Protocols for assessing quality of observational datasets

Jan-Peter Muller,

UCL-Mullard Space Science Laboratory

Thanks to contributions from T. Scanlon, J. Nightingale (NPL), J. Schultz (EUMETSAT), M. Van Roozendael & J-C Lambert (BIRA), N. Gobron (JRC), S. Kharbouche (UCL)

j.muller@ucl.ac.uk

Copernicus Climate Observations Requirements Workshop,
ECMWF, 29 June - 2 July 2015

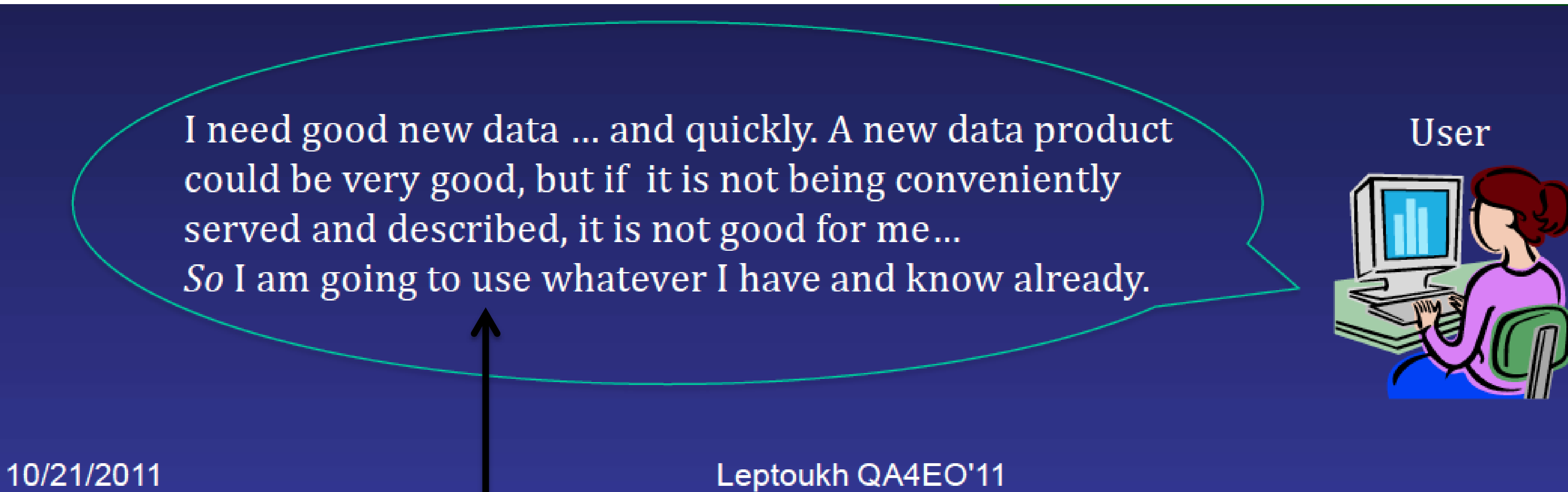
Version: 02.07.2015

Overview

- QA4ECV Objectives
- Metrological definitions of uncertainty, errors, error propagation, validation
- What has been achieved in the past on best practices:
  - land ECVs (CEOS Land cover, LAI)
  - atmospheric chemistry
- Protocols for Validation in QA4ECV
  - Qualitative methods
  - Scaling from aircraft
  - Tower measurements to 1km albedos
  - Biophysical parameters - fAPAR
  - "in situ" data : scene simulation for algorithm validation
  - Atmospheric reference datasets

Motivation

User perspective:



This is where QA4ECV comes in

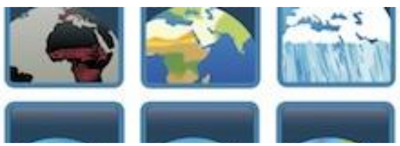
Mission statement QA4ECV

- QA4ECV will show how trustable assessments of satellite data quality can facilitate users in judging fitness-for-purpose of the ECV Climate Data Record.

- QA4ECV will provide quality assured long-term Climate Data Records of several ECVs relevant for policy and climate change assessments.

ESA CCI Aerosol Cloud CMUG Fire GHG Glaciers Ice Sheets Land Cover Ocean Colour Ozone Sea Ice Sea Level S

CCI



ESA Climate Change Initiative

Wed, 2010-09-01 11:03

Climate change is arguably the greatest challenge facing mankind in the twenty-first century. Its importance has been recognised in re reports from the IPCC and from UNFCCC, and the overwhelming economic consequences are set out in the Stern Report.

GCOS Essential Climate Variables

The 50 GCOS Essential Climate Variables (ECVs) (2010) are required to support the work of the UNFCCC and the IPCC. All ECVs are technically and economically feasible for systematic observation. It is these variables for which international exchange is required for both current and historical observations. Additional variables required for research purposes are not included in this table. It is emphasized that the ordering within the table is simply for convenience and is not an indicator of relative priority.

Domain	GCOS Essential Climate Variables
Atmospheric (over land, sea and ice)	**Composition:** Carbon dioxide, Methane, and other long-lived greenhouse gases[3], Ozone and Aerosol, supported by their precursors[4].
Terrestrial	River discharge, Water use, Groundwater, Lakes, Snow cover, Glaciers and ice caps, Ice sheets, Permafrost, Albedo, Land cover (including vegetation type), Fraction of absorbed photosynthetically active radiation (FAPAR), Leaf area index (LAI), Above-ground biomass, Soil carbon, Fire disturbance, Soil moisture.

[4] In particular nitrogen dioxide (NO_2), sulphur dioxide (SO_2), formaldehyde (HCHO) and carbon monoxide (CO).

Target Requirements

Variable/ Parameter	Horizontal Resolution	Vertical Resolution	Temporal Resolution	Accuracy	Stability
Black-sky albedo	1km	N/A	Daily to weekly	max(5%; 0.0025)	max(1%; 0.0001)
White-sky albedo	1km	N/A	Daily to weekly	max(5%; 0.0025)	max(1%; 0.0001)

EO and Climate Data Records

Ideal Harmonisation for Climate Records Over Decades

Requires data that is:

Stable over time
- so data can be compared across decades meaningfully

Insensitive to the method of measurement
- so data from different sensors (and techniques) can be combined

Uniform ‘worldwide’
- so data from different space agencies can be combined

Based on references that can improve
- methods will improve over time as new technologies are available
- harmonisation should not be at the expense of improvements

Metrology Principles

Documentation

Comparisons

Peer review

Auditing

Community defined references (SI)

Changes made cautiously – consistent to old definitions

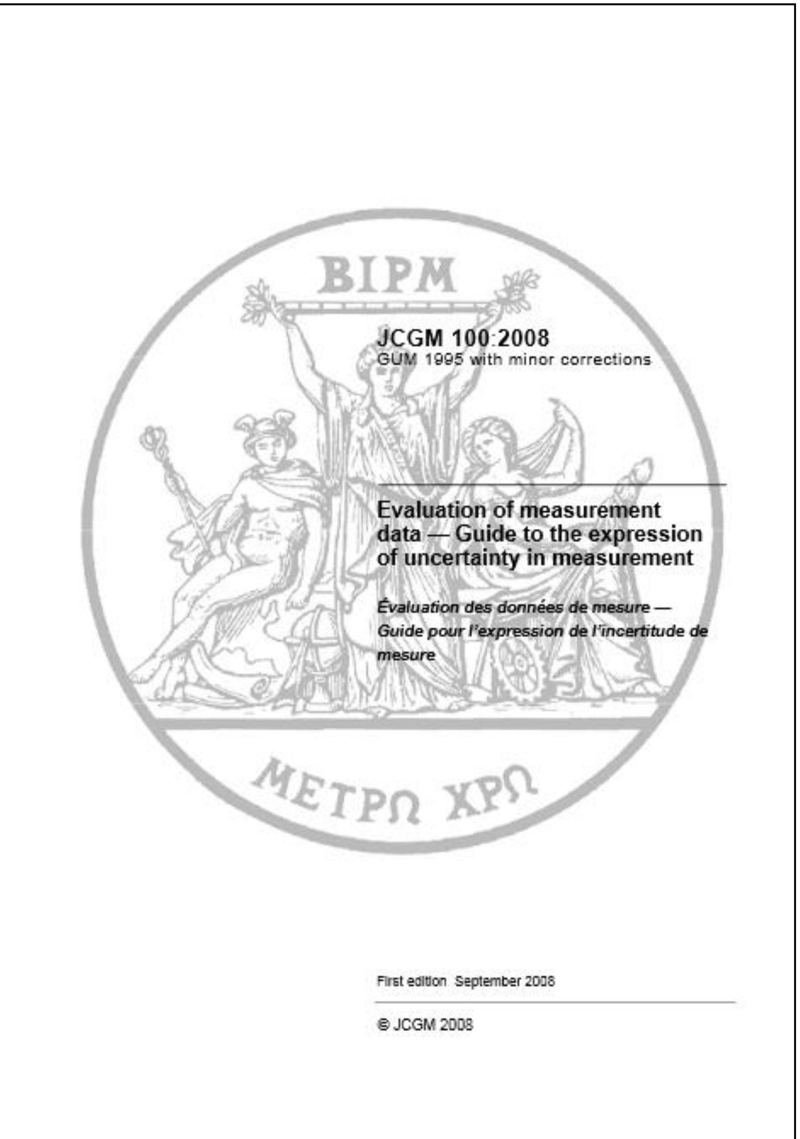
Uncertainty

Traceability

Stable measurements, worldwide consistency, insensitive to methods

j.muller@ucl.ac.uk

Uncertainty and how to deal with it
The GUM


BIPM

JCGM 100:2008
GUM 1995 with minor corrections

Evaluation of measurement data — Guide to the expression of uncertainty in measurement

Évaluation des données de mesure — Guide pour l'expression de l'incertitude de mesure

ΜΕΤΡΩ ΧΡΩ

First edition September 2008

© JCGM 2008

The Guide to the expression of Uncertainty in Measurement (GUM)

- The foremost authority and guide to the expression and calculation of uncertainty in measurement science
- Written by the JCGM and BIPM between 1977 and 1995 (updated 2008)
- Covers a wide number of applications
- Technical with formal mathematics

http://www.bipm.org/en/publications/guides/gum.html

Traceability

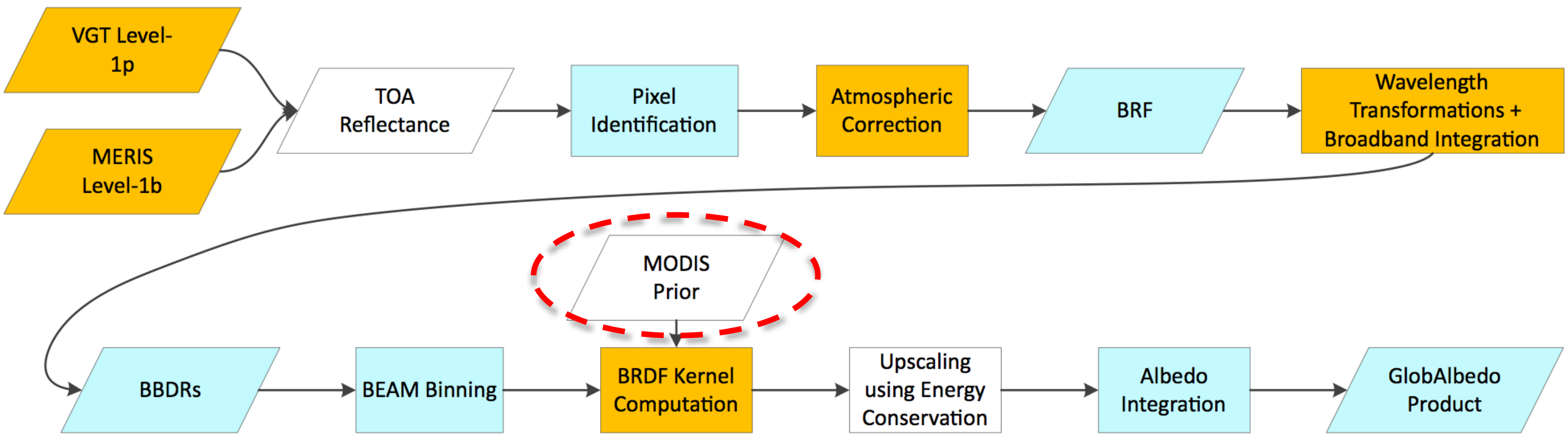
National Physical Laboratory

"Property of a measurement result relating the result to a stated metrological reference (free definition and not necessarily SI) through an unbroken chain of calibrations of a measuring system or comparisons, each contributing to the stated measurement uncertainty"

Committee on Earth Observation Satellites (CEOS)

QA4ECV

Traceability Example: GlobAlbedo processing chain using Optimal Estimation



Uncertainty propagation through the processing chain requires uncertainty estimates which are traceable to reference standards

Error

is NOT the same as

Uncertainty

Uncertainty vs. error

Uncertainty:

- Describes the spread of a probability distribution i.e. standard deviation
  - Uncertainty is the doubt you have on the value

Error:

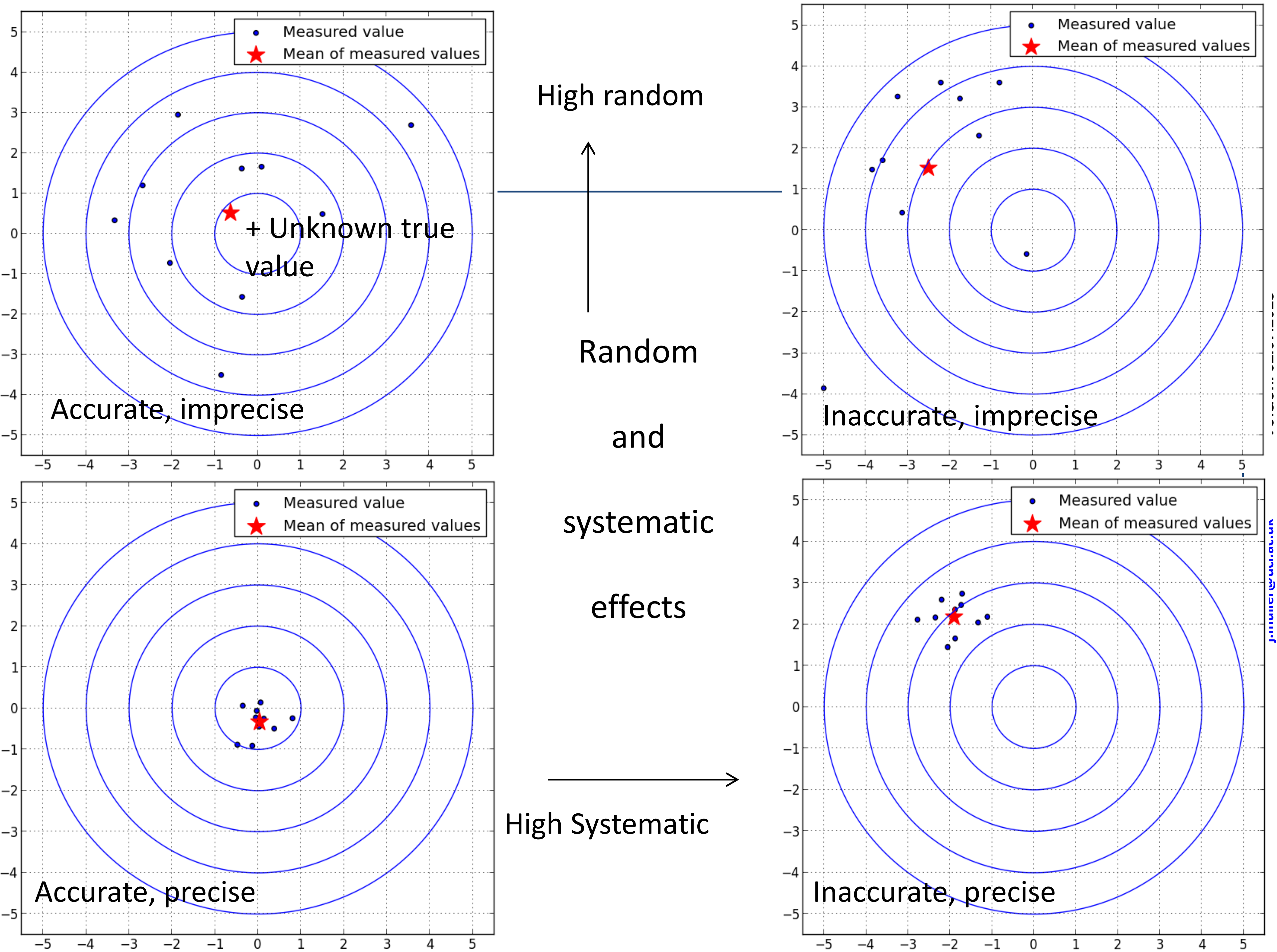
- Difference from truth
  - Result of measurement imperfections
  - From random and systematic effects

Correction

- Where an error is known, it can be corrected by applying a correction
  - There will always be an unknown residual error which adds to the uncertainty

Consistency in terminology is important!

j.muller@ucl.ac.uk



Measured value
Mean of measured values
+ Unknown true value
Accurate, imprecise
High random
Random
and
systematic
effects
Measured value
Mean of measured values
Inaccurate, imprecise
Measured value
Mean of measured values
Accurate, precise
High Systematic
Measured value
Mean of measured values
Inaccurate, precise

Random Effects

- Random effects
  - Different error for every measurement (different random number)
  - Cannot be corrected for even if the measurement is dully understood
  - Can have same associated uncertainty (drawn from same probability distribution)
  - e.g. Detector noise etc.

Note: don't use the incorrect phrase "random uncertainties" – "uncertainty" describes the probability distribution.
Strictly: "uncertainties associated with random effects"

Systematic effects

- Errors which in principle can be corrected for if the cause of the error was fully understood
  - Of course often you don't know what this correction is so you have an uncertainty associated with such systematic effects
  - E.g. Incorrect instrument parameterisation
- With many systematic effects there is a time and space scale which is applicable
  - E.g. Instrument degradation – changes slowly over time
- Local effects
  - Metrology doesn't yet have a formal way of describing these
  - Effects that are local in time and/or space
    - E.g. Atmospheric effects, calibration (solar contamination) etc.

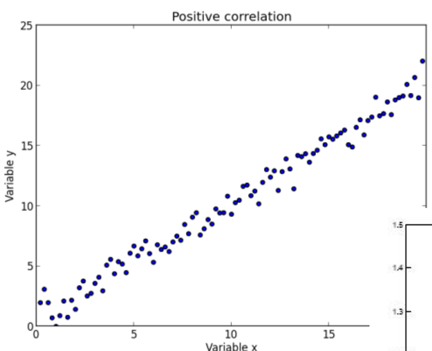
Type A and Type B methods

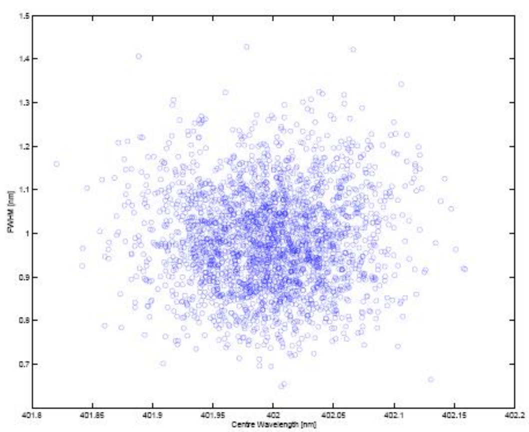
- Two methods of assessing uncertainty
  - Type A
    - Application of statistical methods to a series of repeated determinations (real or simulated)
  - Type B
    - Based on experience and knowledge of physical processes
    - The uncertainty associated with the systematic error is known even though the error itself isn't

How to determine correlation (and covariance)

Type A methods:
From the data (real or simulated)

- Discover correlations





Type B: From knowledge (measurement model)

$$E_i = E_{\mathrm{True}} + S + R_i$$

This is where the correlation comes from!

$$u(x,y) = u(S)$$

Validation

*‘The process of assessing, by **independent means**, the quality of the data products derived from the system outputs’*

(Justice, et al., 2000, p. 3383)

j.muller@ucl.ac.uk
Version: 02.07.2015

International Coordination

Committee on Earth Observation Satellites

Working Group on Cal/Val:

- Synthetic Aperture Radar (SAR)
- Infrared Visible Optical Sensors (IVOS)
- Microwave Sensors (MSSG)
- Terrain Mapping (TMSG)
- **Land Product Validation (LPV)**
- **Atmospheric Composition (ACSG)**

CEOS Validation Hierarchy

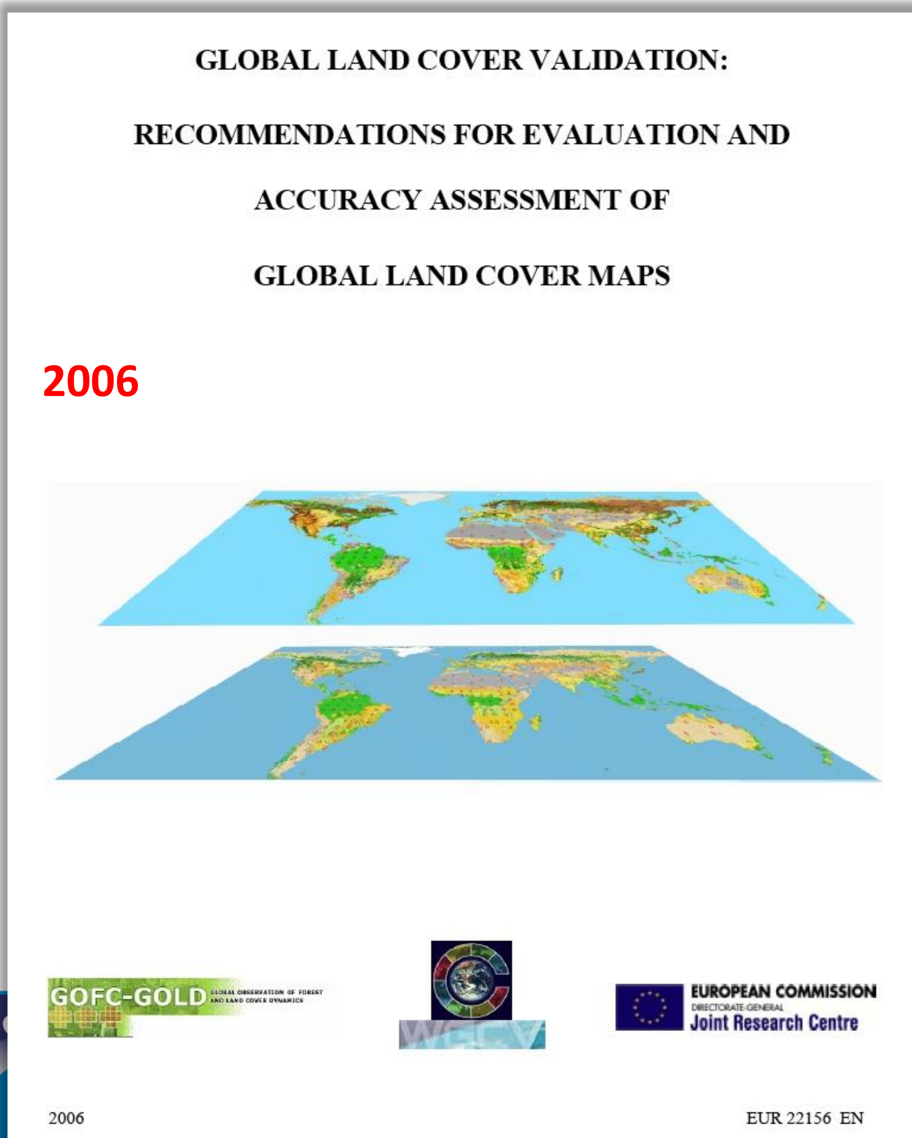
Stage 1 ✓	Product accuracy is assessed from a small (typically < 30) set of locations and time periods by comparison with in situ or other suitable reference data.
Stage 2 ✓	Product accuracy is estimated over a **significant set of locations and time periods** by comparison with reference in situ or other suitable reference data. **Spatial and temporal consistency** of the product with similar products has been evaluated over globally representative locations and time periods.
Stage 3 ?	**Uncertainties** in the product and its associated structure are well quantified from comparison with in situ or other suitable reference data. **Spatial and temporal consistency** of the product with similar products has been evaluated over globally representative locations and time periods. **Uncertainties are characterized in a statistically robust way over multiple locations and time periods** representing global conditions.
Stage 4 ??	Validation results for stage 3 are **systematically updated** when new product versions are released and as the time-series expands.

Alignment with

- Focus areas based on *Essential Climate Variable* (ECV) parameters:
  - Land Cover
  - Land Surface Phenology *
  - Fire (Active/Burned Area)
  - Surface Radiation (Reflectance, BRDF, Albedo)
  - Biophysical – LAI & fAPAR
  - Soil Moisture
  - Snow/Ice
  - Land Surface Temperature *

* Not currently ECV, but large community & product base

Good Practice Documents



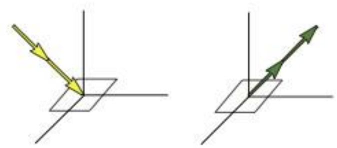
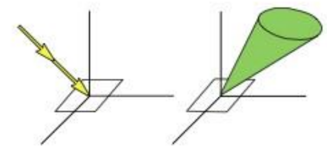
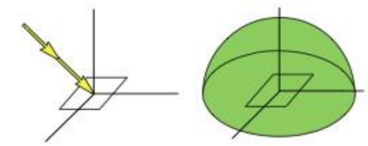
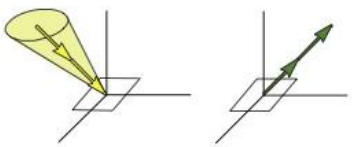
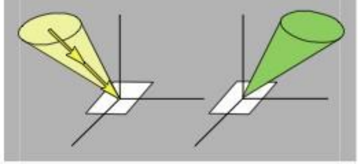
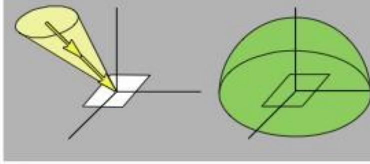
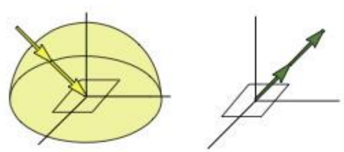
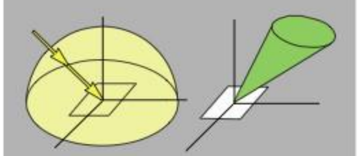
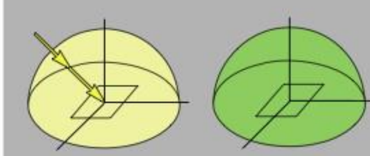
Validation Methods for EO products

- Look at the data using browse products
- From leaf to tower to EO pixel: how to scale
- Uncertainties vs Error
- Assessment of global consistency (e.g. Hovmöller plots)
- Inter-comparison of global EO products using triple collocation
- Reference datasets

j.muller@ucl.ac.uk

Definitions of reflectance

Relation of incoming and reflected radiance terminology used to describe reflectance quantities

