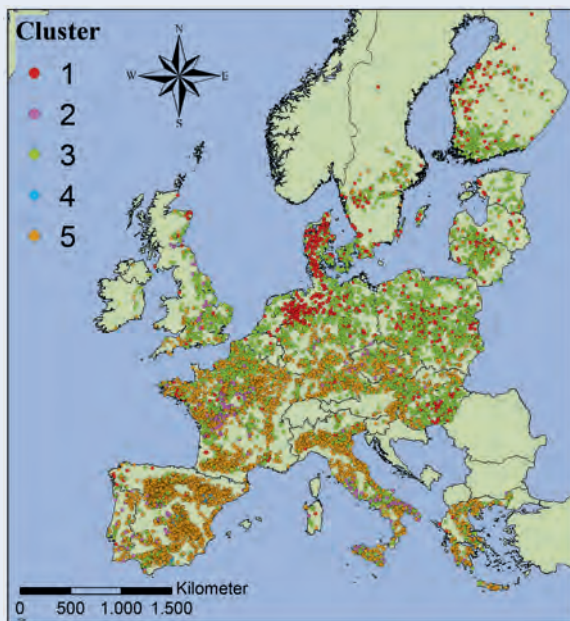


Abb. 3: Europaweite räumliche Verteilung der K-means-Cluster (links) und spektrale Charakteristika für die einzelnen Cluster (rechts) basierend auf der LUCAS-Bodendatenbank

Fig. 3: European wide spatial distribution of k-means cluster (left) and spectral characteristics of the cluster (right) based on the LUCAS soil spectral library

Maschinelle Lernsysteme im Biodiversitätsmonitoring

Die natürliche Vielfalt von Arten und Lebensräumen ist ein wesentliches Merkmal von Ökosystemen und bestimmt deren Funktionsfähigkeit, insbesondere auch für die Bereitstellung von Ökosystemdienstleistungen für den Menschen. So hängt die Verfügbarkeit von Nahrungsressourcen, sauberem Trinkwasser und frischer Luft direkt von den spezifischen Gemeinschaften an Organismen in den unterschiedlichen Biotopen ab. Derzeit wird weltweit ein massiver Rückgang von Arten in sämtlichen Ökosystemen beobachtet (LPR, 2016). Ein Verständnis der zugrunde liegenden Prozesse und die darauf basierende Entwicklung von geeigneten Landschaftspflegemaßnahmen sind ein wichtiges Ziel aktueller Forschung, um den Erhalt der natürlichen Lebensgrundlagen langfristig zu gewährleisten.

In einer am GFZ durchgeführten Studie kamen ML-Verfahren im Rahmen der Fernerkundung zum Einsatz, um auf dem Weg der Vegetationsanalyse die Organisationsform und Zusammensetzung von Arten und Lebensräumen räumlich zu beschreiben. Der Begriff Vegetation umfasst dabei die kontinuierliche Durchdringung der Bodenoberfläche mit unterschiedlichen Pflanzenarten. Diese bilden auf kleiner räumlicher Skala komplexe, multidimensionale Strukturen, die über räumliche Gradienten Ökosysteme charakterisieren.

Die Fernerkundung nutzt die bildgebende Spektroskopie, um solche Strukturen anhand von Reflexionsspektren zu beschreiben. ML-Verfahren werden dabei eingesetzt, um Einzelarten und deren Deckungsgrade in den Bilddaten auszuweisen. So können die Zusammensetzung und die räumliche Verteilung von natürlichen Lebensräumen als Biodiversitätsindikator abgebildet und für eine naturschutzfachliche Bewertung (z. B. europäisches Natura-2000-Netzwerk) nutzbar gemacht werden (Neumann et al., 2015).

Im Rahmen der am GFZ entwickelten fernerkundungsbasierten Methode (Abb. 4) wurden ML-Verfahren speziell für den Zweck genutzt, die Zusammenhänge zwischen den fernerkundlich ableitbaren pflanzenspezifischen Reflexionsspektren und den multidimensionalen Arträumen sowie den biochemischen und strukturellen Pflanzeigenschaften und -zuständen zu quantifizieren. Das Ziel war dabei, jene Eigenschaften des Pflanzenspektrums zu extrahieren, die Gradienten in der Artenzusammensetzung kennzeichnen. Im Ergebnis lieferten die statistischen Modelle eine Vorhersage von Art und Struktur der Vegetation in jedem Bildpixel und bildeten die räumlichen Muster von Pflanzengesellschaften ab (Neumann et al., 2016).

