JPL
Jet Propulsion Laboratory

DOGO Warn Levels: You've got them Let's use them

Lukas Mandrake, Gary Doran

Data version 7

Warn Levels:
A Quick Guide

WL: Usage Definition

A number for each OCO-2 sounding: 0 to 19

Estimates sounding complication / contamination
(lower is better)

WL 0-5 is least confounded

WL 6-15 useful but increasingly confounded

WL 16-19 is likely and increasingly useless

YOU DECIDE WHERE TO DRAW YOUR LINE
WL's make it easy to do so

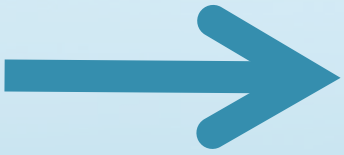
That's what the Quality Flag is for?

NO!

Quality Flag isn't enough

- QF := WL ≤ 15 & outlier_filter & outcome ≤ 2
- Still have 0 ≤ WL ≤ 15 to help make your custom filter
- All passed soundings still continuum of confounding forces

WL's order data rather than just flagging



Using Only Quality Flag ~ Dark Side

Faster, Easier

Less Powerful

Can lead to Pain and Suffering

How to use WL's power

Step 1: Decide what to filter / keep

Require:
popcorn clouds
spot studies
N <50k

Reject:
wild XCO_2
cloud banks
thick aerosols

Require:
clear skies
coastlines
globe coverage

Reject:
any XCO2
complexities

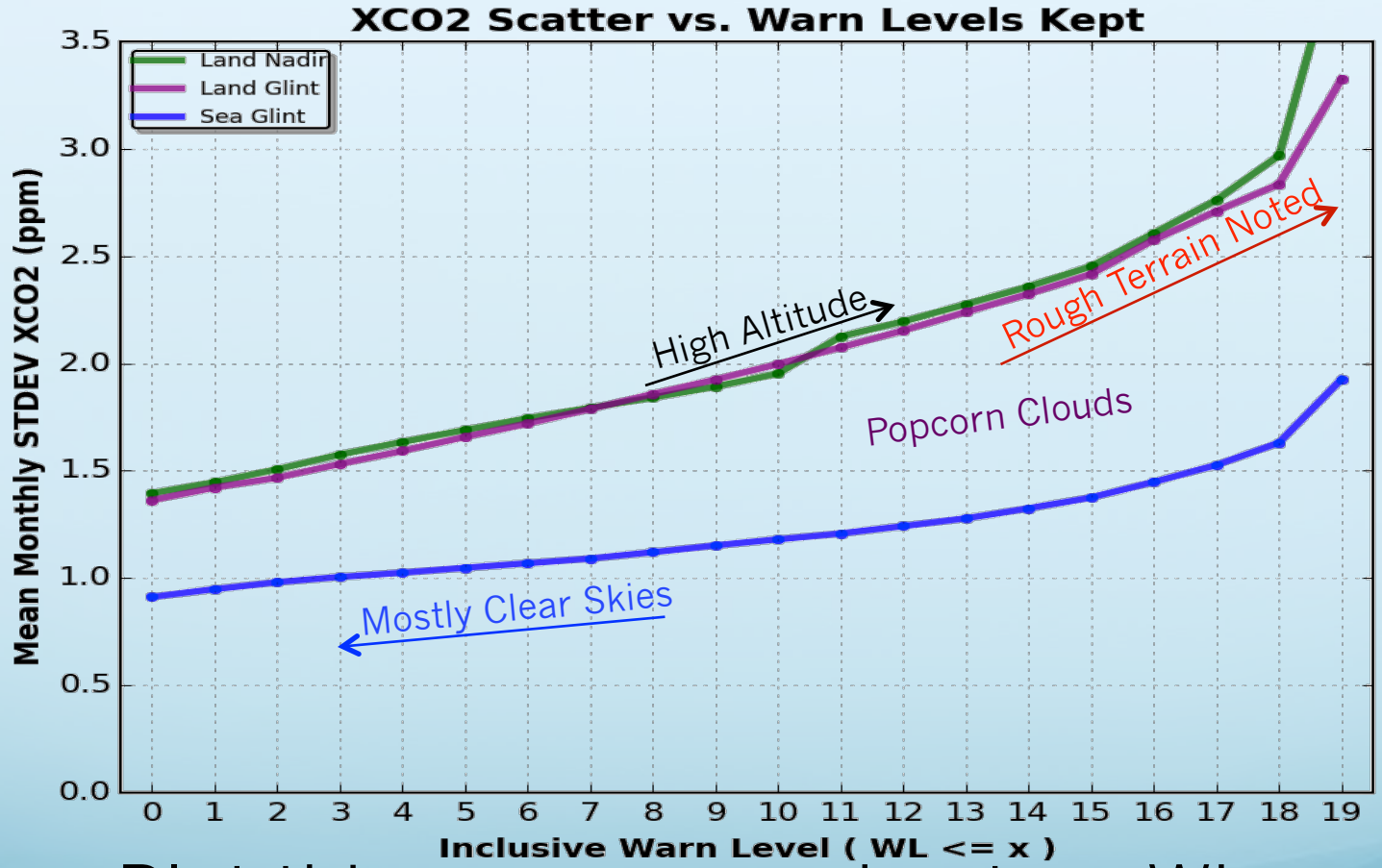
Require:
globe coverage
max soundings

Reject:
Rough terrain

Only you can answer this

How to use WL's power

Step 2: Sweep WL ranges

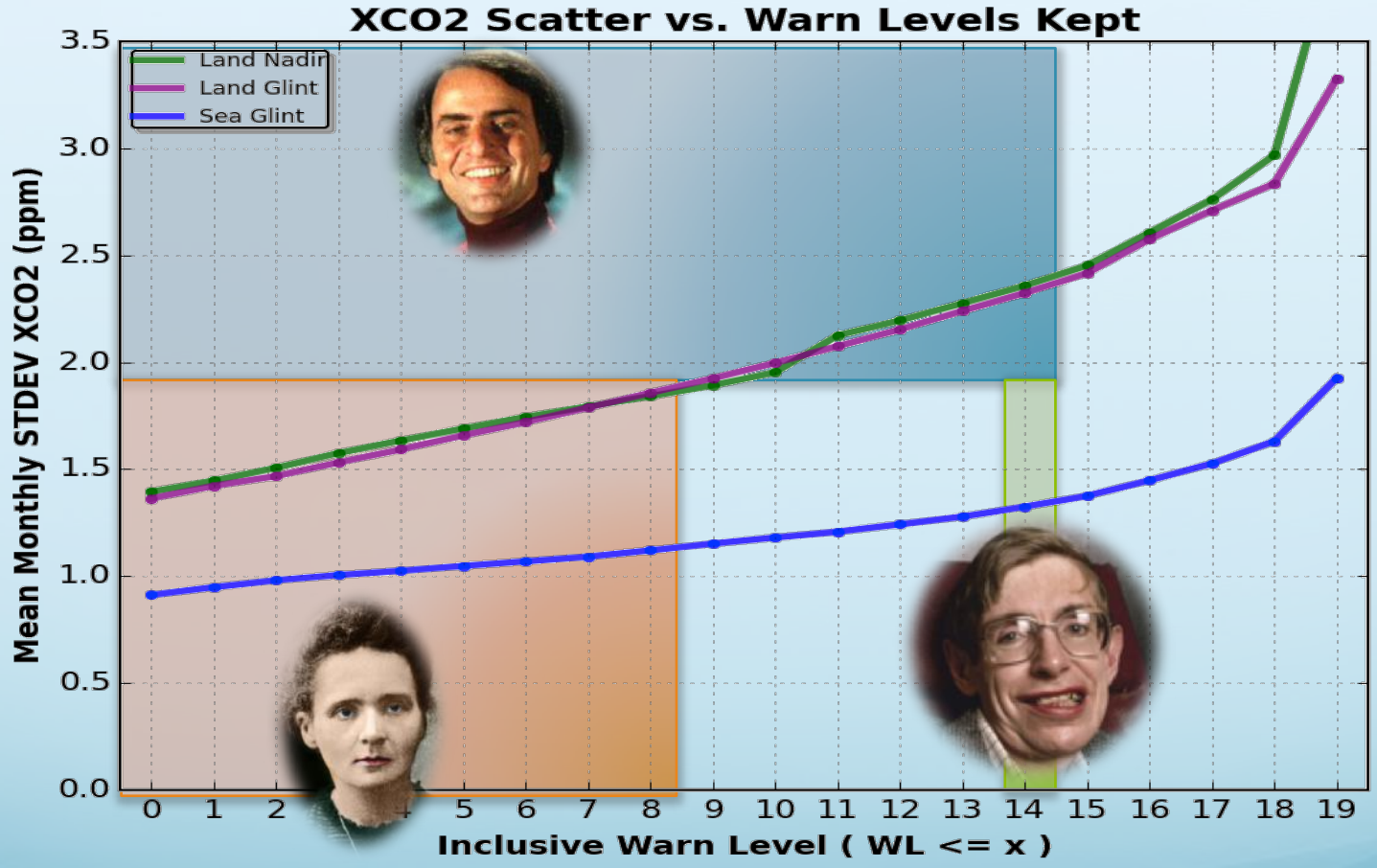


Plot things you care about vs. WL

Explore ranges for useful data

How to use WL's power

Step 3: Fix WL range when satisfied



Each analysis derives its own WL range

Simple to Communicate

A single number does heavy lifting

Cloud study
WL=14
Max. cloud shadows
Min. other effects
Small N obtained

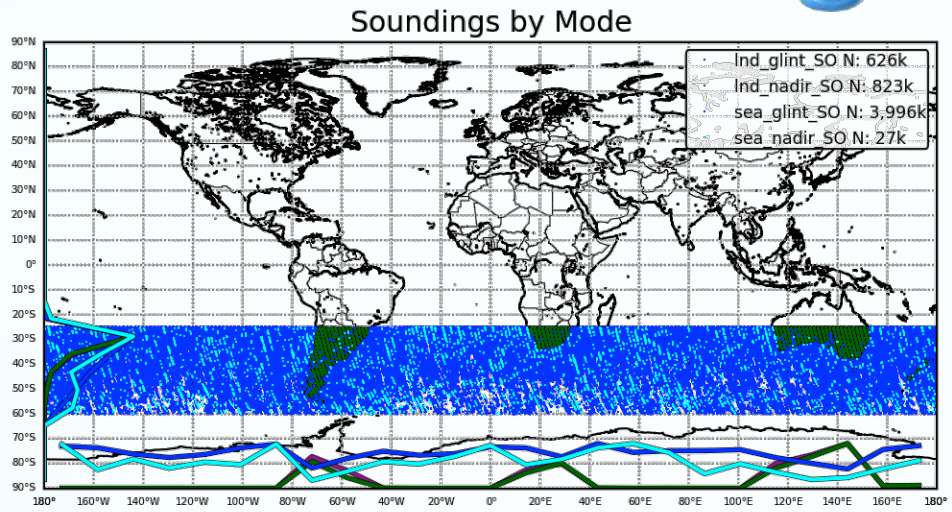
Coastline Study
WL<8
Strongly reduce aerosols
Global coverage
Coastal coverage

