Worksheet 02 · Sentinel-Daten beschaffen und eine NDVI-Reihe berechnen

Datensouveränität - Untersuchungsfläche, ausgewählte Aufnahmen und reproduzierbare Ergebnisse

Auftrag

Wie unterscheidet sich das sommerliche Vegetationssignal im gewählten Burgwald-Ausschnitt zwischen 2018 und 2022?

Sie beschaffen die Satellitendaten selbst, berechnen eine NDVI-Reihe und prüfen, ob die Aufnahmen für einen Vergleich geeignet sind. Sie treffen die Entscheidungen über Untersuchungsfläche und Bildauswahl. Die Skripte stellen den technischen Ablauf bereit.

Aus Worksheet 01 kennen Sie Projektordner, relative Pfade und blockweises Ausführen. Das optionale Mikroklima-Warmup bietet zusätzliche Übung mit Ergebnisrastern. Arbeiten Sie im vorhandenen Projekt weiter. Die bisherige Gebietsgrenze hilft Ihnen bei der räumlichen Orientierung.

Am Ende liegen fünf datierte Sentinel-Aufnahmen, die daraus berechneten NDVI-Raster sowie eine Verlaufsgrafik mit Angaben zur Datenabdeckung vor. Die Reihe besteht aus **einer ausgewählten Sommerbeobachtung pro Jahr**. Sie bildet weder die vollständige Vegetationsperiode noch einen kontinuierlichen Verlauf ab.

Material und Vorbereitung

Installation: 00_install_packages.R.

Die technische Grundlage ist der CDSE-Abruf aus dem aktuellen Kursrepository. Die Lehrfassung übernimmt STAC-Suche und CDSE-Abruf, verwendet Ihr eigenes Gebiet und berechnet den NDVI anschließend lokal. Der große Projekt-Setup des Repositories wird dafür nicht benötigt.

Laden Sie beide Skripte in den Ordner `scripts/` Ihres bestehenden RStudio-Projekts:

- 02_sentinel_abruf.R
- 03_ndvi_reihe.R

Benötigt werden QGIS, RStudio und die R-Pakete `sf`, `terra`, `rstac`, `dplyr` und `CDSE`. Installieren Sie fehlende Pakete vor der Sitzung über die Konsole:

```
install.packages(c("sf", "terra", "rstac", "dplyr", "CDSE"))
```

Für den Rasterabruf richten Sie einmalig Ihren persönlichen CDSE-Zugang und die Anmeldung des R-Skripts ein. Die einzelnen Schritte stehen im folgenden Exkurs.

Exkurs · CDSE-Konto, OAuth-Client und Zugang in R einrichten

Zugangsdaten

Ihr **CDSE-Konto** dient zur Anmeldung im Browser. Ein **OAuth-Client** ermöglicht dem Skript den Zugriff: Er besitzt eine **Client-ID** und ein **Client-Secret**. Diese sind nicht Ihre E-Mail-Adresse und Ihr Kontopasswort. Aus den beiden Werten kann das R-Paket einen zeitlich begrenzten Zugriffstoken beziehen.

1. Konto anlegen und anmelden

Öffnen Sie Copernicus Data Space, registrieren Sie sich und schließen Sie die Bestätigung Ihrer E-Mail-Adresse ab. Falls Sie bereits ein Konto besitzen, verwenden Sie dieses. Öffnen Sie anschließend das Sentinel-Hub-Dashboard für CDSE und melden Sie sich mit diesem Konto an.

2. OAuth-Client erstellen

1. Öffnen Sie im Dashboard **User Settings**.
2. Suchen Sie den Bereich **OAuth clients** und klicken Sie auf **Create**.
3. Vergeben Sie einen Namen, beispielsweise `GI-Kurs-Sentinel`.
4. Wählen Sie ein Ablaufdatum nach dem vorgesehenen Kursende.
5. Lassen Sie die Option für eine **Single-page application (SPA)** ausgeschaltet: Wir verwenden ein lokales R-Skript.
6. Bestätigen Sie mit **Create**.
7. Kopieren Sie **Client ID** und **Client Secret** direkt in die im nächsten Schritt beschriebene Datei. Das Secret ist nach dem Schließen des Dialogs nicht erneut abrufbar. Falls es verloren geht, erstellen Sie einen neuen Client.

3. Werte in R hinterlegen

Öffnen Sie in der RStudio-Konsole Ihre persönliche Umgebungsdatei:

```
file.edit(path.expand("~/.Renviron"))
```

Ergänzen Sie die folgenden beiden Zeilen. Ersetzen Sie die Platzhalter durch die kopierten Werte; vorhandene andere Einträge bleiben erhalten.

```
CDSE_ID=Ihre_Client_ID
CDSE_SECRET=Ihr_Client_Secret
```

Diese Zeilen gehören **in die Datei**, nicht in die R-Konsole. Speichern Sie die Datei als `.Renviron`, ohne zusätzliche Endung `.txt`. Fügen Sie nach der letzten Zeile einen Zeilenumbruch ein. Starten Sie R anschließend über **Session → Restart R** neu.

