Why don't composite studies agree?

An example with cosmic rays and clouds

Benjamin Laken[1]

&

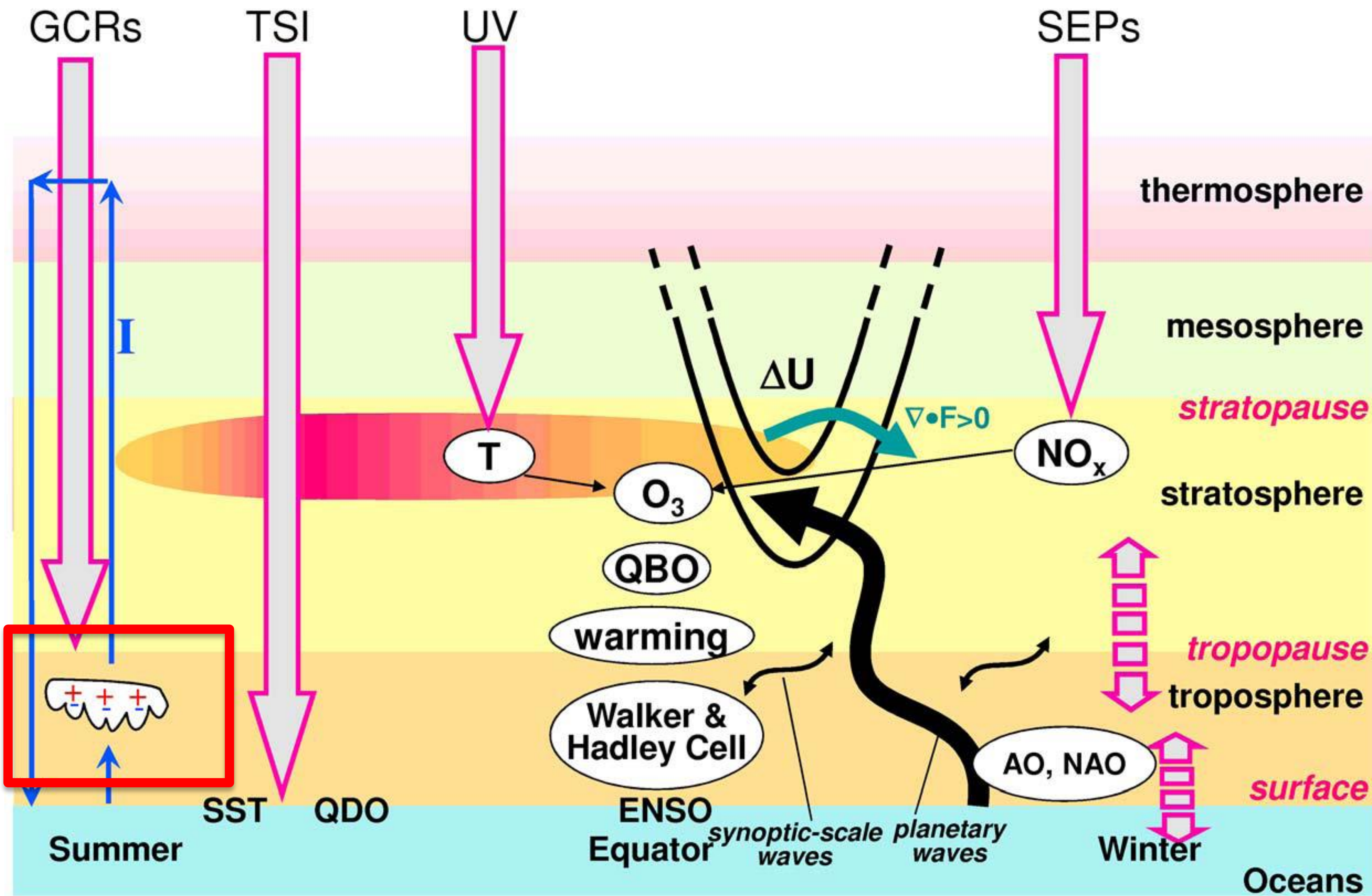
Jasa Čalogović[2]

1. Instituto de Astrofísica de Canarias, Via Lactea s/n, E-38205, La Laguna, Tenerife, Spain
2. Hvar Observatory, Faculty of Geodesy, University of Zagreb, Kačićeva 26, HR-10000, Zagreb, Croatia

