Thinking outside the grid: aggregating data into information

Nadia Smith, Paul Menzel, Elisabeth Weisz

and in conversation with

Bryan Baum, Ralf Bennartz, Andrew Heidinger, Andi Walther,
Shaima Nasiri and Rob Roebeling

CIMSS Science Symposium, UW-Madison, 12/12/12

Smith, N., Menzel, W.P., Weisz, E., Heidinger, A.K. & B.A.Baum (2012) A uniform space-time gridding algorithm for comparison of satellite data products: characterization and sensitivity study. JAMC (in press), online:
http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-12-031.1

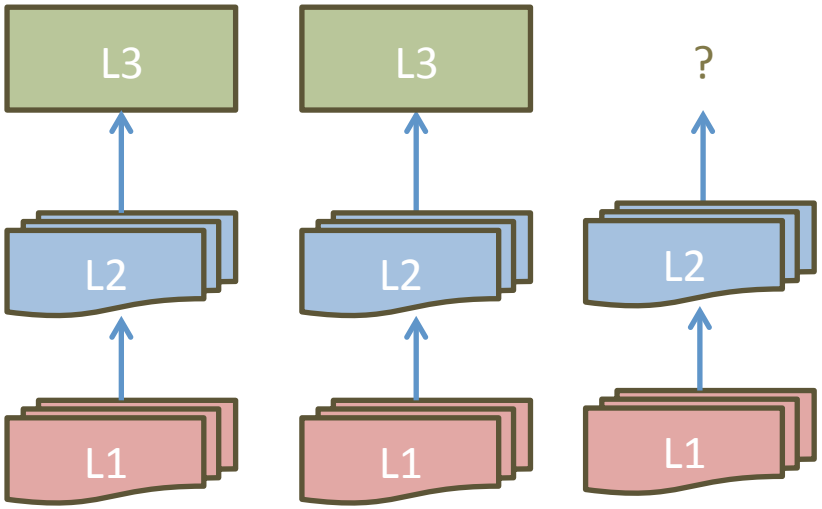
Aggregating Data into Information

- It's not (only) about visualization and data reduction
- Its about performing analysis in a uniform space.
- It's about doing science in a regularized data environment
- It's about a neutralizing instrument differences.
- Composite data products
- Higher level information – environmental monitoring, indicators of change, etc.

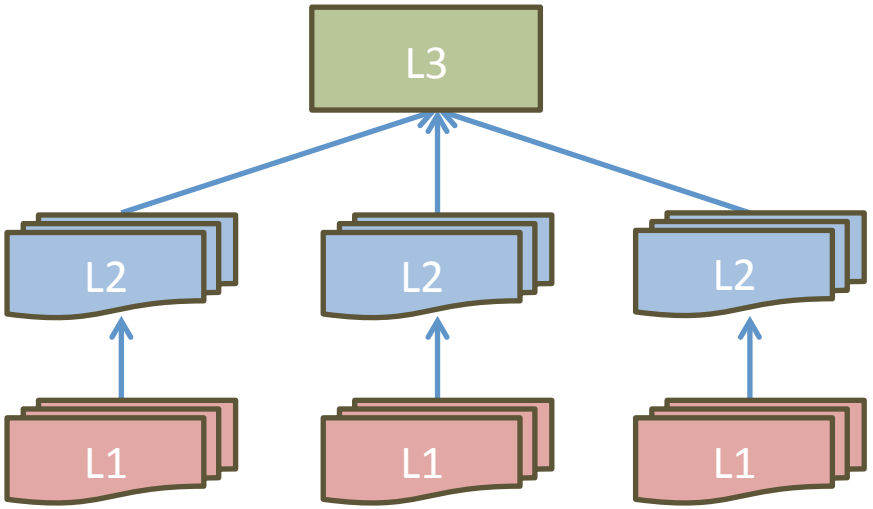
What are the goals of space-time-grid (STG) framework?

- **Reproducibility**
- Instrument independent (sounders, imagers, etc.)
- **Flexibility**; research specific outcomes
- Allowing both linear and non-linear statistical analysis
- **Data exploration** on a global scale
- Fast/**simple** processing of global data
- Support **research** in grid space

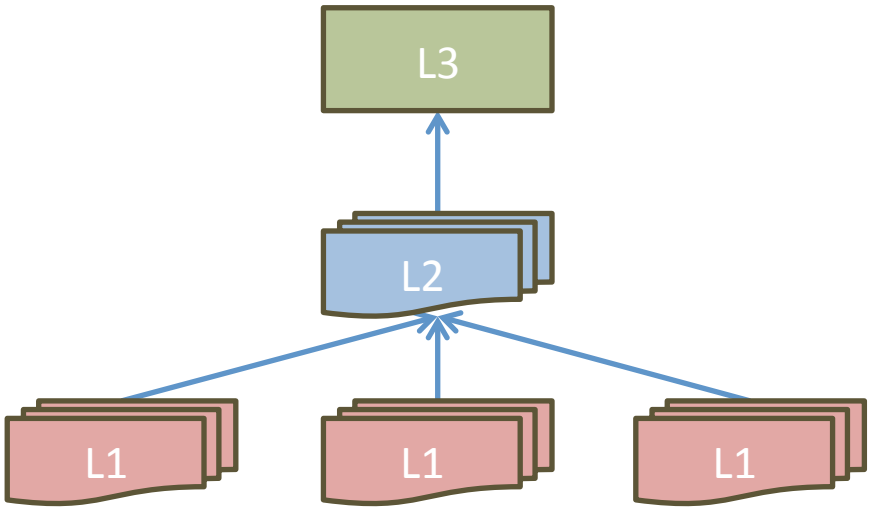
Traditional approach:
Instrument specific algorithms & L2/3 products