Die Komplexität von Arträumen wird dabei aus multidimensionalen „Arten x Flächen-Matrizen“ in einem iterativen Lernprozess – die nicht-metrische multidimensionale Skalierung (NMDS, Kruskal, 1964) – über weniger dimensionale Gradientenräume dargestellt.

Wird nun in diesem Gradientenraum die Verteilung von Pflanzenarten mit deren spektralen Signaturen verknüpft, entsteht ein multikriterielles Optimierungsproblem, welches durch neueste evolutionäre Lernalgorithmen gelöst werden kann. Solche Algorithmen adaptieren die Prozesse des ungerichteten Lernens aus Vorgängen der Vererbung, wie sie für die natürliche Evolution postuliert werden. Demnach kann aus zufälligen Rekombinationen von möglichen Eigenschaften (Modellparametern) unter Einbeziehung lediglich eines Überlebensmechanismus eine optimale Anpassung an die Umgebungskriterien (verschiedene Modelle) erreicht werden. Nach dem Prinzip „survival of the fittest“ werden diejenigen Parameter an eine nächste Generation von Lösungen übergeben, die in den Vorhersagemodellen von Pflanzenarten und -eigenschaften über Reflexionsspektren eine Maximierung der Modellgüte ermöglichen. Für jede Pflanzenart kann somit ein Satz an optimalen Lösungen (Pareto-Satz) ermittelt werden, der bei den vererbten Parameterkombinationen zur Selektion geeigneter Modelle führt (Non-dominated Sorting Genetic Algorithm NSGA-II; *Deb et al., 2002*).

Darauf aufbauend kommt ML bei der Aggregation von Lebensräumen zum Einsatz. Dabei werden auf der Basis von multispektralen Zeitreihen Pflanzengesellschaften a priori zu Biotopen zusammen-

gefasst, indem sie mit Hilfe von Random-Forest- und Support-Vector-Machine-Klassifikatoren anhand ihrer phänologischen Entwicklung unterschieden werden. Diese Verfahren werden mittels im Gelände für größere Flächen erhobenen Spektraleigenschaften trainiert und anschließend auf Bilddatenzeitreihen zur Ableitung von Biotoptypen angewendet. Die Trainingsinformationen werden anschließend in Spektralbibliotheken wie SPECTATION abgelegt (*SPECTATION, 2015*).

Die vorgestellte Methode wurde im Rahmen von zwei Projekten in Zusammenarbeit mit der Heinz Sielmann Stiftung und der Deutschen Bundesstiftung Umwelt (DBU) in ehemaligen und unter Naturschutz stehenden Truppenübungsplätzen in Brandenburg, Döberitzer Heide (DBU gefördert, abgeschlossen) und Kyritz-Ruppiner Heide (BMBF gefördert, laufend), entwickelt und getestet. Im Ergebnis werden für beide Regionen fernerkundungsbasierte Monitoringsysteme zur Beobachtung und Erfolgskontrolle von Pflegemaßnahmen (z. B. Mahd) mit dem Ziel der Erhaltung von Offenland (z. B. Sandtrockenrasen) zur Verfügung gestellt. Dabei ermöglicht die Einbeziehung von ML-Verfahren sowohl die Erstellung von Karten der Habitatverteilung als auch die Gewinnung von tieferen Einblicken in die Ökologie der Organisation von Einzelarten.

