Session 3

Spectral compositing of Sentinel-2 data using SCMaP as input for soil parameter mapping

Uta Heiden¹, Pablo d'Angelo¹, Paul Karlshöfer¹, Adrian Sanz Diaz², Maria Julia Yagüe Ballester² and the whole WorldSoils Team³

¹ DLR
² GMV
³ WorldSoils Team (see below)

ESA Symposium on Earth Observation for Soil Protection and Restoration
Introduction

WorldSoils processing system overview

Pipeline

- **COMPOSITING** (PROC_SCMAP)
- **BARE PREDICTION** (PROC_BARE)
- **MOSAICKING** (PROC_MOS)
- **AGGREGATING** (PROC_AGGR)
- **PERM. VEG. PREDICTION** (PROC_VEG)

SCMaP* Methodology

* Soil Composite Mapping Processor
WorldSoils processing system overview

SCMaP* Methodology

- All Sentinel-2 images in L2A format (Sen2Cor) from 2018 – 2022
- Spectral Index based (e.g. Diek et al. 2017, Rogge et al. 2018, Demattê et al., 2018)
- Used index: PV+IR2 (Heiden et al. 2022, Möller, M. et al. 2022, Dvorakova, K., et al., 2023)

$PV+IR2 = \frac{B8 - B4}{B8 + B4} + \frac{B8 - B12}{B8 + B12}$
Introduction

WorldSoils processing system overview

Pipeline

COMPOSITING
(PROC_SCMAP)

BARE
PREDICTION
(PROC_BARE)

PERM. VEG.
PREDICTION
(PROC_VEG)

MOSAICKING
(PROC_MOS)

AGGREGATING
(PROC_AGGR)

SCMaP* Methodology

Method evaluation

Regionalised Thresholds

Threshold Accuracy

Cloud / Haze

* Soil Composite Mapping Processor
SCMaP Methodology

Flowchart

Input data
- Sentinel-2 Reflectances + Cloud Masks
- ESA WorldCover
- HISET
- European Threshold Database

Threshold definition

SCMaP Compositing
- Index calc. + thresholding
- Cloud and haze filter
- Averaging
- Landcover Masking
- Averaging
- Statistics

Output products
- Mean + STD reflectance composite
- MEAN + STD Bare surface reflectance composite (SRC)
- Bare surface pixel count
- Valid pixel count
- Bare surface frequency
- Bare surface mask

Method Evaluation
Threshold Validation
Criteria for large scale areas

- Generic and globally applicable
- Allows for regionalized threshold derivation
- Accounts especially for spectral similarity between bare soils (crops) and non-photosynthetic vegetation (grassland)
- Spectral index independent
- Fully automated
Regionalised Thresholds

Threshold Definition Concept - HISET

1. Index calculation -> Index minimum composite
2. Selection of specific LC classes (e.g. WorldCover - 10m)
3. Temporal behaviour of LC classes (normalised histogram)
4. Threshold definition

**HISET**

Histogram SEparation Threshold
Temporal minimum of a vegetation index

- **Cropland:** Likely bare soil
- **Grassland:** Likely NPV

Best separation of both classes → Index threshold
Challenges – Land Cover/Use Map

1. Areas with limited or missing pixels of the two LC types interpolation and extrapolation

2. Refinement of LC cropland
   • LC class includes spectral mixtures (border pixels)
   • LC definition – pasture land not actively managed, do not show bare soils
   • Assessment of activity of surfaces

   Temporal Variability Index: \[
   TVI = \sum_{i=0}^{N} \left[ \frac{M_{i+1} - M_i}{d_{i+1} - d_i} \right],
   \]
   where \( M_i \) is the \( i \)-th of \( N \) bimonthly minimum index composites, that is centered at date \( d_i \).

   • Activity map is used to clean up the crop layer
- PV+IR2 thresholds range between – 0.1 and 0.6
- Correlated with bioclimatic zones

Karlshöfer et al., in preparation
Multiple steps:

1. Selection of single scenes with < 80 % cloud cover
2. Masking of clouds, haze, snow etc. using Scene Classification Layer (SCL 4/5/6) of Sen2Cor processing
3. Bare soil specific cloud and haze masking

noticed remaining clouds and haze

Masked areas with cloud probability of 60 % (yellow)
NIR – SWIR difference (clouds)

Distinct difference in NIR and SWIR behavior between clouds and almost all soils

- Soils: B11 > B8, Clouds: B11 < B8
- \( \frac{(B11-B8A)}{(B11+B8A)} > 0.02 \)
- Only very few misclassifications: 0.1% of all LUCAS spectra (some nut tree orchards in southern Spain)

“Blue outliers” (haze / thin clouds)

Atmospheric effect strongest in blue band

- detect remaining haze and thin cloud contamination based on higher blue reflectance
- Local statistics based outlier filter:

\[
 b \leq \text{median}(B) + 3\sigma \\
 \sigma = 1.48 \times \text{median}(|B - \text{median}(B)|)
\]
Iceland – without additional cloud filtering
Iceland – with additional cloud filtering
Method Evaluation

Evaluating the soil reflectance composite product

- How to evaluate the used compositing methodology:
  - Selection of indices
  - Thresholds
  - Universal versus regional approach?
- What is the reference for the soil composite spectra?
- Can we evaluate for large areas (e.g. Europe) instead of small test areas?
At each LUCAS points, **ideal thresholds** can be computed, that minimize angular distance.

Performance: 
\[
\min_t \frac{1}{N} \sum_{i=0}^{N} SAM(l_i, C_i(t)), \text{ for } N \text{ LUCAS points and the Composite } C \text{ based on thresholds } t.
\]

**PV+IR2 outperforms** established indices.

Significant range [-0.05, 0.4] of ideal thresholds and local patterns indicate that a **regionalized** thresholds is crucial.

- Also evident by the bad performance of universal thresholds.

\[\rightarrow\] **Regional PV+IR2 good choice for an index.**

* Universal thresholds taken from literature: \(-0.25 < \text{NDVI} < 0.25, -0.1 < \text{NBR2} < 0.3, \text{VNSIR} < 0.9\)
Summary:

• SCMaP – fully automated processor for enhanced image products for soil mapping
• PV+IR2 suitable to select bare soil surfaces by reducing NPV influence to a minimum
• Technique for regionalised threshold definition developed, tested and evaluated
• Tested at continental scale (Europe)
• Approach evaluated against other soil compositing strategies using LUCAS spectrum as reference

Outlook:

• Reduce dependencies (Land cover map, thresholds)
• Produce pixel-based spectral uncertainty maps
Many thanks for your attention!
uta.heiden@dlr.de

ESA Symposium on Earth Observation for Soil Protection and Restoration