High-level algorithm and product harmonization for upcoming global imaging spectroscopy missions

Phil Townsend, University of Wisconsin / Caltech Jet Propulsion Laboratory

Contributors:

People with whom I have discussed this topic in the past several years!

- Kerry Cawse-Nicholson (JPL)
- Michelle Gierach (JPL)
- Adam Chlus (JPL)
- Phil Brodrick (JPL)
- David Thompson (JPL)
- Rob Green (JPL)
- Chip Miller (JPL)
- Ben Poulter (NASA GSFC)
- Marco Celesti (ESA)
- Jens Nieke (ESA)
- Jenny Adams (U. Zurich, previously ESA)
- Michael Schaepman (U. Zurich)
- Mike Rast
- Dave Schimel (JPL)
- Fabian Schneider (JPL)

CHIME High Priority Prototype Products

DOMAIN	THEMATIC AREA	VARIABLES CHPPP	CHIME Candidate Algorithms
AGRICULTURE / FOOD SECURITY	 Assessment of biophysical and biochemical variables related to the crops and of agronomic interest	Leaf/Canopy Pigment Content	Semi-empirical modelling based on narrow-band vegetation indices; Hybrid methods based on ANN/LUT or other machine learning algorithms applied to vegetation canopy radiative transfer models outputs (e.g. PROSAIL).
		Leaf/Canopy Nitrogen Content	Narrow-band vegetation indices; Hybrid methods based on ANN/LUT or other machine learning algorithms e.g. GPR methods applied to vegetation canopy reflectance models (e.g. PROSAIL).
		LAI	
		Canopy Water Content	
		Leaf/Canopy Pigment Content	
		Leaf Mass/Area	
	 Topsoil properties	Soil organic carbon content	Chemometrics modelling (e.g. PLSR); Spectral analysis; Spectral indices; Machine learning (e.g. Random Forest)
		Soil texture (clay, silt, sand)	
GEOLOGY & MINERALS	 Raw material detection	Mineral identification / classification (Kaolinite, Smectite, Jarrosite, Dolomite)	Sub-pixel linear unmixing
		Hematite – Goethite distribution	
		Ferric oxide content	
		Kaolin Cristallinity	

SBG Key Product Suites

SBG Algorithm Class	SBG Algorithm Products (examples)
CORE Algorithms	
Earth Surface Temperature and Emissivity	Land Surface Temperature* and Emissivity
VSWIR Reflectance	Land and Water Reflectances, BRDF Corrections, Albedo
Cover Classifications	Cloud, Water, Land Cover, Plant Functional Types, etc.
PRODUCT Algorithms	
Terrestrial Ecosystems	
Vegetation Traits	Nitrogen, LMA, Chlorophyll, Canopy water
Evapotranspiration	ET*, Evaporative stress index
Proportional Cover	GV, NPV, Substrate, Snow/Ice, Burned Area
Geology/Earth Surface	
Substrate Composition	Mineral type*, Fractional abundance*, Soil types and constituents
Volcanic Gases and Plumes	SO2, Volcanic ash
High Temperature Features	Volcanic temperature anomalies (lava temperature), Forest fires
Aquatic and Coastal Ecosystems	
Water Biogeochemistry	Pigments, CDOM, Suspended particulate matter
Water Biophysics	Diffuse light attenuation, Inherent optical properties, Euphotic depth, PAR
Aquatic Classification	Phytoplankton functional types, Floating vegetation, Benthic cover, Wetlands
Snow and Ice	
Snow albedo	Albedo, Grain size, SSA, Light absorbing particles, Fractional cover

Additional Products

- Wildfires
- Greenhouse gases
- Biodiversity (α, β)
- Vegetation species composition

**Leverages ECOSTRESS and EMIT algorithms = already "Global" → many (or most) of the remaining products do not have algorithms that have been validated or are mature for global application.*

Crosswalk-Vegetation

CHIME

- Core Products
  - Surface Reflectance

SBG

- Core Products
  - Surface Reflectance
  - Land Cover
    - Plant functional Types (PFTs)

- Vegetation → ***Focus Agriculture***
  - Leaf / Canopy Nitrogen Content
  - Leaf / Canopy Pigment Content
  - Leaf Area Index
  - Canopy Water Content
  - Leaf Mass per Area (LMA, SLA)

- Vegetation → ***Focus Natural Vegetation***
  - Foliar Nitrogen
  - Chlorophyll
  - Canopy Water Content
  - Leaf Mass per Area (LMA, SLA)
  - Fractional Cover (GV, NPV)
    - Substrate, Soils, Water, Snow, Ice
  - Biodiversity Metrics

