

# SpectroAnalyst – Management von Hyperspektraldaten für die Grünlandforschung

ANDREAS SCHAUMBERGER<sup>1</sup> & MANUEL ADELWÖHRER<sup>1</sup>

*Zusammenfassung: Die Messung der Entwicklung von Pflanzenbeständen setzt nicht destruktive Beobachtungen voraus. Dabei stellt die Reflexionsspektroskopie eine Methode dar, mit der viele Vegetationscharakteristiken aus dem elektromagnetischen Spektrum des sichtbaren und nahen Infrarotbereiches ausgewertet werden. Wirtschaftsgrünland entwickelt sich in mehreren, durch Schnitte voneinander getrennten, relativ kurzen Wuchsperioden, in denen die vertretenen Arten den Pflanzenbestand unterschiedlich stark dominieren. Für Grünland braucht es deshalb dichte Zeitreihen mit vielen Messwiederholungen. Schnell fallen dabei mehrere tausend Spektralsignaturen an. Das hier vorgestellte Werkzeug SpectroAnalyst unterstützt den gesamten Arbeitsprozess von der Speicherung über Rohdatenverarbeitung und Korrektur bis hin zur Berechnung von Vegetationsindizes.*

## 1 Einleitung

Die Pflanzenbestände einer Grünlandfläche entwickeln sich in Abhängigkeit der Nutzungsintensität und der Standortbedingungen sehr unterschiedlich. Anders als im Ackerbau besteht das Grünland aus mehreren Arten mit verschiedenen Ansprüchen und Entwicklungsphasen und wird meist mehrmals im Jahr geerntet. Sowohl die Art der Bewirtschaftung als auch der Standort mit Klima, Boden und Topographie wirken auf die Zusammensetzung des Pflanzenbestandes und in weiterer Folge auf den Ertrag und die Futterqualität (PÖTSCH 2009). Ein zentrales Thema der Grünlandforschung ist die Wirkung dieser Einflussfaktoren auf die Leistungsfähigkeit eines Grünlandbestandes. Versuchsernten im Rahmen wissenschaftlicher Feldexperimente werden meist in Abstimmung mit der zu untersuchenden Schnitthäufigkeit, also wenige Male pro Jahr vorgenommen und das so gewonnene Pflanzenmaterial im Labor untersucht. Für die Grünlandwirtschaft können auf diese Weise Aussagen über den Zusammenhang zwischen Höhenstufen und Futterertrag (BUCHGRABER & PÖTSCH 1999), Schnittzeitpunkt und Futterinhaltsstoffe (BUCHGRABER 1997), Bewirtschaftungsintensität und Stickstoffkreislauf, Bewirtschaftungssysteme und Futterqualität (PÖTSCH et al. 2006) sowie Düngung und Futterinhaltsstoffe (PÖTSCH 1995) getroffen werden, um einige Beispiele zu nennen.

Eine kontinuierliche, nicht destruktive Beobachtung von Pflanzenbeständen während eines Aufwuchses ermöglicht im Gegensatz zur Ernte die Erfassung von Entwicklungsparametern, da hier die oberirdische Biomasseentwicklung nicht beeinflusst wird. Neben traditionellen Erhebungen in Grünlandexperimenten wie zum Beispiel die Schätzung der Artengruppenverhältnisse, phänologische Beobachtungen oder Höhenmessungen des Pflanzenbestandes, bietet die Reflexionsspektroskopie unter Einsatz eines Feldspektrometers eine weit darüber hinaus gehende Fülle an strukturellen Informationen. Der hohe Informationsgehalt von Spektralsignaturen wird unter

---

<sup>1</sup> Höhere Bundeslehr- und Forschungsanstalt für Landwirtschaft Raumberg-Gumpenstein, Altdrning 11, A-8952 Irdning-Donnersbachtal, E-Mail: andreas.schaumberger@raumberg-gumpenstein.at

anderem dazu genutzt, Modelle zur Ertrags- und Qualitätsschätzung zu entwickeln bzw. Vegetationskennwerte abzuleiten (vgl. BIEWER et al. 2009; DORIGO 2008; KLINGLER et al. 2019; PSOMAS et al. 2011; PUNALEKAR et al. 2018; THENKABAIL et al. 2014). Mit der Umformung von Hyperspektraldaten in Sentinel-2 entsprechende Multispektraldaten mit Hilfe von Spectral Response Functions (S2-SRF) (ESA 2018) besteht zudem die Möglichkeit, Modelle zu entwickeln, die durch Einbindung von Sentinel-2-Szenen regional anwendbar sind. Auf diese Weise können Grünlandexperimente als Datengrundlage verwendet werden, deren Parzellengröße in der Regel nicht ausreicht, um darauf Sentinel-2-Daten sinnvoll auszuwerten und für die Modellierung zu verwenden.

Um die einzelnen Entwicklungsschritte eines Pflanzenbestandes differenzieren und sie nach Möglichkeit mit Umwelteffekten, wie z. B. der Witterung, in Beziehung setzen zu können, ist eine Beobachtung in kurzen Zeitabständen notwendig. Zudem dauert einer von meist mehreren Grünlandaufwüchsen nur wenige Wochen, in denen eine relativ rasche Entwicklung erfolgt. Wird auch die räumliche Variabilität eines Pflanzenbestandes erfasst, müssen innerhalb der beobachteten Flächen Messungen an möglichst vielen Positionen durchgeführt werden. Die hohe zeitliche und räumliche Beobachtungsdichte mit einem Feldspektrometer führt zwangsläufig und sehr rasch zu einer erheblichen Datenmenge. Die Verarbeitung der Rohspektren, deren Verwaltung und Auswertung erfordert ab einem gewissen Umfang zwangsläufig technische Unterstützung, ohne die eine effektive und effiziente Arbeit mit Hyperspektralsignaturen nicht mehr möglich ist. Der hier als Beispiel vorgestellte SpectroAnalyst ist ein Managementwerkzeug, das dem Benutzer erlaubt, die gesamte Spektraldatenverarbeitung mit möglichst geringem Aufwand zu bewerkstelligen.

