COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Ecosystem spectral measurements: best practice metadata/ancillary dataset

Shari Van Wittenberghe (1), Andy Hueni (2), Helge Aasen (3), Miriam Machwitz (4), Zbyněk Malenovský (5), Kaniska Mallick (4), Eugenie Paul-Limoges (2), Martin Schlerf (4), Fabian Schneider (6), Lola Suarez (7), Sebastian Wieneke (8), Sebastian Wolf (9)

(1) Laboratory for Earth Observation, University of Valencia, Burjassot, Spain
(2) Remote Sensing Laboratories, University of Zurich, Zurich, Switzerland
(3) USYS, ETHZ, Zurich, Switzerland
(4) LIST, Luxembourg
(5) Czech Globe, Brno, Czech Republic
(6) The University of Melbourne, Melbourne, Australia
(7) University of Antwerp, Antwerp, Belgium
(8) ETHZ, Zurich, Switzerland

OPTIMISE Final Conference, 21-23 February 2018, Sofia, Bulgaria

OPTIMISE
Innovative Optical Tools for Proximal Sensing of Ecophysiological Processes

COST is supported by the EU Framework Programme Horizon 2020

cost
EUROPEAN COOPERATION
IN SCIENCE AND TECHNOLOGY

# Why metadata?

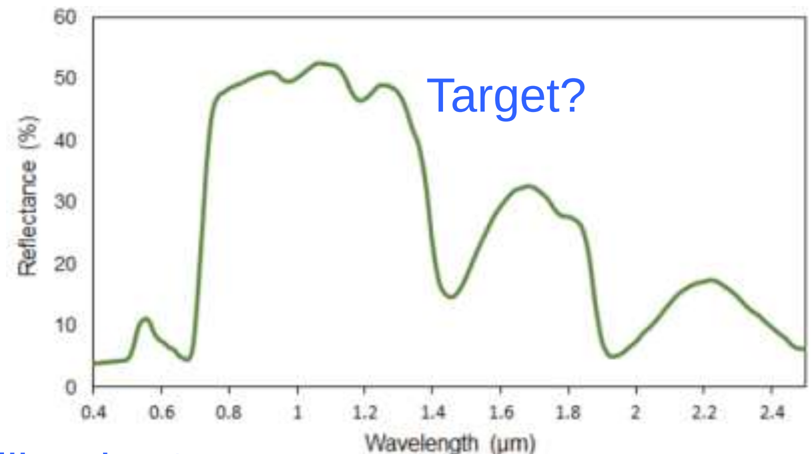
- Reproducibility
- Data interpretation
- Quality assessment
- Data Exchange
- Long-term usability

Illumination?

File name?

WR panel?

Target?




Calibration?

Instrument settings?

# Ancillary data

- Required to interpret the primary measurement
- Mandatory, but also case-dependent

For repeatability, any (primary as well as ancillary) measurement requires basic *mandatory* metadata:

- Method (direct/lab analysis, indirect/optical, etc.), standards and definitions
- **Uncertainty**
- Instrument spec. (type, model name, producer)
- Operational protocol (description of data-to-info transformation)

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Metadata challenges

**Instrumentation & Measurement protocols**

Standardisation
User-friendly
***[Previous presentations]***

**Processing/Scaling**
**Leaf-Tree-Canopy-Site-Ecosystem**

Tuning of algorithms/Model parameterization (priories/assumptions/boundary settings): important metadata!

*Lack of metadata= source of uncertainty*

Metadata

**Mandatory vs. additional Ancillary data**

~temporal, spatial scale
→ WS topic

**Storage & User-friendliness**

Automatic registration
Spectral libraries

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Core metadata vs. ancillary data

**Core Metadata** (spectroradiometer) ↔ **Ancillary data**

**Instrument // Instrument settings // Reference standard // Calibration // Illumination // Viewing Geometry // Atmospheric conditions // Coordinates // Target // Location // ···**

Do not (or limited) account for spatiotemporal variability (target)

↓

Spectral libraries (e.g. USGS)
Spectral information system (SPECCHIO)

Additional measurements, accounting for spatiotemporal variability due to

- Target Structure
- Physiology

Have own mandatory metadata!