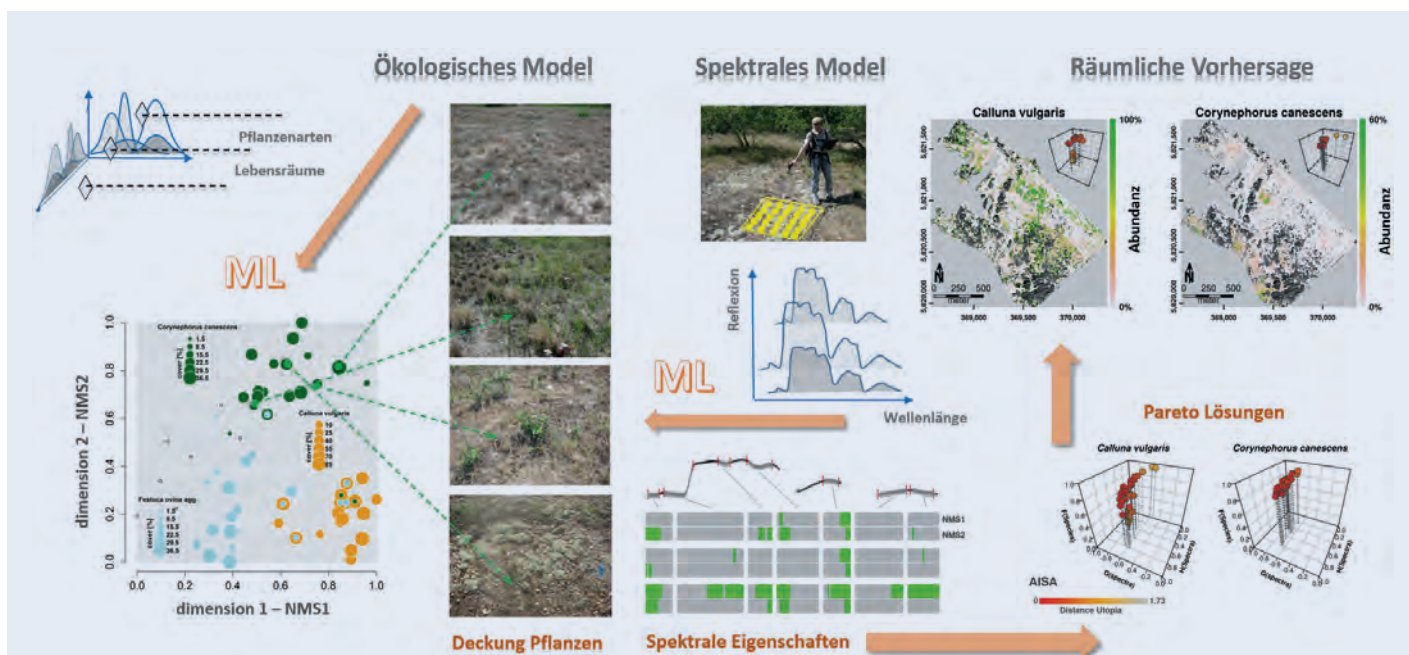


Abb. 4: Konzeptionelles Schema zur Darstellung des Einsatzes von maschinellem Lernen (ML) bei der Modellierung der Verteilung von Pflanzenarten mittels spektralen Reflexionssignaturen und deren räumlicher Vorhersage über multikriterielle evolutionäre Optimierungsräume. Dargestellt sind die räumlichen Verteilungen der Deckungsgrade (Abundanzen) von Silbergras (*Corynephorus canescens*) und Heidekraut (*Calluna vulgaris*) auf einem trockenen Offenland in der Döberitzer Heide.

Fig. 4: Conceptual scheme to illustrate the use of machine learning (ML) in modeling the distribution of plant species using spectral reflection signatures and their spatial prediction using multi-criteria evolutionary optimization spaces. The spatial distribution of the degrees of cover (abundances) of silver grass (*Corynephorus canescens*) and heather (*Calluna vulgaris*) is shown for a dry open land in the Döberitzer Heide.

Satellitengestützte Schätzung der Brutto-Primärproduktivität

Die Schätzung der Photosyntheseleistung von landwirtschaftlichen Nutzflächen ist wichtig, um den physiologischen Zustand der Pflanzen zu überwachen und die landwirtschaftliche Produktion zu prognostizieren. Die durch die Photosynthese von Pflanzen fixierte Kohlenstoffmenge wird durch die Brutto-Primärproduktivität (BPP) ausgedrückt und ist ein wichtiger Indikator für die Gesundheit und Leistungsfähigkeit von Pflanzengemeinschaften, der insbesondere auch große Bedeutung vor dem Hintergrund eines ständig wachsenden Bedarfs an Lebensmitteln hat.

