SPATIAL SCALING CHALLENGE
DESCRIPTION OF THE DATASETS

[1] Background spatial and scaling challenge study

The mechanisms driving plant ecophysiological status (e.g., rapid adaptation to stress conditions and photosynthetic rate) occur at the leaf level. However, we need to rely on top-of-canopy remote sensing observations to monitor vegetation in a fast and continuous mode, which poses a scale problem. The incoming radiation reaching the leaf, which is reflected and emitted, carries information on the ecophysiological status of the leaf. However, the light suffers a distortion through the plant's vertical profile, which makes it difficult to relate the leaf level physiological process with the top-of-canopy spectral signals.

The scientific community has developed different down-scaling methods, traditionally focusing on studying biophysical properties. However, recent advances in remote sensing of hyperspectral, fluorescence, and thermal signals are bringing more attention to physiological variables. These have weaker and more complex links with the spectral signals and can be highly dynamic, posing new retrieval, scaling, and interpretation problems.

In this regard, in the frame of the "Optical synergies for spatiotemporal SENsing of Scalable ECOphysiological traits" (SENSECO) COST Action (CA17134), we would like to challenge the scientific community to down-scale relevant biophysical and physiological variables from a simulated dataset of remote reflectance factors, chlorophyll sun-induced fluorescence, thermal radiance imagery, and field measurements. In particular, the objectives of this exercise are:

(1) to down-scale or retrieve relevant biophysical and plant physiological variables, e.g., leaf chlorophyll content (C_{ab}), leaf area index (*LAI*), maximal carboxylation rate ($V_{max,25}$), and non-photochemical quenching (*NPQ*) from hyperspectral imaging spectroscopy data.

(2) based on the retrieved variables, we would like to encourage the participants to evaluate the remote sensing capability to diagnose and discriminate between different plant ecophysiological states and translate it into low or high vegetation efficiency or stress.

Finally, based on the output of points (1) and (2), we aim,

(3) to gather the community knowledge and experience to provide clear, evidence-based guidelines of scaling protocols in terms of measurement protocols and data treatment; to ensure consistent, reproducible, and comparable results between different models.

The maps of variable estimates produced by the participants will be evaluated with the simulated values not known by the participants. Participants will be invited to contribute and coauthor this SENSECO manuscript summarizing the retrieval uncertainties, scaling methods, and lessons learned regarding handling remote sensing and field data and metadata. Individual results will be anonymous in the final publication.

[2] Exercise instructions

1. In case you decide to participate in the spatial scaling challenge, send an email to scalingchallenge@gmail.com with the following information:
   - Subject: Joining spatial scaling challenge
   - Name:
   - Institution:

   Example:
   - *Subject: Joining spatial scaling challenge*
   - *Name: Moana Brave*
   - *Institution: Ocean University*

2. Open the SSC_SENSECOWG1 folder. There you will find the following information:
   - **1_SSC_data**: this folder contains both the airborne imagery and field data set. Read section 5 of this document for further details.
   - **2_SSC_templates**: this folder contains the templates to be filled in by the exercise participant with the description of the approach/model you used. Read section *[5] Data set description* of this document for further information.
   - **3_SSC_results**: in this folder, you (or rather the scripts provided) should locate the outputs of your approach/model.
   - **SSC_Script.m // SSC_Script.r //SSC_Script.py**: matlab (.m), R (.r), and phyton (.py) scripts to open/import the provided data set as well as to save in a standardized format the results. The scripts make data ready for analyses; their use is strongly recommended. Read section *[5] Data set description* of this document for further information.

IMPORTANT!! Do not modify the SSC_SENSECOWG1 folder organization

3. Open the SSC_Script.* file and read the commented lines carefully. Using this script, you will be able to:
   - **[step 2 on SSC_Script]:** Open/Import the airborne imagery and field data.
   - **[step 3 on SSC_Script]:** Link field and imagery data.
   - **[step 4-7 on SSC_Script]:** Save the results of your model output in the requested format.

4. Implement your approach/model (empirical, semi-empirical, statistical, physically based, …) to down-scale and retrieve the requested biophysical and plant physiological variables (e.g., *LAI*, C_{ab}, $V_{cmax,25}$, and *NPQ*).

5. Provide your diagnosis of the actual vegetation stauts and describe your approach/model. Use any of the MS Word (.doc or .docx) or OpenOffice Writer (.odt) templates: /2_SSC_templates/SSC_report.xxx. Save the completed templates in the "**3_SSC_results"** folder.

6. On the SSC_Script, steps 5-7, fill up the requested information.

7. On the SSC_Script, execute code between steps 5-7. This will generate a "surname_name_4sub.zip" file in the "**3_SSC_results"** folder with the results of your exercise to be submitted.