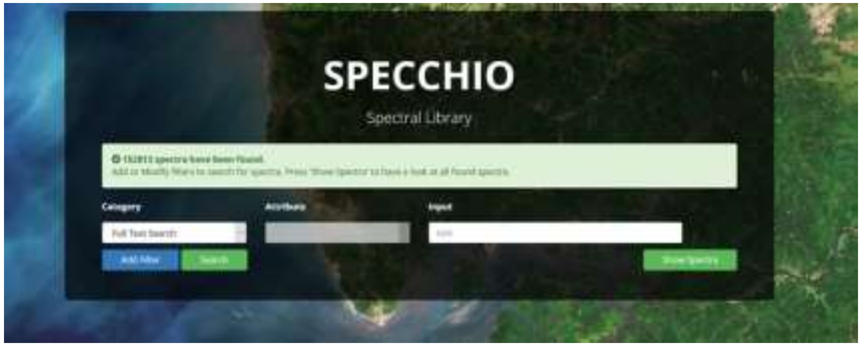
**Customized spectral libraries?**

- “Pure” spectra needed (~case)
- Metadata/AD-driven queries

# Core metadata: described in the SPECCHIO data model (http://www.specchio.ch/)

| Metadata variable | Type | Automation |
|---|---|---|
| Auto number | C | SF |
| User comment | S | SF |
| Capturing date and time | Q | SF |
| Spectral file name | S | SF |
| Number of spectra averaged internally by the instrument | Q | SF |
| Sensor | C | SF |
| File format | C | SF |
| Instrument | C | SF |
| Instrument calibration number | C | SF |
| Foreoptic | C | SF |
| Illumination source | C | |
| Sampling environment | C | |
| Measurement type (single, directional, temporal) | C | |
| Measurement unit (Reflectance, DN, radiance, absorbance) | C | SF |
| Target homogeneity | C | |
| Spatial position (latitude, longitude, altitude) | Q | SF |
| Landcover (based on CORINE land cover (8)) | C | |
| Cloud cover (in octas) | C | |
| Ambient temperature | Q | WS |
| Air pressure | Q | WS |
| Relative humidity | Q | WS |
| Wind speed (Qualitative description) | C | WS |
| Wind direction (categories in 45 degree steps) | C | WS |
| Sensor zenith angle | Q | CA (Goniom.) |
| Sensor azimuth angle | Q | CA (Goniom.) |
| Sensor distance | Q | |

Instrument settings and methodology
Measurement geometry
Environmental drivers
Target description
Data storage
Pictures



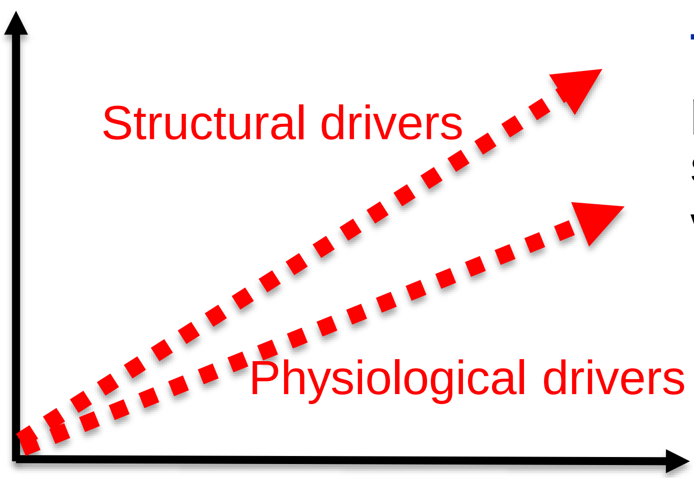
## Metadata Issues

*Lack of metadata/ancillary data = source of uncertainty*

# Drivers
# Spectral variability

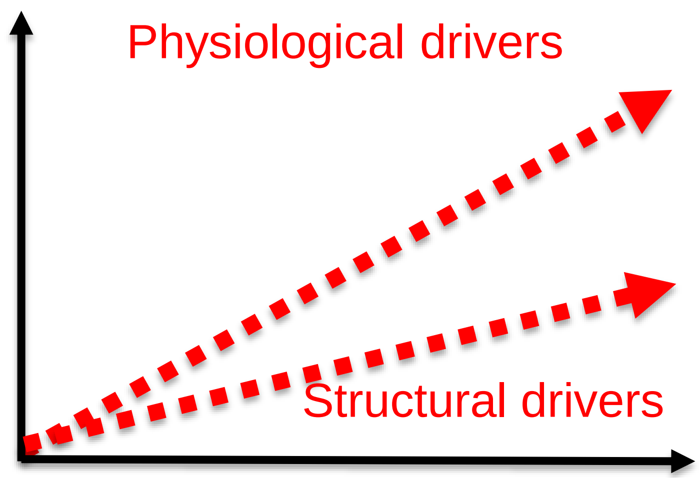
~ set-up/case
~ spatial & temporal scale

**Spatial point** spectral variability



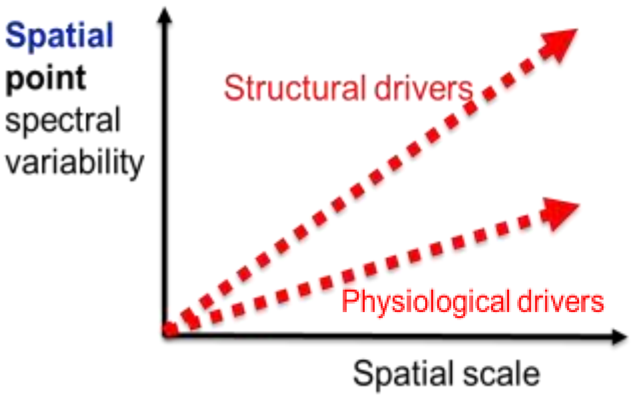

Spatial scale (ecosystem)