## 2 Problemstellung

Mit dem Feldspektrometer können innerhalb kürzester Zeit viele Beobachtungen durchgeführt werden, da der Messaufwand bei günstigen Witterungsbedingungen vergleichsweise gering ist. Vor allem bei komplexen Grünlandversuchen, in denen die Entwicklung von Pflanzenbeständen innerhalb kurzer Zeitabstände beobachtet werden soll, steigt die Zahl der erfassten Spektren kontinuierlich an. Eine manuelle Verarbeitung der Daten würde dabei viel Zeit und Energie erfordern. Ausgehend von den Rohdaten bis hin zur Anwendung der hyperspektralen Daten sind viele Arbeitsschritte notwendig, die ohne softwaretechnische Unterstützung und konsistentem Datenmanagement kaum zu bewältigen sind. Vor allem dann, wenn eine allgemeine Spektraldatenbank als Infrastruktur für beliebige Versuche und verschiedene Projekte entstehen soll.

Ein wichtiger Aspekt für die langfristige Verwertbarkeit von Daten ist deren Verknüpfung mit Metadaten. Sie spielen eine entscheidende Rolle bei der Beurteilung der Qualität von Spektren, indem sie Einblicke in die Entstehungsgeschichte geben (RASAIAH et al. 2015). Ohne Metadaten ist weder eine systematische Speicherung der Spektren noch eine Kriterien basierte Selektion und Auswertung möglich. Neben der Bereitstellung von allgemeinen Informationen über Messzeitpunkt, Wetter, Art des Pflanzenbestandes, seine phänologische Entwicklungsphase sowie Messpersonal muss jede Spektralsignatur einem Experiment und Messgerät zugeordnet werden können. In der Regel bestehen Grünlandversuche aus Varianten mit mehrfachen Wiederholungen. Auch hier ist eine entsprechende Zuordnung der einzelnen Spektren notwendig, um Auswertungen nach Faktoren und Faktorstufen effizient durchführen zu können. Zusätzlich zur Beschreibung

empfiehlt es sich, fotografische Aufnahmen der gemessenen Oberfläche in die Metadaten einzugliedern; sie erleichtern bei der späteren Datenanalyse die Identifikation unerklärbarer Abweichungen und Ausreißer in den Spektren (MACARTHUR 2011).

Um eine vollständige Verknüpfung zwischen Metadaten, Rohdaten, weiter verarbeiteten Daten und den daraus abgeleiteten Statistiken und Indizes zu realisieren, ist der Einsatz eines Datenbankmanagementsystems sinnvoll und notwendig. Eine wichtige Anforderung an Datenbanken im Allgemeinen und einer Spektraldatenbank im Besonderen ist die Skalierbarkeit, um die mit der Zeit wachsende Datenmenge weiterhin effizient bewältigen zu können. Ebenso wichtig ist eine umfassende Erweiterbarkeit des Systems, ohne Änderungen an den Strukturen vornehmen zu müssen. Bei der Implementierung eines relationalen Datenschemas ist in diesem Zusammenhang auf die vollständige Einhaltung der Normalisierungsregeln zu achten. Eine Anlehnung an Richtlinien und Anforderungsdefinitionen für wissenschaftliche Datenbanken, wie sie unter anderem PORTER (2000), VAN EVERT et al. (1999a), VAN EVERT et al. (1999b) oder WIGGINS et al. (2013) vorschlagen, ist hilfreich und vermindert die Wahrscheinlichkeit, nachträgliche und mit hohem Aufwand verbundene Änderungen des Datenschemas vornehmen zu müssen.

Um Spektralsignaturen evaluieren zu können, ist deren Visualisierung als Graph von großem Vorteil. Neben einem geeigneten Datenmodell ist daher die Benutzerschnittstelle, welche sowohl die grafische Ausgabe der Spektren als auch die Eingabe von Daten unterstützt, eine wichtige Voraussetzung für die effiziente Verwaltung und Bearbeitung von Spektren. Im SpectroAnalyst wurde der Programmoberfläche große Aufmerksamkeit geschenkt, da nur mit ihrer Hilfe ein gemeinsamer Zugriff auf Spektraldaten und all ihren Verknüpfungen (Metadaten, Sensoren, Varianten, usw.) möglich ist. In den folgenden Abschnitten wird das Konzept dieses Werkzeuges vorgestellt. Es soll beispielhaft zeigen, wie die Ansprüche an eine wissenschaftliche Datenbank mit den Anforderungen an das Management von Hyperspektraldaten kombiniert werden.

### **3 Material und Methoden**

#### **3.1 Datenmodell**

Bei der Entwicklung einer Software für das Management von Hyperspektraldaten muss im Vorfeld entschieden werden, ob die Spektren dateibasiert oder in einer Datenbank gespeichert werden. Ein dateibasiertes System muss mit Funktionen ausgestattet sein, die eine Zusammenstellung und dateiübergreifende Auswertung unterstützen. Beispiele für diese Art der Verwaltung sind HyperCor von AASEN et al. (2014) oder 3S von MERONI & COLOMBO (2009). Die Software GRAMS von ThermoFisher Scientific geht einen Schritt weiter und verknüpft die Oberfläche für Management und Auswertung mit einer Datenbank, in der die Spektren vorgehalten werden. Hier wird bereits eine direkte Beziehung von Metadaten und einzelnen Spektren realisiert und zudem eine breite Palette an Auswertungsmöglichkeiten angeboten. Die Entscheidung, dennoch ein eigenes System wie den SpectroAnalyst zu entwickeln, entstand aus der Notwendigkeit heraus, den gesamten Workflow von der Rohdatenverarbeitung bis hin zur Auswertung in ein System zu integrieren und dabei jeden Arbeitsschritt zusammen mit den Rohdaten im Sinne einer umfassenden Nachvollziehbarkeit zu speichern. Das Erreichen eines Höchstmaßes an Flexibilität

und die Möglichkeit einer beliebigen Erweiterbarkeit von Funktionen waren dabei die wichtigsten Motive. Anforderungen an das System und speziell an das Datenmodell sind dabei folgende:

- Der Import verschiedener proprietärer Formate (z. B. ASD FieldSpec, tec5 HandySpec, OceanOptics, Carl Zeiss Corona, usw.) mit unterschiedlichen Wellenlängenbereichen (VNIR, SWIR1, SWIR2) in eine einheitliche, homogene Datenstruktur wird unterstützt.
- Die Rohdaten werden im Originalzustand gespeichert und können jederzeit für eine neuerliche Datenverarbeitung bereitgestellt werden.
- Zu jeder Spektralsignatur werden Metainformationen gespeichert.
- Spektren werden Experimenten und in weiterer Folge einzelnen Varianten zugeordnet.
- Glättungen, Interpolationen und Korrekturen (Detector Offset) werden getrennt gespeichert und können jederzeit entfernt bzw. geändert werden.
- Spektralsignaturen können geglättet und zur Erstellung von Mittelwertspektren beliebig zusammengefasst werden.
- Die Gleichungen für die Berechnung von Vegetationskennwerte werden in der Datenbank abgelegt und können über die Benutzerschnittstelle auf beliebige Spektren angewendet werden.

Um die hier angeführte Funktionalität zu unterstützen, wurde das folgende relationale Datenmodell (Abb. 1) entwickelt und im Datenbankmanagementsystem MS SQL Server implementiert:

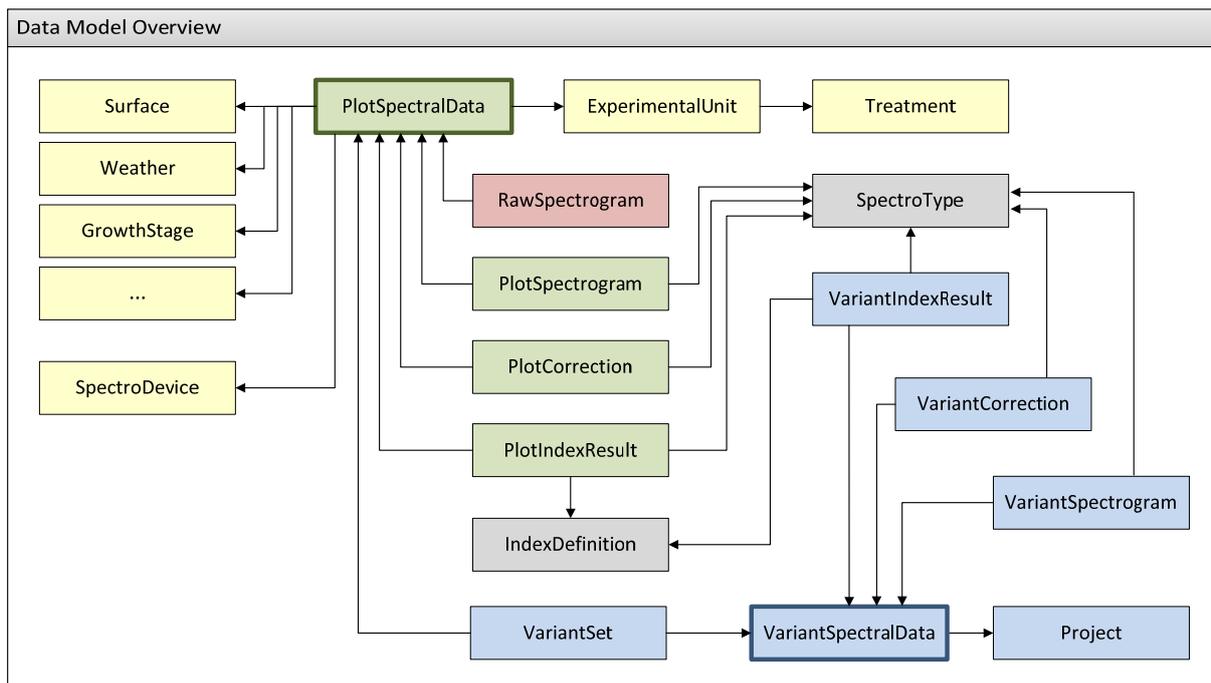


Abb. 1: Vereinfachtes Datenmodell des SpectroAnalyst auf Relationenebene

Die mit den verschiedenen Feldspektrometern erfassten Daten werden in einer schlanken Tabelle *RawSpectrogram* mit nur vier Feldern gespeichert. Diese vier Felder beinhalten ein Indikator-Attribut (ID), das auf die Tabelle *PlotSpectralData* verweist, eine weitere ID, welche alle Reflexionswerte einer einzelnen Spektralsignatur als zusammengehörend kennzeichnet, die

Wellenlänge sowie den Reflexionswert selbst. Indem für diese Tabelle ein dreifach zusammengesetzter Primärschlüssel mit der *PlotSpectralData*-ID als Fremdschlüssel, der Rohsignatur-ID sowie dem Wellenlängenwert definiert ist, können sämtliche gespeicherte Reflexionswerte exakt zugeordnet werden. Auf diese Weise ist es möglich, Spektren aus verschiedenen Quellen und Wellenlängenbereichen in einer gemeinsamen Struktur zu speichern.

Werden beispielsweise auf einer Versuchsparzelle Spektren in vierfacher Wiederholung gemessen, ergeben sich für die Eingabe in die Datenbank folgende Arbeitsschritte: Zunächst wird ein neuer Datensatz in der Tabelle *PlotSpectralData* mit einer laufenden ID angelegt. Dieser Datensatz enthält eine Reihe von Fremdschlüsselattributen, die auf Tabelleneinträge verweisen, in denen Metadaten definiert sind (Wetter, Oberfläche, phänologische Phase, usw.). Ebenso wird auf die Parzellendefinition (*ExperimentalUnit*) mit der jeweiligen dort stattfindenden Behandlung (*Treatment*) verwiesen. Ein Tupel in *PlotSpectralData* bildet sozusagen die Klammer über Spektral- und Metainformationen. Die Tabelle *RawSpectrogram* enthält die einzelnen Reflexionswerte in einer jeweils eigenen Zeile. Alle vier erfassten Spektralsignaturen verweisen auf das gleiche Tupel in *PlotSpectralData*, damit werden sie als zusammengehörende Wiederholungen gekennzeichnet. Die vier Spektren selbst unterscheiden sich durch die in einem weiteren Feld gesetzte *RawSpectrogram*-ID, indem alle Reflexionswerte einer Spektralsignatur die gleiche ID erhalten. Innerhalb des Spektrums sind die Einzelwerte der jeweiligen Wellenlänge zugeordnet.