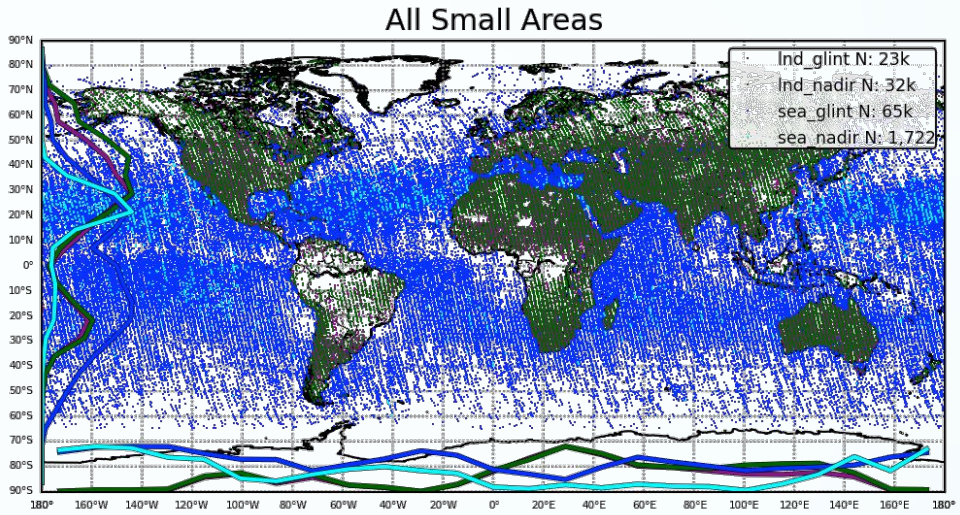
Flux Inversion
WL<14
Global Coverage
Avoid Clouds
Avoid rough terrain

Decide on well-founded filtration in an afternoon rather than days

What Metrics Derive WL's?

Surrogate Truth





Southern Hemisphere

Pros:

- Large-scale confounders
  airmass, seasonal, solar zenith…
- Small-scale confounders
  Surface parameters, snr, etc.

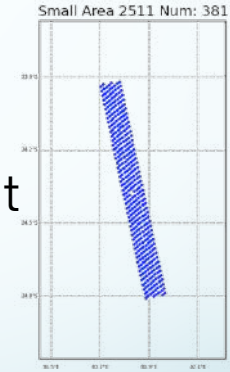
Cons:

- Real latitude XCO_2 gradient
- Not much land
- XCO_2 isn't truly constant in time

Small Areas

Pros:

- XCO_2 really is nearly constant
- Global coverage, not just SH
- HD Small-scale confounders
  Surface parameters, snr, etc.



Span < 100km
N > 100
Contiguous

Cons:

- Large-scale confounders inaccessible
- Small(ish) N statistics ~100-300

Data Ordering Genetic Optimization (DOGO)

- Desire to minimize STD in low-XCO2 variation regions (SH / SA)
- Use Genetic Algorithm to explore effect of all possible filters
- Filters can be based on any number of features (out of hundreds)
- End up with list of "most powerful" features
- Also thresholds for the chosen features that define WL's

Winner JPL Software of the Year
Runner-Up NASA Software of the Year

OCO-2 Training Set

Chosen Features

Aerosol_total_aod
Surface_pressure_delta
Roughness
Rel_resid_mean2_SCO2
CO2_vertical_gradient_delta