**Temporal point** spectral variability

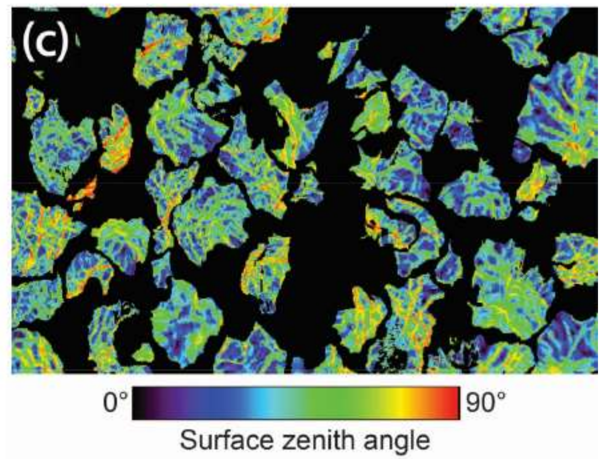


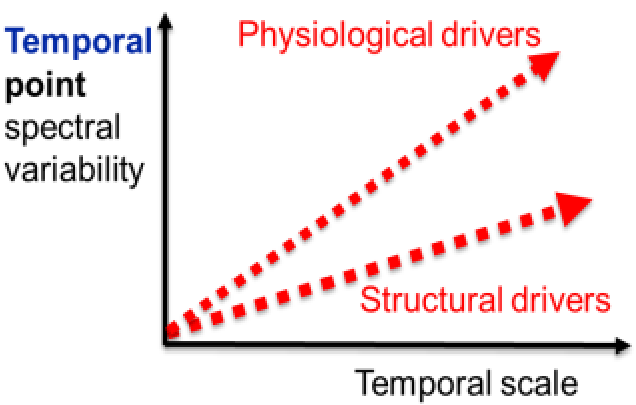

Temporal scale (leaf)

# Drivers spectral variability

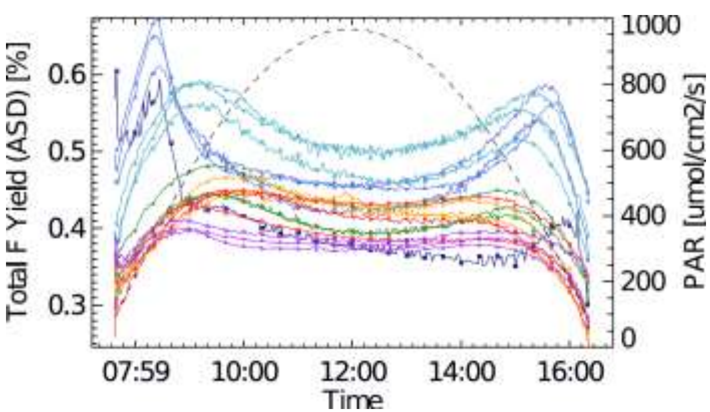



Pinto et al. 2017 *Rem. Sens.*




Alonso et al. 2017 *Rem. Sens.*




## Structural drivers

- Highly influenced by spatial scale, growth rate,…
- **Leaf average angle** or angle distribution, plant/leaf area index, … but also: wind