Um aus Messwiederholungen eine für die erfasste Versuchsfläche repräsentative Spektralsignatur zu generieren, werden die vier Spektren gemittelt und als Mittelwertsignatur in die Tabelle *PlotSpectrogram* gespeichert. Sie ist ebenso wie *RawSpectrogram* mit nur wenigen Feldern ausgestattet: Ein Fremdschlüsselattribut verweist auch hier auf *PlotSpectralData*, mit dem Effekt, dass die zunächst gesetzte Klammer um Rohspektren und Metainformationen auf alle davon abgeleiteten Daten erweitert wird. In *PlotSpectrogram* werden nicht nur Mittelwertspektren aus den Wiederholungen, sondern auch weitere Ableitungen gespeichert. Aus diesem Grund wird der Primärschlüssel dieser Tabelle ebenfalls aus mehreren Attributen zusammengesetzt und enthält neben der Wellenlänge auch eine Identifikation des Typs, indem mit einem ID-Feld auf die Tabelle *SpectroType* verwiesen wird. Auf diese Weise können Mittelwertspektren, geglättete Spektren, trendbereinigte Spektren, Spektren aus statistischen Kennzahlen, usw. voneinander unterschieden und in einer gemeinsamen Tabelle mit einheitlicher Struktur abgelegt werden. Über die *PlotSpectralData*-ID bleiben sie mit den dazugehörigen Basisdaten verknüpft.

Im Übergangsbereich verschiedener Sensoren kommt es häufig zu kleineren und größeren Verschiebungen (Detector Offset). Der SpectroAnalyst bietet die Möglichkeit, diesen Versatz zu eliminieren und die Spektralsignatur in den betroffenen Wellenlängenbereichen zu korrigieren. In der Tabelle *PlotCorrection* werden die berechneten Korrekturfaktoren gespeichert und den jeweiligen Spektren zugeordnet. Gespeichert werden keine korrigierten Signaturen, sondern lediglich Faktoren, die über definierte Wellenlängenbereiche additiv oder multiplikativ bei jedem Zugriff auf die Originaldaten dynamisch angebracht werden.

In der SpectroAnalyst-Datenbank werden neben Spektren und Metadaten auch Index-Definitionen in der Tabelle *IndexDefinition* gespeichert. Sie enthält eine Spalte mit allen Formeln der erfassten Indizes, die in der Benutzeroberfläche angezeigt und bei entsprechender Auswahl auf beliebige Spektren angewendet werden können. Der Vorteil dieser Vorgehensweise besteht darin, dass die Indexliste über die Datenbank beliebig erweitert werden kann und über die Benutzeroberfläche für

die Anwendung sofort zur Verfügung steht. In der Tabelle *PlotIndexResult* werden die Berechnungsergebnisse abgelegt und auch hier besteht über die *PlotSpectralData*-Tabelle eine direkte Verknüpfung mit allen dazugehörigen Spektral- und Metainformationen.

Die in Abb. 1 in blauer Farbe markierten Relationen weisen die gleiche Funktionalität wie die grün markierte Plot-Struktur auf, beziehen sich jedoch auf Spektren, die zu Varianten zusammengefasst wurden. In der Tabelle *VariantSet* sind beliebig viele *PlotSpectralData* miteinander verbunden und können über *VariantSpectralData* auf die gleiche Weise wie die Spektren auf Parzellenebene verarbeitet und ausgewertet werden. Über die Beziehung zu *PlotSpectralData* bleiben die Varianten weiterhin mit den Rohdaten in *RawSpectrogram* verbunden, wobei die Tabelle *VariantSpectrogram* die davon abgeleiteten Spektren enthält.

### 3.2 Korrektur der Rohspektren

Bevor Hyperspektraldaten ausgewertet und verwendet werden, müssen die Rohdaten einer Kontrolle und gegebenenfalls einer Bearbeitung unterzogen werden. Im SpectroAnalyst werden zu diesem Zweck alle neu importierten Spektren in ihrem Originalzustand grafisch dargestellt. Weichen einzelne Messwiederholungen innerhalb einer beobachteten Versuchsparzelle stark voneinander ab, wird in der Regel das vom Messzeitpunkt verfügbare Bild des Pflanzenbestandes herangezogen, um mögliche Ausreißer innerhalb einer bestimmten Parzelle zu identifizieren und diese vor der Bildung eines Mittelwertspektrums auszuschließen.

In einem nächsten Schritt wird die Korrektur eines möglicherweise aufgetretenen Detector Offsets durchgeführt. Eine Verschiebung der Signatur im Übergangsbereich verschiedener im Feldspektrometer verbauter Sensoren, wie beispielsweise bei etwa 1000 nm von VNIR zu SWIR1, ist ein immer wieder auftauchendes Problem und sollte bei der Rohdatenbearbeitung unbedingt berücksichtigt werden (AASEN et al. 2014; DORIGO 2008; DORIGO et al. 2006; GARFAGNOLI et al. 2013; MACARTHUR et al. 2012; MILTON et al. 2009). Die Ursachen dafür sind nach eigenen Erfahrungen mit den HandySpec von tec5 nicht exakt zu identifizieren, dürften nach HUENI & BIALEK (2017) aber ihren Grund im unterschiedlich starken Einfluss von Umweltbedingungen (Temperatur und Luftfeuchte) auf die verschiedenen Sensoren haben. Aufgrund des häufigen Auftretens dieser Situation wird mit dem SpectroAnalyst jeder Detector Offsets vor der Bildung des Mittelwertspektrums korrigiert. Die Herausforderung dabei ist, eine vom Messrauschen in diesem Übergangsbereich unabhängige Differenz zwischen den beiden Spektralbereichen zu ermitteln. Dabei wird folgender Algorithmus angewendet: Zunächst erfolgt eine Savitzky-Golay-Glättung in einer variablen Breite (smoothing region) und Polynom 2. Grades (SAVITZKY & GOLAY 1964) für jenen Bereich des Spektrums, dessen Niveau an das des benachbarten angeglichen werden soll. Auf dem letzten Abschnitt der geglätteten Spektralsignatur, also unmittelbar vor dem Sensorübergang, wird eine Forecast- bzw. Trend-Funktion angewendet, um den Basiswert für die anschließende Korrektur zu ermitteln. Dieser Basiswert wird entweder absolut oder relativ in Beziehung zum unmittelbar benachbarten Wert des nicht zu korrigierenden Spektralbereiches gesetzt und so ein Korrekturwert ermittelt, der in Abhängigkeit der Benutzerangaben additiv oder multiplikativ auf den zu korrigierenden Spektralbereich angewendet wird. Es besteht die Möglichkeit, VNIR an SWIR1 oder SWIR1 an VNIR anzupassen und, als dritte Option, beide Bereiche zu gleichen Teilen aufeinander hin zu korrigieren. Bevorzugt erfolgt die