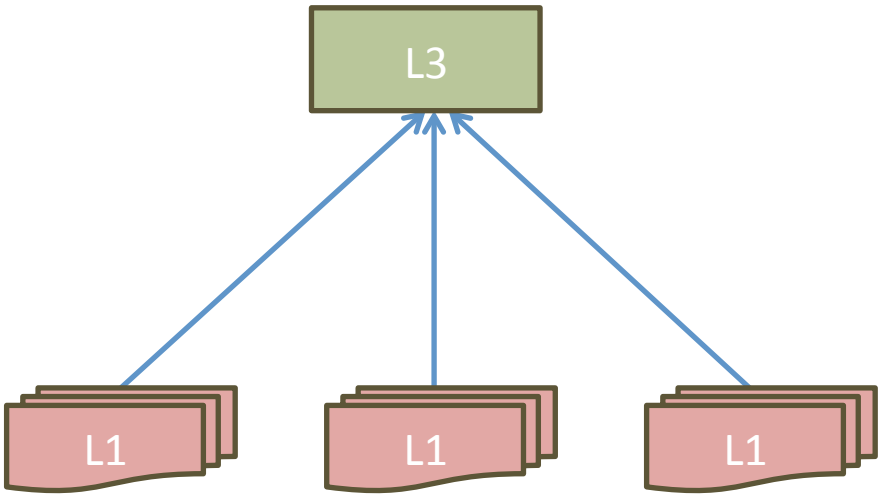
Single/shared aggregation algorithm



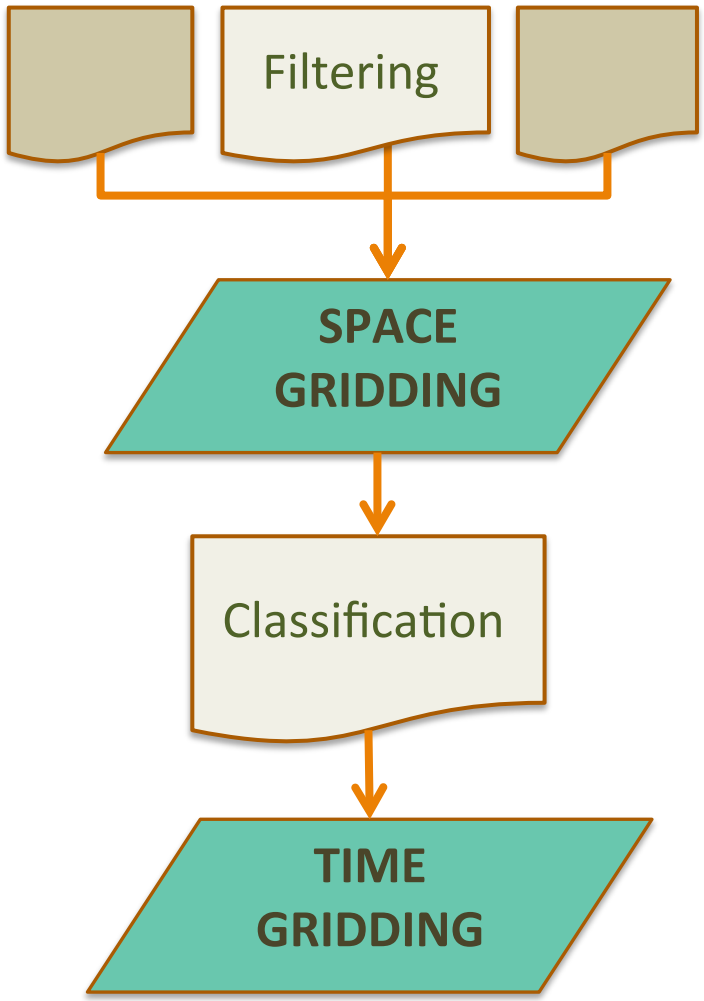
Single/shared retrieval and aggregation



Level 1 to Level 3 directly



Space-Time-Gridding: Data-to-Information flow



Geophysical sub-setting

Lat/Lon Binning,
equal-angle grid
Geo-Histogram

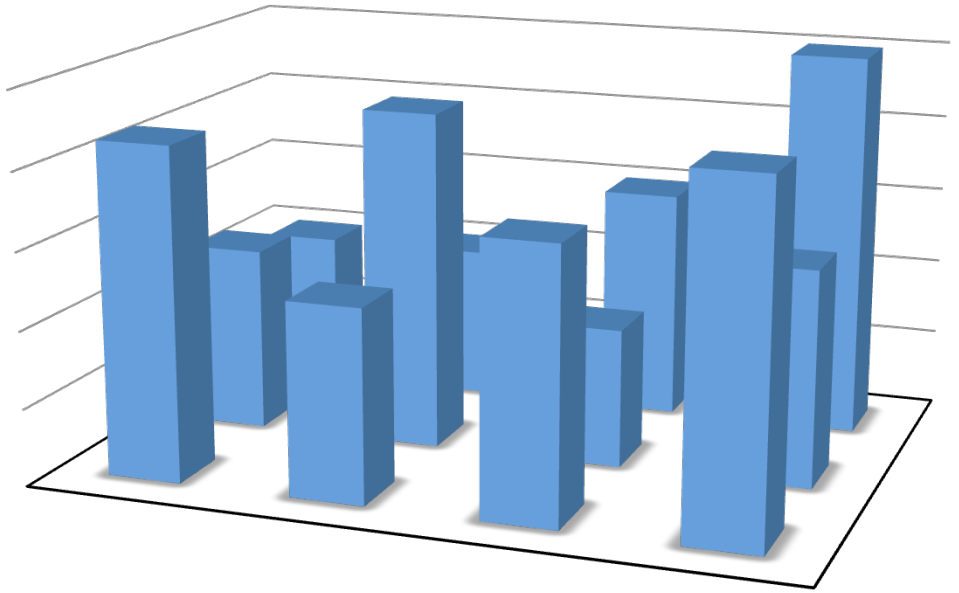
Nominal/Ordinal re-grouping

Statistical analysis

Two phases in the Space-Time-Gridding Framework

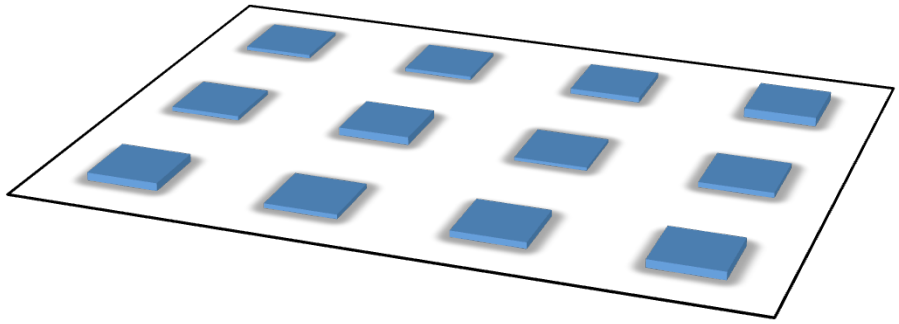
(1) **Space aggregation**

- Parameter filtering (based on ancillary data
- Physical data reduction
- Choice of grid size