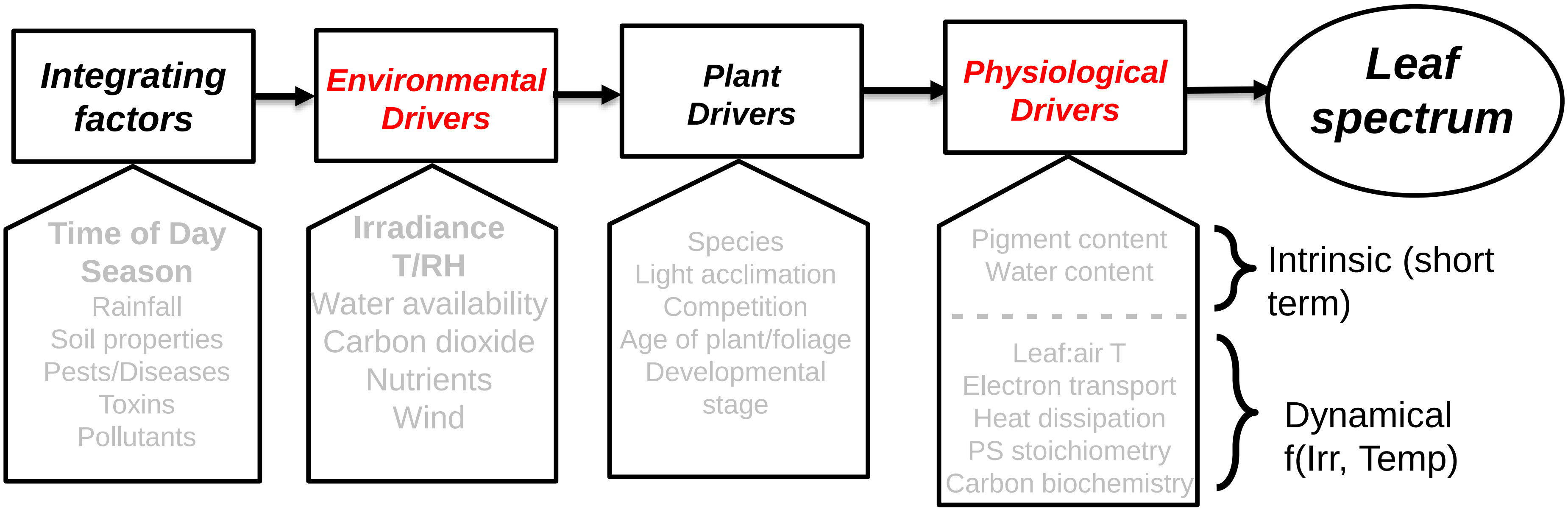
## Physiological drivers

- Importance↑ Time/ Stress events
- Intrinsic drivers: pigment content, water content;
- Photosynthesis drivers: f(Irr, T), energy dissipation balance (heat, F)

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

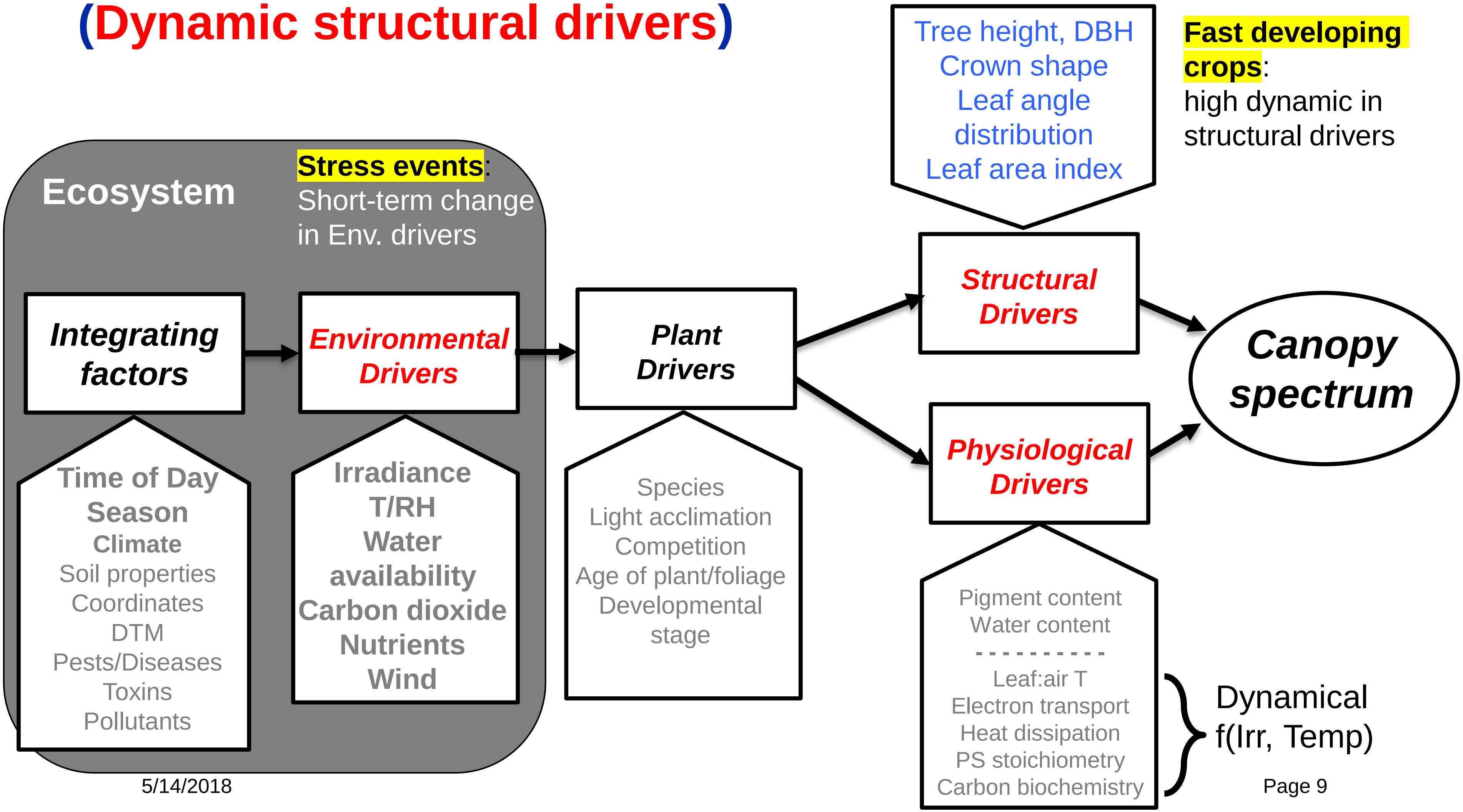
# Ancillary data at the leaf scale (Dynamic physiological drivers)