Crosswalk-Soils and Geology

CHIME

- Soils → *Focus agriculture*
  - Soil Organic Carbon
  - Soils Texture
    - Sand, Silt, Clay (mixes)
- Geology
  - Mineral Identification / Classification
  - Hematite / Goethite Distribution
  - Ferric Oxide Content
  - Kaolin Crystallinity

SBG

- Soils
  - Soil Types and Constituents
    - Soil Texture (sand, silt, clay)
- Geology
  - Mineral Type
  - Mineral Type Fractional Abundance

*Snow, Methane and Aquatic/Coastal Biochemistry/Physics/Classification likely have similar crosswalks but are not explicitly listed as CHPPPs.

Assessment of Maturity (SBG Key Products)

PRODUCT	MATURITY	GREATEST NEED
Snow products	High	In situ data in glaciers and below-canopy snow
Evapotranspiration	High	Data fusion and improved latency
High Temp Features	High	High spatial resolution (<5 m) thermal data over lava
Substrate Composition	High (minerals)	VSWIR [mature] / TIR [low] fusion
Proportional Cover	Medium	Complimentary combination of algorithms from different fields
Volcanic Gas&Plumes	Medium	Improvements in computational efficiency
Water Biogeochem	Medium	Analysis of applicability and compatibility of PACE algorithms for coastal and inland waters, at SBG GSD
Vegetation Traits	Low	Global in situ and remote sensing data
Substrate Composition	Low (soils)	Global in situ and remote sensing data
Water Biogeophysics	Low	In situ water column data
Aquatic Classification	Low	Global datasets; build upon biogeochem & biogeophysics products to produce applications-ready data

Assessment of Maturity (exploratory products)

PRODUCT	MATURITY	SUMMARY
Wildfires	High	In-situ/validation data
Glint	Medium	Sky-glint vs. sun-glint, need aerosol optical thickness, cannot be entirely removed through tilt, leverage PACE, spatial resolution considerations, validation needed
Greenhouse gases	Medium	In-situ/validation data, significant dependency on instrument characteristics
Biodiversity	Low	Global field and airborne data, consistent data acquisition standards, seasonal spectral libraries, leaf-to-canopy scaling, angular dependence
Plant species	Low	Community composition, or “oak-ness” more achievable. Traceability between other products and species composition, uncertainty quantification

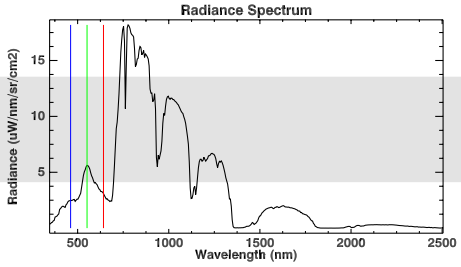
Perspectives on Harmonization

- Current algorithms will likely be superseded for most if not all products
- Machine learning (ML) algorithms will likely be employed for many algorithms to generate products rapidly at global scale
  - ML to emulate physically-based and even data-driven approaches
- How to harmonize?
  - Harmonize the input data – use common models [resample the data: spectroscopy purist's nightmare]
  - Harmonize (tune) the models to differences in input data [resample the model: multiple flavors of algorithms, *goal: sensor-agnostic models*]
    - Retain spectroscopic [reflectance] data at all levels, resample at end
  - Harmonize the products post-hoc [LSH approach]
    - Raises many questions (base data, nonlinearities, etc.)

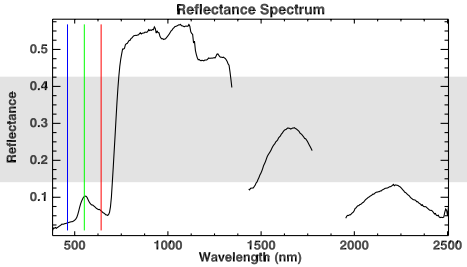
Perspectives on Harmonization

- We need to be user-aware
- With CHIME, SBG and other missions having global coverage, it is likely that our user base will increase by >1 order of magnitude
- Average users will not be imaging spectroscopy specialists, many will be naïve users
- They will want to download comparable products from each mission and many will assume the data are equivalent (or want it to be)
- And there will be many users who will want to merge surface reflectance data from CHIME and SBG (or other sensors)
  - This will require spectral resampling
  - Decisions we make as a community now can ensure best practices by future users → avoid worst case scenarios