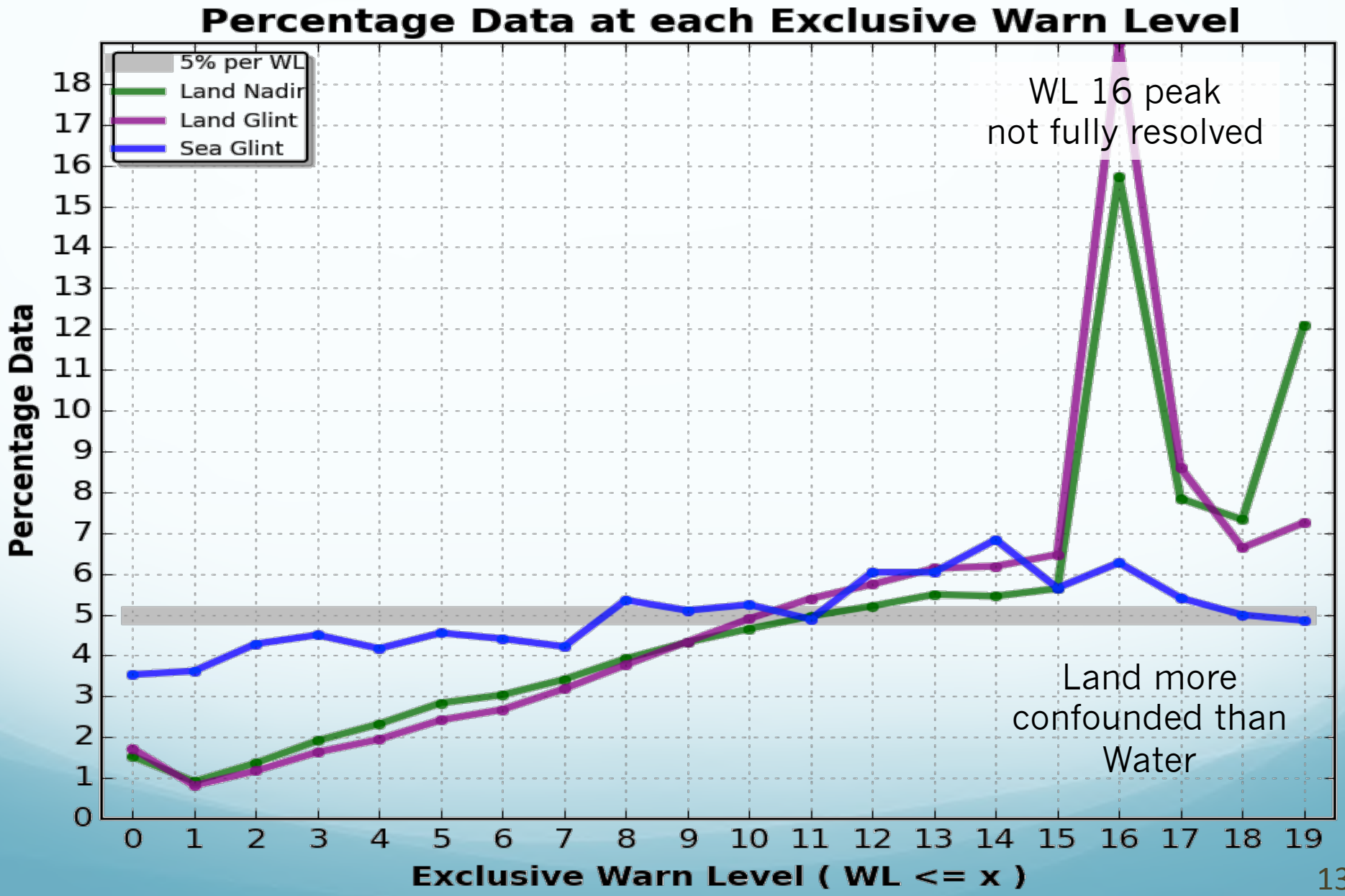
WL Thresholds

1	0.046	1.4E-05	-6.6E-06	-1.0E-05
2	0.051	1.3E-05	-6.0E-06	-1.0E-05
3	0.056	1.3E-05	-5.4E-06	-1.1E-05
4	0.061	1.2E-05	-4.8E-06	-1.1E-05
5	0.065	1.1E-05	-4.2E-06	-1.2E-05
6	0.070	1.1E-05	-3.5E-06	-1.2E-05
7	0.075	1.0E-05	-2.9E-06	-1.3E-05
8	0.080	9.5E-06	-2.2E-06	-1.3E-05