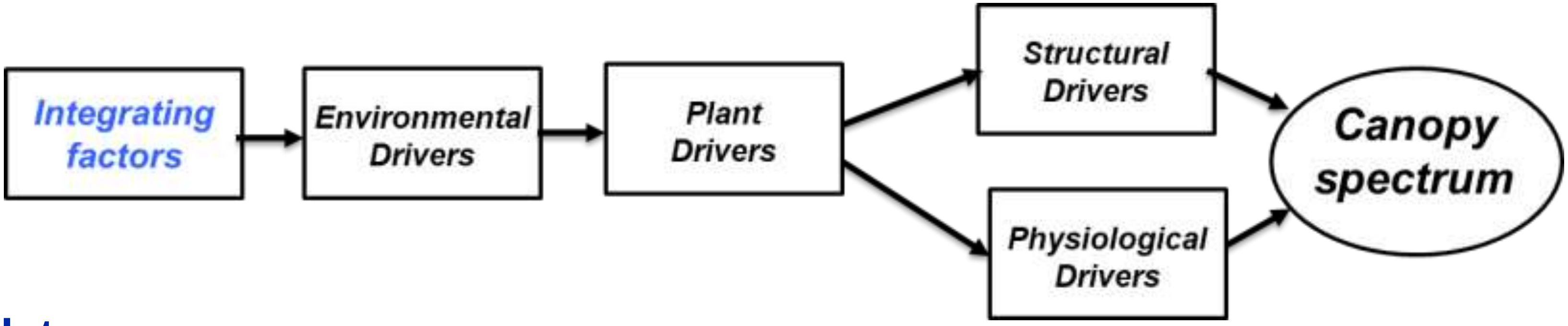

→ Case-specific: Which are the significant drivers?

(after Miller et al. 2005)

# Ancillary data at the canopy scale
# (Dynamic structural drivers)



# Site (*Integrating factors*) Metadata and Ancillary Data
## → Sensor ground footprint & BRDF localization and modelling



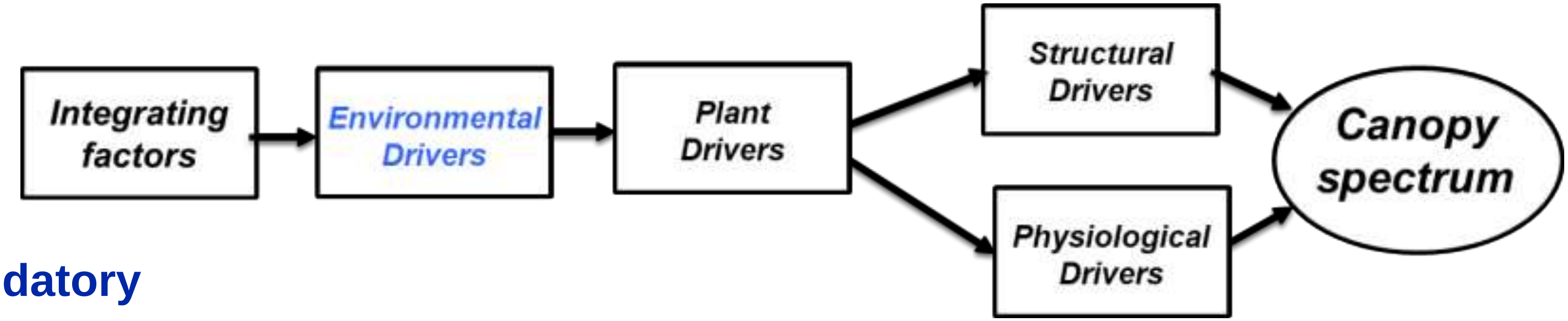

## Mandatory

- DOY
- Time of the day (UTC)
- Coordinates (lat/lon/alt)
- DTM – 3D coordinates & geo-projection system
- Understory & bare ground (optical) description