Korrektur multiplikativ, um sie im Verhältnis zur Höhe des Reflexionswertes anzubringen. Beim tec5-HandySpec sollte laut Herstellerangabe der SWIR1-Bereich unverändert gehalten und andere mit dem Detector Offset abgesetzte Spektralbereiche darauf hin korrigiert werden.

### **3.3 Glättung der Rohspektren**

Nachdem die Rohspektren korrigiert und um den Detector Offset bereinigt wurden, wird die Spektralsignatur geglättet, um Messrauschen weitgehend zu eliminieren (KING et al. 1999; RUFFIN & KING 1999). Damit bei dieser Prozedur spektrale Informationen bestmöglich erhalten bleiben, wurde eine dynamische Savitzky-Golay-Glättung nach BARAK (1995) implementiert, bei der abschnittsweise Polynome verschiedenen Grades in Abhängigkeit der Kurvencharakteristik mit mehrfacher Iteration angewendet werden.

Bei den mit dem HandySpec (tec5) gemessenen Spektralsignaturen von 400 bis 1690 nm wird der Bereich der Wasserabsorption zwischen 1350 bis 1450 nm sowie jener am Ende des Messbereiches ab 1650 nm bei der dynamischen Glättung nicht berücksichtigt. Hier kann der SpectroAnalyst-Benutzer optional eine kubische Spline-Interpolation aus der Programmbibliothek ALGLIB anwenden, um die am Ende des Messbereichs vermehrt auftretenden Schwingungen bestmöglich zu glätten und im Wasserabsorptionsband einen ebenfalls geglätteten Übergang zu schaffen. Dieser Bereich zwischen 1350 und 1450 nm ist für eine Auswertung zwar nicht relevant, schafft jedoch durch die Interpolation aus optischen Gründen und alternativ zum Herausschneiden dieses Kurvenabschnittes einen glatten Übergang zwischen den beiden relevanten SWIR1-Bereichen 1000 bis 1350 und 1450 bis 1690.

### **3.4 Werkzeuge zur Analyse und Auswertung**

Für sämtliche gespeicherte Spektren können Vegetationskennwerte berechnet werden, deren Definitionen beim Start des SpectroAnalyst aus der Datenbank eingelesen werden und dem Benutzer über die Programmoberfläche zur Verfügung stehen. Die Berechnungen können auf die originalen bzw. korrigierten und/oder auf die noch zusätzlich geglätteten Spektralsignaturen angewendet werden.

Im Zuge der Speicherung der korrigierten Spektren (vgl. Abschnitt 3.2) werden automatisch Sentinel-2-Bänder (S2-SRF) (ESA 2018) generiert und in einer eigenen Tabelle abgelegt. Auf diese Weise entsteht neben den Hyperspektraldaten auch ein multispektraler Datensatz, der mit einem Bestimmtheitsmaß von 99 % (KLINGLER et al. 2019) jenen von Sentinel-2-Bändern entspricht und daher in satellitenbasierten, regionalen Modellen Anwendung finden kann. So fließt zum Beispiel dieser Datensatz in ein neuronales Netz zur Berechnung des LAI nach BARET et al. (2010) ein, das mit den Strahlungstransfermodellen PROSPECT und SAIL (JACQUEMOUD et al. 2009) speziell auf Sentinel-2 abgestimmt wurde. Ausgeführt wird das externe Programm zur Berechnung des LAI über eine vom SpectroAnalyst aufgerufene MATLAB-Runtime. Dabei werden acht S2-SRF-Bänder (B3, B4, B5, B6, B7, B8A, B11 und B12) sowie Sonnenwinkelparameter, die in Abhängigkeit des Zeitstempels der Spektren in R (Package solarpos) berechnet wurden, mitgegeben.

Für ein beliebig zusammenstellbares Set von Spektren besteht die Möglichkeit, statistische Auswertungen zu rechnen. Realisiert wird dies über die im SpectroAnalyst implementierte Schnittstelle zum Statistikprogramm R und der Anwendung des Packages `prospectr` (STEVENS & LOPEZ 2014). Auf diese Weise können gegebenenfalls weitere R-Packages integriert und der Funktionsumfang des SpectroAnalyst angepasst und erweitert werden.

## 4 Implementierung

### 4.1 Systemkomponenten

Das Softwaresystem des SpectroAnalyst entspricht dem Konzept einer Dreischichtenarchitektur. Auf Ebene der Datenhaltung wurde ein relationales Datenmodell im Datenbankmanagementsystem MS SQL Server implementiert und auf die im Abschnitt 3.1 dargestellte Weise strukturiert. Die Art und Weise der Speicherung von Spektren und der damit verbundenen und daraus abgeleiteten Daten ist die Voraussetzung für die im SpectroAnalyst implementierte Funktionalität. Die Anwendungslogik befindet sich in MS C# erstellten Klassen des .NET-Frameworks und ist über ADO.NET Data Provider Objects mit der Datenhaltungsebene verbunden. Neben den eigenen Klassen wird weitere Funktionalität über externe Prozesse eingebunden. Das für die Auswertung von Spektraldaten entwickelte R-Package `prospectr` wird beispielsweise über einen R-Prozess ausgeführt, welcher mit der Bibliothek R.NET über NuGet in den SpectroAnalyst eingebunden wurde. Auf diese Weise können auch weitere Packages verwendet und der SpectroAnalyst effizient erweitert werden. Die vom Benutzer in der Anwendungsoberfläche ausgewählten Spektren werden in einer zweidimensionalen Array-Struktur an den R-Prozess übergeben, dort verrechnet und die Ergebnisse mit gleicher Struktur wieder entgegengenommen und gespeichert. Ebenso wie R können auch MATLAB-Algorithmen über Einbindung der MATLAB-Runtime und deren Aufruf mittels `PROCESS`-Objekt verwendet werden. Im SpectroAnalyst wird die Berechnung des LAI aus den Sentinel-2-Multispektraldaten auf diese Weise durchgeführt.