4. Prüfen, ob R die Werte eingelesen hat

```
nzchar(Sys.getenv("CDSE_ID"))
nzchar(Sys.getenv("CDSE_SECRET"))
```

Beide Ausgaben müssen `TRUE` lauten. Damit ist zunächst nur geprüft, ob Werte vorhanden sind. Ihre Gültigkeit wird beim authentifizierten Zugriff geprüft.
Falls `FALSE` erscheint, kontrollieren Sie Dateiname, Schreibweise der Variablen und Speicherung. Sie können die persönliche Datei auch ausdrücklich einlesen und die Prüfung wiederholen:

```
readRenviron(path.expand("~/.Renviron"))
```

5. Verbindung im Abrufskript verwenden

Der folgende Abschnitt ist bereits in `02_sentinel_abruf.R` enthalten:

```
client <- CDSE::GetOAuthClient(
  id = Sys.getenv("CDSE_ID"),
  secret = Sys.getenv("CDSE_SECRET")
)
```

Das Objekt `client` wird anschließend an `CDSE::GetImage()` übergeben. Führen Sie zunächst den vorgesehenen Einzelabruf aus. Bei einer Meldung wie `invalid_client` prüfen Sie, ob ID und Secret zum selben aktiven Client gehören und vollständig kopiert wurden.
Vertraulich behandeln: Geben Sie Secret, Zugriffstoken und persönliche `.Renviron` nicht weiter. Sie gehören nicht in das Projektarchiv, ein Git-Repository, Screenshots oder KI-Prompts.
Die Menüfolge und die Hinweise zum Secret sind in der offiziellen CDSE-Anleitung zur OAuth-Registrierung beschrieben.

1 · Eine eigene Untersuchungsfläche festlegen

Öffnen Sie die Burgwald-Gebietsgrenze aus Worksheet 01 in QGIS. Wählen Sie anhand einer Hintergrundkarte einen kleinen Ausschnitt mit Wald und angrenzendem Offenland. In Worksheet 03 vergleichen Sie darin eine Wald-, eine Offenland- und eine auffällige Teilfläche. Ein Gebiet mit ungefähr 1–2 km Kantenlänge reicht; die maximale Kantenlänge dieser Übung beträgt 5 km.

> **i** Legen Sie einen neuen GeoPackage-Layer an:
>
> Erstellen Sie die Untersuchungsfläche in QGIS:
>
> 1. Öffnen Sie Layer → Layer erstellen → Neuer GeoPackage-Layer.
> 2. Wählen Sie als Datei data/aoi.gpkg in Ihrem RStudio-Projekt, als Tabellennamen aoi, als Geometrietyp Polygon und als Koordinatensystem EPSG:32632 – WGS 84 / UTM zone 32N.
> 3. Bestätigen Sie mit OK und markieren Sie den neuen Layer.
> 4. Aktivieren Sie den Bearbeitungsmodus über das Stiftsymbol. Wählen Sie Polygonobjekt hinzufügen.
> 5. Setzen Sie mit Linksklicks die Eckpunkte Ihrer Untersuchungsfläche. Ein Rechtsklick schließt die Erfassung ab; bestätigen Sie den Attributdialog mit OK.
> 6. Speichern Sie die Layeränderungen und beenden Sie den Bearbeitungsmodus.
>
> Zeichnen Sie genau ein Polygon. Diese Fläche wird für alle Aufnahmetermine verwendet.

- Datei: `data/aoi.gpkg` im RStudio-Projekt;
- Geometrietyp: Polygon;
- Koordinatensystem: EPSG:32632;
- genau ein Polygon als Untersuchungsfläche.

Zeichnen Sie das Polygon, speichern Sie die Bearbeitung und beenden Sie den Bearbeitungsmodus. Verwenden Sie über die gesamte Reihe dieselbe Fläche.

Dokumentation: Halten Sie in `README.md` Lage, Größe und erkennbare Landbedeckung der gewählten Fläche fest.

2 · Aufnahmen suchen und eine Auswahl treffen

Öffnen Sie `02_sentinel_abruf.R`. Ergänzen Sie den Skriptkopf. Führen Sie zunächst nur die Abschnitte 1–3 aus.