Time of the day and Coordinates → Sun angle

DTM → Deriving topographic variables - Radiative regime in simulation

## Environmental drivers Metadata and Ancillary Data
## → Dynamical drivers of ecophysiological processes

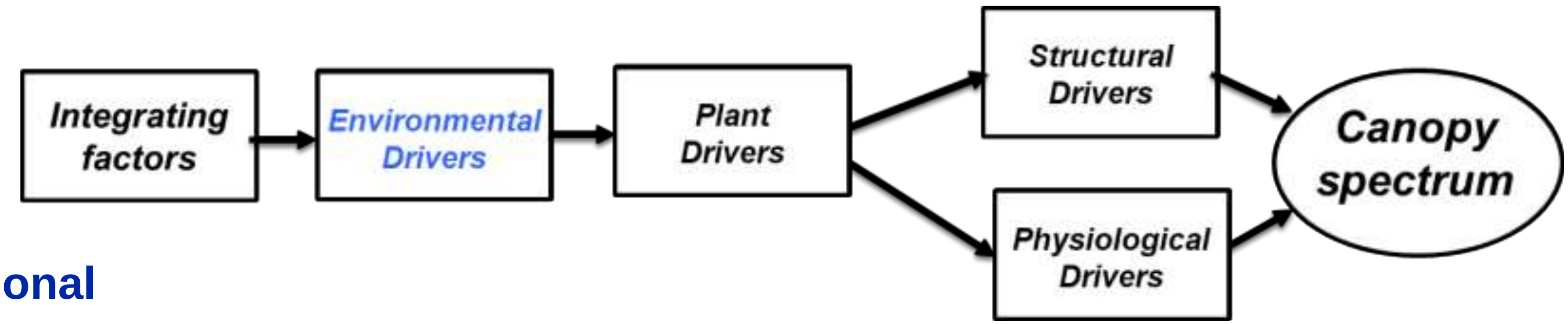



### Mandatory

- Downwelling PAR (400-700 nm) or VNIR (300-2500 nm)
- Upwelling PAR
- Irradiance (spectrally resolved diffuse/direct) → High precision reflectance, APAR, F
- Air T/RH
- Soil moisture → Evaporative flux, drought stress
- Hemispherical photos → Resolve cloud cover
- Precipitation
- Wind speed & direction

# Environmental drivers Metadata and Ancillary Data

## → Dynamical drivers of ecophysiological processes




### Optional

- Angularity, spectrally resolved irradiance
- Polarization of incoming light
- Atmospheric properties (AOT, wáter column vapor)
- Cloud height and thickness
- Soil T & evaporation
- Soil nutrients

→ Detailed radiative simulations

→ Upscaling to airborne or spaceborne at-sensor radiance, SIF retrieval

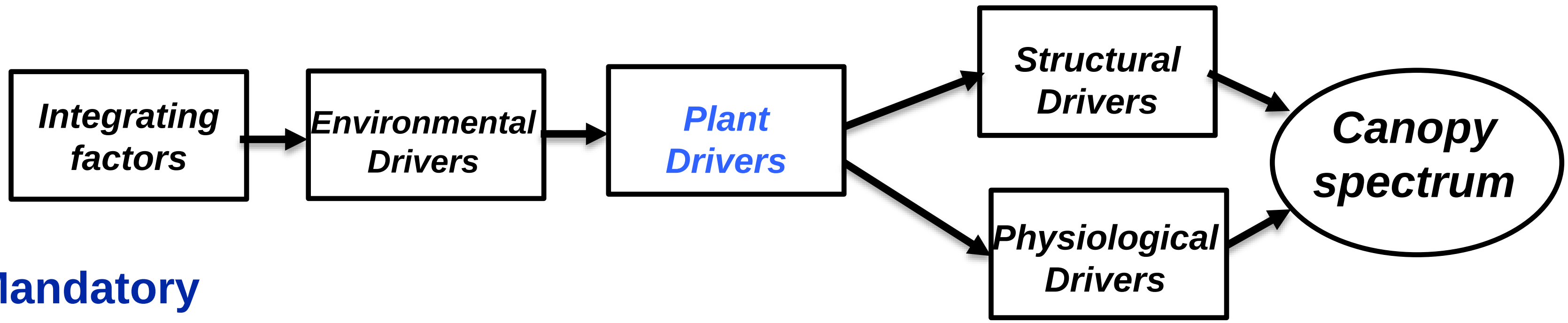
→ Input for evaporation flux

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Target drivers Metadata and Ancillary Data

## → Dynamical drivers of ecophysiological processes




## Mandatory

- Species identification (species/family/plant functional type)
- Plant positions (rooting location) in 3D coordinates

→ Target specification

- Upward photo (and side photo if possible)

→ General visualization, “unquantifiable” metadata,

## Optional

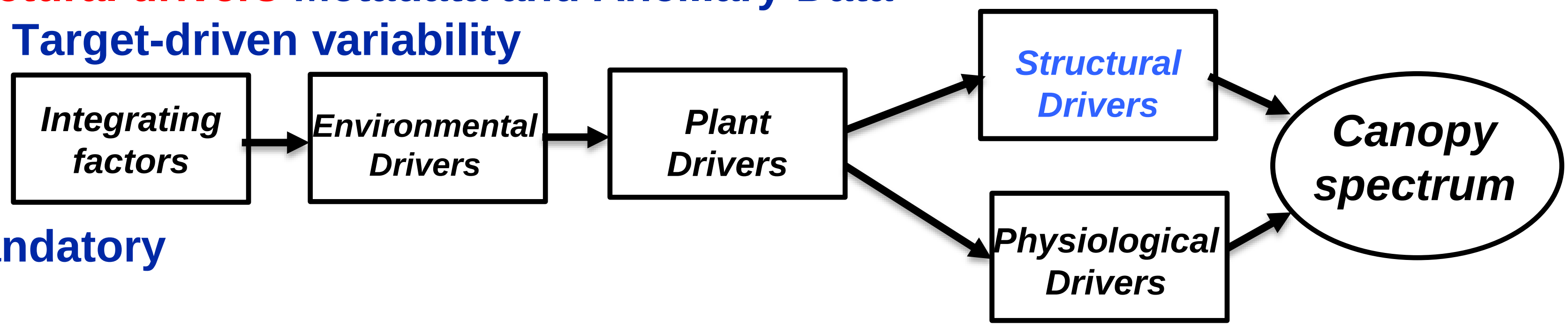
- Age of plant/foliage, developmental stage, visual description (moratlity, evident defoliation,…)