Die Präsentationsschicht als Benutzerschnittstelle wurde mit C#.NET GUI-Klassen erstellt und bietet dem Benutzer die Möglichkeit, geladene Spektren dynamisch zu visualisieren, Kurvenausschnitte zu vergrößern und zu verkleinern sowie Werte abzufragen. Neben der grafischen Darstellung kann im linken Bereich (siehe Abb. 2) navigiert und zwischen den Aggregationsebenen (Plot und Variant) gewechselt werden. Die Oberfläche bietet die Möglichkeit, Spektraldaten zusammen mit Metainformationen zu importieren, aus der Datenbank nach Meta-Kriterien aufzurufen und zu exportieren. Ebenso ist der Export der mittels SRF berechneten Sentinel-2-Bänder sowie der Indizes als Tabellen vorgesehen. Die Benutzeroberfläche erlaubt die Auswahl einzelner Spektren, um sie beispielsweise aus einer Zusammenstellung für die Mittelwertbildung herauszunehmen bzw. überhaupt aus der Datenbank zu löschen. Mit „Show Picture“ kann für die Erhebung auf einer Versuchsparzelle das zugehörige Foto eingeblendet werden, um eine Evaluierung der Spektralsignaturen zu erleichtern. Für die Metadateneingabe und -änderung, die Optionen für die Rohdatenverarbeitung (Korrektur, Mittelwertbildung, Glättung) sowie die Optionen für die Index-Berechnung und für die statistische Analyse (`prospectr`) können über Subfenster ein- und ausgeblendet werden. Abb. 2 zeigt die wichtigsten Elemente der Oberfläche.

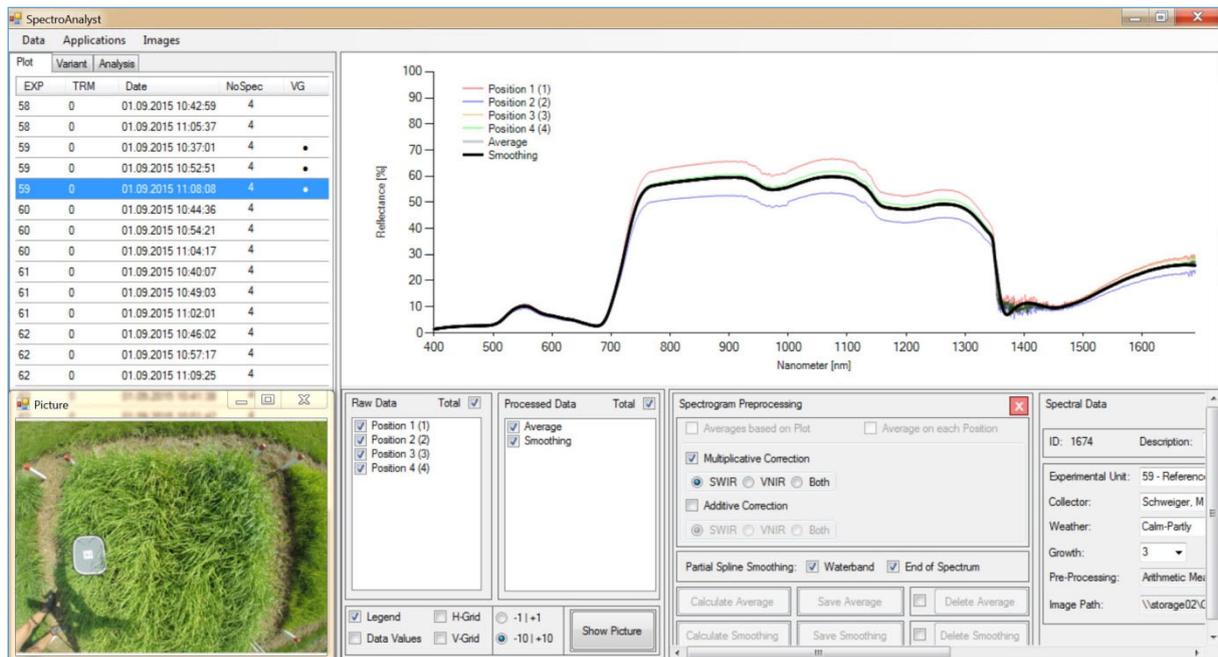


Abb. 2: Benutzeroberfläche des SpectroAnalyst

## 4.2 Erfassung und Abfrage von Daten

Beim Import neuer Spektren wird über Zeitstempel geprüft, ob diese bereits in der Datenbank vorhanden sind. Ist dies nicht der Fall, werden alle sich innerhalb eines ausgewählten Verzeichnisses befindlichen Spektren importiert und in die Datenbank gespeichert. Beim Import werden auch die Metainformationen vom Benutzer eingegeben, die als Template für alle im aktuellen Import-Set enthaltenen Spektren verwendet, bei Änderungsbedarf für einzelne Spektren aber nachträglich jederzeit angepasst werden können. In den SpectroAnalyst-Settings können einige, meist gleichbleibende Informationen wie Pfadangaben für Import, Feldspektrometer, Versuchsdetails usw. bereits vorgegeben werden. Ebenso enthalten die Settings änderbare Parameter für fixe und dynamische Savitzky-Golay-Glättungen, für Spline-Interpolationen und für die Sensorkorrektur (beispielsweise die Position des Sensorübergangs, welcher gerätespezifisch unterschiedlich sein kann). Wenn eine größere Menge an Spektren importiert wird, können Korrektur und Glättung in einem Batch-Prozess für alle markierten Spektren durchgeführt und damit Zeit gespart werden. Dies ist allerdings nur dann sinnvoll, wenn kaum Ausreißer in den Daten vorhanden sind.