Outline

- The theorized cosmic rays - cloud link
- Datasets
- Composites explained
- Using Monte Carlo to determine significance and design composites
- How to get a false-positive
- Wider context & conclusions

Schematic diagram of theoretical solar influence on climate

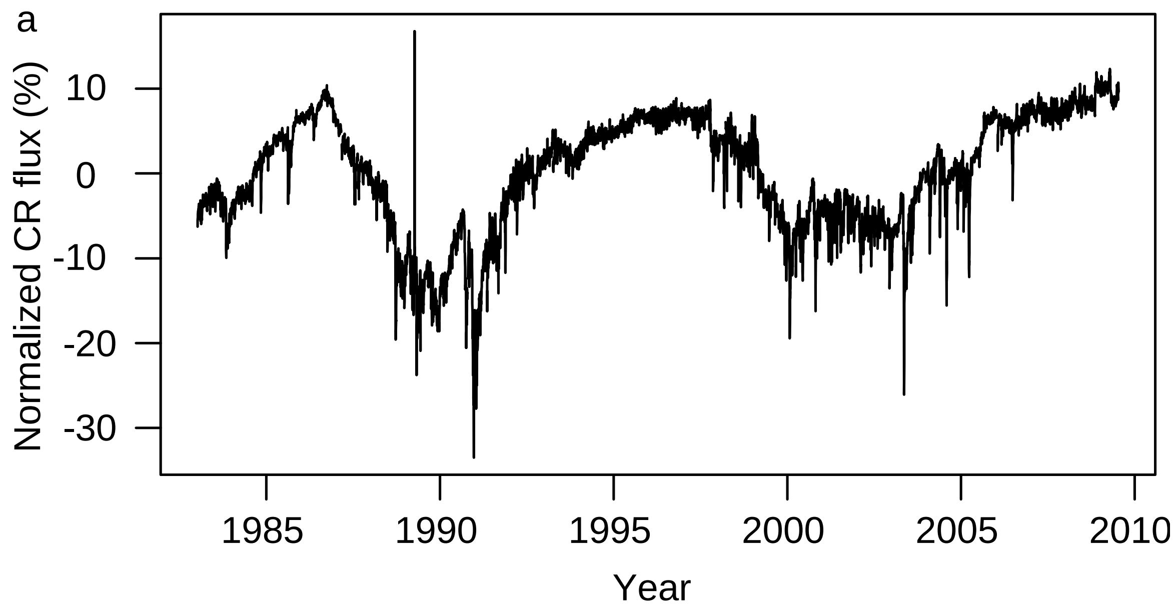


Kodera & Kuroda , 2002

Working from real data
daily averaged values (1983–2010):

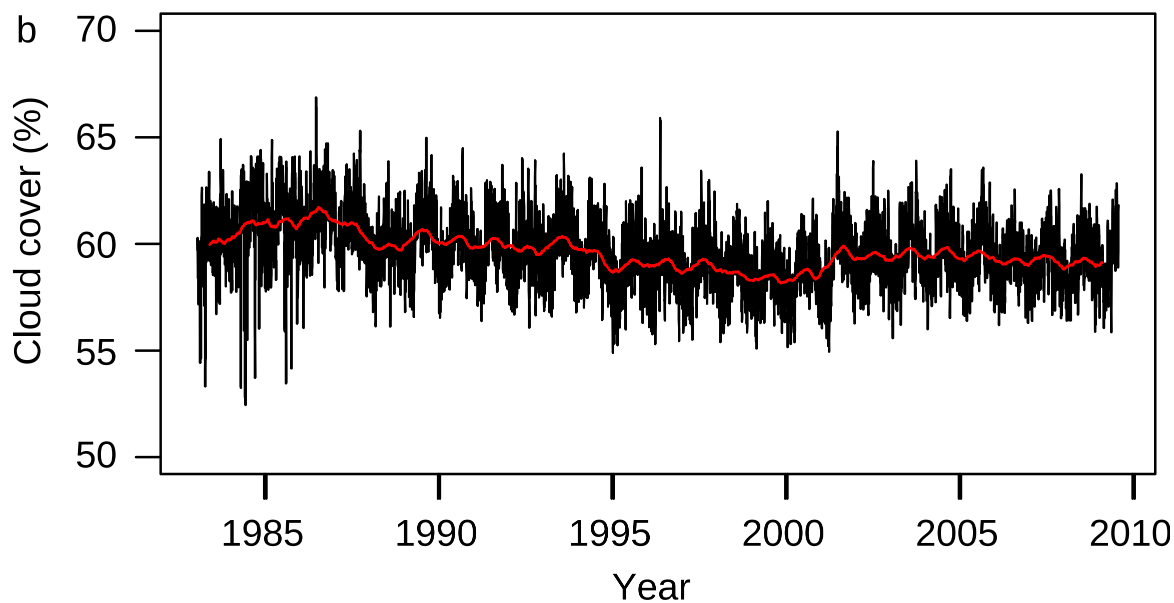
Cosmic ray flux (%)