8. Send an email to scalingchallenge@gmail.com with the following information:
   - Subject: Results spatial scaling challenge
   - Name:

- Institution:
- Attach: generated "surname_name.zip" file

Example:
- *Subject: Results spatial scaling challenge*
- *Name: Moana Brave*
- *Institution: Ocean University*
- *Attach: generated " BraveMoana.zip" file*

9. Deadline to submit your results: 31st of May 2022, included

[3] Questions

In case you have any questions, please, send an email to scalingchallenge@gmail.com with the following information:

- Subject: Question spatial scaling challenge
- Name:
- Institution:
- Question:

Example:
- *Subject: Question spatial scaling challenge*
- *Name: Moana Brave*
- *Institution: Ocean University*
- *Question: xxxx*

[4] Exercise expected outputs

Mandatory results:
- *LAI*, C_{ab}, $V_{cmax,25}$, and *NPQ* maps.
- Diagnosis of vegetation status and methods description fusing the MS Word (.docx or .doc) or OpenOffice Writer (.odt) provided templates.

Optional (BONUS!):
- Provide an uncertainty map for each of the requested parameters.
- Based on your retrieved results, discriminate between different plant ecophysiological states and translate it into a low-or-high vegetation efficiency/stress map. To simplify the comparison between models, please, normalize the results of your "vegetation efficiency/stress model" between [0-1].

Important information:
- Maps format: NetCDF files using the script given in the exercise
- Output map dimensions: same as the dimensions of the provided input data
- Units of the outputs: *LAI* [m m^{-2}], C_{ab} [µg cm^{-2}], $V_{cmax,25}$ [µmol CO2 m^{-2} s^{-1}], *NPQ* [-]

[5] Data set description

<table>
<tr><td>Study area and plant types</td><td>
The study focuses on two experimental crop fields that are approximately 100 m long and 50 m wide and align in the North-South direction (Fig. 1). The fields contain Wheat (Triticum aestivum, on the West) and Maize (Zea mays, on the East). The experimental farm was located in southwestern Spain (39.95279º, -5.77705º).
At the time of the flight, campaign wheat was about to start booting when leaves had been almost erectophile, and some began to lay more horizontal. Maize was in the growth stage V14, still vegetative, before tasseling.

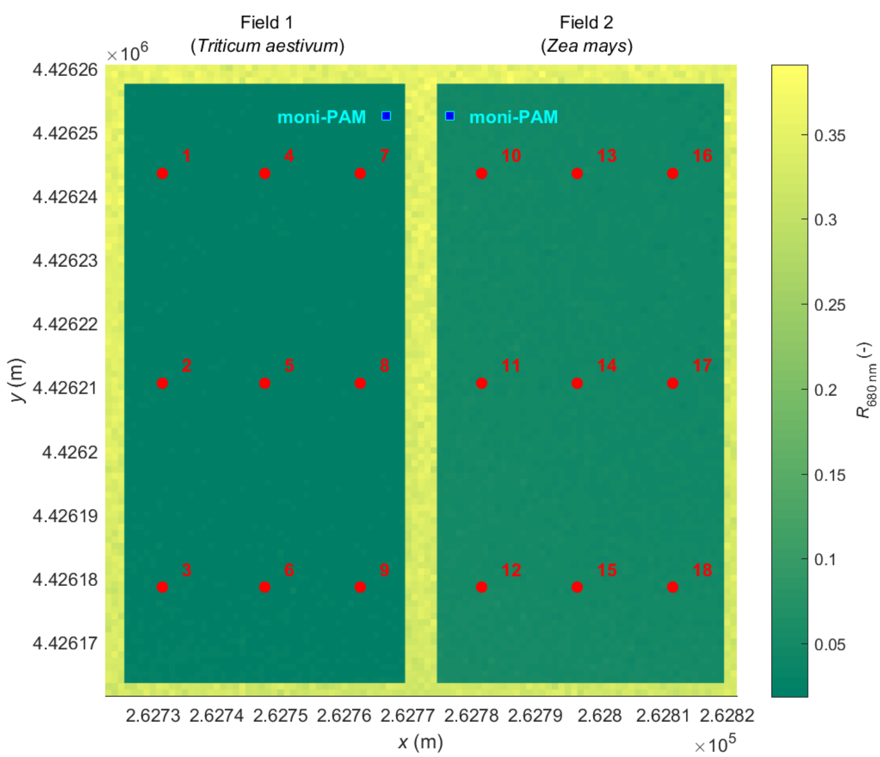

Figure 1. Red reflectance map of the study area presenting the field sampling scheme
</td></tr>
<tr><td>Imagery</td><td>
Airborne imagery acquisition occurred on 01-Jul-2019 at 11.45 UTC (13.45 local time). The imagery was acquired simultaneously by three airborne sensors with different spectral and spatial characteristics (VSWIR for reflectance factors, very high spectral resolution for fluorescence retrieval, and a thermal hyperspectral imager). The plane flew over the central area of the fields 350 m above the surface, close to solar azimuth (SAA = 149º). Since the study area was small, a single overpass was enough to cover it completely. The campaign day was sunny and warm, incoming shortwave incoming radiation reached more than 1000 W m^{-2} around midday, and the maximum air temperature was 35.6 ºC.