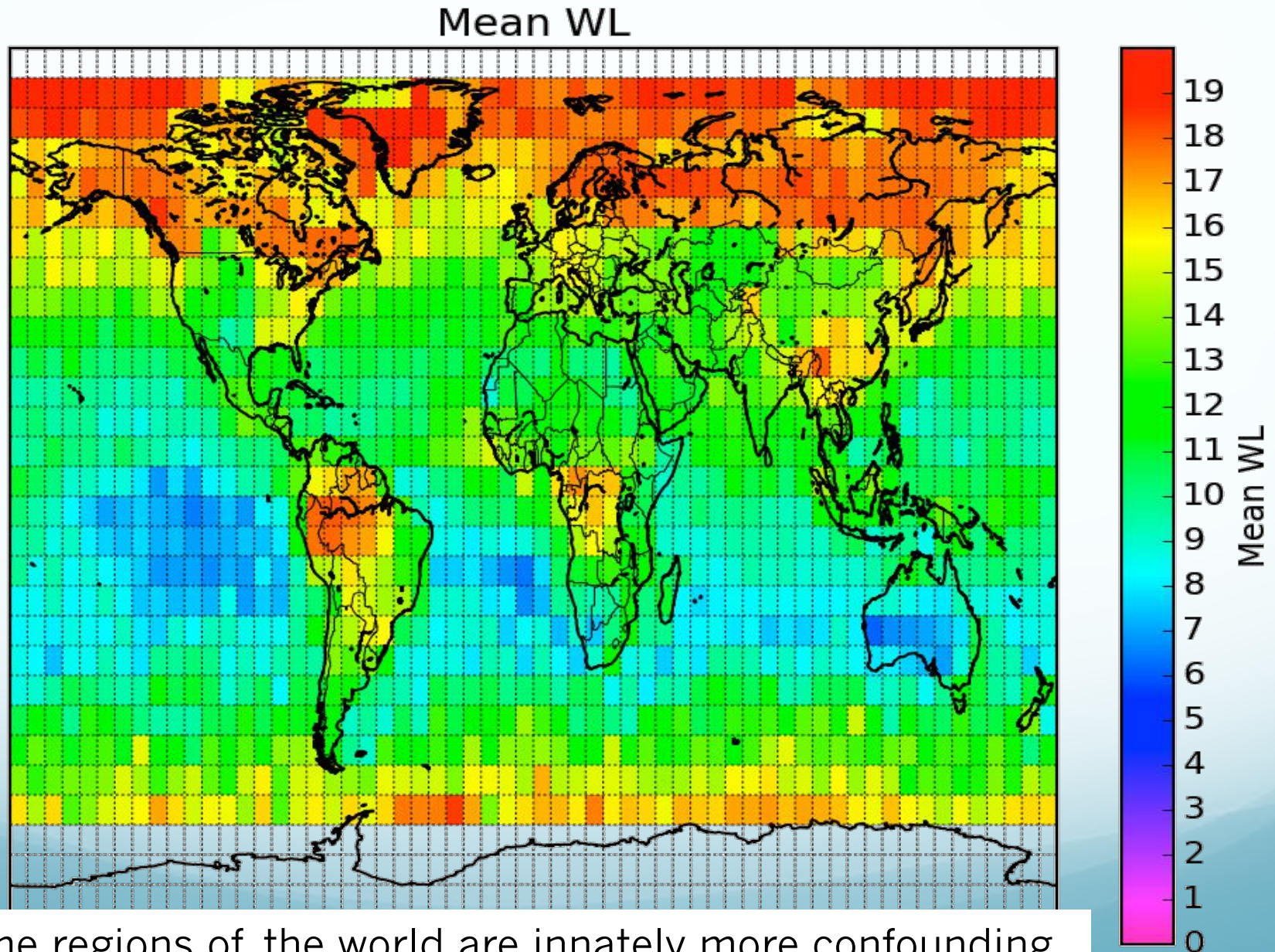
Warn Levels: Effects & Distributions

Sep 2014 – Sep 2015

Warn Level Distribution

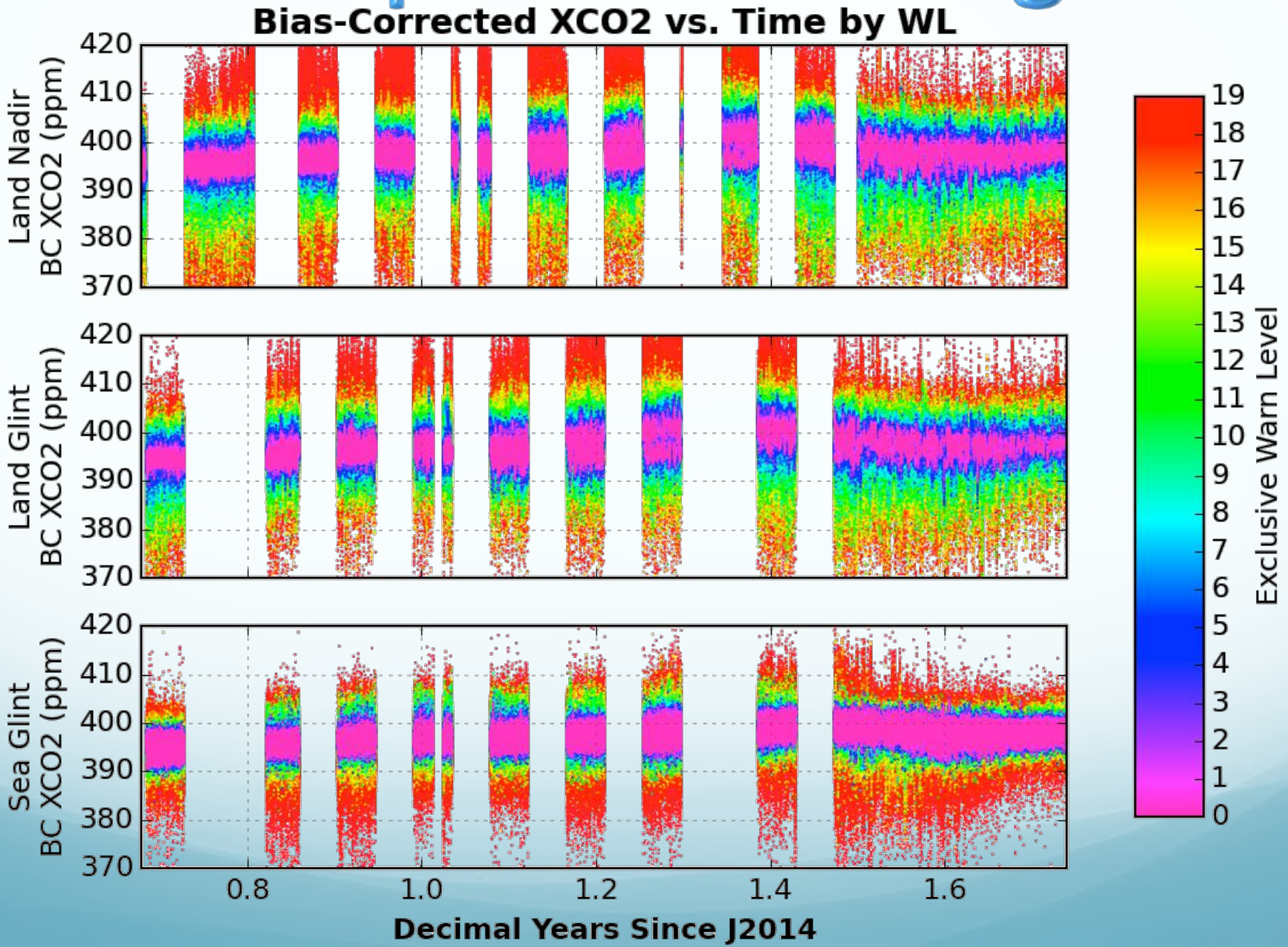


Spatial WL Coverage