Die Einstellungen lauten:

```
jahre <- 2018:2022
monate <- 6:8
anzahl_abrufe <- 1
```

Die Suche liefert Metadaten der Sentinel-2-L2A-Aufnahmen, die Ihr Gebiet schneiden. Sie erzeugt:

Datei	Bedeutung
`data/stac_treffer.rds`	Gespeicherte Antwort des Katalogs.
`output/aufnahmekandidaten.csv`	Verfügbare Sommertage mit Szenenzahl und mittlerer Szenenbewölkung.
`data/auswahl.csv`	Ihr Arbeitsstand: ein ausgewählter Aufnahmetag je Jahr.

Der erste Vorschlag wählt pro Jahr den Tag mit der geringsten mittleren Szenenbewölkung. **Diese Prozentangabe bezieht sich auf die gefundenen Szenen, nicht auf Ihre kleine Untersuchungsfläche.** Ein solcher Tag kann über Ihrem Gebiet trotzdem bewölkt sein.

Vergleichen Sie die fünf vorgeschlagenen Aufnahmetage nach ihrer Lage innerhalb des Sommers. Sehen Sie die Aufnahmen bei Bedarf im Copernicus Browser an. Wenn Sie einen anderen Tag auswählen, übernehmen Sie die vollständige zugehörige Zeile aus der Kandidatentabelle in `data/auswahl.csv`.

Die Auswahl-Datei wird beim erneuten Ausführen nicht überschrieben. Für eine neue Suche mit geändertem Gebiet oder Zeitraum archivieren Sie die alte Auswahl und lassen eine neue erzeugen.

3 · Einen Abruf prüfen, dann die Reihe laden

Führen Sie im Abrufskript die Abschnitte 4–6 aus. Mit `anzahl_abrufe <- 1` wird zunächst nur der erste ausgewählte Tag geladen.

Öffnen Sie die erzeugte TIFF-Datei aus `data/sentinel/` in QGIS. Die Bänder sind fest angeordnet:

Band	Inhalt	Verwendung
1–3	B02, B03, B04	Blau, Grün, Rot; Echtfarbdarstellung mit Rot = Band 3, Grün = Band 2, Blau = Band 1.

Band	Inhalt	Verwendung
4	B08	Nahes Infrarot; zusammen mit B04 für NDVI.
5	SCL	Szenenklassifikation zur Qualitätsmaskierung.
6	dataMask	Kennzeichnet verfügbare Bilddaten.

Kontrollieren Sie Lage, Abdeckung, 10-m-Rasterweite sowie erkennbare Wolken und Schatten. Alle Termine liegen auf demselben 10-m-UTM-Gitter. Reflexionen werden bilinear, SCL und dataMask mit dem nächsten Nachbarn übertragen. Die SCL besitzt dadurch keine zusätzliche ursprüngliche Detailinformation. Für die AOI-Auswertung zählen Zellen, deren Mittelpunkt innerhalb des Polygons liegt.

Wenn der erste Abruf stimmt, setzen Sie `anzahl_abrufe <- Inf` und führen Sie Abschnitt 6 erneut aus. Vorhandene Tagesdateien werden nur bei passender AOI und passenden Abrufeinstellungen übersprungen. Zu jedem TIFF wird ein Abrufmanifest `.tif.rds` gespeichert. Ältere TIFFs ohne Manifest oder mit geänderten Einstellungen archivieren Sie zusammen mit dem bisherigen Ordner `data/sentinel/` und laden die ausgewählten Tage erneut. Kontrollieren Sie auch die übrigen Aufnahmen.

Zwischenergebnis: Für jedes ausgewählte Datum existiert eine lokale TIFF-Datei. Ungeeignete Tage ersetzen Sie in `data/auswahl.csv` durch einen anderen Kandidaten. Führen Sie danach Abschnitt 6 des Abrufskripts erneut aus; er liest die Auswahl bei jedem Start neu ein. Alte Dateien dürfen zur Dokumentation liegen bleiben; die Auswertung verwendet ausschließlich die aktuelle Auswahl-Datei.

Der Abruf verwendet jeweils einen Tag. Überlappende Kacheln können dabei zu einem Bild mosaikiert werden; eine einzelne Produkt-ID wird damit nicht festgeschrieben.

4 · Aus den Aufnahmen eine NDVI-Reihe berechnen

Öffnen Sie `03_ndvi_reihe.R`. Führen Sie die Abschnitte 1 und 2 aus. Das Skript prüft, ob alle ausgewählten Aufnahmen vorhanden sind und dasselbe Rastergitter besitzen.

Die Berechnung lautet:

$$NDVI = \frac{B08 - B04}{B08 + B04}$$

Vor der Auswertung werden ungültige Daten und bestimmte SCL-Klassen ausgeschlossen. In der Vorlage bleiben Vegetation, unbewachsene Flächen, Wasser und unklassifizierte Pixel erhalten; Wolken, Wolkenschatten, Schnee und weitere ausgeschlossene Klassen werden zu `NA`.