Incoming/Reflected	Directional	Conical	Hemispherical
Directional	Bidirectional CASE 1 	Directional–conical CASE 2 	Directional–hemispherical CASE 3 
Conical	Conical–directional CASE 4 	Biconical CASE 5 	Conical–hemispherical CASE 6 
Hemispherical	Hemispherical–directional CASE 7 	Hemispherical–conical CASE 8 	Bihemispherical CASE 9 

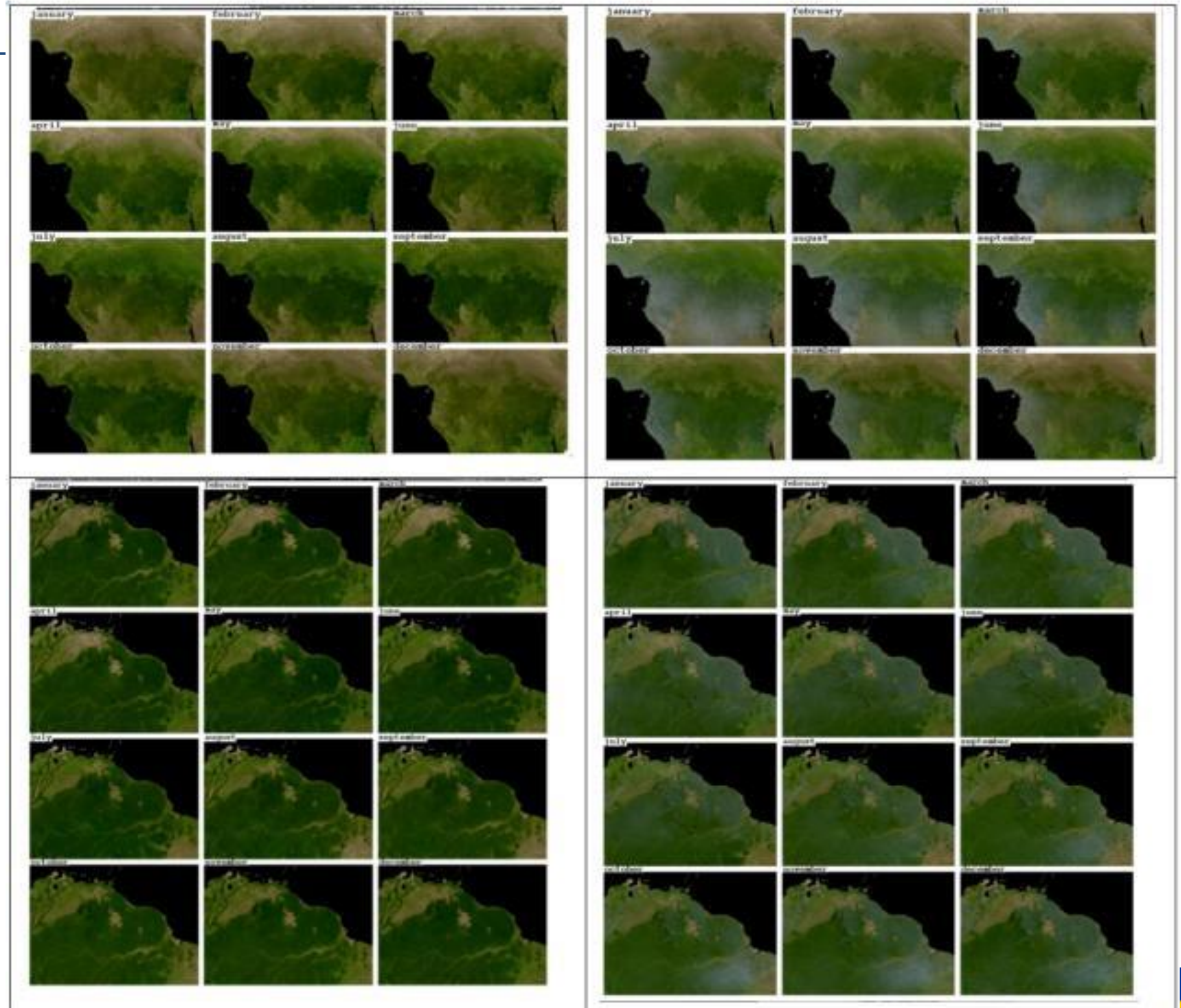
The labeling with 'Case' corresponds to the nomenclature of Nicodemus et al. (1977). Grey fields correspond to measurable quantities (Cases 5, 8), the others (Cases 1–4, 6, 7, 9) denote conceptual quantities.

Schaepman-Strub, G. and Schaepman, M. (2006). Reflectance quantities in optical remote sensing—definitions and case studies. Remote Sensing of Environment vol. 103 pp. 27-42.

j.muller@ucl.ac.uk

Atmospheric contamination of MCD43 products

- Africa (upper panel) and S. America (lower panel)
- True colour composite of isotropic reflectance component in RGB channels
- Cloud-cleared (left from U of Lille)
- Note the whitish hue due to uncorrected aerosol/cloud contamination issues in GlobAlbedo product



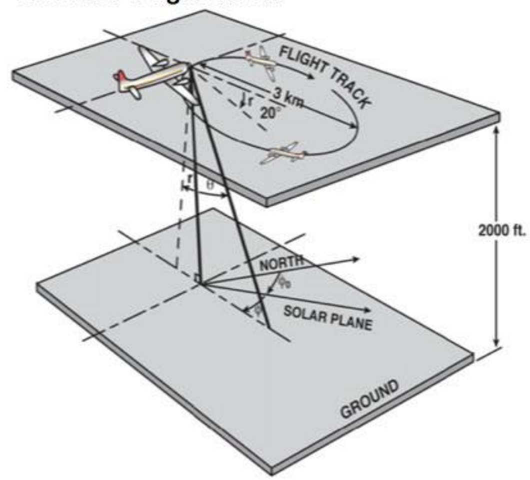
Fonds des Sols   GlobAlbedo

http://adam.noveltis.com/

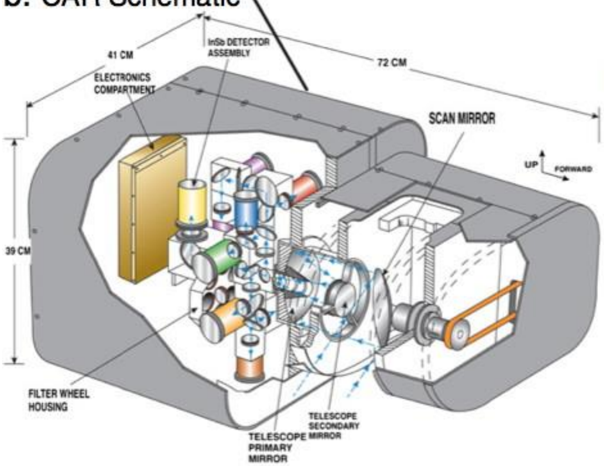
NASA CAR (Cloud Absorption Radiometer) : an excellent source for BRDF

a. Jetstream-31 Aircraft

c. BRDF Flight Track



b. CAR Schematic

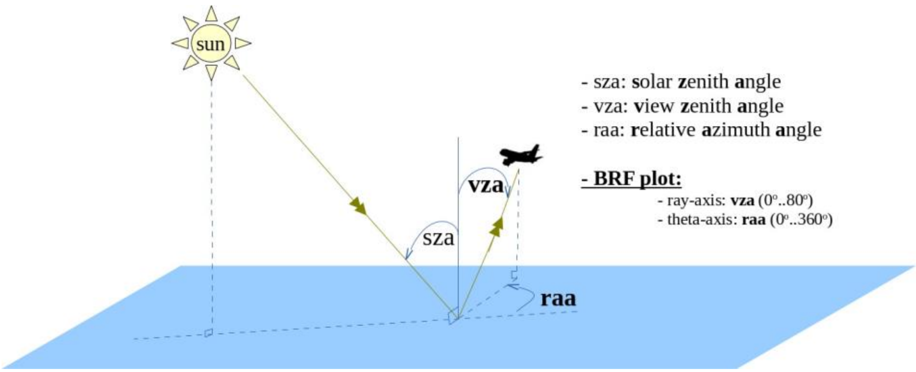


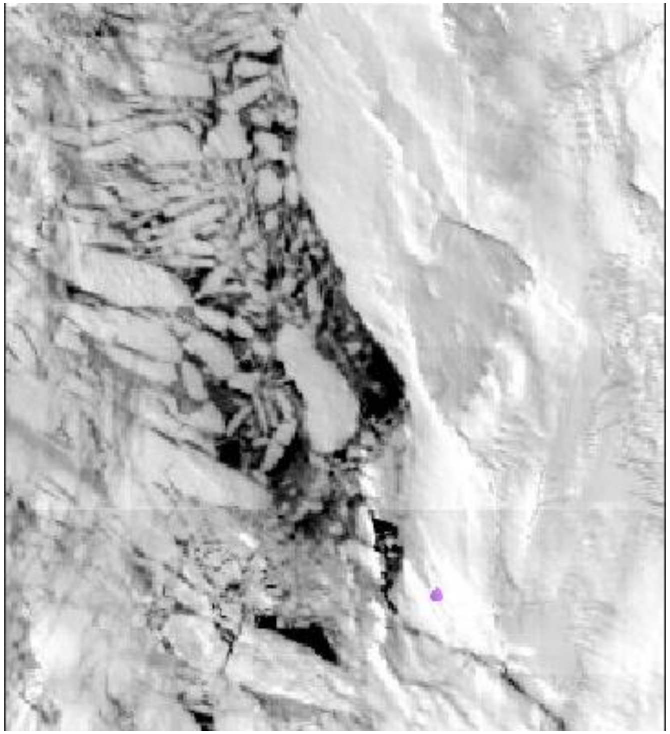
d. Cloud Absorption Radiometer (CAR) Parameters

Angular scan range	190°
Instantaneous field of view	17.5 mrad (1°)
Pixels per scan line	382
Scan rate	1.67 scan lines per second (100 rpm)
Spectral channels (μm; bandwidth (FWHM))	14 [a] (8 continuously sampled and last six in filter wheel): 0.340(0.009), 0.381(0.006), 0.472(0.021), 0.682(0.022), 0.870(0.022), 1.036(0.022), 1.219(0.022), 1.273(0.023), 1.556(0.032), 1.656(0.045), 1.737(0.040), 2.103(0.044), 2.205(0.042), 2.302(0.043)

Taken from Gatebe, C. K., et al. (2010). *ACP*, *10*(6), 2777–2794.