Neutron counts from Climax and Moscow NM combined: (daily counts/mean count of all data) centered on zero.



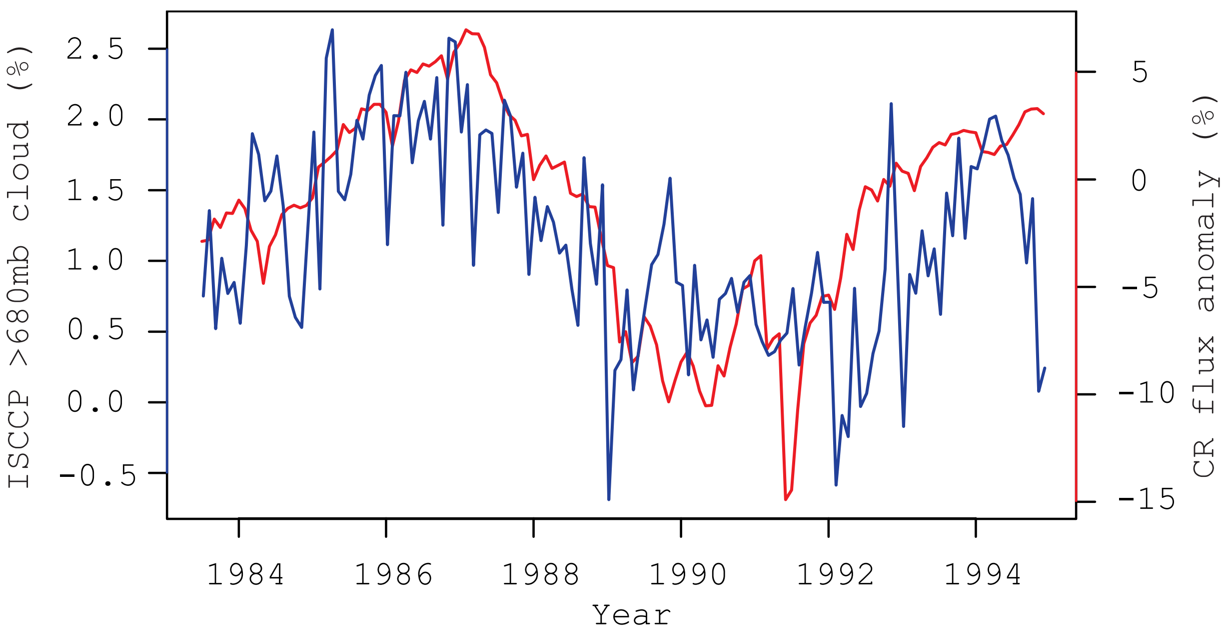
Cloud over (%)

International Satellite Cloud Climatology Project D1 Infrared-detected. Globally averaged.



From Laken *et al.*, 2012, *JGR*

The hypothesized connection between cosmic ray flux and cloud cover



ISCCP IR-cloud cover at low (>680mb) levels

Cosmic Ray Flux (from Moscow and Climax NM

R = 0.6 (*df* in CR flux = 7, *df* in ISCCP = 4)

From Laken et al. 2012, SWSC, reproducing analysis from Marsh & Svensmark & 2000

History

Link suggested b/w solar activity and weather (Herschel, 1801)