Die Klassen stehen in `data/qualitaetsmaske.csv`; Skript 03 legt die Datei beim ersten Lauf an. Skript 05 liest dieselbe Datei. Ändern Sie die Maske dort und führen Sie danach Skript 03, 05 und die weiteren Auswertungen erneut aus. Die visuelle Bildprüfung bleibt erforderlich.

```
read.csv("data/qualitaetsmaske.csv")
# Vorgabe in der Spalte scl: 4, 5, 6, 7
```

Das Skript berechnet einen NDVI pro Datum, begrenzt ihn auf Ihr Polygon und schreibt die Ergebnisse nach `output/ndvi/`. Zusätzlich entstehen `output/ndvi_reihe.tif` und `output/ndvi_banddaten.csv`. Die CSV ordnet jedem Band sein tatsächliches Aufnahmedatum zu.

Prüfung: Öffnen Sie zwei NDVI-Raster in QGIS. Verwenden Sie dieselbe Farbskala von -1 bis 1. Prüfen Sie auffällige Stellen im ursprünglichen RGB-Bild. Fehlende Werte sind keine NDVI-Werte von null.

5 · Einen belastbaren Verlauf darstellen

Führen Sie die Abschnitte 3 und 4 aus. Das Skript erzeugt eine Tabelle und eine Grafik mit:

- mittlerem NDVI über die an jedem Termin gültigen Pixel;
- Anteil gültiger Pixel innerhalb Ihrer Untersuchungsfläche;
- einem zweiten Verlauf ausschließlich für die Pixel, die an **allen** ausgewählten Terminen gültig sind.

Der zweite Verlauf macht sichtbar, ob wechselnde räumliche Abdeckung den Vergleich beeinflusst. Bleiben nur sehr wenige gemeinsame Pixel, ist auch dieser Vergleich räumlich schwach gestützt. Prüfen Sie dann die Bildauswahl. Die Grafik verbindet einzelne Beobachtungstage; über den Verlauf zwischen ihnen liegen keine Aussagen vor.

Ergebnis: Nennen Sie das Datumspaar mit dem größten Unterschied und geben Sie die NDVI-Differenz für beide räumlichen Bezugsflächen an. Legen Sie einen zusätzlichen Aufnahmetermin fest, mit dem Sie Vegetationsveränderung, sommerliche Entwicklung oder Bildstörungen unterscheiden können.

6 · Eine Entscheidung überprüfen

Wählen Sie für ein Jahr einen zweiten geeigneten Sommertag aus der Kandidatentabelle. Sichern Sie die bisherigen Ergebnisdateien in einem Unterordner `output/vergleich_1/`. Ersetzen Sie den entsprechenden Eintrag in `data/auswahl.csv`, laden Sie den neuen Tag und berechnen Sie die Reihe erneut.

Vergleichen Sie den betroffenen NDVI-Wert, die Datenabdeckung und Ihre ursprüngliche Interpretation. Dokumentieren Sie, welche Tagesauswahl Sie für die weitere Arbeit verwenden und warum.

Bei technischen Problemen können Sie Dokumentation oder KI nutzen. Geben Sie den betroffenen Abschnitt und die Fehlermeldung an, verlangen Sie eine begrenzte Korrektur und überprüfen Sie danach Datum, Rastergitter und Werte. Halten Sie wesentliche Änderungen im Skript fest.

Ergebnis der Sitzung

Ihr Projekt enthält die selbst gewählte Fläche, die Aufnahmekandidaten, Ihre dokumentierte Auswahl, die abgerufenen Raster und die berechnete NDVI-Reihe. Ergänzen Sie in `README.md`:

- Fragestellung und Begründung der Untersuchungsfläche;
- gewählte Aufnahmetage, Auswahlgründe und Abrufdatum;
- verwendete Qualitätsmaske;
- beobachteten Unterschied und Ergebnis des zweiten Tagesvergleichs;
- Ausführungsreihenfolge der beiden Skripte.

Abgabe: eine Verlaufsgrafik, die zugehörige CSV und fünf bis acht Sätze zur Aussagekraft der Reihe. Eine beobachtete Abnahme wird zunächst als verändertes Vegetationssignal beschrieben. Ihre Ursache muss anhand weiterer Beobachtungen geprüft werden.

In Worksheet 03 vergleichen Sie NDVI und kNDVI auf Wald-, Offenland- und auffälligen Vergleichsflächen. Die Hausaufgabe zu Pixelverfügbarkeit und Zeitreihe bleibt als Zusatz zu dieser Sitzung verfügbar; ihre Ergebnisse werden in Worksheet 10 weiterverwendet.