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Structural drivers Metadata and Ancillary Data

## → Target-driven variability




## Mandatory

- Tree height, DBH
- Crown height and projection diameters
- Height of first branching point
- Trunk optical properties
- Leaf average angle or angle distribution
- Total plant area index (i.e. leaf/branch density related to crown projection)
- Leaf length, Shoot length and density (for coniferous trees)
- Optical properties (reflectance & transmittance) per generation

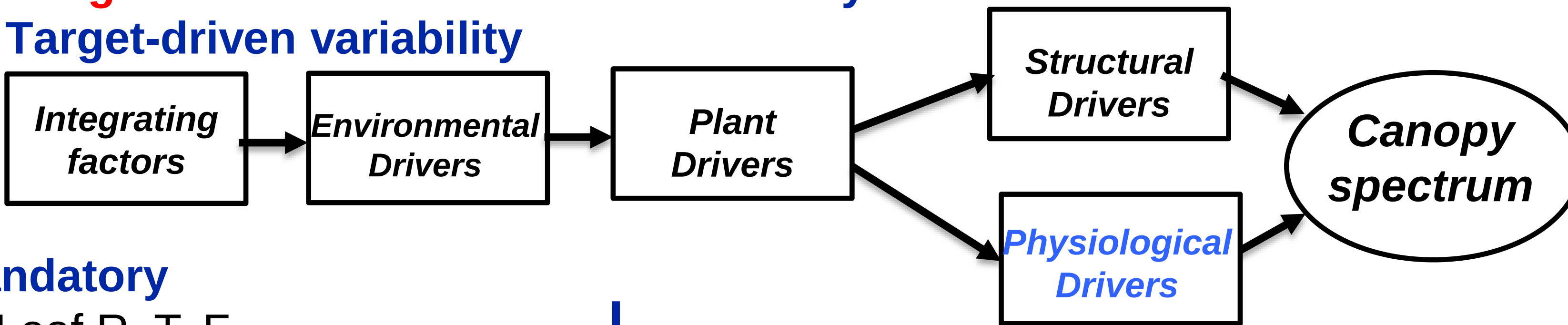
→ 3D target reconstruction and parameterization BRDF modelling

## Optional

- Branch diameter function, Branch optical properties, leaf shape, …

## Physiological drivers Metadata and Ancillary Data

### → Target-driven variability




### Mandatory

- Leaf R, T, F
- Irradiance
- Diffuse to direct ratio
- Pigment content (Chls, Cars)
- Leaf /canopy Temp
- Water content
- Dry matter content

→ Intrinsic leaf variables

### Optional (photosynthesis modeling): measured or estimated

- N content (Vcmo estimate)
- Vcmo
- ETRmax
- NPQ
- Ambient O2, ….

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

# Target variability as prerequisite to sampling design (sensor placement, parameters, frequency, ..)

| Water stress | Fluorescence | Pigmentation/composition | Damage/Infection | Functional diversity modelling |
|---|---|---|---|---|
| - **Stem Water Potential***<br>- **Stomatal conductance***<br>- Soil vertical profile<br>- Surface description (aspect, gradient and roughness)<br>- Crown relative position (emergent, co-dominant, supressed)<br>- **Photosynthetic rate*** | - **Leaf Reflected and Transmitted Radiance****<br>- **Irradiance****<br>- Light source<br>- Filter type<br>- Light history<br>- APAR Green (APARChl)<br>- **Leaf NPQ****<br>- Electron Transport Rate<br>- Chl content<br>- **Xanthophyll pigments (violaxanthin, antheraxanthin, zeaxanthin)**** | - **Leaf spectra****<br>- Chlorophyll A, B<br>- Carotenoids<br>- Anthocyanins<br>- N, P, K, C<br>- Wax content<br>- Leaf thickness (N parameter)<br>- **VAZ****<br>- Soil composition (N, P, K, C)<br>- Fertilising/management history | - Cause (virus, bacterial, fungus, Insects, bugs, Frost, wind break, snow break, unknown)<br>- **% damage^**<br>- **Pictures of damage^**<br>- Treatment/management history | - Soil vertical profile<br>- Surface characteristics (aspect, gradient, roughness)<br>- Meteorological history (T, precipitation, G Rad)<br>- Tree age<br>- Soil moisture<br>- Coarse Woody debris |
| **^days elapse, *20 min elapse, **Seconds elapse** | | | | |

Tree sampling metadata protocol: Coordinates, spp, heigh, Diameter at breast height, sample position (relative height and orientation), sampling size, storage conditions)