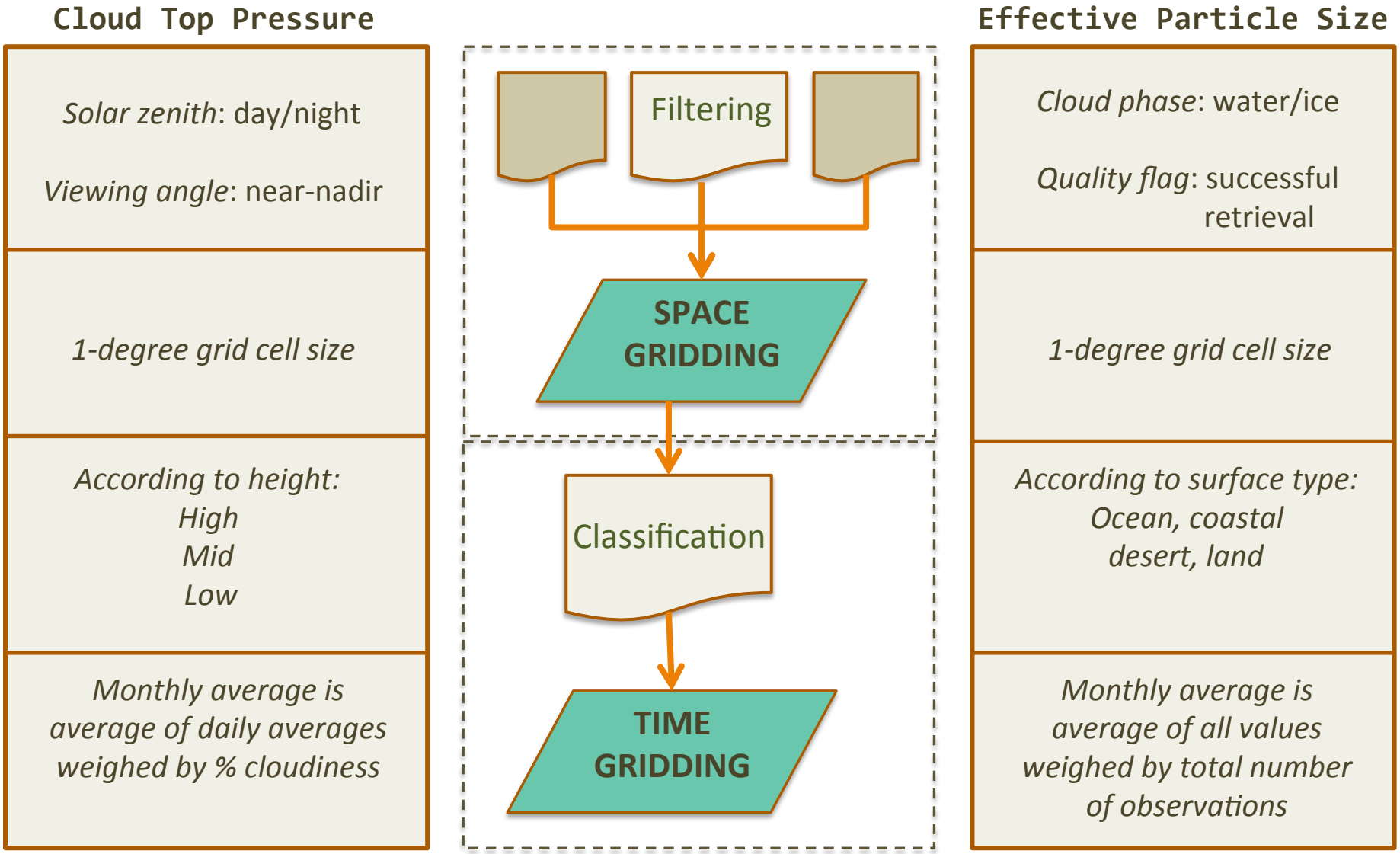


(2) **Time aggregation**

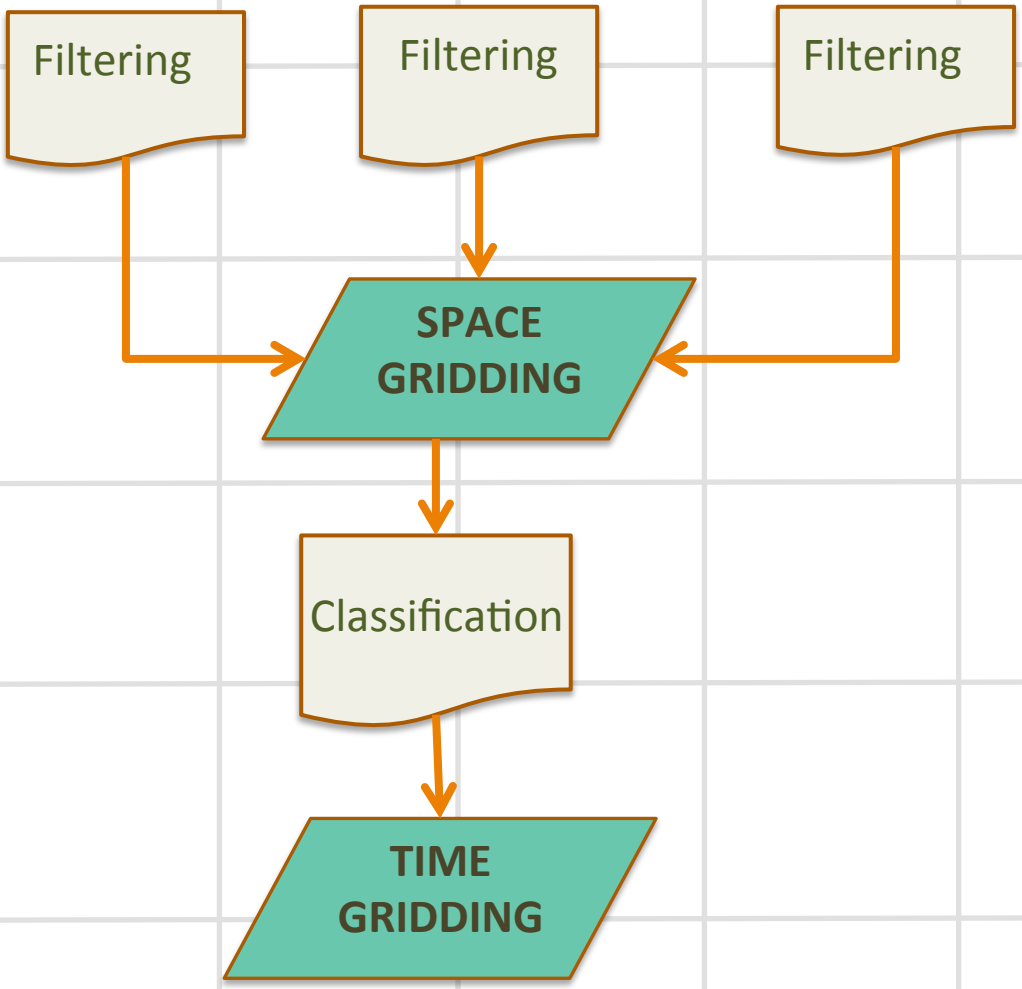
- Min sample size testing
- Data classification (e.g., nominal/ordinal scales)
- Statistical data reduction (mean, mode, std-dev, etc.)
- Daily gridded values/statistics aggregated into time. Choice of number of days.