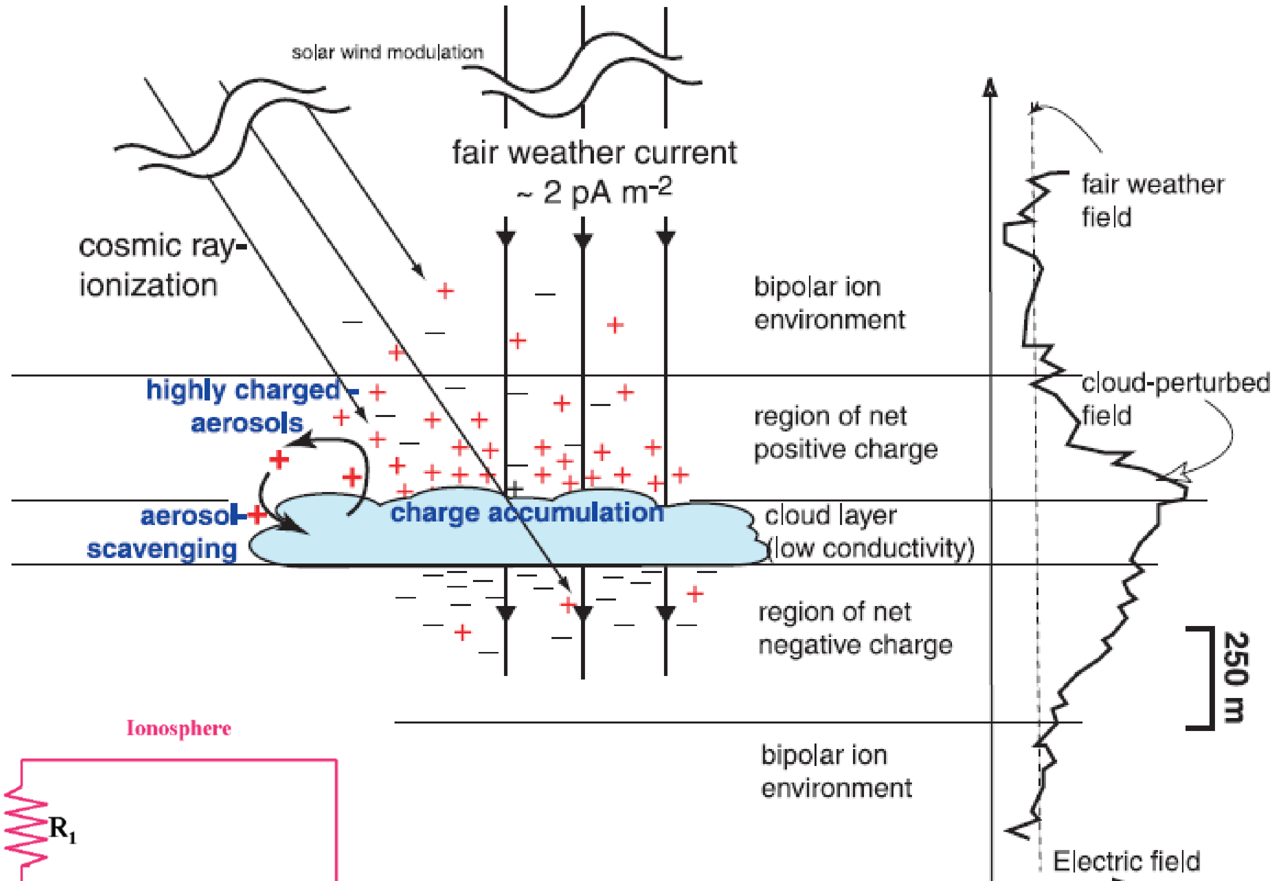
Suggested correlation b/w weather and CR flux (Ney, 1959)

Proposed link between CR induced atmospheric ionization from CR flux, and cloud microphysics (Dickinson, 1975)

Composites suggests link b/w storm properties and CR flux (Tinsley & Deen, 1991)