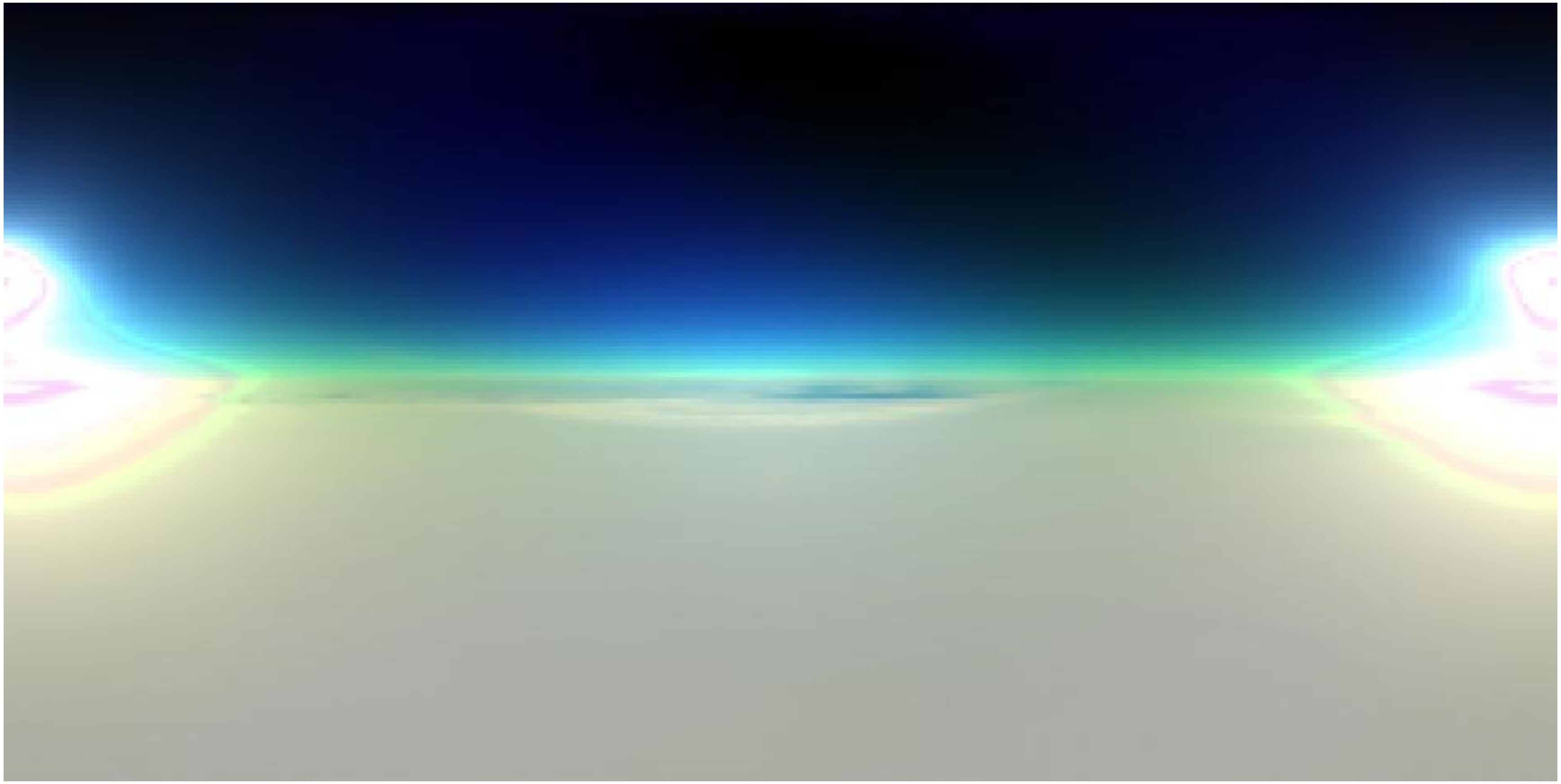
NASA CAR sampling and common location





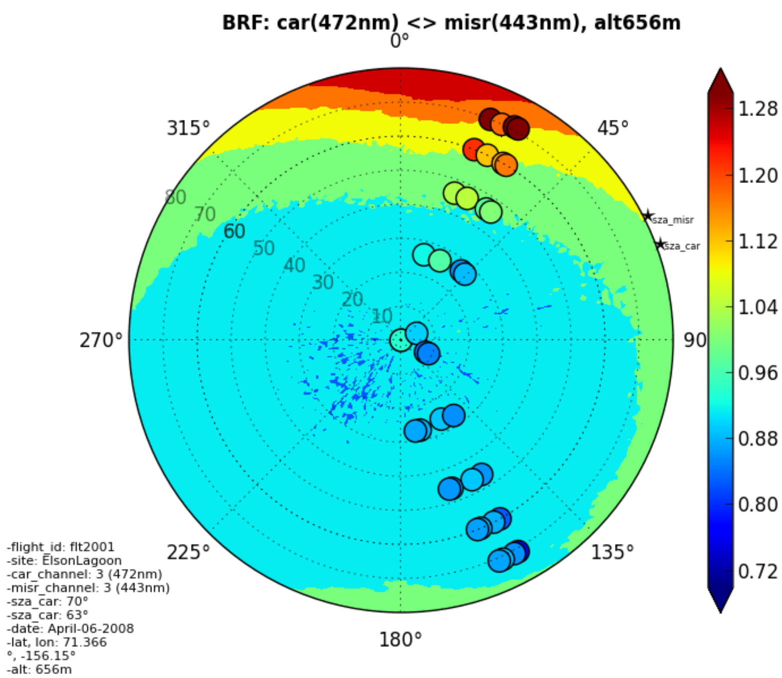
Roman et al. (2011)

CAR observations of sea-ice during ARCTAS (full sky and ground views)

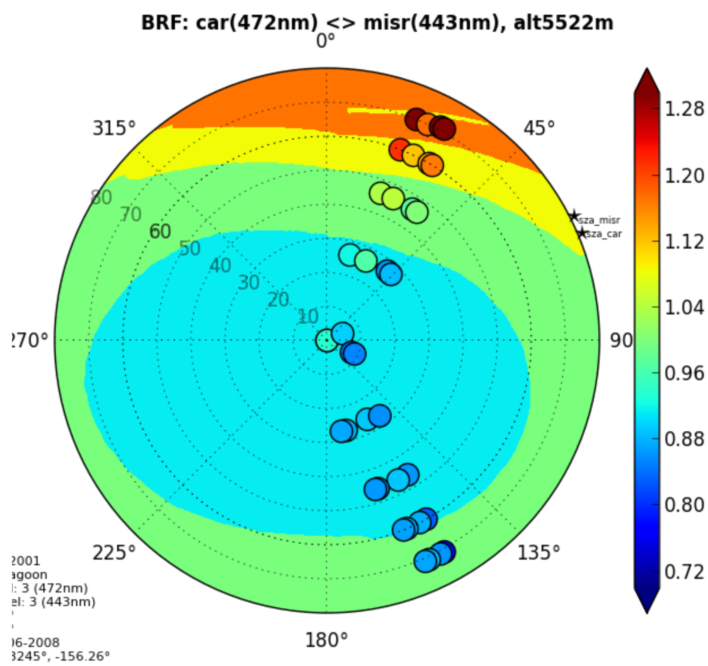


j.muller@ucl.ac.uk
Version: 02.07.2015

MISR vs CAR for ARCTAS @ 2 altitudes



IFoV≈5m

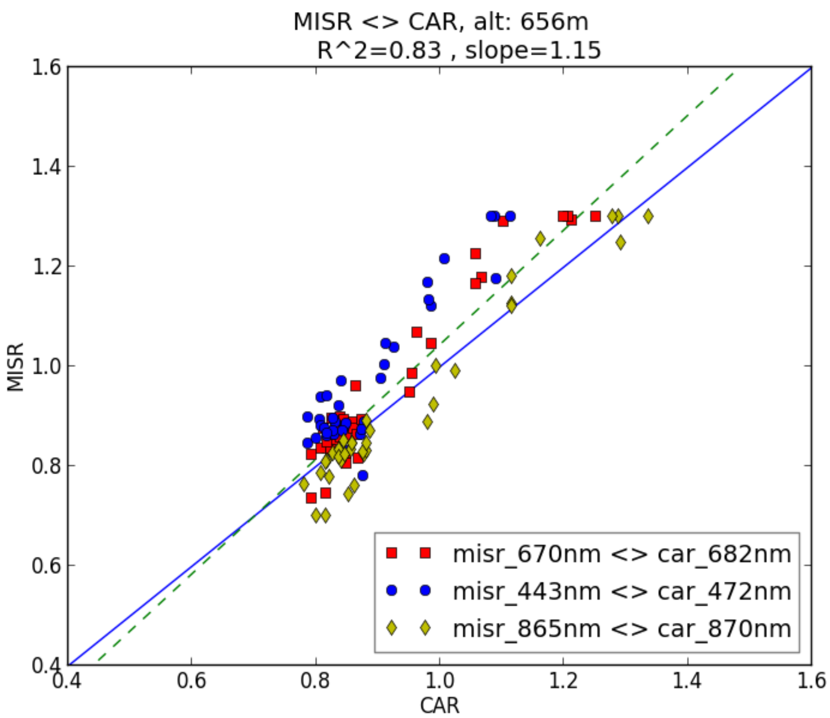


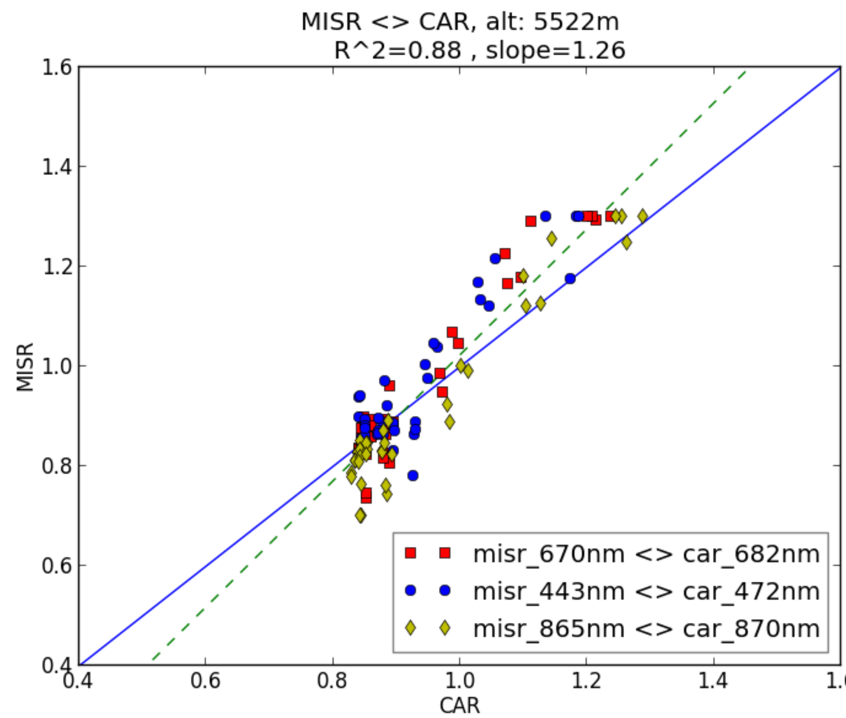
IFoV≈50m

NASA CAR data courtesy of C. Gatebe (NASA (GSFC)

j.muller@ucl.ac.uk

MISR vs CAR for all spectral bands





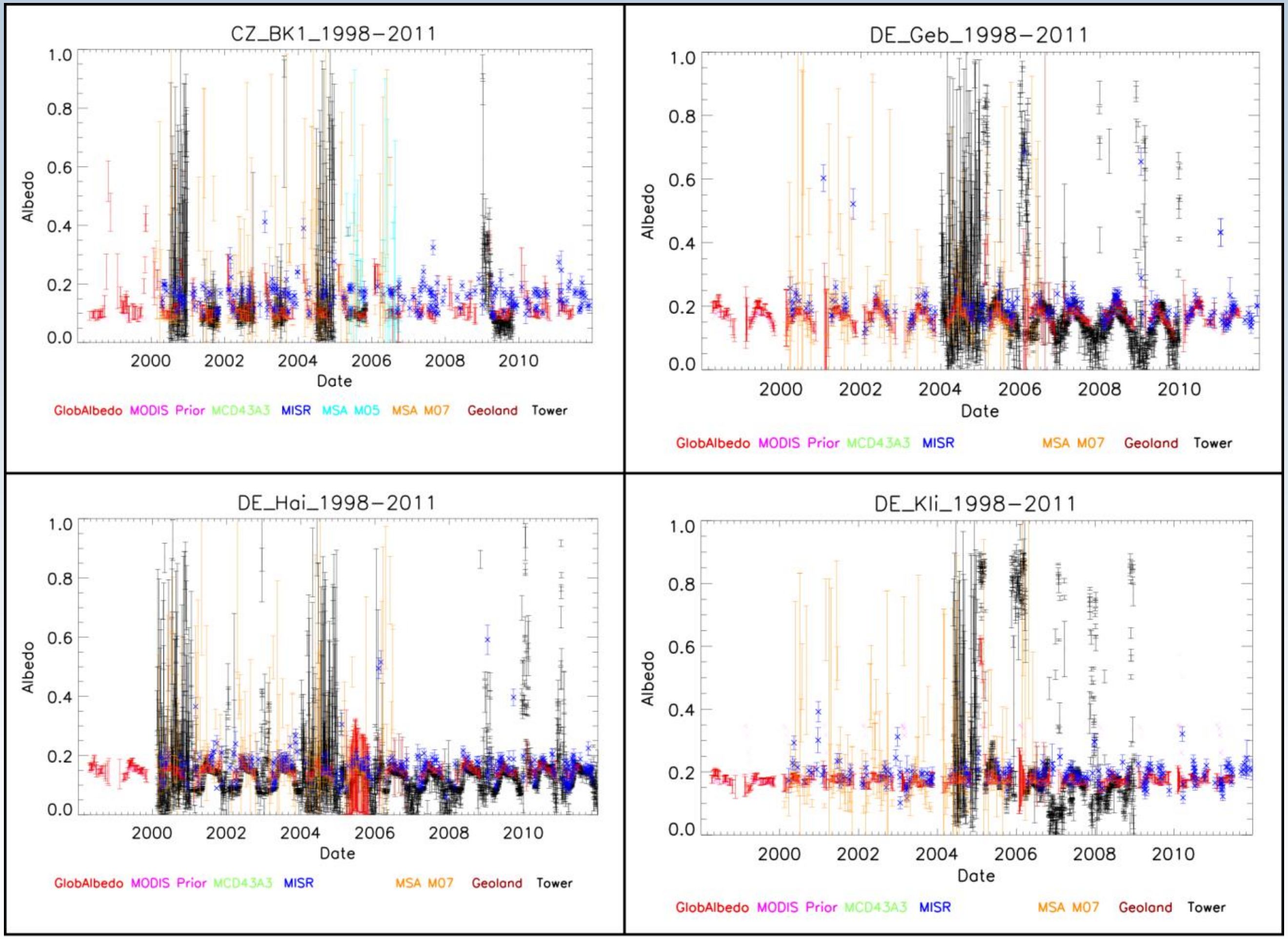
NASA CAR data courtesy of C. Gatebe (NASA (GSFC)

European FLUXNET/BSRN test sites (19 FLUXNET, 1 BSRN)