Die BPP kann aus optischen Fernerkundungsdaten mit auf photosynthetischen-biochemischen Prozessen basierenden Modellen abgeschätzt werden. Ihre Anwendung wird jedoch durch ihre Komplexität und die hohe Anzahl von notwendigen Eingabeparametern erschwert, die oft nicht verfügbar oder mit starken Unsicherheiten behaftet sind. Alternativ können prozessbasierte Strahlungstransfermodelle zusammen mit ML-Verfahren mit dem Ziel der Modellinversion im Rahmen eines hybriden Ansatzes genutzt werden (Camps-Valls et al., 2011). Eine am GFZ durchgeführte Studie hat sich mit dieser Vorgehensweise befasst. Anstatt mehrere die Vegetationsstruktur kennzeichnende Parameter aus Satellitendaten abzuleiten, wurden bei diesem Ansatz mit Hilfe von ML-Verfahren die spektralen Reflexionsdaten und die meteorologischen Daten direkt mit der BPP der Pflanzen in Beziehung gesetzt. Auf diese Weise konnte der Zwischenschritt der Ableitung der für die Ermittlung der Photosyntheseleistung notwendigen Eingangsvegetationsparameter umgangen werden, während die komplexen Prozesse des ursprünglichen Modells berücksichtigt blieben.

Mehrere künstliche neuronale Netze und Random-Forest-Klassifikatoren kamen zum Einsatz, die mit einem umfangreichen über das Energiebilanzmodell SCOPE (van der Tol et al., 2009) abgeleiteten Trainingsdatensatz trainiert wurden. Anschließend wurden sie auf die meteorologischen Informationen und Sentinel-2-Satellitendaten, die eine hohe räumliche und zeitliche Auflösung sowie mehrere Spektralkanäle aufweisen, angewendet. Die Validierung der erzielten Ergebnisse erfolgte mit Daten von Messtürmen, die eine kontinuierliche Überwachung der Netto-CO₂-Ausgasung ermöglichten. Sie hat gezeigt, dass durch die ML-basierte Methode die BPP mit sehr guter Genauigkeit abgeschätzt wird (Abb. 5).

Da der erzeugte Trainingsdatensatz auch leicht an andere optische Fernerkundungssensoren angepasst werden kann, ist die Übertragbarkeit zwischen Systemen gewährleistet. Darüber hinaus arbeitet die entwickelte Methode schnell und effizient. Da sie gleichzeitig auf einem prozessbasierten Modell aufbaut, ist sie auf andere Standorte übertragbar. Eine globale Anwendung dieses ML-Verfahrens erfordert jedoch eine sehr hohe Rechenleistung. Zur Erprobung wurden daher die Cloud-Computing-Ressourcen der Plattform Google Earth Engine (GEE) genutzt, die auch alle verfügbaren Satellitendaten bereitstellt. Abbildung 6 zeigt einen Ergebnisausschnitt für ein Testgebiet in den USA sowie Unterschiede zwischen Blattflächenindex und BPP. Mit dieser neu entwickelten Methode ist es somit möglich, weltweit die Photosynthese von Nutzpflanzen zu einem bestimmten Zeitpunkt und für eine bestimmte Region zu bestimmen.

Ausblick

Maschinelles Lernen hat einen festen Platz in den Algorithmen zur Auswertung von Fernerkundungsdaten. Es wird in fast allen