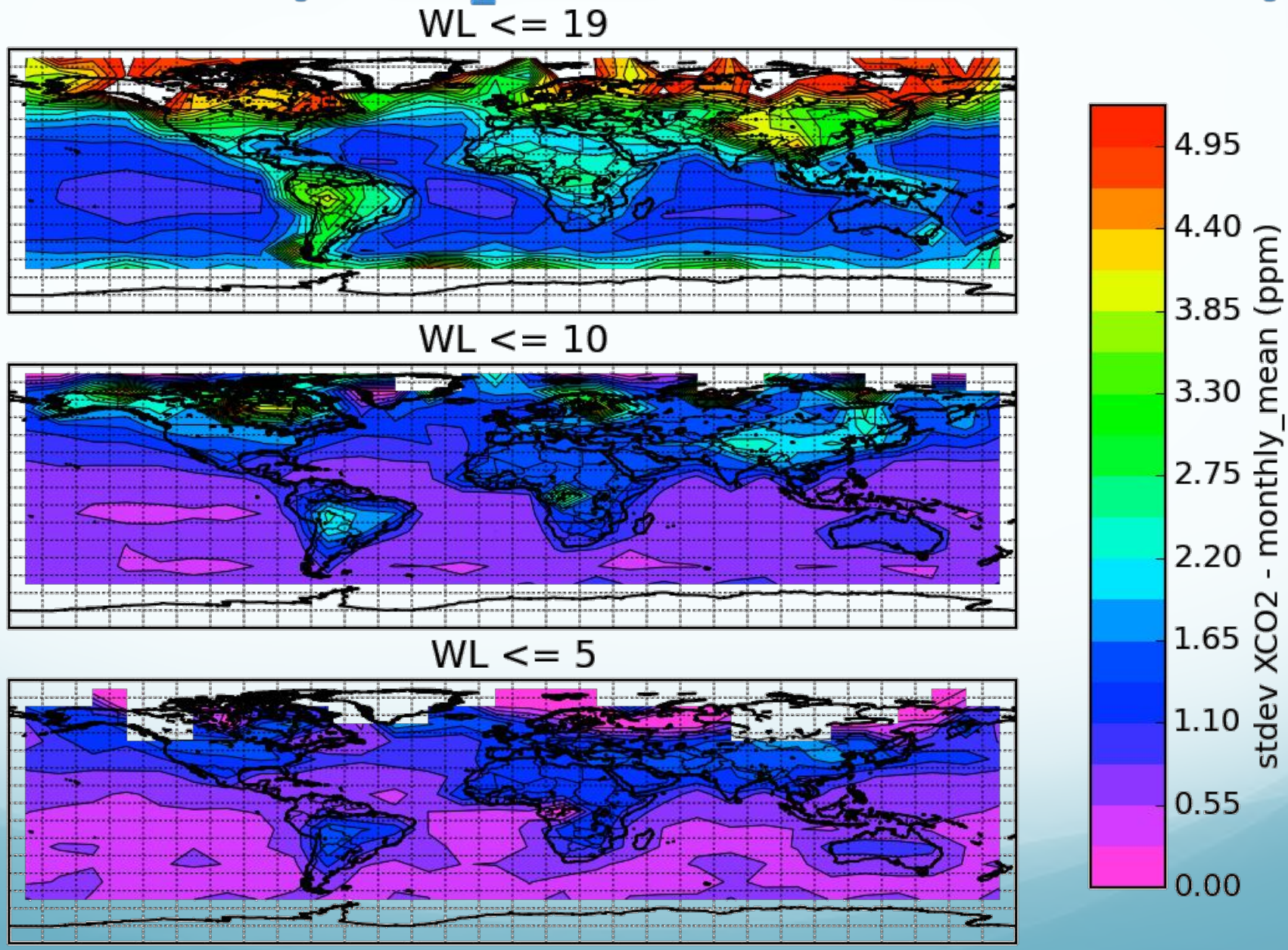


Some regions of the world are innately more confounding

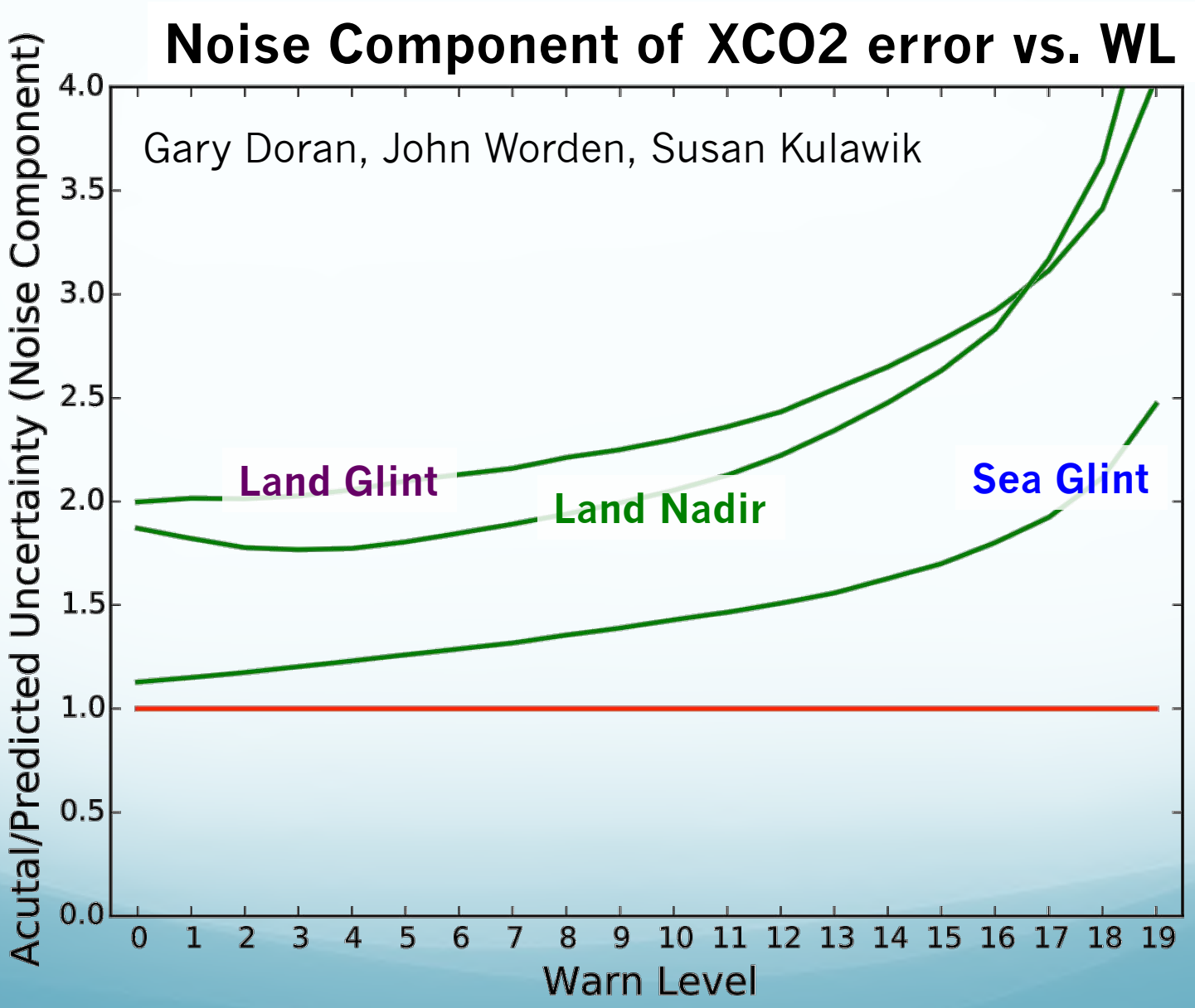
Temporal Coverage



STDEV(xCO_2 – bin monthly mean)