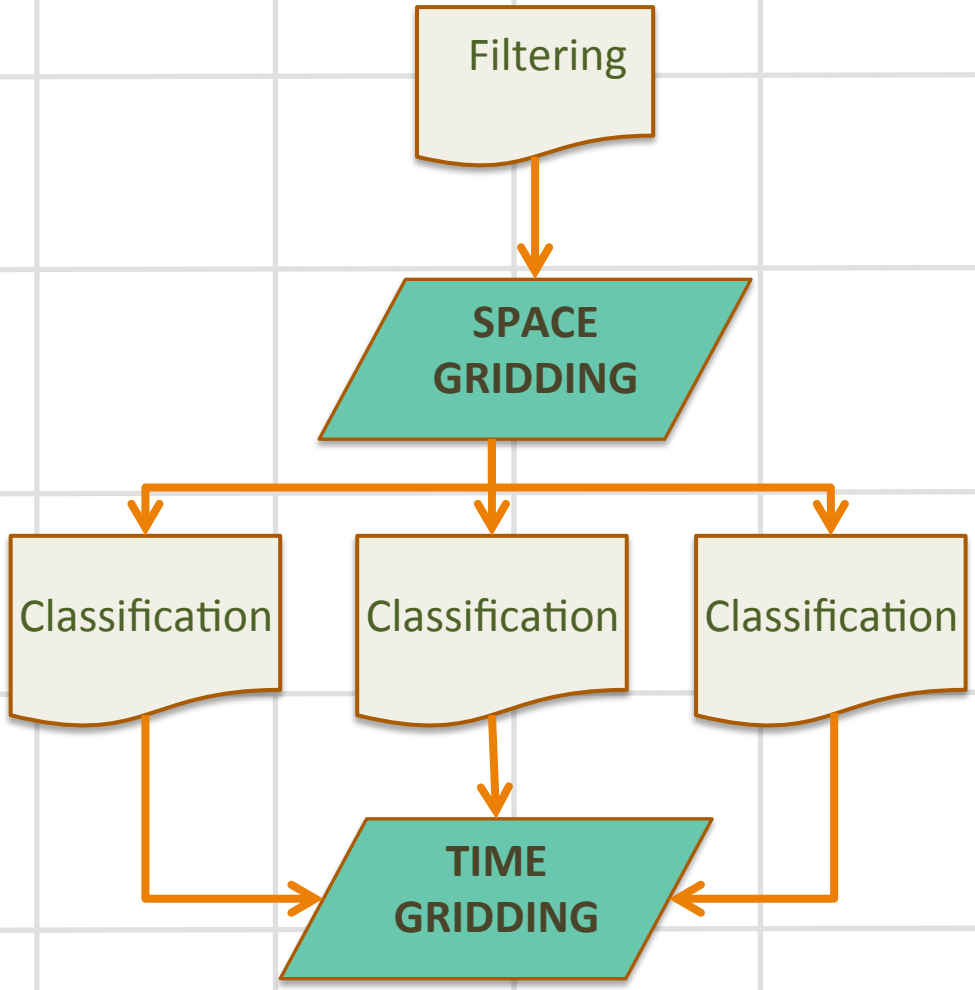
STG requirements for two Cloud Properties