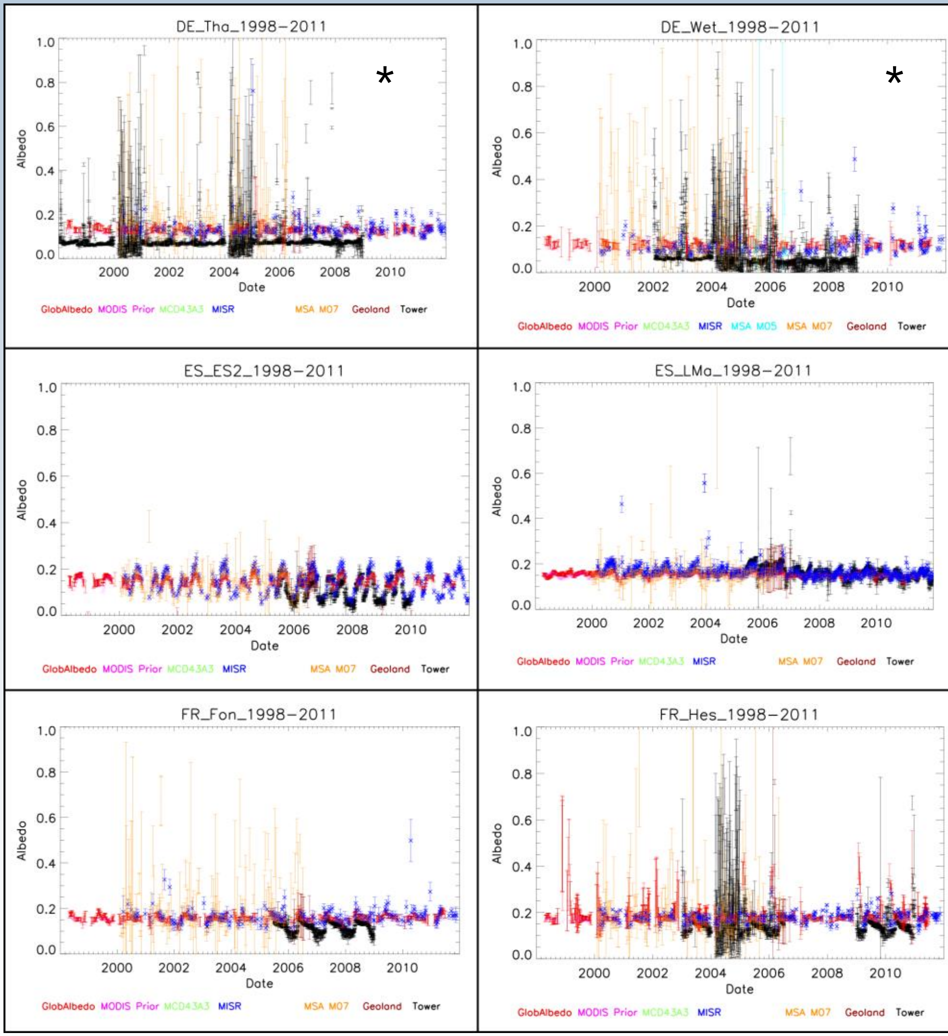
N.B. Open/fair sites which are homogeneous at ≤3km

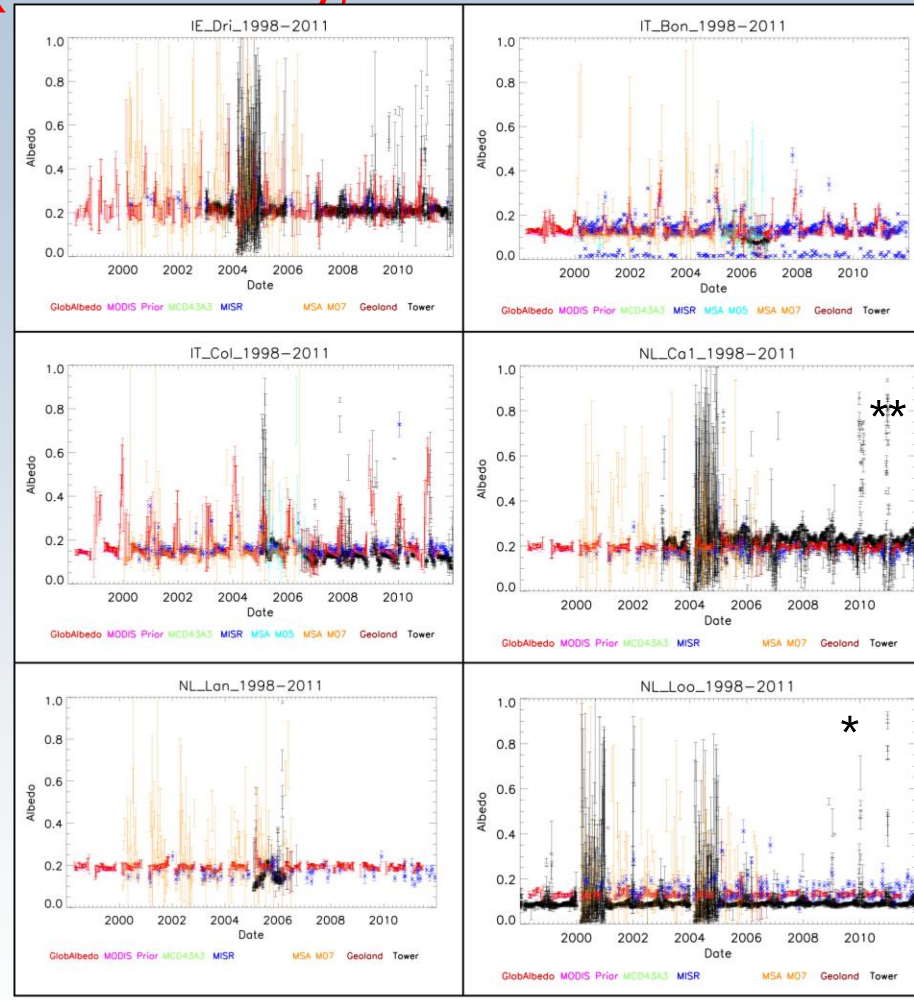
Europe – Validation (1998-2011), Part 1

MISR>>
other EO
& tower



Europe – Validation (1998-2011), Part 2

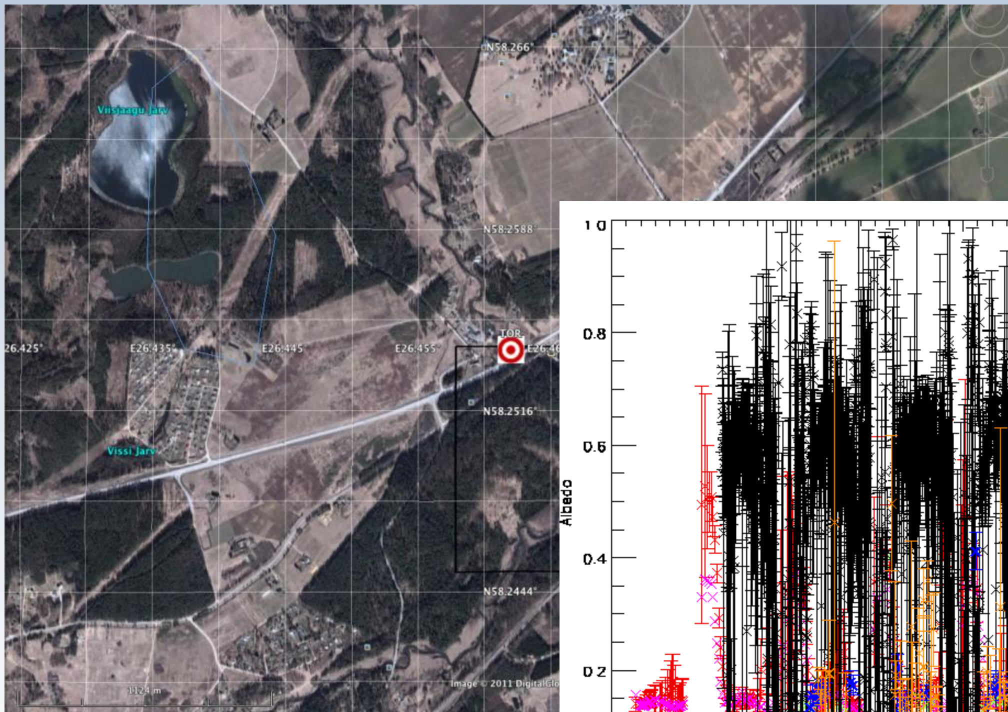


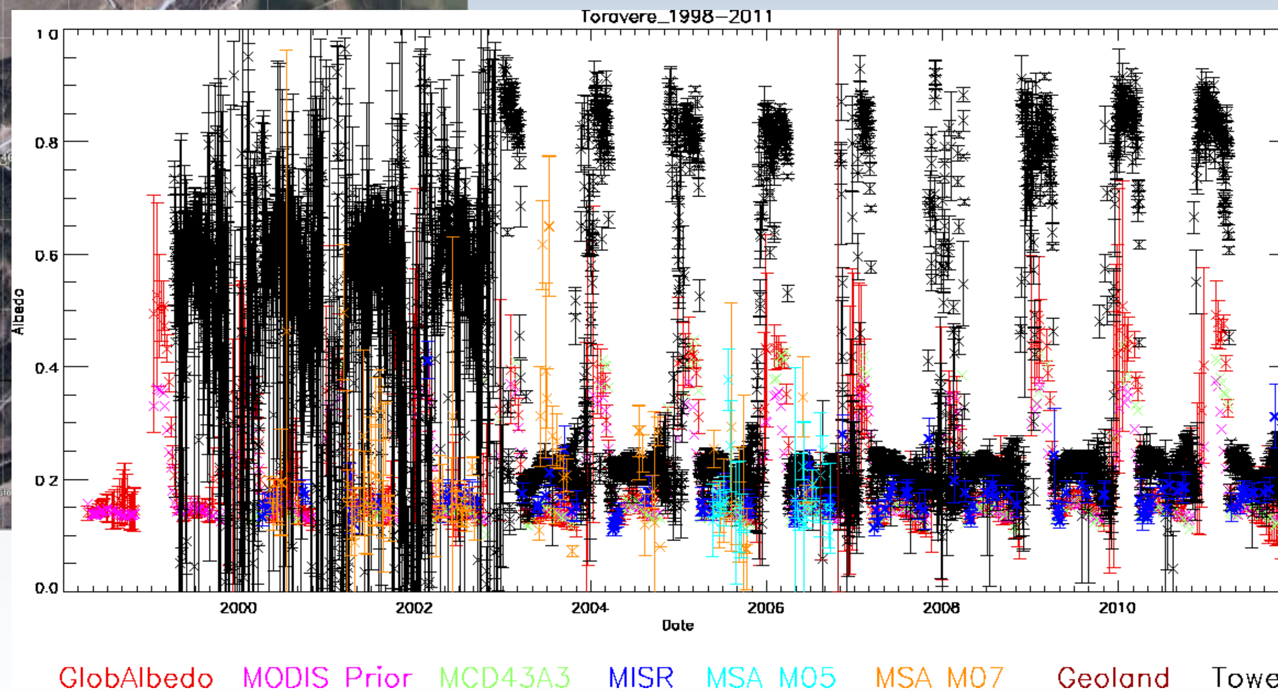


* All EO higher than tower

** Tower albedos higher than EO

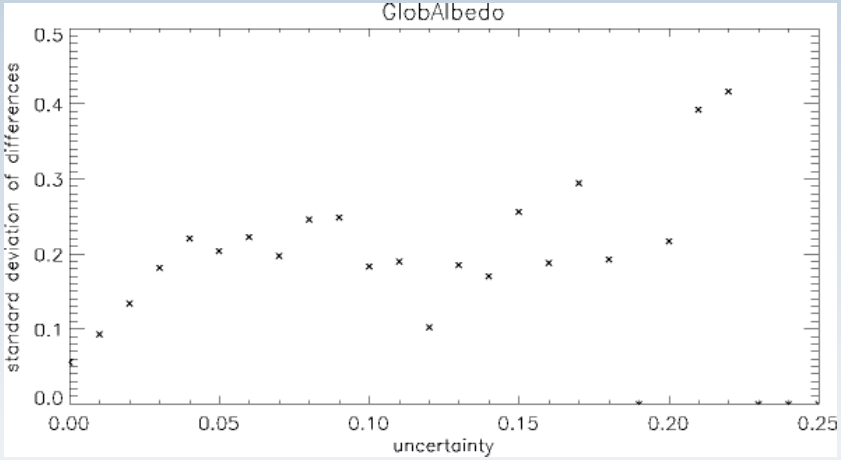
BSRN Toravere (2m footprint)



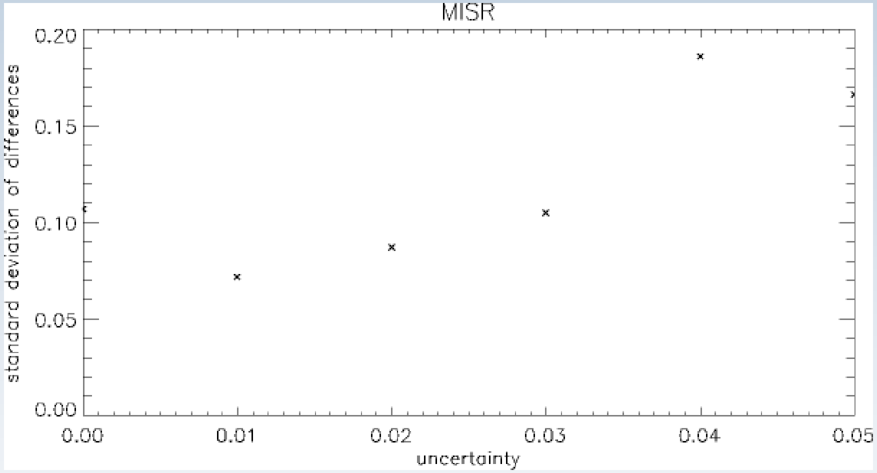


N.B. Very noisy tower albedometer data, much higher values from tower cf all other EO values

GlobAlbedo & MISR uncertainties vs Standard Deviations with Tower Blue-Sky albedo measurements