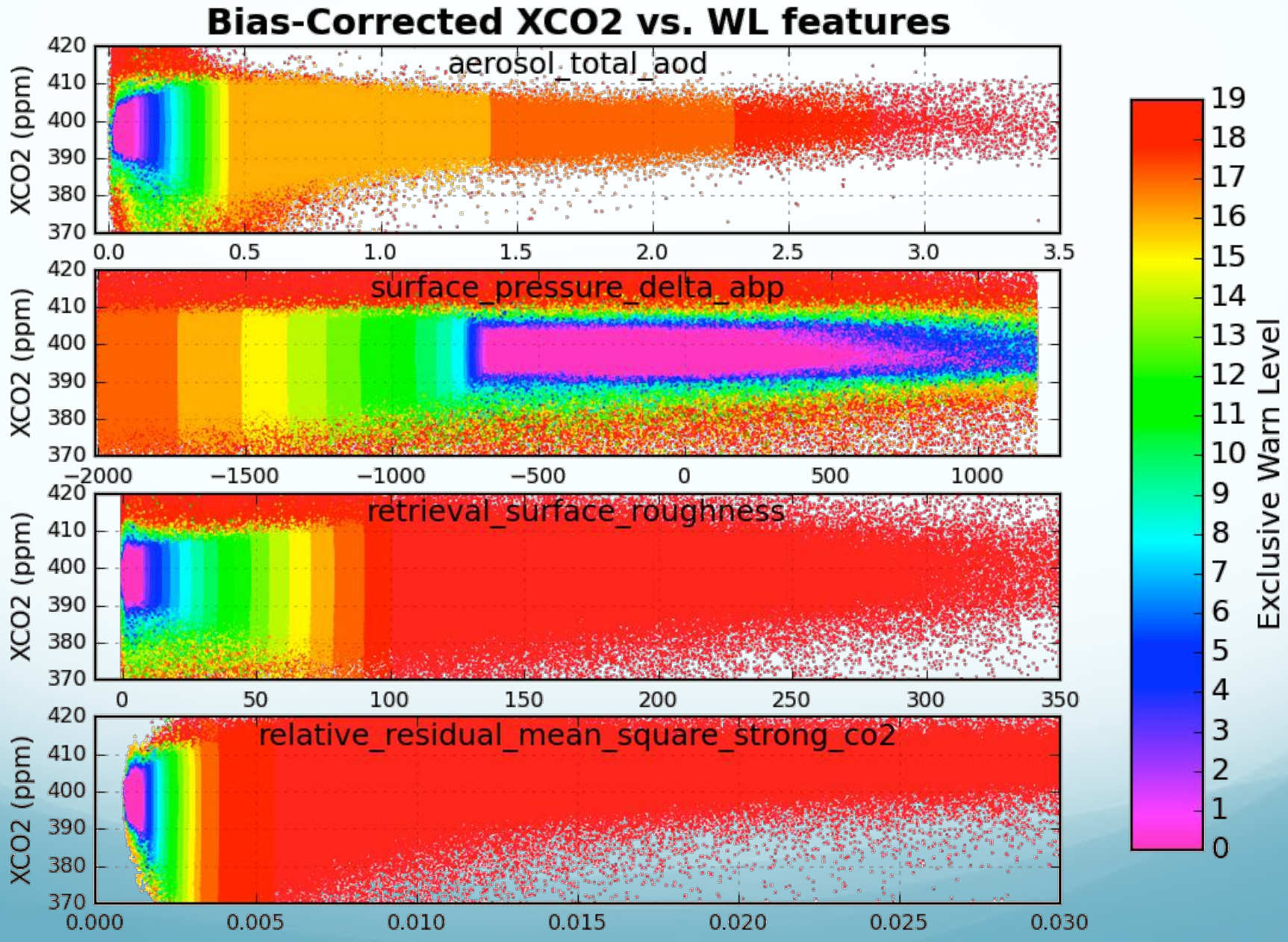


WL vs. XCO2 Uncertainty

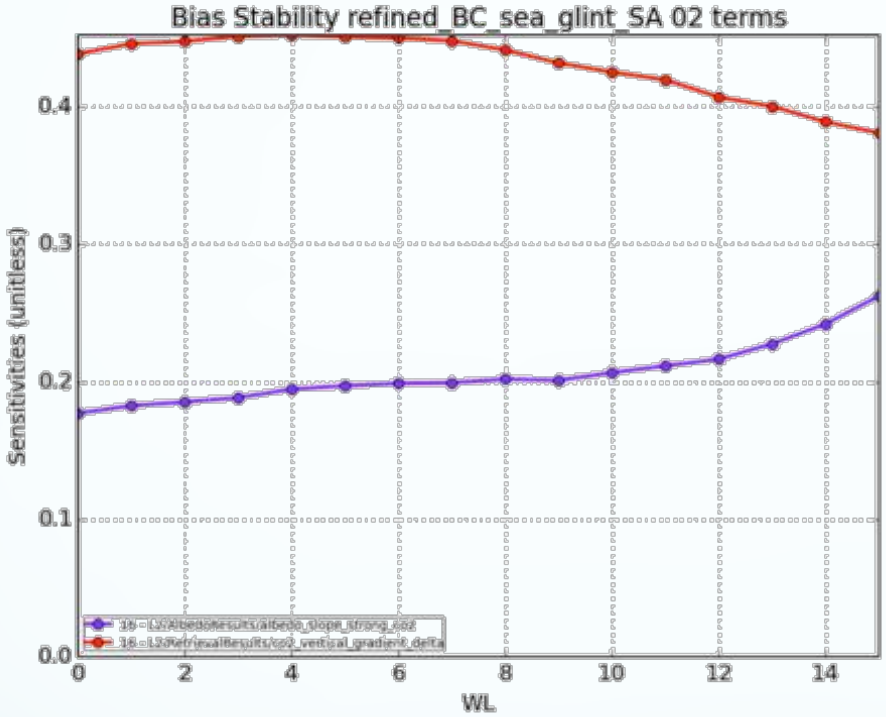


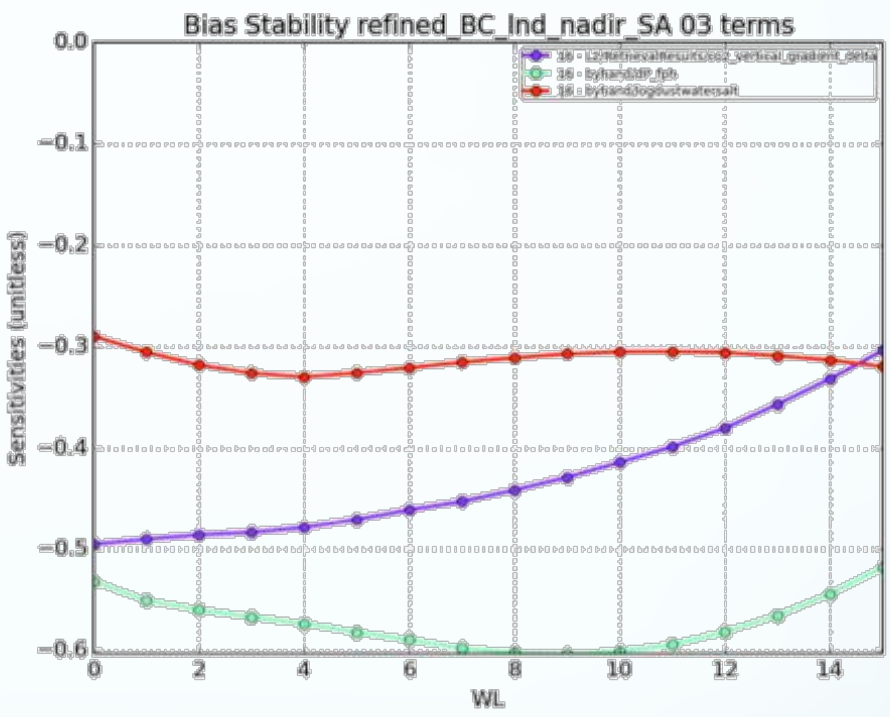
Land Nadir Example

Features WL's Use



Bias & Filtration





- Most XCO_2 derived quantities are function of Warn Level
- Must perform "sweep" of WL to look for stable regions
- XCO_2 bias is similarly sensitive

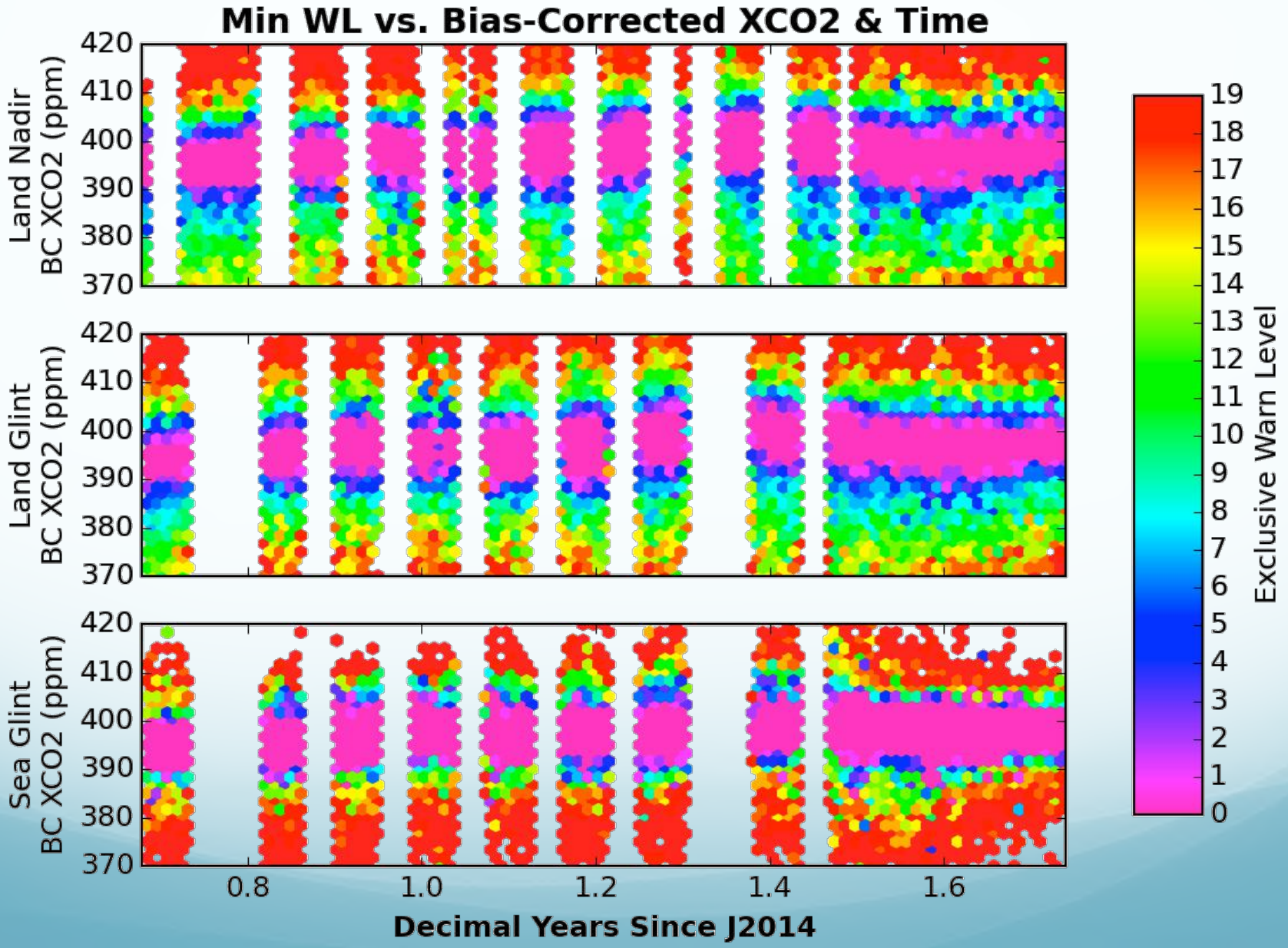
- For v7, bias was determined fixing WL <= 10 (middle of usage range)
- Some bias terms are insensitive to WL, others not so
- An excellent example of WL/filtration effects on science results

Warn Level Summary

- WL's order OCO-2 data by level of confounding forces present
- Derived by reducing spurious XCO_2 deviation in ~homogenous regions
- Sweeping WL's to look for (un)desirable traits makes custom filter
- Sweeping science results as function of WL tests robustness of claims
- Single parameter does "heavy lifting" of filtration definition
- Correlate with independent measures of retrieval quality (uncertainty)

Backup Slides

Binned Temporal Coverage



Spatial Coverage