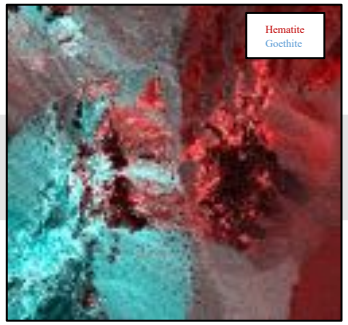
Data Levels



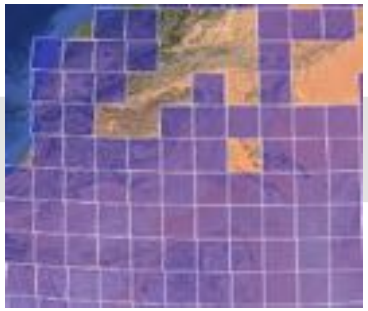
L1b: Radiance at Sensor



L2a: Surface Reflectance (HRDF)

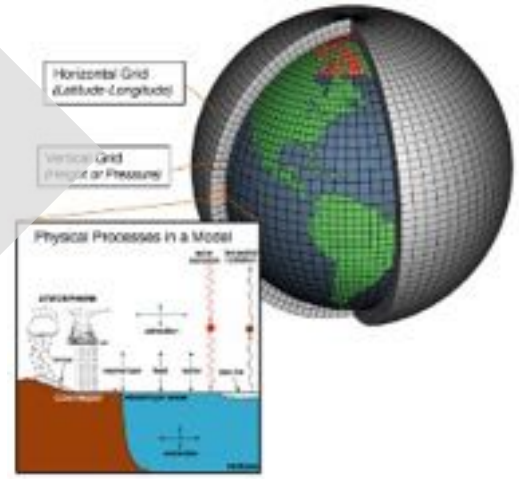