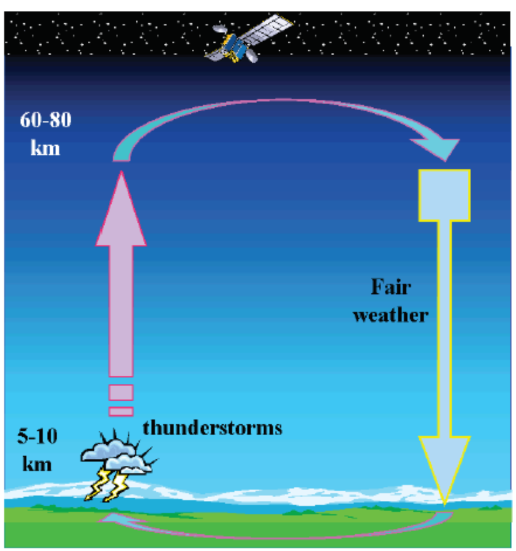
Obs. of cloud from satellite correlated to CR variations (Svensmark & Friis-Christensen 1997)

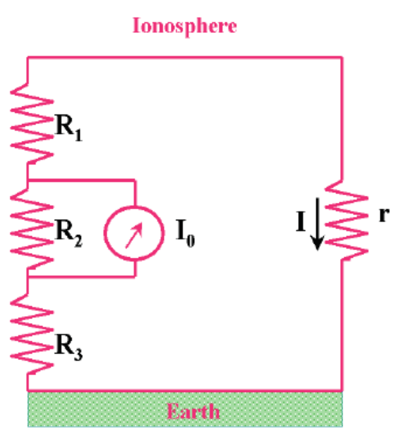
Current density-cloud hypothesis
Global Electric Circuit