Der Benutzer hat die Möglichkeit, Spektren nach Kriterien wie Datum, Versuch, Feldspektrometer und anderen Metadaten auszuwählen und aus der Datenbank in die Anwendung zu laden. Hier können sämtliche bereits durchgeführte und gespeicherte Verarbeitungsschritte und Dateneingaben wieder rückgängig gemacht und/oder ergänzt werden.

### 4.3 Datenaufbereitung für Auswertungen

Im Navigationsbereich der in Abb. 2 angeführten Spektren sind drei Zeilen in der Spalte VG (= VariantGroup) mit einem Punkt markiert. Sie zeigen an, dass sie Bestandteil einer Variante sind. Die Zusammensetzung von Varianten kann in einem eigenen Fenster verwaltet werden, in dem auch Metadaten auf Variantenebene eingegeben werden. Wird der Reiter im Navigationsfenster auf Variante umgestellt, werden alle zu einer Variante gehörenden Spektren angezeigt und der Benutzer kann sämtliche Funktionen (Mittelwert, Glättung, Indexberechnung, Statistik, usw.) in gleicher Weise wie auf Parzellenebene anwenden und die Ergebnisse in die Datenbank ablegen (vgl. blau markierte Relationen in Abb. 1). Beim Export ist es dem Benutzer möglich, sämtliche Spektren und davon abgeleitete Daten sowohl auf Parzellen- als auch auf Variantenebene auszugeben. Das Format der Exportdateien ist CSV; sie können damit problemlos in Tabellenkalkulations- oder Statistikprogramme für weitere Analysen übernommen werden.

## 5 Fazit & Ausblick

Der SpectroAnalyst wurde mit dem Anspruch entwickelt, ein offenes System zur Verwaltung von Hyperspektraldaten aus verschiedenen Quellen bereitzustellen. In den letzten vier Jahren wurden knapp 28.000 Spektren für mehrere Grünlandexperimente erhoben, verarbeitet und gespeichert. Die systematische Bearbeitung dieser Menge an Spektren erfordert ein Werkzeug, das den Benutzer bei allen Arbeitsschritten bestmöglich unterstützt und den dafür erforderlichen Zeitaufwand auf ein Mindestmaß beschränken kann. Für etwa 200 Spektren (400 bis 1690 nm in 1-nm-Auflösung), das entspricht in etwa jener in der Vegetationsperiode wöchentlich anfallenden Datenmenge, können innerhalb von maximal 30 Minuten alle Metadateneingaben, Korrekturen, die Mittelwertbildung für Parzellen und Varianten sowie die gesamte statistische Auswertung inklusive der Berechnung von ca. 40 verschiedenen Indizes bewerkstelligt und vollständig nachvollziehbar gespeichert werden.

Damit der SpectroAnalyst auch für künftige Forschungsprojekte und Anforderungen gerüstet bleibt, wurde zum einen die Erweiterbarkeit des Datenmodells durch die streng relationale Struktur sichergestellt und zum anderen Schnittstellen zu extern entwickelten Analysewerkzeuge, wie beispielsweise R-Packages oder MATLAB, implementiert.

Der SpectroAnalyst wurde für den Eigenbedarf entwickelt und wird im Rahmen dieser Arbeit als Implementierungsbeispiel vorgestellt. Die Kombination von eigener C#.NET-Implementierung und externen Komponenten wie R, MATLAB, Klassenbibliotheken sowie die Verwendung einer kommerziellen Datenbankmanagementsoftware erfüllt zwar die Anforderungen an ein offenes System und deckt den Bedarf der Entwickler bestmöglich ab, macht aber eine Distribution der Software schwierig. Das primäre Ziel der künftigen Weiterentwicklung des SpectroAnalyst ist die Bereitstellung der Software zur allgemeinen Benutzung.

## 6 Literaturverzeichnis

- AASEN, H., GNYP, M.L., MIAO, Y. & BARETH, G., 2014: Automated Hyperspectral Vegetation Index Retrieval from Multiple Correlation Matrices with HyperCor. *Photogrammetric Engineering & Remote Sensing* **80**(8), 785-795.
- BARAK, P., 1995: Smoothing and Differentiation by an Adaptive-Degree Polynomial Filter. *Analytical Chemistry* **67**(17), 2758-2762.
- BARET, F., WEISS, M., BICHERON, P. & BERTHELOT, B., 2010: Sentinel-2 MSI Products WP1152 Algorithm Theoretical Basis Document for Product Group B, INRA-EMMAH, Avignon, France.
- BIEWER, S., FRICKE, T. & WACHENDORF, M., 2009: Determination of Dry Matter Yield from Legume-Grass Swards by Field Spectroscopy. *Crop Science* **49**(5), 1927-1936.
- BUCHGRABER, K., 1997: Der Einfluss des Schnittzeitpunktes auf den Mineralstoffgehalt des Grünlandfutters. Conference on nutrition of domestic animals „Zadravec-Erjavec-Days“, Radenci, Zivinorejsko-Veterinarski zavod za Pomurje Murska Sobota, 27.-28. Oktober 1997, 57-62.
- BUCHGRABER, K. & PÖTSCH, E.M., 1999: Einfluss der Höhenstufen auf Futterertrag, Futterqualität, Tierbesatz und Milchleistung: Landkalender 2000, 63. Jahrgang, *Leopold Stocker Verlag*, Graz, 133-138.
- DORIGO, W.A., 2008: Retrieving canopy variables by radiative transfer model inversion - a regional approach for imaging spectrometer data. Dissertation, Technische Universität München, Wissenschaftszentrum Weihenstephan, 230 S.
- DORIGO, W.A., BACHMANN, M. & HELDENS, W., 2006: AS Toolbox & Processing of field spectra, User's manual, German Aerospace Center (DLR), Oberpfaffenhofen, 31 S.
- ESA, 2018: Sentinel-2 Spectral Response Functions (S2-SRF), 5 S., [https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset\\_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses](https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses), (15.01.2020).
- GARFAGNOLI, F., MARTELLONI, G., CIAMPALINI, A., INNOCENTI, L. & MORETTI, S., 2013: Two GUIs-based analysis tool for spectroradiometer data pre-processing. *Earth Science Informatics* **6** (4), 227-240.
- HUENI, A. & BIALEK, A., 2017: Cause, Effect, and Correction of Field Spectroradiometer Interchannel Radiometric Steps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **10**(4), 1542-1551.
- JACQUEMOUD, S., VERHOEF, W., BARET, F., BACOUR, C., ZARCO-TEJADA, P.J., ASNER, G.P., FRANÇOIS, C. & USTIN, S.L., 2009: PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment* **113**, 56-66.
- KING, R.L., RUFFIN, C., LAMASTUS, F.E. & SHAW, D.R., 1999: The analysis of hyperspectral data using Savitzky-Golay filtering-practical issues (Part 2). *Geoscience and Remote Sensing Symposium, IGARSS '99 Proceedings, Hamburg, IEEE*, 1, June 28 - July 2, 1999, 398-400.
- KLINGLER, A., SCHAUMBERGER, A., VUOLO, F. & POETSCH, E.M., 2019: Vergleich unterschiedlicher Methoden zur Schätzung von Blattflächenindex in Dauergrünlandbeständen. Dreiländertagung der DGPF, der OVG und der SGPF, Vienna, DGPF, 28, Februar 20 - 22, 301 - 309.