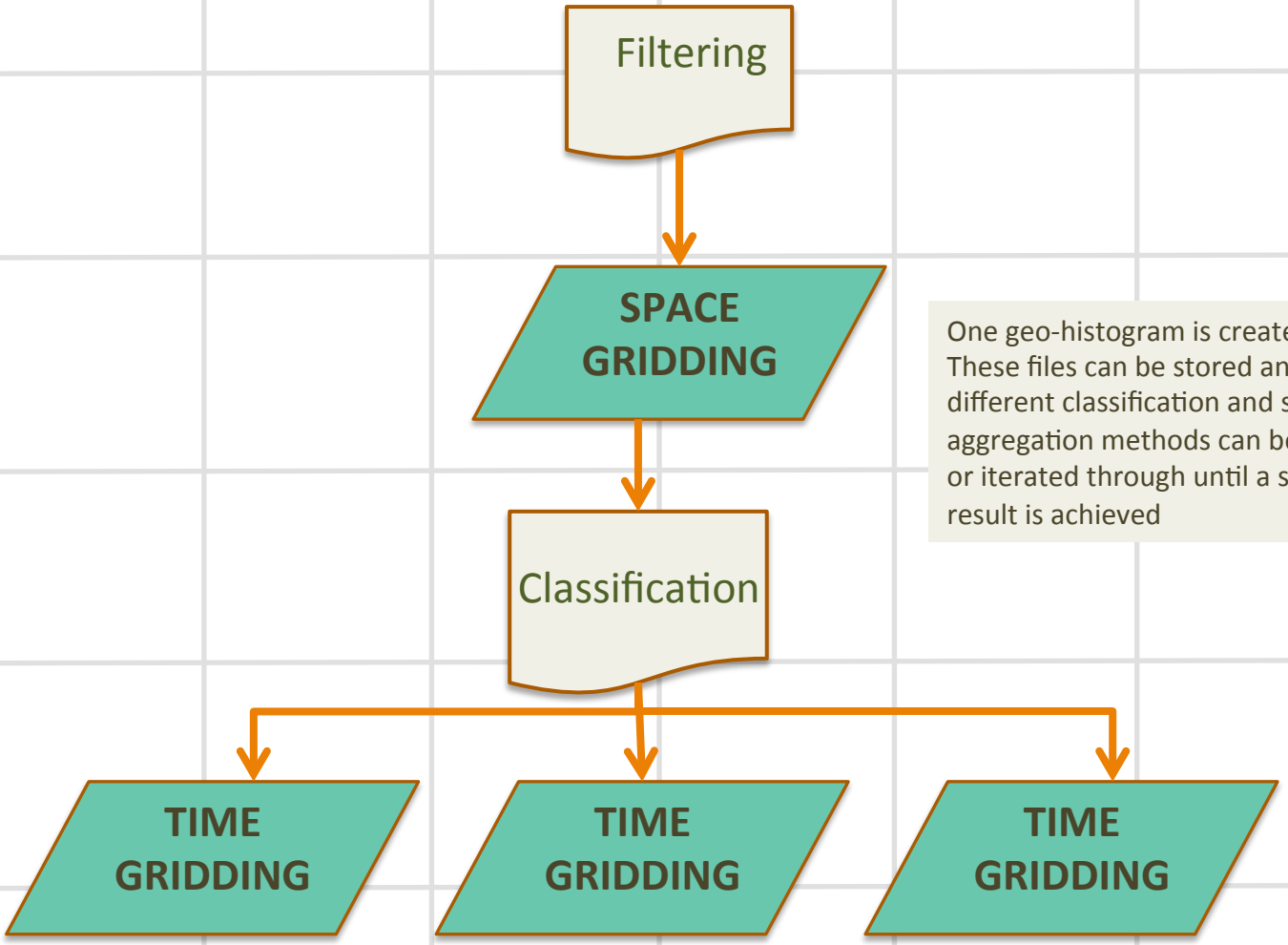
Different approaches to STG



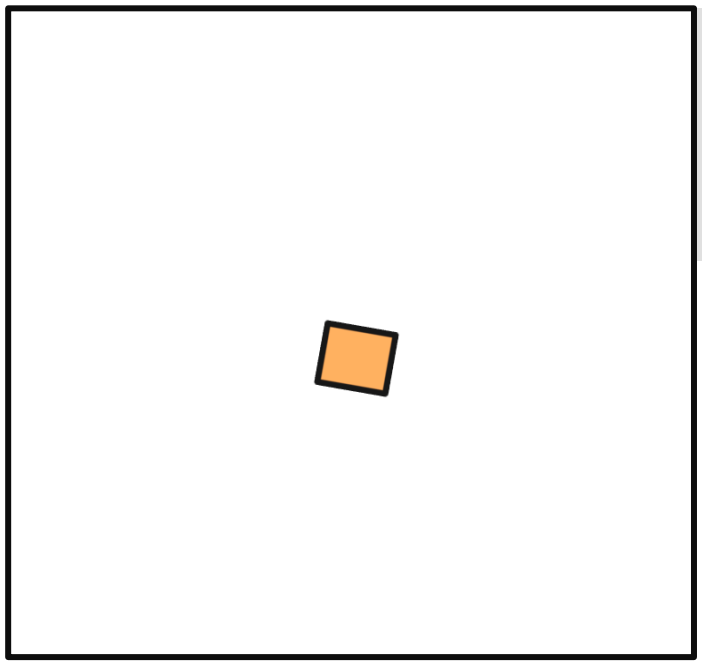
Different approaches to STG



Different approaches to STG

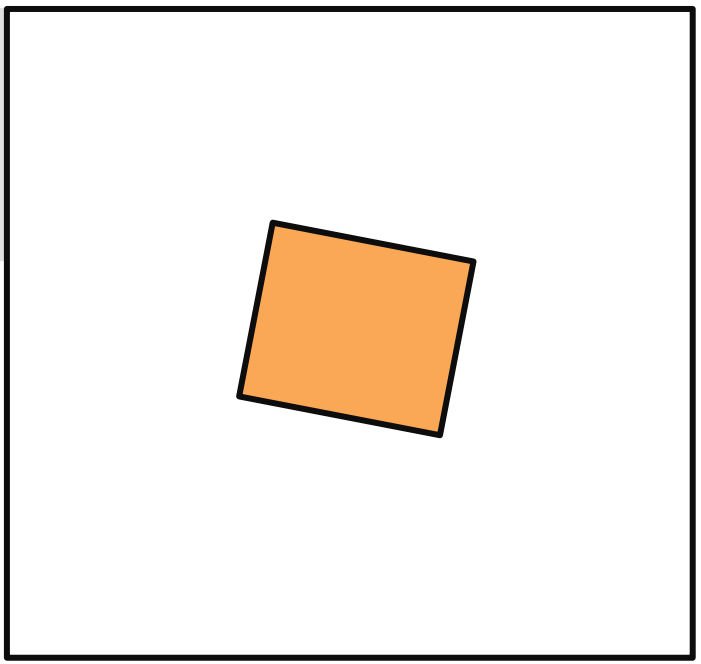


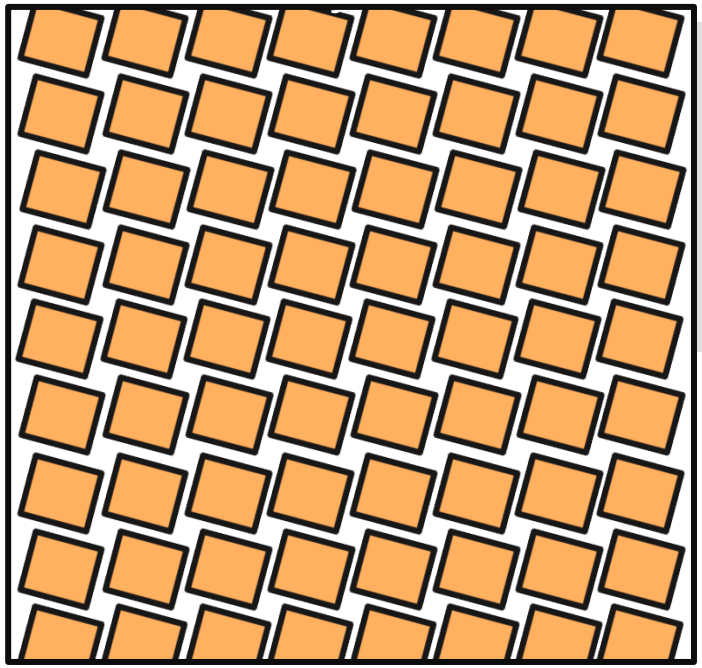
One geo-histogram is created per day. These files can be stored and a number of different classification and statistical aggregation methods can be implemented or iterated through until a suitable final result is achieved