Carslaw, Harrison et al., 2002

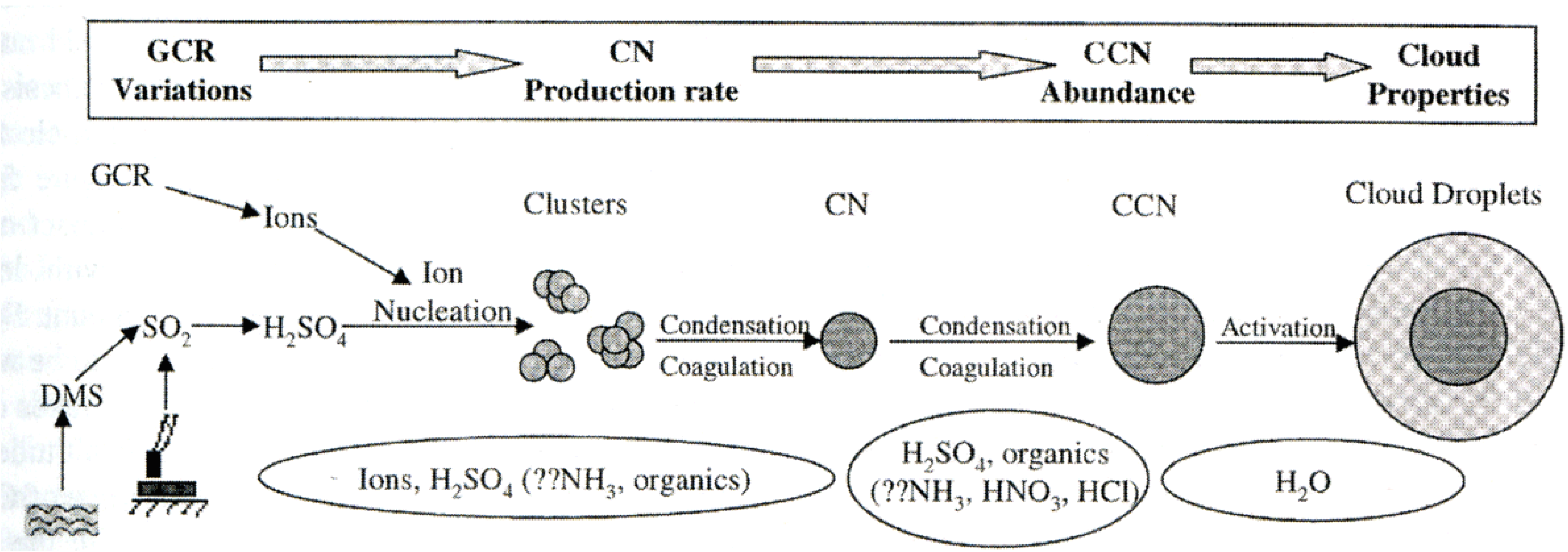
Global electric circuit



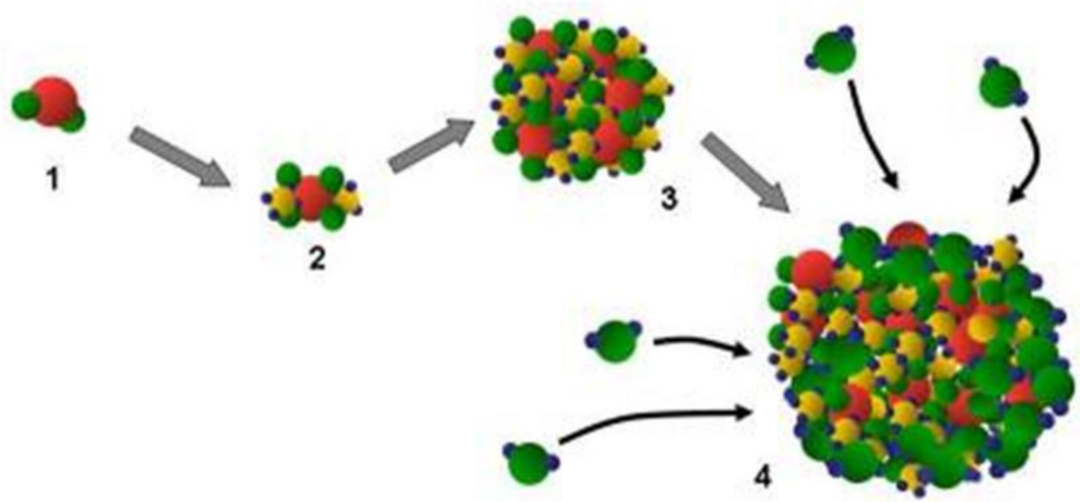


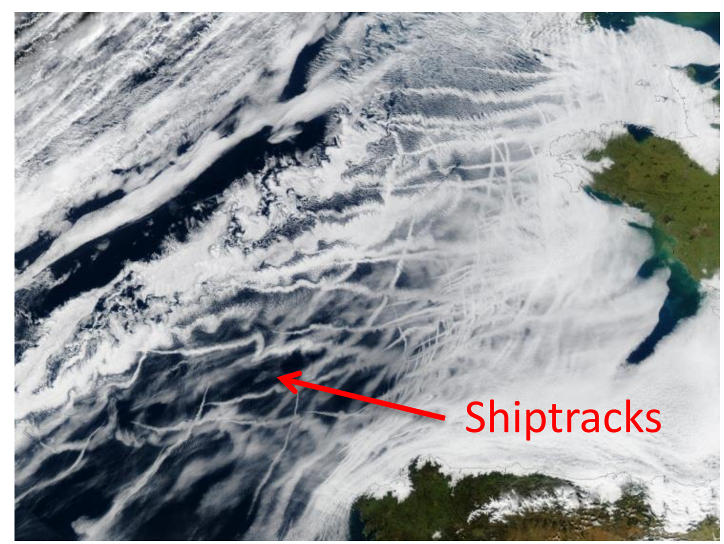
Makino and Ogawa, 1984

GCR-CN-CCN-Cloud Hypothesis



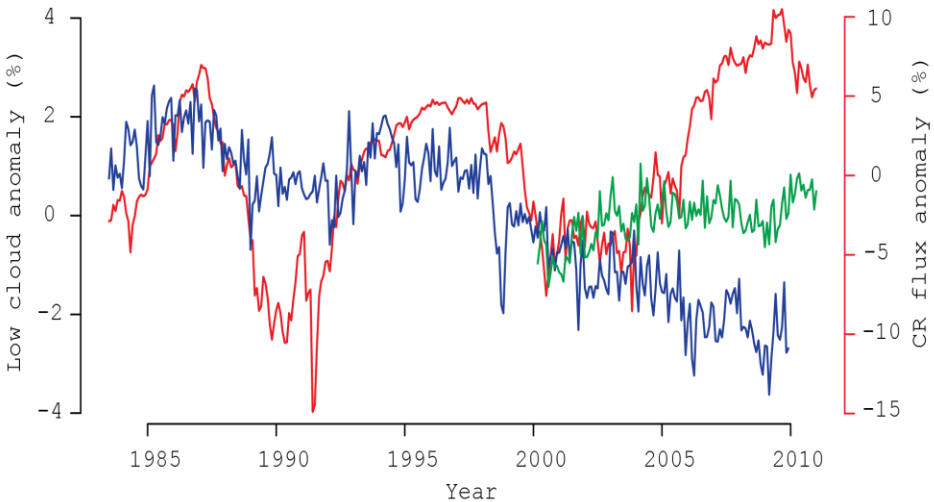
Yu, 2002





Long-term cloud data doesn't support GCR-cloud link

MODIS & ISCCP low clouds (1983 – 2012)