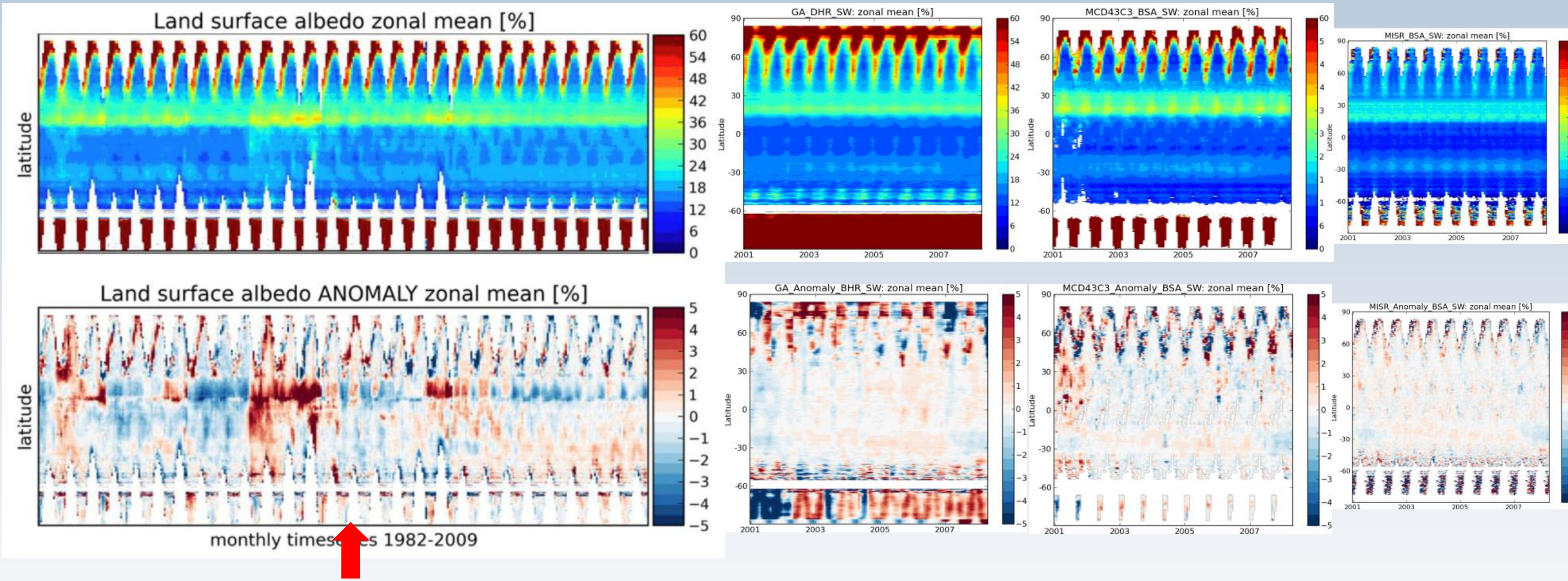


GlobAlbedo



MISR

Hovmöller Plots of CLARA-SAL vs EO-derived DHR Albedos



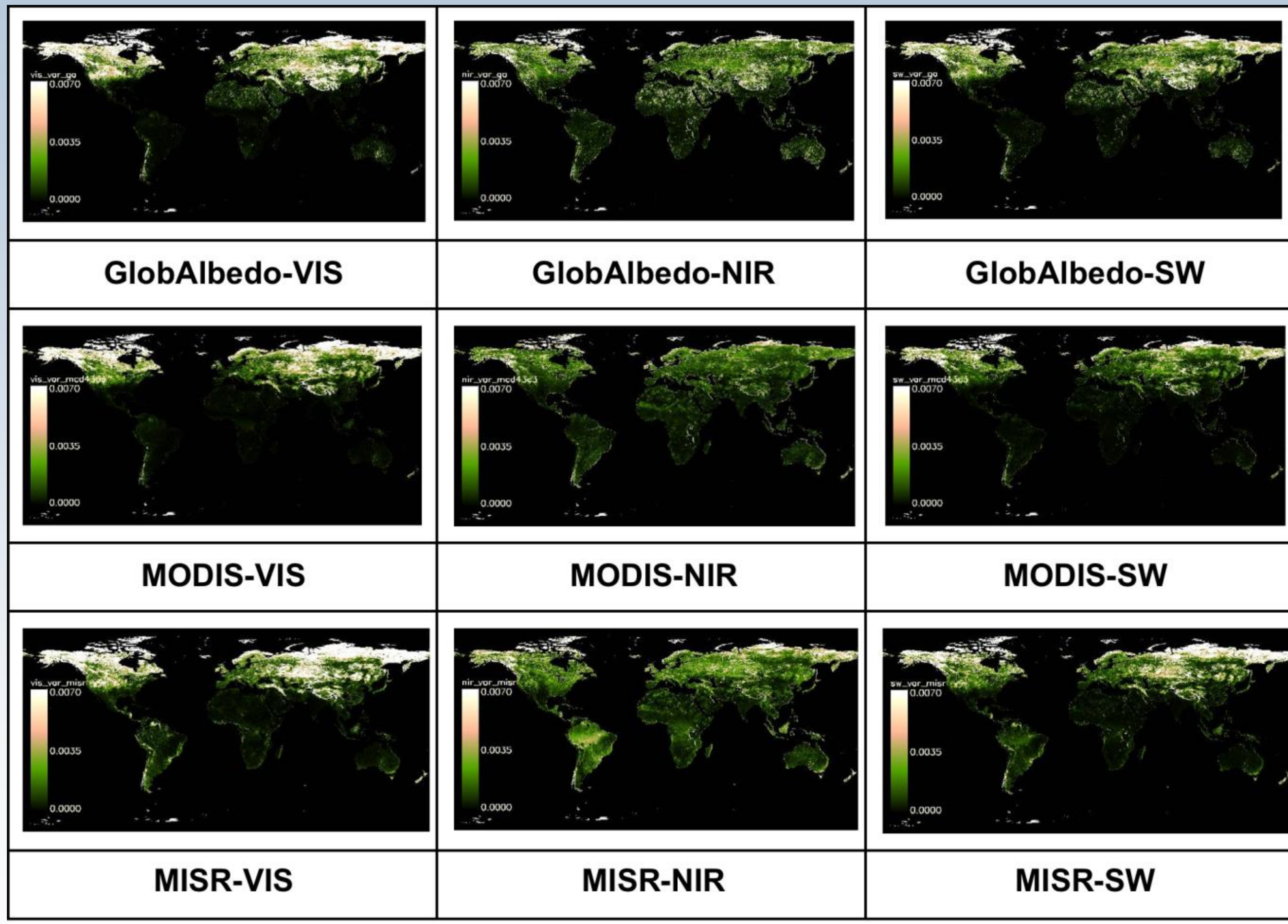
AVHRR - CLARA/SAL
Courtesy of Alexander Loew, MPI

GlobAlbedo

MODIS

MISR

Triple Collocation Variance maps



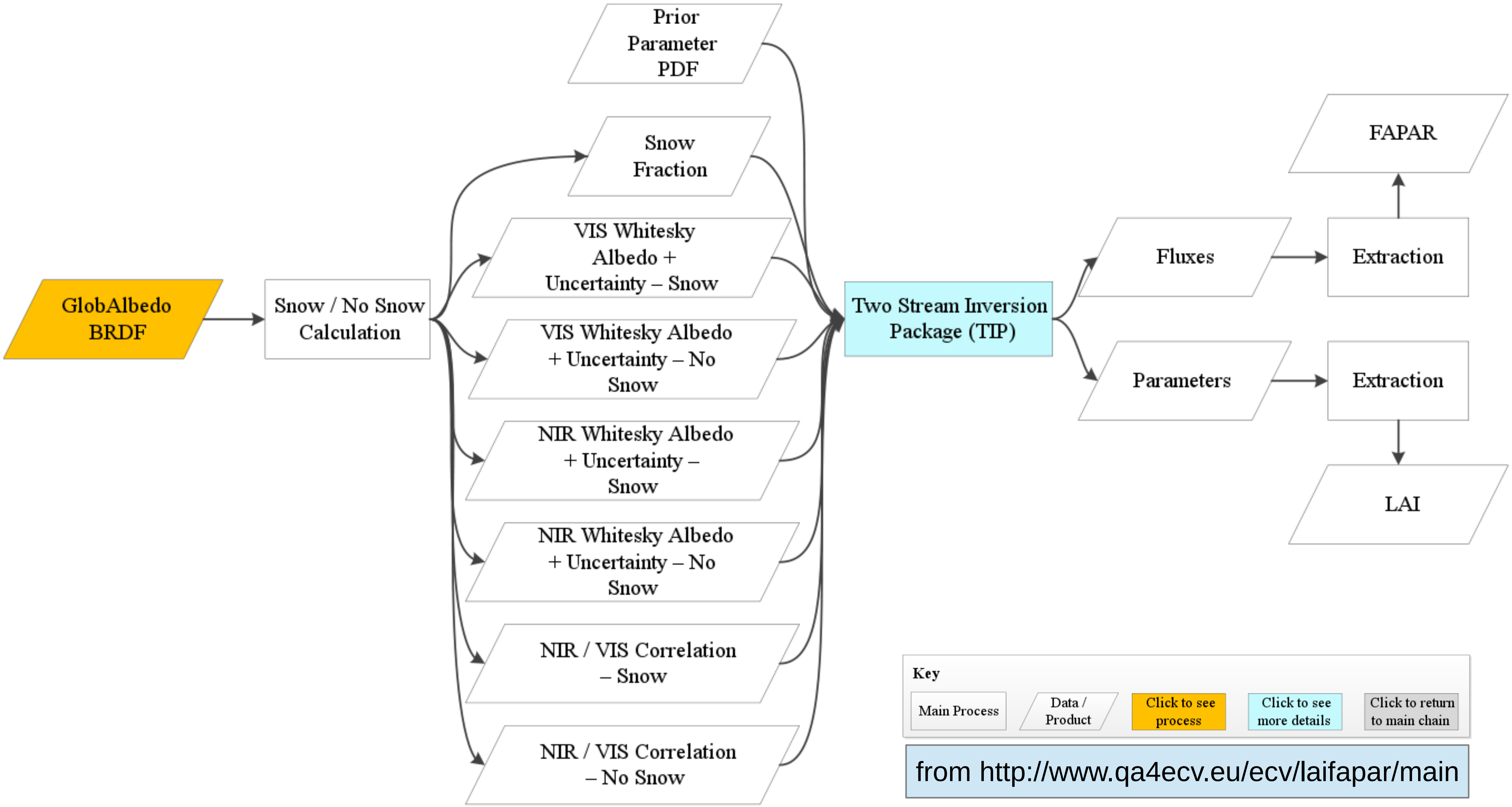
GlobAlbedo variance with MISR and MODIS

Overview LAI/fapar Space Products

Projects/Institution Sensors/Period	Input data	Output product	Retrieval Method	References
JRC-FAPAR SeaWiFS ESA MERIS (07/97-04/12)	Top of Atmosphere (TOA) BRFs in blue, red and near-infrared bands	Daily Instantaneous green FAPAR based on direct incoming radiation	Optimization Formulae based on Radiative Transfer Models	Gobron et al (2000, 2006, 2008)
NASA MODIS LAI/FPAR (00-on going)	Surface reflectance in 7 spectral bands and land cover map.	8-days FAPAR with direct and diffuse incoming radiation	Inversion of 3D Model versus land cover type with backup solution based on NDVI relationship)	Knyazikhin et al. (1998b)
NASA MISR LAI/FPAR (00-on going)	Surface products BHR, DHR & BRF in blue, green, red and near-infrared bands + CART	8-days FAPAR with direct and diffuse incoming radiation.	Inversion of 3D Model versus land cover type with backup solution based on NDVI relationship)	Knyazikhin et al. (1998a)
GLOBCARBON	Surface reflectance red, near infrared, and shortwave infrared	Instantaneous FAPAR (Black leaves)	Parametric relation with LAI as function as Land cover type.	Plummer et al. (2006)
CYCLOPES VEGETATION	Surface reflectance in the blue, red, NIR and SWIR bands	FAPAR at 10:00 solar local time	Neural network based on 1D model	Baret et al (2007)
JRC-TIP MODIS/MISR (00-On going)	Broadband Surface albedo in visible and near-infrared bands.	8-(16) days Standard FAPAR or/& Green FAPAR for direct or/& diffuse incoming radiation	Inversion of two-stream model using the Adjoint and Hessian codes of a cost function.	Pinty et al. (2007)
GEOLAND2/GLS VEGETATION/PROBA-V (99-2012/on going)	Normalized surface reflectance in red and near-infrared bands	FAPAR at 10:00 solar local time	Neural network based on CYCLOPES and MODIS products	Baret et al (2010)

FastOpt

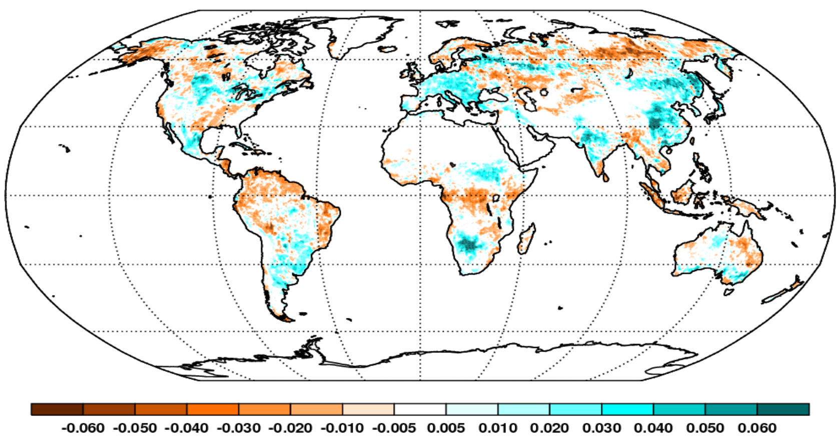
TIP fapar/LAI Processing chain

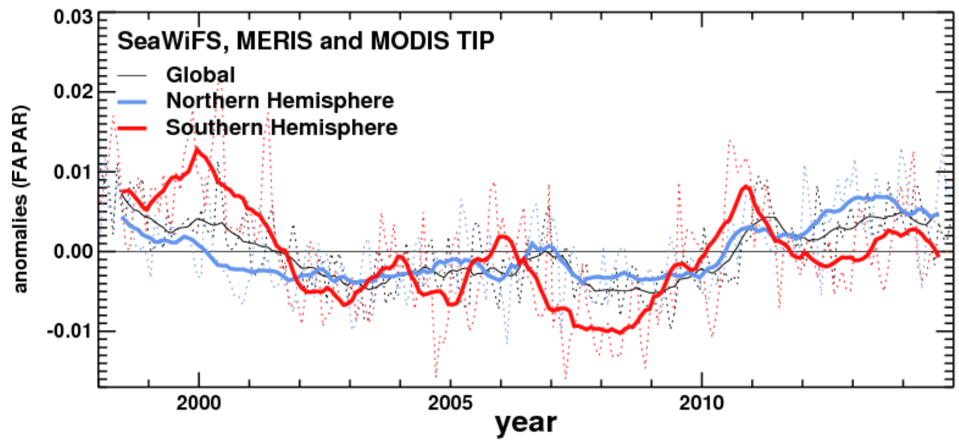


from http://www.qa4ecv.eu/ecv/laifapar/main

LAI/FAPAR/ALBEDO are linked !