Nearest Neighbor (NN)

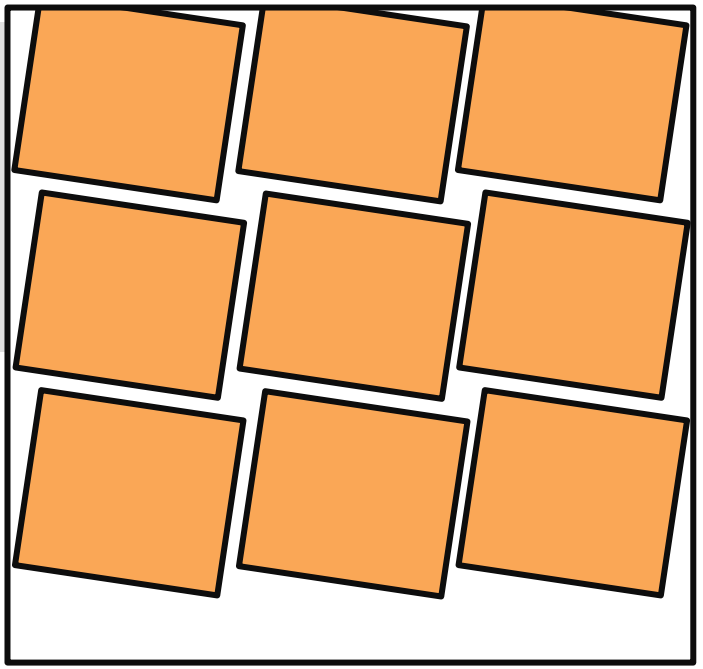
Finding a way to compare instruments with different instrument resolutions





Space Time Gridding (STG)

Accumulating a sample of measurements per grid cell allows statistical comparison/analysis