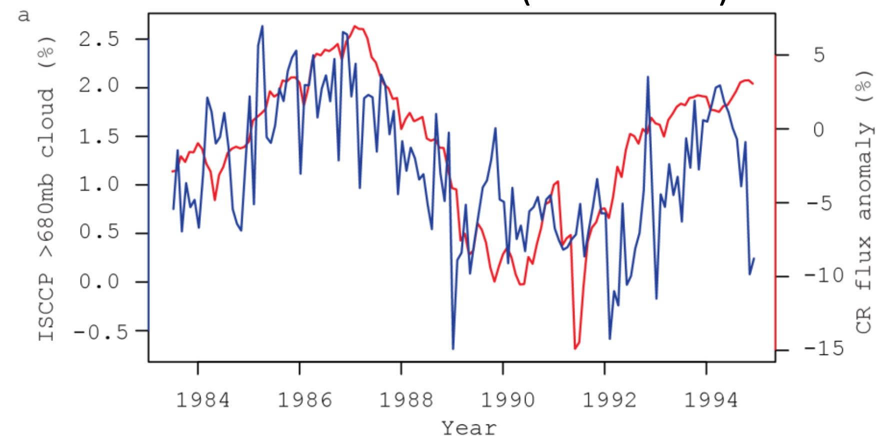


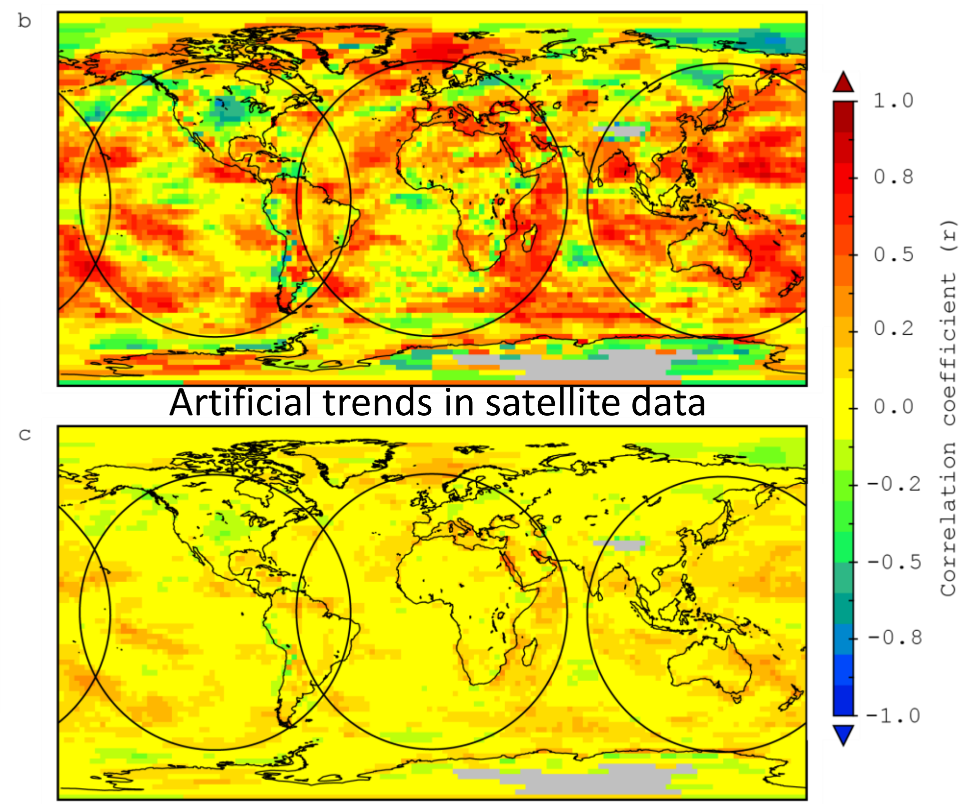
Cosmic ray flux (from N.M. data)