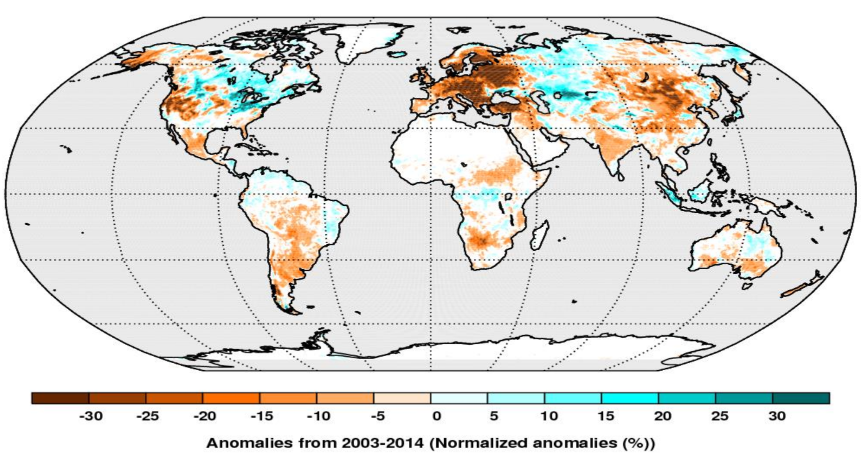
Fraction of Absorbed Photosynthetically Active Radiation

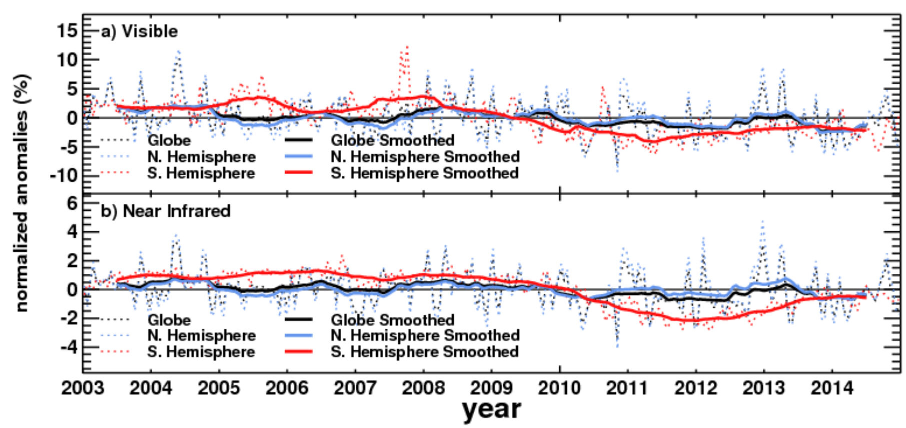




Gobron, 2015: [Global Climate] Terrestrial vegetation dynamics during 2014 [in "State of the Climate in 2014"]. Bull. Amer. Meteor. Soc, in print.

Land Surface Albedo in the Visible





Pinty & Gobron, 2015: [Global Climate] Land surface albedo dynamics 2014 [in "State of the Climate in 2014"]. Bull. Amer. Meteor. Soc, in print.

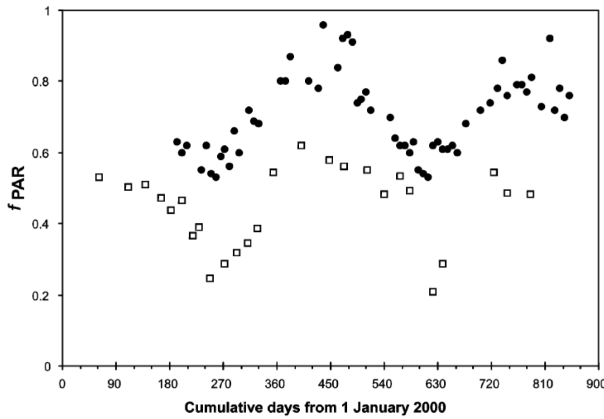
Inter-comparison with similar products: local scale.

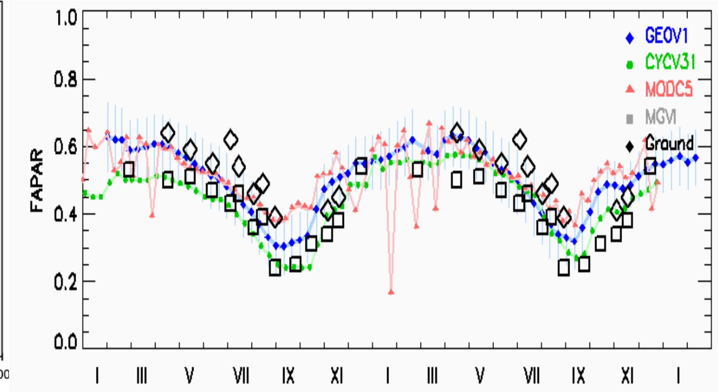
Scale	What do we get?
Global/Continental 10-days/monthly Long time series	+: Provide information on products stability and performance when same retrieval algorithm is used with different sensors. +: fast check of difference of products (spatial and seasonal) -: aggregation method and time composite may be different
Regional scale at nominal resolution (~1km)	+: Provide information, if disagreements, on quality of input data or/and pre-processing step. -: remapping method, geo-reference.
Site Level at nominal resolution (~1km) - Daily	+: Provide information on products stability and performance. +: Provide accuracy only with 'validation' step information

EO Validation

Mongu: Shrubland/woodland

Camacho et al., 2011



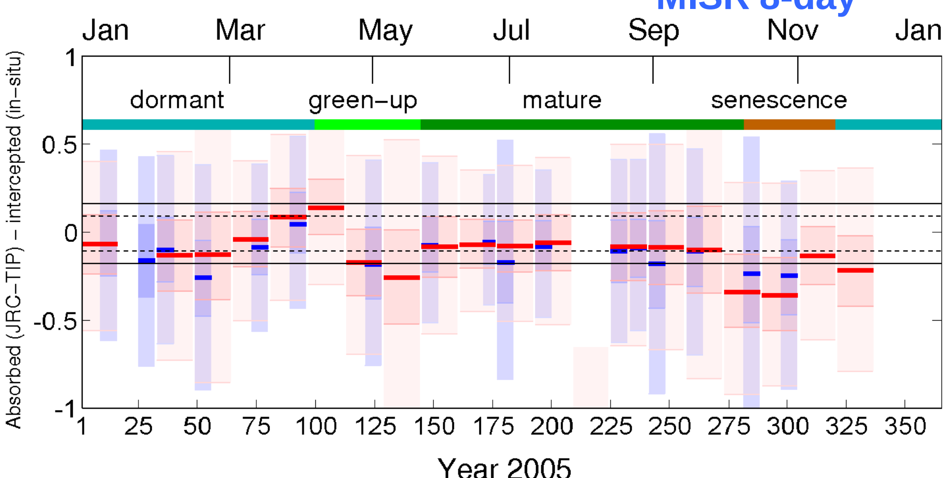


Hummerich et al., 2005

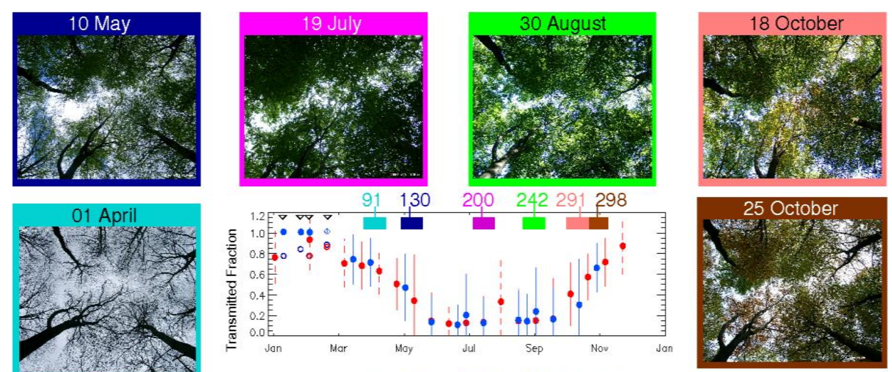
Figure 1. f_{PAR} values for the Mongu site over time. f_{PAR} data from ground measurements (□) and derived from MODIS observations (●).

MODIS 16-day

MISR 8-day

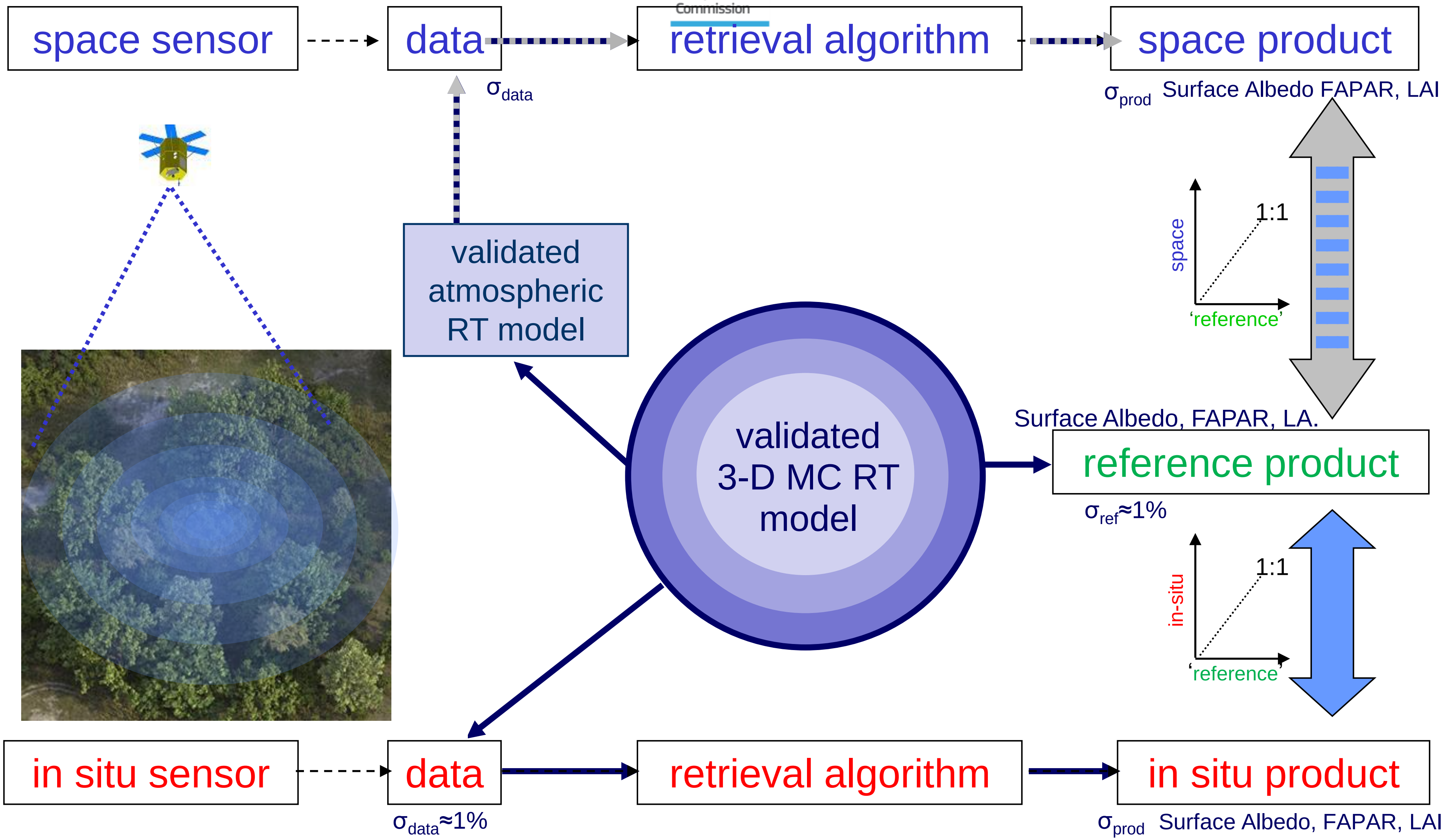


Hainich: deciduous forest



Pinty et al., 2011

QA4ECV Framework



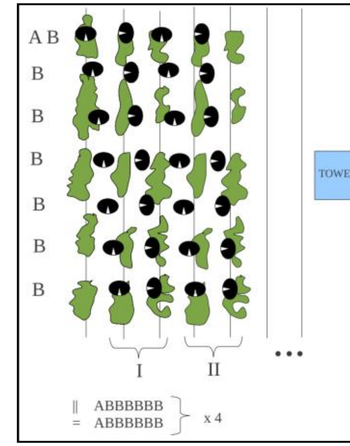
Model-Based Approach

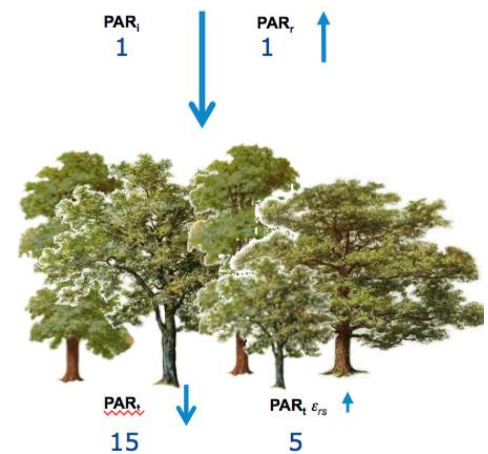
European Commission

A series of 'virtual' validation sites are being constructed using specific scenes (spectral values of each elements, 3-D architecture etc …)