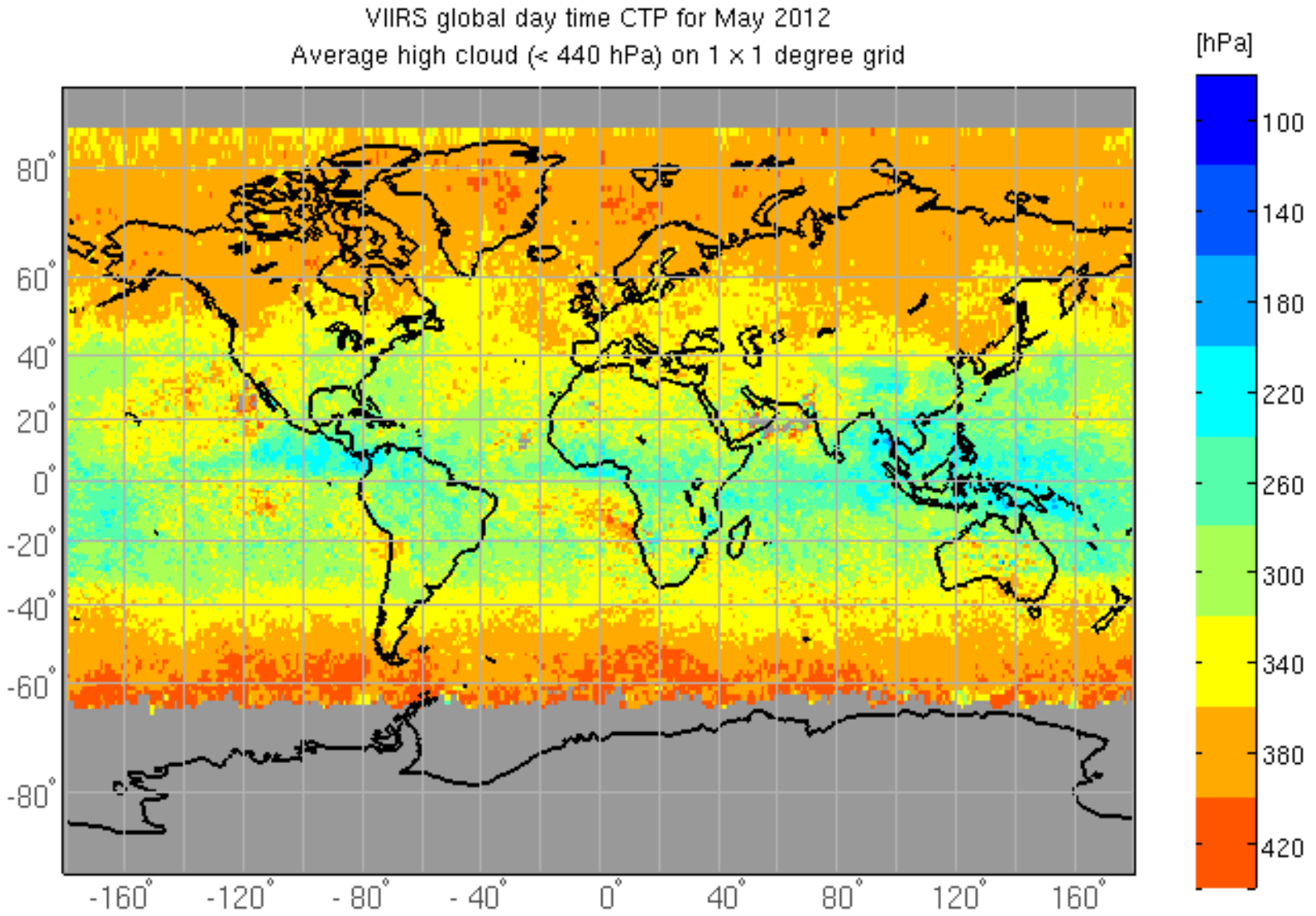


STG Processing: 1 day of global data

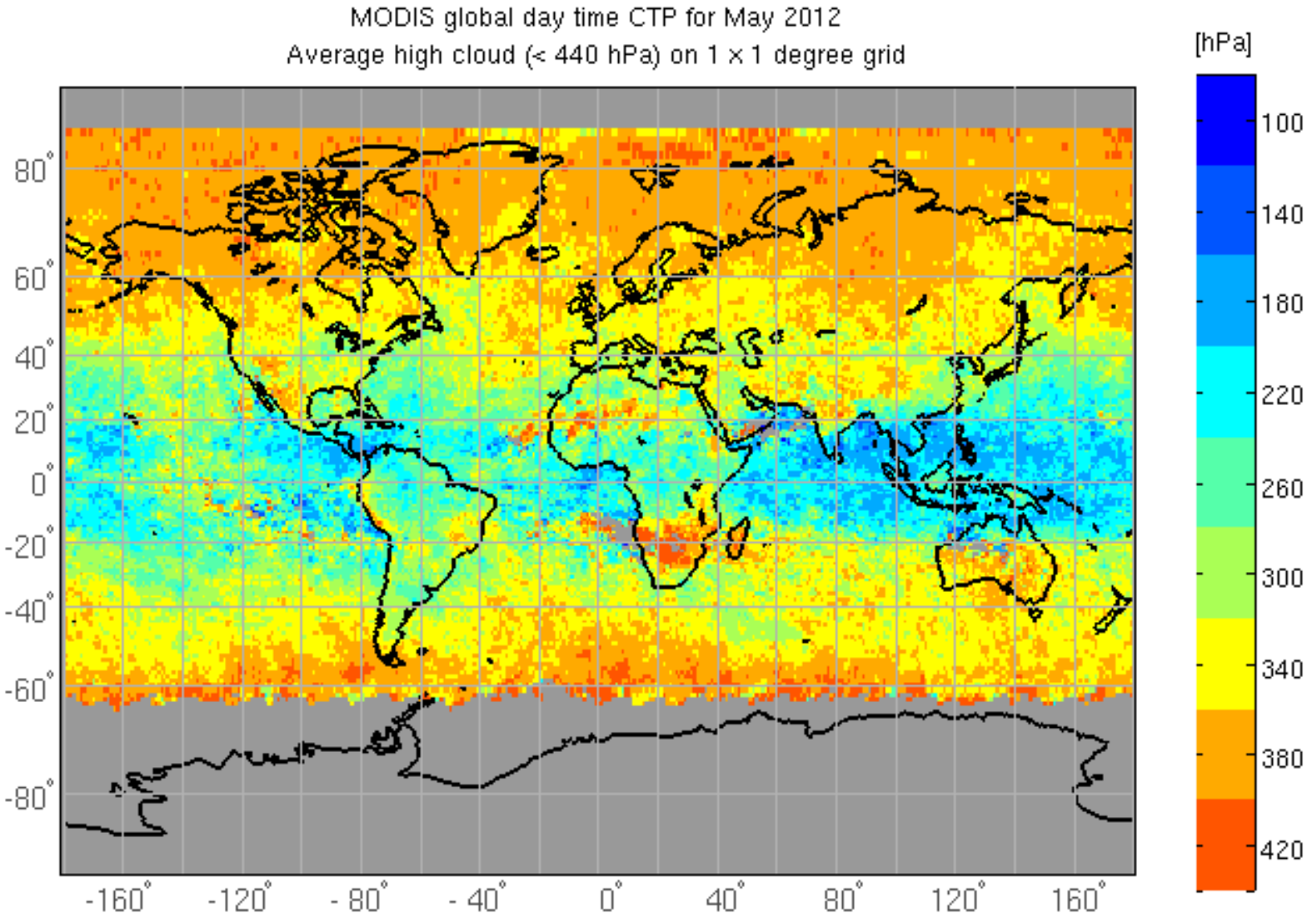
File size = 20-80MB

Instrument	Gridding approach	Native resolution	Grid resolution	Processing time
MODIS	NN	5km 284 granules/day	1x1 degree [180 x 360]	13 min
MODIS	STG	5km 284 granules/day	1x1 degree [180 x 360 x 500]	20 min
VIIRS	STG	5km 1014 granules/day	1x1 degree [180 x 360 x 1000]	60 min
CrIS	STG	14km 180 granules/day	1x1 degree [180 x 360 x 100]	15 min

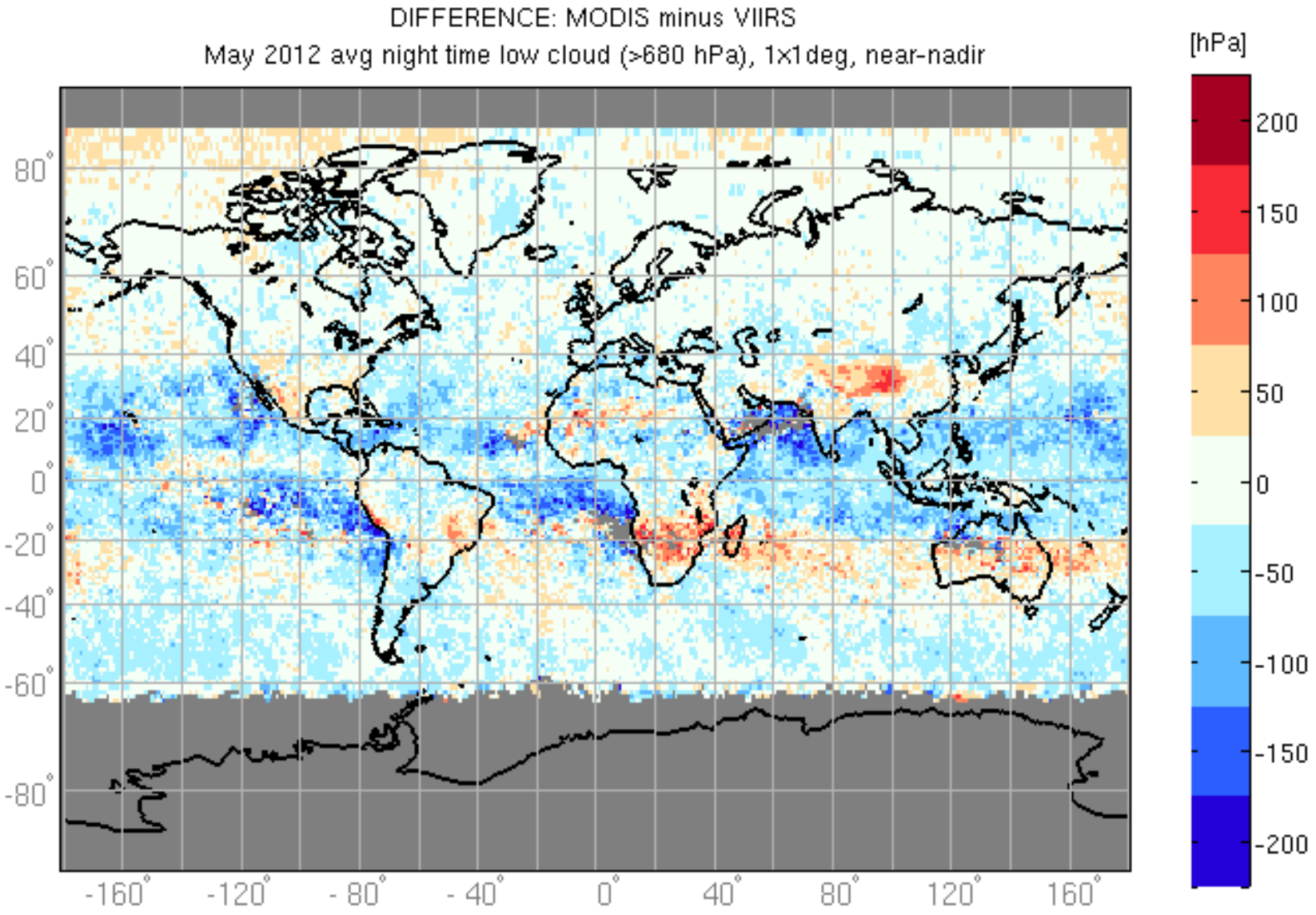
VIIRS EDR: Daytime near-nadir CTP average for May 2012