The imagery generated for this study is described below:
</td></tr>
</table>

<table>
<tr><td></td><td>

Hemispherical Directional Reflectance Factor (*HDRF*)

- Spectral features: 405.36 to 2386.62 nm with 335 bands of 3.65 nm spectral resolution (full width at half maximum, *FWHM*) in the VNIR (< 970 nm) spectral range and 250 bands of 10.55 nm spectral resolution in the NIR/SWIR spectral ranges (> 970 nm). Noisy data from the atmospheric water absorption bands around 1450 and 1950 nm were removed, leaving a spectral dataset of 526 bands.
- Spatial resolution: 1 m.
- Units: [-]

Sun-induced chlorophyll fluorescence radiance (*F*)

- Spectral features: Fluorescence radiance retrievals at 687 and 760 nm.
- Spatial resolution: 4 m, gridded to 1 m without modifying its values to facilitate integration with the rest of the imagery.
- Units: [mW m^{-2} sr^{-1} nm^{-1}]

Land surface temperature (*LST*)

- Spectral features: LST retrieval from a hyperspectral thermal imager assuming emissivities of 0.98 and 0.94 for vegetation and bare soil, respectively. Any pixel with a normalized difference vegetation index (*NDVI*) lower than 0.2 was classified as bare soil
- Spatial resolution: 2 m, gridded to 1 m without modifying its values to facilitate integration with the rest of the imagery.
- Units: [K]

Each imagery netCDF file includes two additional layers with the view zenith angle (VZA) and the phase angle between sun and view azimuth angle (SVAA).

</td></tr>
<tr><td>

Field ancillary data

</td><td>

Simultaneously to the flight campaign, a ground team took measurements of the following biophysical and physiological variables. All the variables were measured as much as possible in the same 1 x 1 m plots, 9 points per crop. Sampling schemes (order and, therefore, times) were separately randomized for each plot.

Leaf Chlorophyll Content (C_{ab})

- Methods: 10 top-of-the-canopy leaves were measured with a chlorophyll meter at each sampling plot. Laboratory calibrated models were used to convert the chlorophyll meter values to C_{ab}.
- Estimated uncertainties: the relative uncertainty of each leaf's C_{ab} prediction (measurement + model) is ~5 %
- Sampling plots: 18 (9 points x 2 crop fields)
- Units: [µg cm^{-2}]

Absorbed photosynthetic active radiation (*aPAR*)

- Methods: 10 ceptometric measurements of incoming photosynthetic active radiation (PAR_{in}), transmitted PAR below the canopy PAR_{bottom}, PAR by reflected the top-of-the-canopy ($PAR_{refl,top}$) and at the soil ($PAR_{refl,soil}$) were acquired at each sampling plot. From the averaged values, aPAR was calculated as $aPAR = PAR_{in} - PAR_{bottom} - PAR_{refl,top} + PAR_{refl,soil}$; and the fraction of absorbed PAR (f_{aPAR}) as $aPAR$ / PAR_{in}.

</td></tr>
</table>

Measurements were fast (~5 min per plot); they took place between 12:45 and 14:00 the day of the campaign; timestamps are reported together with the data.
- Estimated uncertainties: the relative uncertainty of each individual ceptometric measurement is estimated to be ~2.5 %.
- Sampling plots: 18 (9 points x 2 crop fields)
- Units: [mol m^{-2} s^{-1}]

Leaf area index (*LAI*)
- Methods: the day after the flight campaign, destructive sampling was conducted at each sampling plot of 1 m^2. Leaves were sampled and scanned the day after the campaign, and the leaf area index was determined as the ratio of leaf area and plot area.
- Estimated uncertainties: the relative uncertainty of leaf area scanning is estimated ~1.5 %.
- Sampling plots: 18 (9 points x 2 crop fields)
- Units: [m^2 m^{-2}]