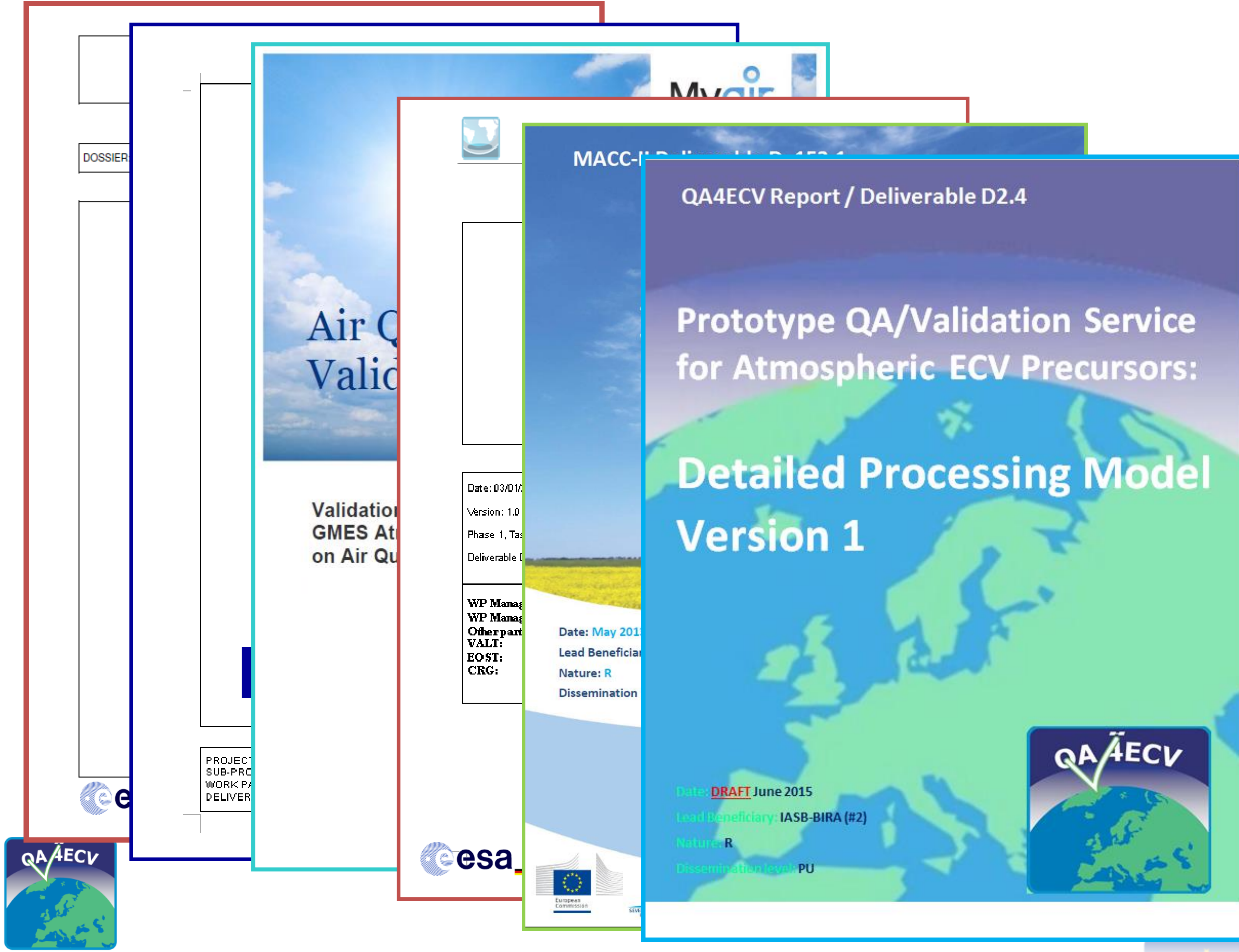
1-Simulate Top Of Atmosphere Bidirectional Factors for several sensor/satellite data
→ **Benchmark land variables retrieval algorithms**

2-Reproduce various measuring protocols of in-situ measurements
→ **Assess error budget of in-situ products**





Joint Research Centre



j.muller@ucl.ac.uk

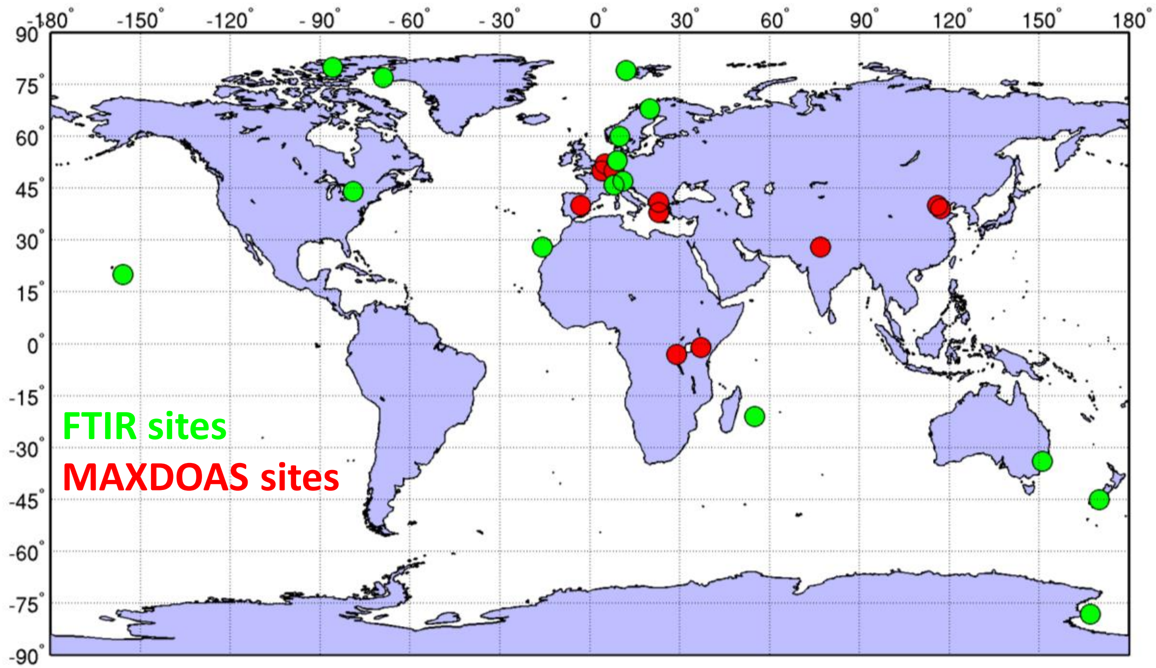
Confrontation with atmospheric reference data

- 1st choice: Ground-/balloon-/aircraft-based measurements of documented quality
  - Monitoring networks with official measurement and QA protocols: WMO GAW, NDACC, TCCON, SHADOZ, EARLInet…
  - Dedicated campaigns: SAUNA, AVE, CINDI, AROMAT…
- Other satellite data, of documented quality
- Modelling support
  - Knowledge of atmospheric context, better interpretation
  - Detection of patterns, striping, internal inconsistencies…

Reference data for atmosphere in QA4ECV

What is needed?

- Quality-assured long-term reference data for NO_2, HCHO, and CO column measurements in the troposphere
- This includes traceable metrics on their precision, accuracy, (horizontal) representativity, and preferred use



j.muller@ucl.ac.uk

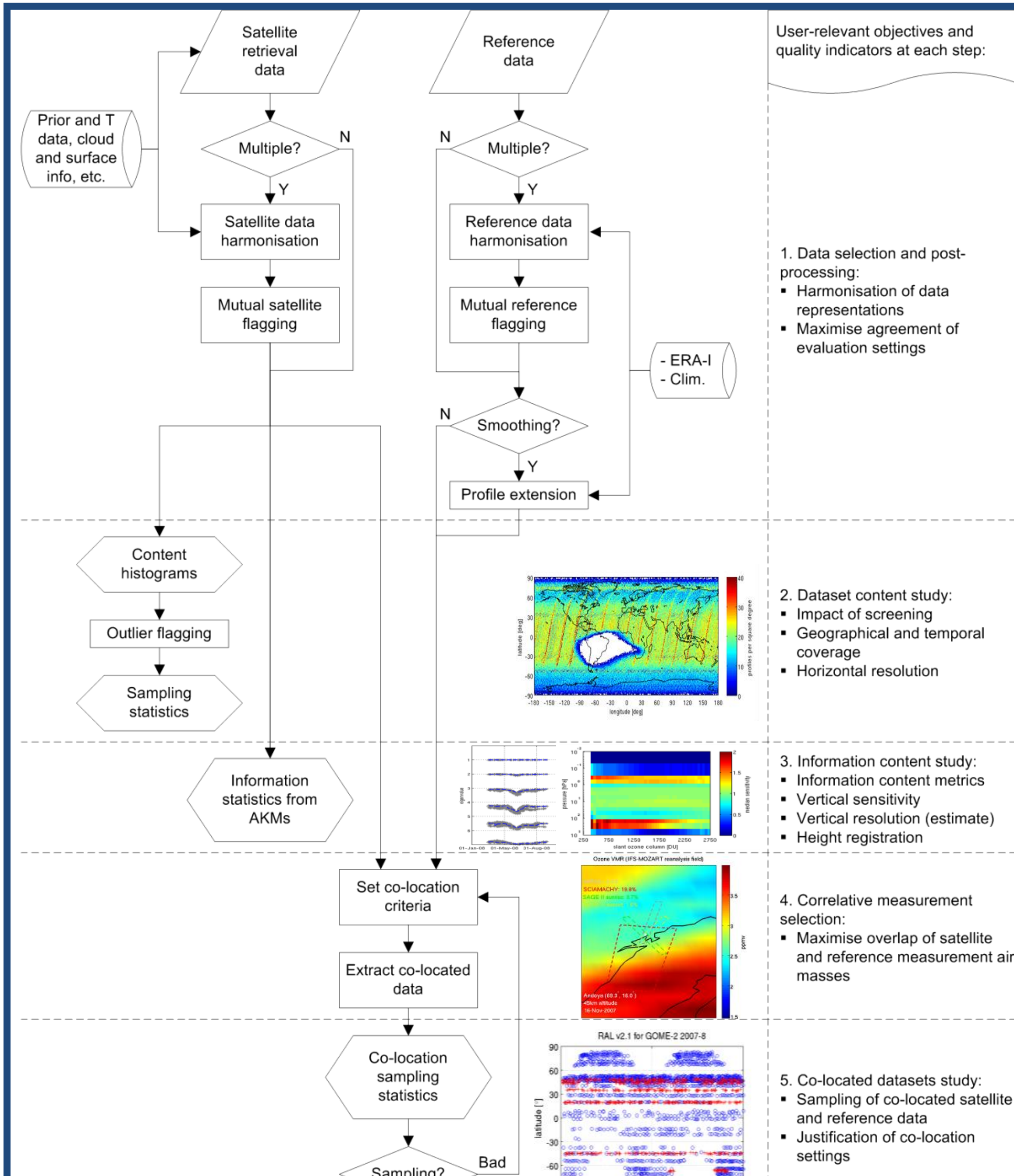
Standardisation of retrieval methods and error analysis

- Common algorithm selection → standardized MAXDOAS and FTIR retrieval methods. Ideally processing should be centralised (beyond QA4ECV, but in progress)
- Common data quality indicators and data flagging
- Verification of consistency between instruments, based on regular campaigns and use of travelling standards
- Fully traceable algorithms description and error characterization → ATBDs
- Characterization of spatial representativeness of measurement, desirable for each site and reported trace gas (still to be developed)

Chain of QA / Validation process

1. Translation of user requirements to validation requirements
2. Satellite data selection, filtering and post-processing
3. Data content study (DCS) of satellite dataset
4. Information content study (ICS) of satellite dataset
5. Selection & characterisation of correlative reference data
6. Identification and characterisation of co-located data pairs
7. Homogenization: Resampling, smoothing, and conversions of representation systems and units
8. Data comparisons: bias, spread, stability, dependences…
9. Derivation of appropriate Quality Indicators
10. Discussion of compliance with user requirements

Traceability chain of the validation process



Keppens *et al.*, AMTD 2014

Conclusions & Future Work

- Conclusions:
  - Protocols are being developed by CEOS-WGCV for land and atmospheric EO observational product validation
  - QA4ECV is developing new paradigms for validation of "in situ", airborne and EO derived ECVs including validation of algorithms as well as products
- Future Work:
  - Develop protocols for validation of all 6 ECVs in QA4ECV
  - Assess the potential of these validation methods for other ECVs
  - Assess the use of triple collocation and Hovmoeller decompositions to hunt for systematic errors