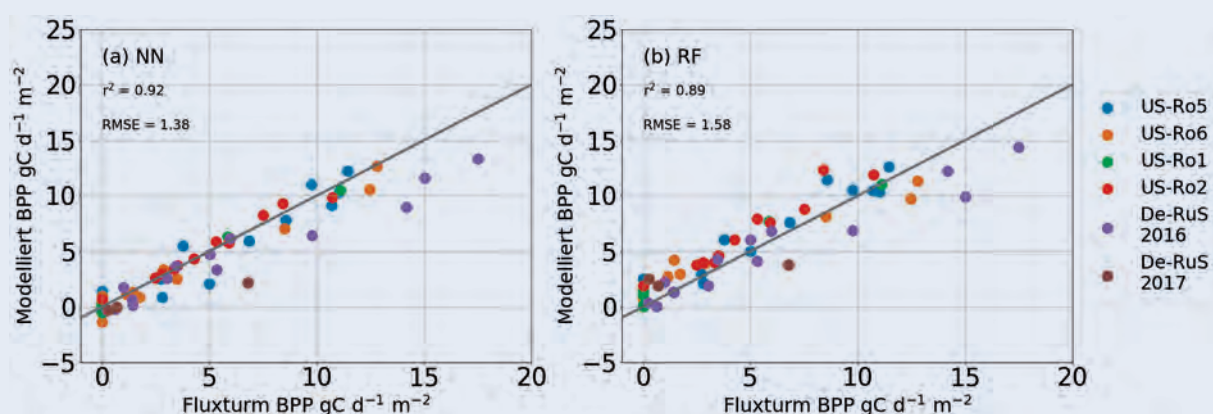


Abb. 5: Beziehung zwischen der täglich gemessenen BPP (Messturm) und der BPP abgeleitet durch (a) neuronales Netz (NN), und (b) Random Forest (RF) für sechs Turmstandorte (US-Ro1-5 und De-Rus)

Fig. 5: Relationship between daily flux tower gross primary productivity (BPP) and gross primary productivity estimated using (a) neural network, NN, and (b) random forest, RF, established for five flux tower sites (US-Ro1-5 and De-Rus).

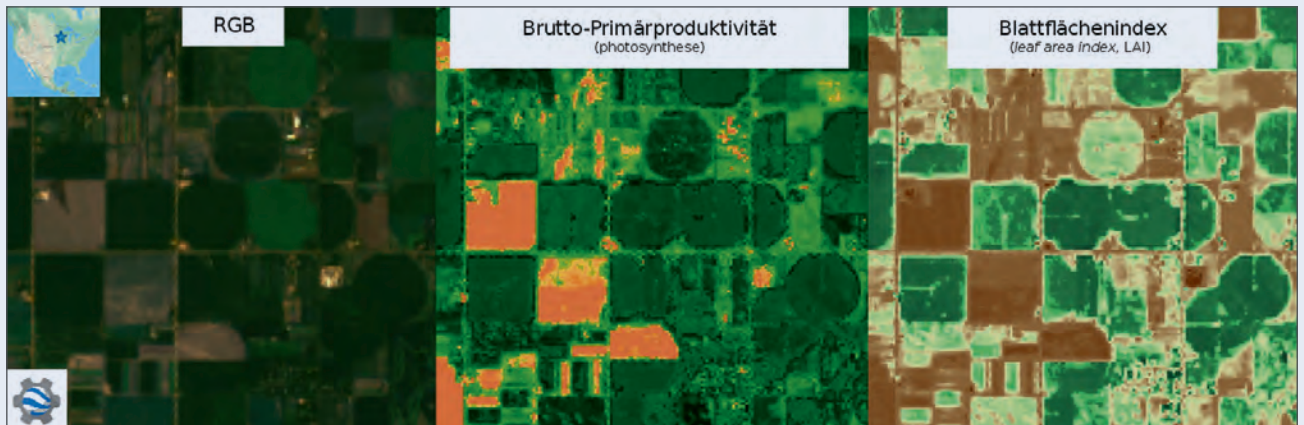


Abb. 6: Echtfarben-RGB-Bild auf der Basis von Landsat-8-Daten, abgeleitete BPP und LAI – berechnet und visualisiert in der Google Earth Engine

Fig. 6: RGB image from Landsat 8, together with estimated gross primary productivity (BPP) and leaf area index, calculated and visualized in the Google Earth Engine.

Bereichen zur Klassifizierung, Quantifizierung und Modellierung von Oberflächenparametern angewendet. Neben verbesserten Sensordaten tragen neue Entwicklungen im Bereich des ML dazu bei, die Genauigkeit der Fernerkundungsergebnisse zu erhöhen und initiieren neue Fernerkundungsprodukte für vielfältige Anwendungen in unserer modernen Gesellschaft.