Leaf maximum carboxylation rate at 25 ºC ($V_{cmax,25}$)
- Methods: two portable photosynthesis measurement systems featuring leaf chambers were used simultaneously to measure C_i/C_a curves and derive $V_{cmax,25}$ from one leaf at the top of the canopy in each location. Measurements were limited to the morning, before noon, to prevent stomatal closure. Since measurements require long times (~40 minutes), only 12 plots could be measured on the morning of the flight, while the 6 plots remaining were measured the morning of the following day, expecting non-significant changes from one day to another. The timestamps of the center of the measurement period time (~20 min after the start) are provided with the data.
- Estimated uncertainties: the relative uncertainty of each $V_{cmax,25}$ estimate is ~2.5 %.
- Sampling plots: 18 (9 points x 2 crop fields)
- Units: [µmol CO_2 m^{-2} s^{-1}]

Non-Photochemical Quenching (*NPQ*)
- Methods: 5 leaves were measured with a portable pulse-amplitude-modulation fluorometer within the top half of the canopy. Maximum fluorescence in dark-adapted leaves (F_m) measurements were taken before dawn. Since *NPQ* can be highly dynamic during the diel cycle, measurements of maximum fluorescence of light-adapted leaves (F_m') were repeated during the morning of the flight campaign. The plots were measured twice in the same order, between 8.45 and 12.30, approximately. Timestamps of each light-adapted measurement are provided together with the *NPQ* estimate. *NPQ* was calculated as (F_m - F_m') / F_m'.
- Estimated uncertainties: the relative uncertainty of each individual *NPQ* estimate is ~2.5 %.
- Sampling plots: 18 (9 points x 2 crop fields)
- Units: [-]

	Non-Photochemical Quenching (*NPQ*) time series • Methods: An automated monitoring pulse-amplitude-modulation fluorometer with ten leaf probes was deployed between the two fields during the day of the flight. In each crop, 5 leaves from the top half of the canopy were continuously monitored in an additional plot. Nighttime data before dawn was used as the F_m reference, whereas daytime measurements of F_m' were taken with a frequency of 30 minutes. • Estimated uncertainties: The relative uncertainty of each individual *NPQ* estimate is ~2.5 %. • Sampling plots: 2 (1 point x 2 crop fields) • Units: [-] • Location: sensors were placed near the track separating two fields due to the length of the probe cables. No other samples were taken in these spots, which were located in the coordinates 262766.5, 4426169.5 UTM 30S (wheat) and 262776.5, 4426169.5 UTM 30S (maize). **Meteorological data time series*** • Methods: Half-hourly values of incoming short (SW_{in}, [W m^{-2}]) and longwave radiation (LW_{in}, [W m^{-2}]), incoming *PAR* (PAR_{in}, [µmol m^{-2} s^{-1}]), air temperature (T_{air}, [ºC]), atmospheric pressure (*Pres*, [hPa]) and wind speed (W_{speed}, [m s^{-1}]) and atmospheric vapor pressure (e_a, [hPa], were provided from a nearby eddy covariance tower for the day of the flight. Also, solar zenith (θ_{sun} [º]) and azimuth (ϕ_{sun} [º]) angles were included in the dataset. The timestamps corresponding to the half of each 30 minutes period are also provided. • Estimated uncertainties: unknown. • Sampling plots: 1 (outside the research station). • Units: see methods. *Meteorological data from Majadas de Tiétar experimental station were provided by the Max Planck Institute for Biogeochemistry (Germany) and Fundación CEAM (Spain).
Data formats and files	Imagery data is provided in the following in NetCDF files: • /SSC_data/Airborne_HDRF.nc • /SSC_data/Airborne_F.nc • /SSC_data/Airborne_LST.nc Field ancillary data is provided in two CSV files: • /SSC_data/FieldData.csv. It contains the spatial data acquired in the field sampling plots (18) • /SSC_data/FieldData_hh.csv. It includes the time series of *NPQ* (2 plotS) and meteorological variables measured in a nearby Eddy Covariance tower.
Scripts	Python, Matlab, and R scripts are provided to facilitate the import of the datasets and the generation of standardized outputs, including the participant's affiliation and contact data. Participants are welcome to use any software of their choice; however, only results in the format produced by these scripts will be accepted for submission to the Spatial Scaling Challenge. In addition, the scrips include comments with instructions and hints to complete the analyses

	and generate the results. These might be useful even for users of different languages or software. • / SSC_script.py • / SSC_script.m • / SSC_script.R The scripts will place the standardized output files in the folder /SSC_results/. In the publication summarizing the Spatial Scaling Challenge, results and their scores will be anonymized.
Templates for the description of the methods	Microsoft Word and OpenOffice Document empty templates that the participant will need to fill up with your diagnosis of vegetation status and a brief description of the methods used. • /SSC_report.docx • /SSC_report.doc • /SSC_report.odt Once completed, the participant will have to place the .docx, .doc, or .odt file in the folder /SSC_results/. The methods and results will be anonymized in the publication summarizing the Spatial Scaling Challenge.