L2b: Mineralogical Maps



L3: Aggregated Mineralogy

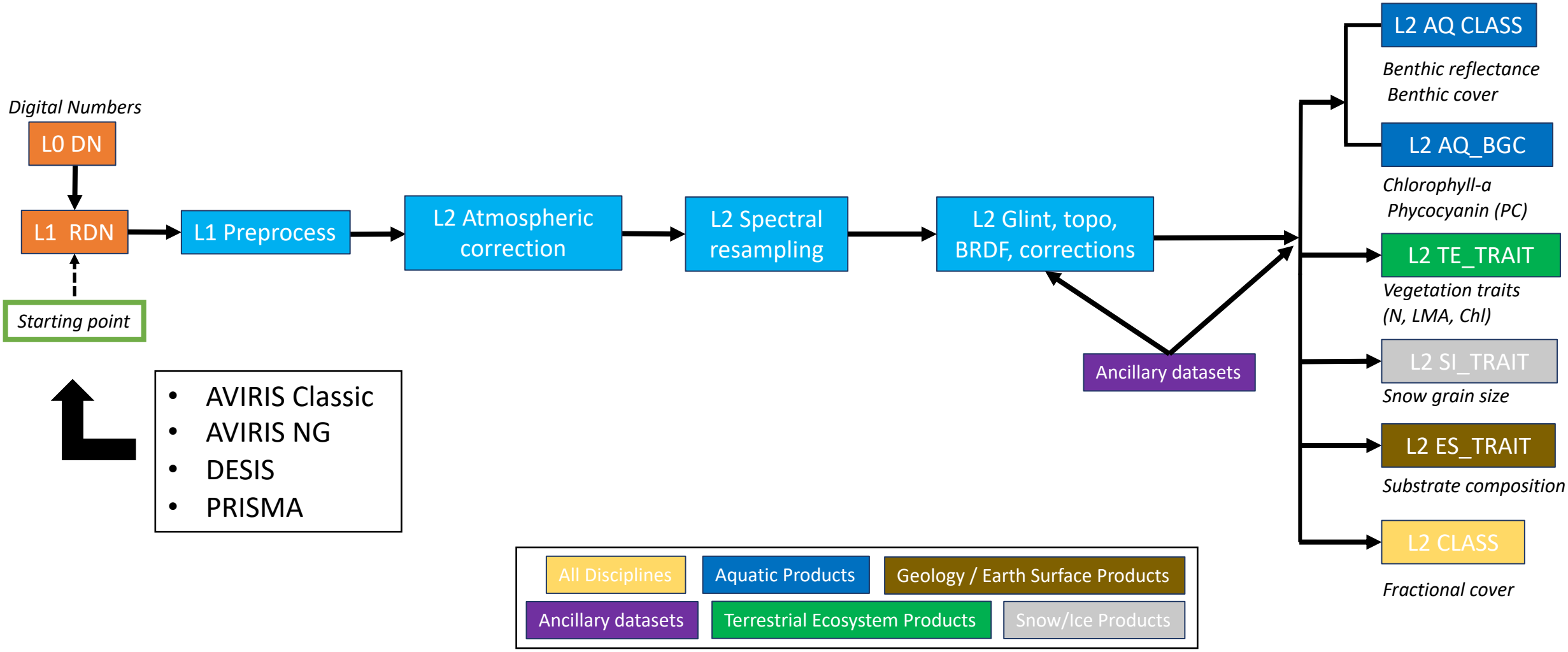
L4: CESM, GISS Model Runs



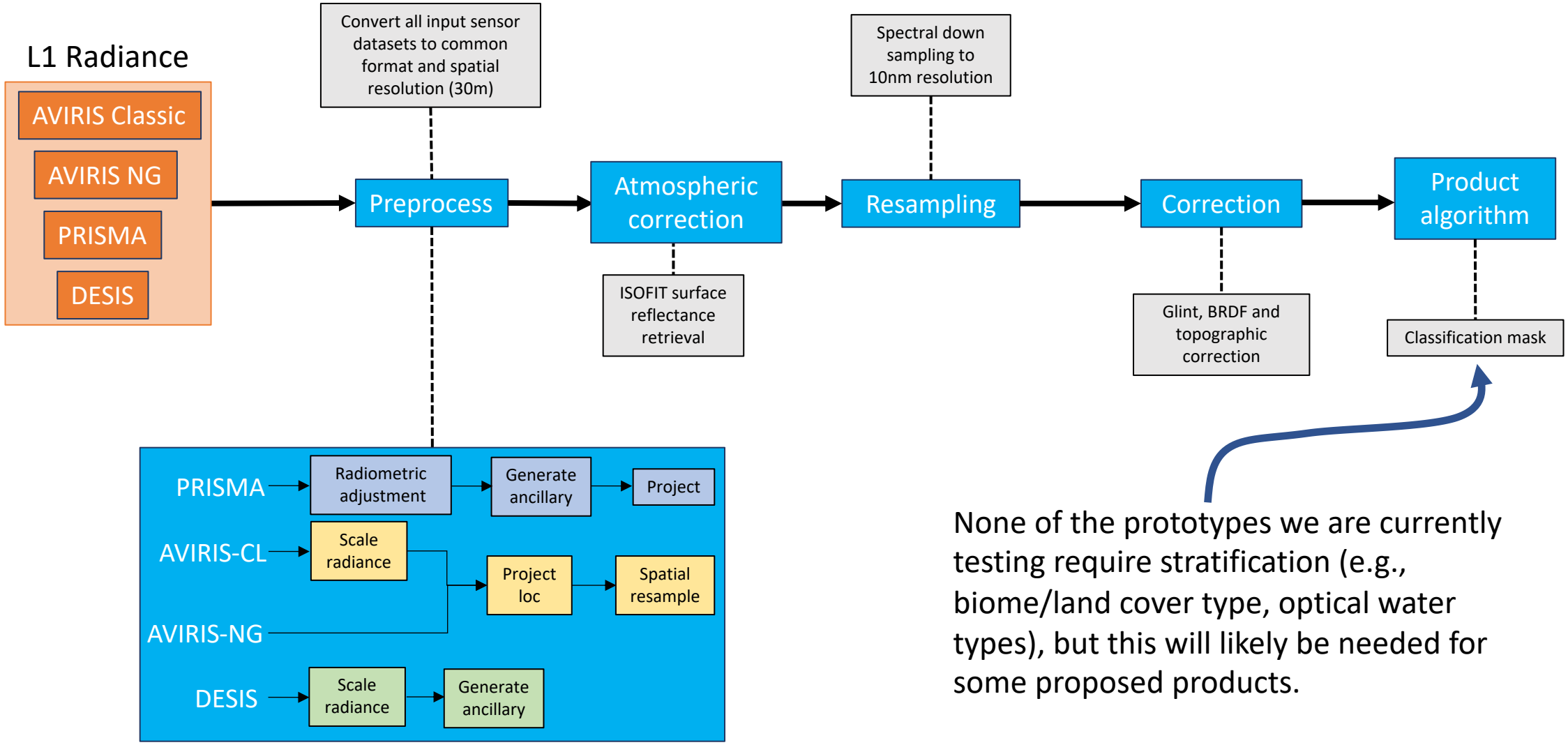
This is where harmonized data reside.