# The bigger picture: Physiological/Eddy Covariance Metadata and Ancillary Data (ecosystem level)

- Radiative energy fluxes, evapotranspiration flux, carbon flux
  - Complementary information on the ecosystem energy budget
  - Metadata on their representativeness:

- **Footprint** metadata:
  - **Optical footprint** of point spectrometer with FOV 25degrees is only covering a couple of few % of the EC
  - **EC footprint is** dynamic (i.e. wind)

- Contributors:
  - below ground (soil) vs. aboveground to decouple
  - Understory vs. upper canopy// different land cover classes

# The bigger picture: Physiological/Eddy Covariance Metadata and Ancillary Data (ecosystem level)

- **Evapotranspiration**:
  - Tricky to parameterize (leaf conductance highly variable) and upscale
  - Unknown effects (e.g. canopy wetness)
  - Hyperspectral thermal signatures (link latent and sensible heat transfer)

- **Timing of meteorological effects and associated (possibly time-delayed) responses**: e.g. timing of drought (temporal compensation)

- **Geographical heterogeneity (local patterns of stochastic variations)**: At ecosystem level the local energy aspects can be levelled out which we don't see at local spot

# Practical suggestions for the community

- **Be aware** of metadata-AD complexity during spectral sampling of vegetation as **highly spatiotemporal dynamic target**
- For upscaling: matching measurement time (format), and measurements when possible
- Keep **protocols** simple and documented for future usability and repeatability
- Document specific **anomalies** or incidents in the field
- Practical user interface or digital device for registering metadata in a digital format (e.g. with standardized parameter ranges, quality flag)
- **Parameterization and validation**: use of same measurement protocol to collected data for both uses
- Processing of metadata/AD prone to many **assumptions** (keep raw data)
- Keep in mind **error propagation** during scaling

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

| WG1: |
| --- |
| Definition of a mandatory metadata set, aligned with current international efforts in the spectroscopy community<br>**Status/action**<br>• EcoSIS, SPECNET, CSIRO and Geo Science Australia collaborations ongoing<br>• Metadata Workshop: Luxembourg Sept 2017 -> Talk by Shari at final conference |
| Develop an on-line instance of a spectral information database to serve as demonstration and testing platform for data sharing and information building<br>**Status/action**<br>• SPECCHIO Online System, training material, OPTIMISE summer school data |
| Develop a wireless automated dataflow from in-situ and **UAV** sensor for the database system<br>**Status/action**<br>• Wireless interaction with FloX/RoX Spectrometers -> Software available online<br>• ASD iPad App -> Software available on request<br>• Tower mounted instruments to database data flow: concept and code<br>• FloX to SPECCHIO automated dataflow: Talk by S. Trim at final conference |
| Definition and implementation of data pre-processing and metadata augmentation algorithms and routines including quality checks and flagging and data assimilation<br>**Status/action**<br>• Complete: Matlab example of processing tool presented at EARSeL |
| Definition and implementation of system interfaces and algorithms for data retrieval allowing the building of products using sources such as biogeochemical modelling, flux data specialisation and space-sensed data<br>**Status/action**<br>• Complete: Interface example implemented for Fluspect |

# Online documentation of Aims and Deliverables

optimise.dcs.aber.ac.uk

**WG1 Deliverables:**

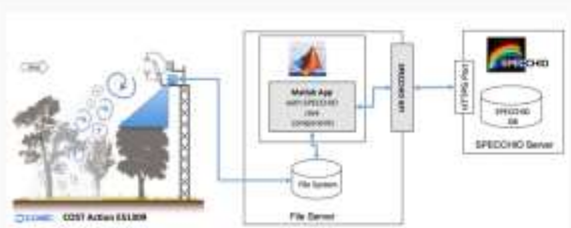
1. Definition of a mandatory metadata set, aligned with current international efforts in the spectroscopy community
2. Develop an on-line instance of a spectral information database to serve as demonstration and testing platform for data sharing and information building

-> Public online database system: see www.specchio.ch
-> SPECCHIO API to program SPECCHIO from e.g. Matlab or R or Java
-> SPECCHIO online spectral database holding the data of OPTIMISE training schools (ABEL and SWAMP): SPECCHIO_Web_Interface

3. Develop a wireless automated dataflow from in-situ and UAV sensor for the database system

-> ASD iPad App to control an ASD field spectroradiometer, see also ASD iPad App Poster presented at EARSeL 2017, Zurich, Switzerland

-> Concept and Code for tower mounted instruments.




Data flow from tower mounted sensor to spectral database

-> Wireless User Interface for Flox/Rox point spectrometers: Open Source Code and compiled binaries are available on GitHub.


http://optimise.dcs.aber.ac.uk/working-groups/wg1-spectral-information-system/

COST is supported by the EU Framework Programme Horizon 2020

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

**Thank you!**


*Acknowledgements:*
*OPTIMISE - COST*
*Luxembourg Institute of Science and Technology (LIST)*