Zukünftige Arbeiten in der GFZ-Sektion „Fernerkundung“ zielen insbesondere auf die Entwicklung von Algorithmen zur Verarbeitung von Big-Data-Fernerkundungsdaten ab, wie sie im Rahmen des EU/ESA-Copernicus-Programms bereits heute bereitgestellt werden. Hier gilt es insbesondere sogenannte Emulatoren zu entwickeln, die rechentechnisch aufwendige und komplexe Modelle einfach und effizient nachbilden können. Damit lassen sich sehr hohe Rechengeschwindigkeiten erzielen, die zur Verarbeitung der großen Datenmengen dringend benötigt werden.

Literatur

- Camps-Calls, G., Tuia, D., Gómez-Chova, L., Jiménez, S., Malo, J. (2011): Remote Sensing Image Processing. - *Synthesis Lectures on Image, Video, and Multimedia Processing*, 5, 1, pp. 1–192. DOI: <https://doi.org/10.2200/S00392ED1V01Y201107VM012>
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T. (2002): A fast and elitist multiobjective genetic algorithm: NSGA-II. - *IEEE Transactions on Evolutionary Computation*, 6, 2, pp. 182–197. DOI: <https://doi.org/10.1109/4235.996017>
- Guanter, L., Kaufmann, H., Segl, K., Förster, S., Rogaß, C., Chabrillat, S., Küster, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M., Mühle, H., Müller, R., Habermeyer, M., Ohndorf, A., Hill, J., Buddenbaum, H., Hostert, P., van der Linden, S., Leitão, P. J., Rabe, A., Doerffer, R., Krusemann, H., Xi, H., Mauser, W., Hank, T., Locherer, M., Rast, M., Staenz, K., Sang, B. (2015): The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. - *Remote Sensing*, 7, 7, pp. 8830–8857. DOI: <https://doi.org/10.3390/rs70708830>
- Hollstein, A., Fischer, J., Carbajal Henken, C., Preusker, R. (2015): Bayesian cloud detection for MERIS, AATSR, and their combination. - *Atmospheric Measurement Techniques*, 8, pp. 1757–1771. DOI: <https://doi.org/10.5194/amt-8-1757-2015>
- Hollstein, A., Segl, K., Guanter, L., Brell, M., Enesco, M. (2016): Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. - *Remote Sensing*, 8, 8, 666. DOI: <https://doi.org/10.3390/rs8080666>
- Kruskal, J. B. (1964): Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. - *Psychometrika*, 29, 1, pp. 1–27. DOI: <https://doi.org/10.1007/BF02289565>
- Neumann, C., Weiss, G., Schmidlein, S., Itzerott, S., Lausch, A., Doktor, D., Brell, M. (2015): Gradient-Based Assessment of Habitat Quality for Spectral Ecosystem Monitoring. - *Remote Sensing*, 7, 3, pp. 2871–2898. DOI: <https://doi.org/10.3390/rs70302871>
- Neumann, C., Itzerott, S., Weiss, G., Kleinschmit, B., Schmidlein, S. (2016): Mapping multiple plant species abundance patterns - A multiobjective optimization procedure for combining reflectance spectroscopy and species ordination. - *Ecological Informatics*, 36, pp. 61–76. DOI: <https://doi.org/10.1016/j.ecoinf.2016.10.002>
- SPECTATION [Client] (2015): SPECTATION – SPECTral database for vegeTATION, verfügbar unter <http://www.gfz-potsdam.de/spectation> [letzter Zugriff: 18.04.2018]
- Steinberg, A., Chabrillat, S., Stevens, A., Segl, K., Förster, S. (2016): Prediction of Common Surface Soil Properties Based on Vis-NIR Airborne and Simulated EnMAP Imaging Spectroscopy Data: Prediction Accuracy and Influence of Spatial Resolution. - *Remote Sensing*, 8, 7, 613. DOI: <https://doi.org/10.3390/rs8070613>
- van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., Su, Z. (2009): An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. - *Biogeosciences*, 6, 12, pp. 3109–3129. DOI: <https://doi.org/10.5194/bg-6-3109-2009>
- WWF (2016): Living Planet Report 2016 : Risk and resilience in a new era, Gland : WWF International, 144 p., verfügbar unter http://awsassets.panda.org/downloads/lpr_living_planet_report_2016.pdf [letzter Zugriff: 18.04.2018]