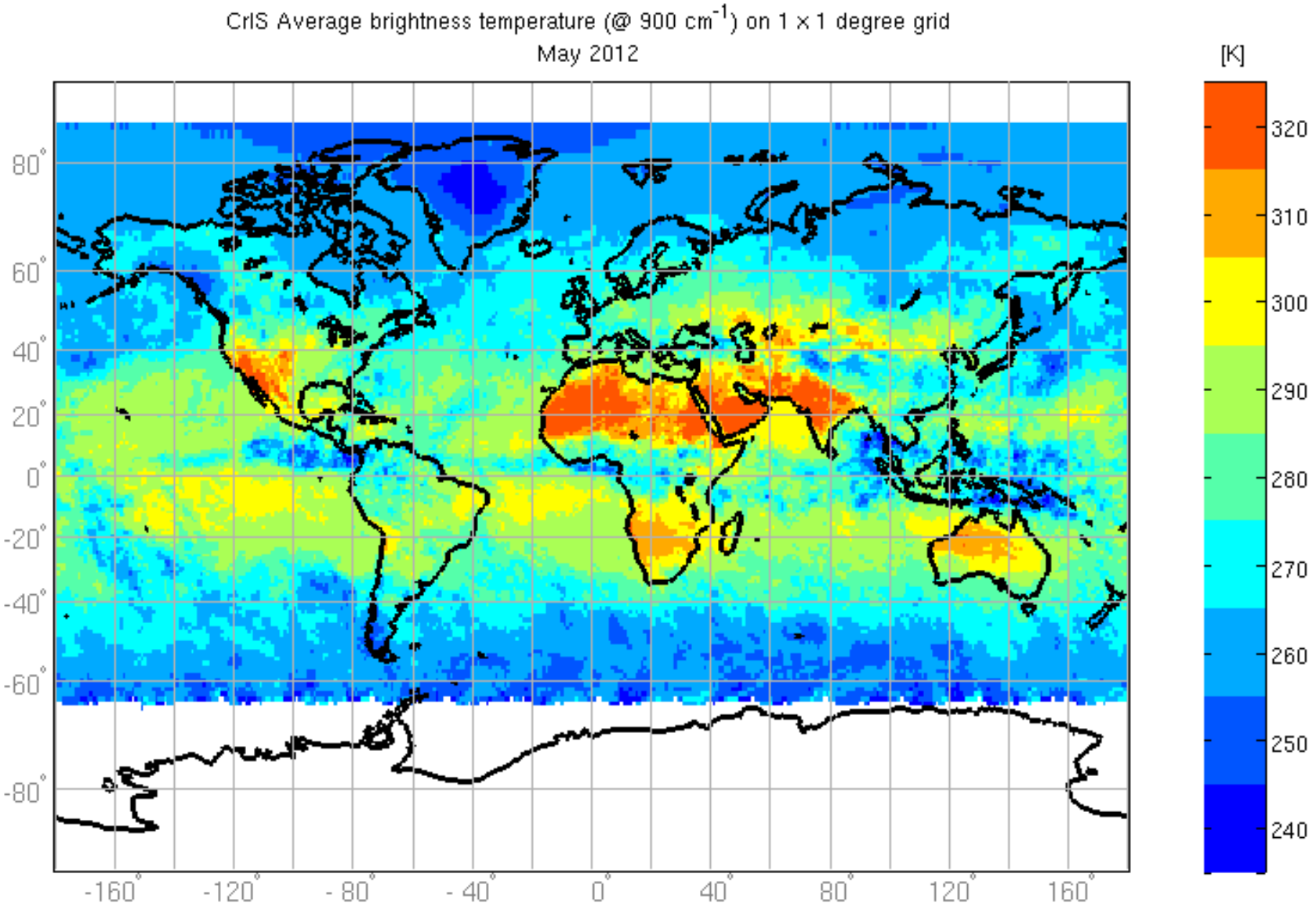
MODIS C6: Daytime near-nadir CTP average for May 2012



Difference: MODIS-C6 minus VIIRS EDR



CrIS: Daytime near-nadir BT @ 900 cm^{-1} avg for May 2012



Current STG work and collaborations

Cloud Retrieval Evaluation Workshop (**CREW**): Evaluation of different L3 aggregation methods with **EUMETSAT**, **KNMI**, **SMHI** (and us at CIMSS) (preparation for CREW-4)

Supporting VIIRS science teams (**NASA**); implementation on the **PEATE**

Shaima Nasiri and Katie Pitts (**Texas A&M**) are studying how differences between MODIS and VIIRS cloud retrievals affect the vertical distribution of liquid water and ice clouds (AMS 2013 poster presentation)

Helen de Klerk (**Stellenbosch University**, South Africa) is studying how biomass burning and CO emissions correlate in space given data from imagers and sounders (Fynbos Forum presentation)