- MACARTHUR, A., 2011: Introduction to Field Spectroscopy, NERC Field Spectroscopy Facility, Monte Bondone, Italy, Course Handbook - COST Action ES0903, 63 S.
- MACARTHUR, A., MACLELLAN, C.J. & MALTHUS, T., 2012: The Fields of View and Directional Response Functions of Two Field Spectroradiometers. *IEEE Transactions on Geoscience and Remote Sensing* **50**(10), 3892-3907.
- MERONI, M. & COLOMBO, R., 2009: 3S: A novel program for field spectroscopy. *Computers & Geosciences* **35**(7), 1491-1496.
- MILTON, E.J., SCHAEPMAN, M.E., ANDERSON, K., KNEUBÜHLER, M. & FOX, N., 2009: Progress in field spectroscopy. *Remote Sensing of Environment* **113**, S92-S109.
- PORTER, J.R., 2000: Scientific Databases. In MICHENER & BRUNT (Eds.): *Ecological Data: Design, Management and Processing, Methods in Ecology*, Blackwell Science Ltd, 48-69.
- PÖTSCH, E.M., 1995: Einfluss der Düngung und Nutzung des Grünlandes auf den Gehalt an Futterinhaltsstoffen. Conference on nutrition of domestic animals „Zadravec-Erjavec-Days“, Radenci, Zivinorejsko-Veterinarski zavod za Pomurje Murska Sobota, 26.-27. Oktober 1995, 85-95.
- PÖTSCH, E.M., 2009: Einflussfaktoren auf Ertrag und Qualität von Grünlandfutter. Fortbildungsveranstaltung "Tierärztliche Bestandsbetreuung von Milchviehbetrieben in Österreich", Gumpenstein, LFZ Raumberg-Gumpenstein, 4.-6. Juni 2009, 5-13.
- PÖTSCH, E.M., RESCH, R., SCHAUMBERGER, A. & BUCHGRABER, K., 2006: Effect of different management systems on quality parameters of forage from mountainous grassland. 21st EGF General Meeting "Sustainable Grassland productivity", Badajoz, Spain, *Grassland Science in Europe*, 11, April 3-6, 2006, 484-486.
- PSOMAS, A., KNEUBÜHLER, M., HUBER, S., ITTEN, K.I. & ZIMMERMANN, N.E., 2011: Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats. *International Journal of Remote Sensing* **32**(24), 9007-9031.
- PUNALEKAR, S.M., VERHOEF, A., QUAIFFE, T.L., HUMPHRIES, D., BERMINGHAM, L. & REYNOLDS, C.K., 2018: Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. *Remote Sensing of Environment* **218**, 207-220.
- RASAI AH, B., JONES, S., BELLMAN, C., MALTHUS, T. & HUENI, A., 2015: Assessing Field Spectroscopy Metadata Quality. *Remote Sensing* **7**(4), 4499-4526.
- RUFFIN, C. & KING, R.L., 1999: The analysis of hyperspectral data using Savitzky-Golay filtering-theoretical basis (Part 1). *Geoscience and Remote Sensing Symposium, IGARSS '99 Proceedings, Hamburg, IEEE*, 2, June 28 - July 2, 1999, 756-758.
- SAVITZKY, A. & GOLAY, M.J.E., 1964: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**(8), 1627-1639.
- STEVENS, A. & LOPEZ, L.R., 2014: An introduction to the prospectr package, 22 S., <http://antoinestevens.github.io/prospectr/>, (15.01.2020).
- THENKABAIL, P.S., GUMMA, M.K., TELUGUNTLA, P. & MOHAMMED, I.A., 2014: Hyperspectral Remote Sensing of Vegetation and Agricultural Crops. *Photogrammetric Engineering & Remote Sensing* **80**(8), 697-709.
- VAN EVERT, F.K., SPAANS, E.J.A., KRIEGER, S.D., CARLIS, J.V. & BAKER, J.M., 1999a: A Database for Agroecological Research Data: I. Data Model. *Agronomy Journal* **91**(1), 54-62.

- VAN EVERT, F.K., SPAANS, E.J.A., KRIEGER, S.D., CARLIS, J.V. & BAKER, J.M., 1999b: A Database for Agroecological Research Data: II. A Relational Implementation. *Agronomy Journal* **91**(1), 62-71.
- WIGGINS, A., BONNEY, R., GRAHAM, E., HENDERSON, S., KELLING, S., LEBUHN, G., LITTAUER, R., LOTT, K., MICHENER, W., NEWMAN, G., RUSSELL, E., STEVENSON, R. & WELTZIN, J., 2013: *Data Management Guide for Public Participation in Scientific Research*. GROUP, University of New Mexico, Albuquerque, USA, 15 S.