Graphic courtesy of David Thompson (EMIT example)

SISTER VSWIR Prototyping Workflow

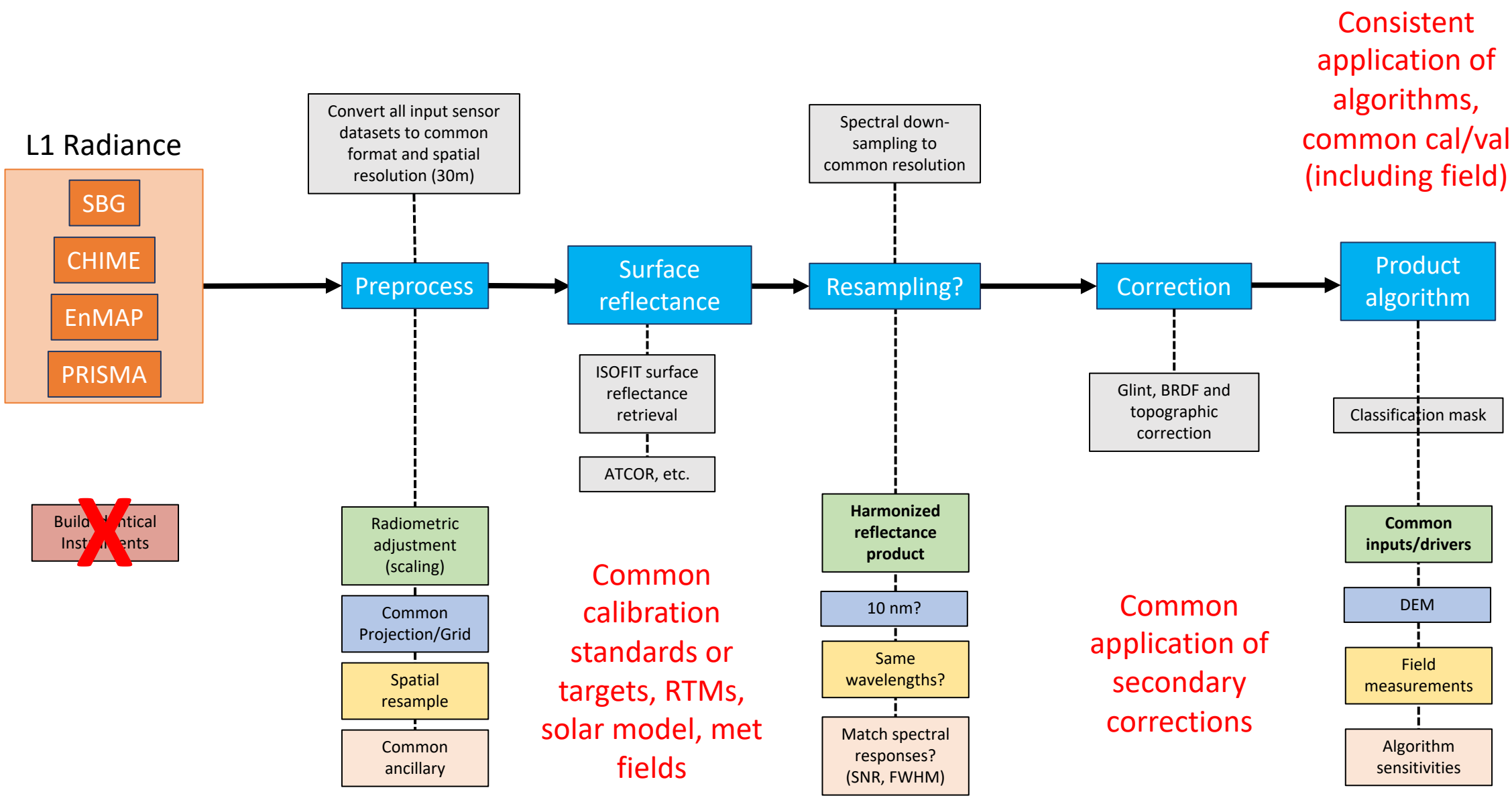


SISTER VSWIR Workflow Production Run #1



None of the prototypes we are currently testing require stratification (e.g., biome/land cover type, optical water types), but this will likely be needed for some proposed products.

Harmonisation framework



Final Thoughts

- Effective harmonization:
  - Modular transparent, open-source algorithm development
    - Objective: sensor-agnostic algorithms
  - Emphasize intercomparison activities in lead-up to ATBDs
  - Maximum use of common inputs:
    - If not possible: have the workflows be adaptable so that common data sets can be tested across platforms
    - Common cal/val/training/ancillary data
    - Common processing workflows
- Consider on-demand processing approaches when these diverge
- Remember the user needs
  - Many users may trust what's "under the hood"