ISCCP low-level cloud

MODIS low-level cloud

ISCCP low clouds (1983-1995)

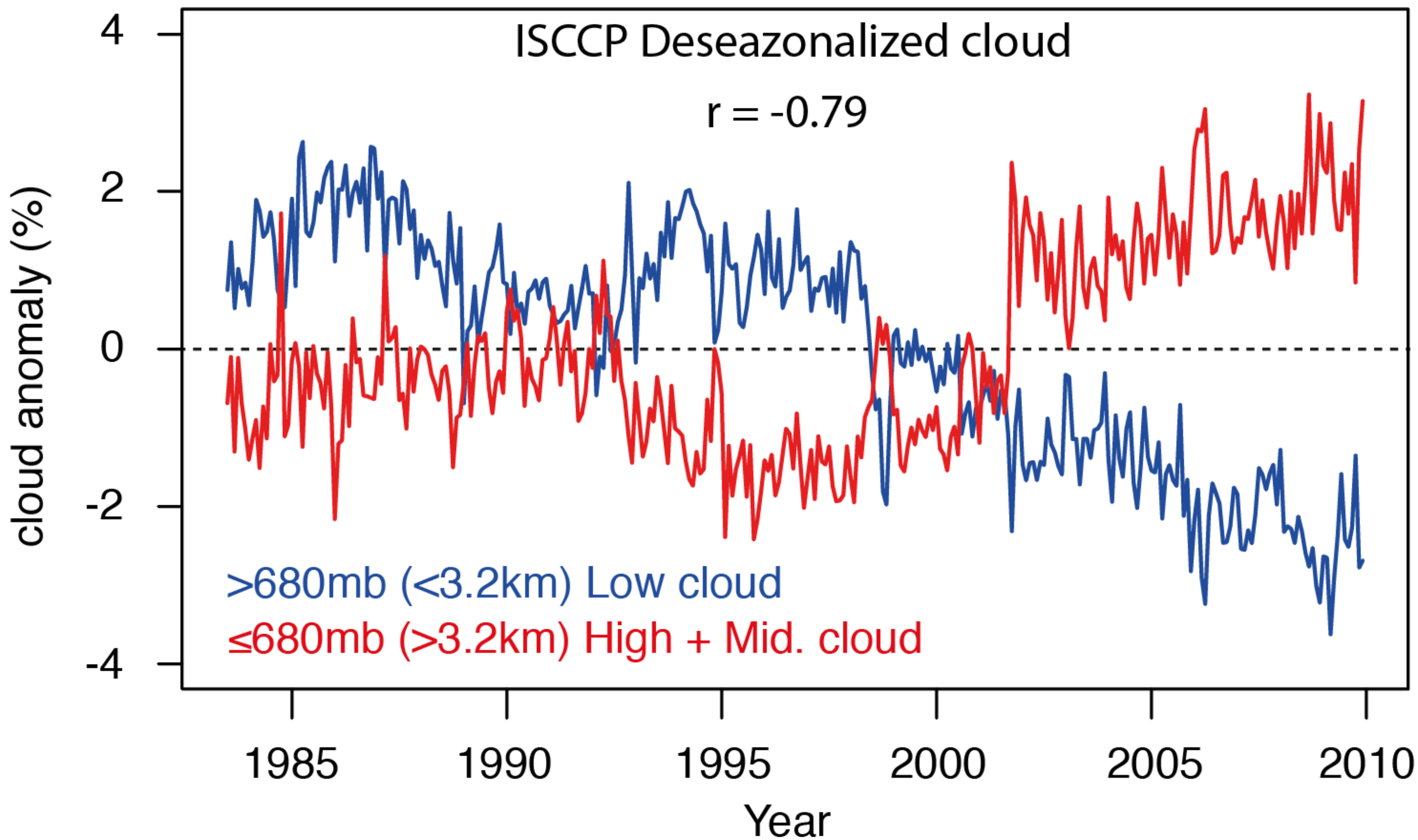




- Correlation only in low (<3.2km) cloud, 1983–1995.
- High correlation from 12-month smoothed data (df=4).
- Low (non-significant) correlation from unsmoothed data.

Laken et al. 2012, SWSC

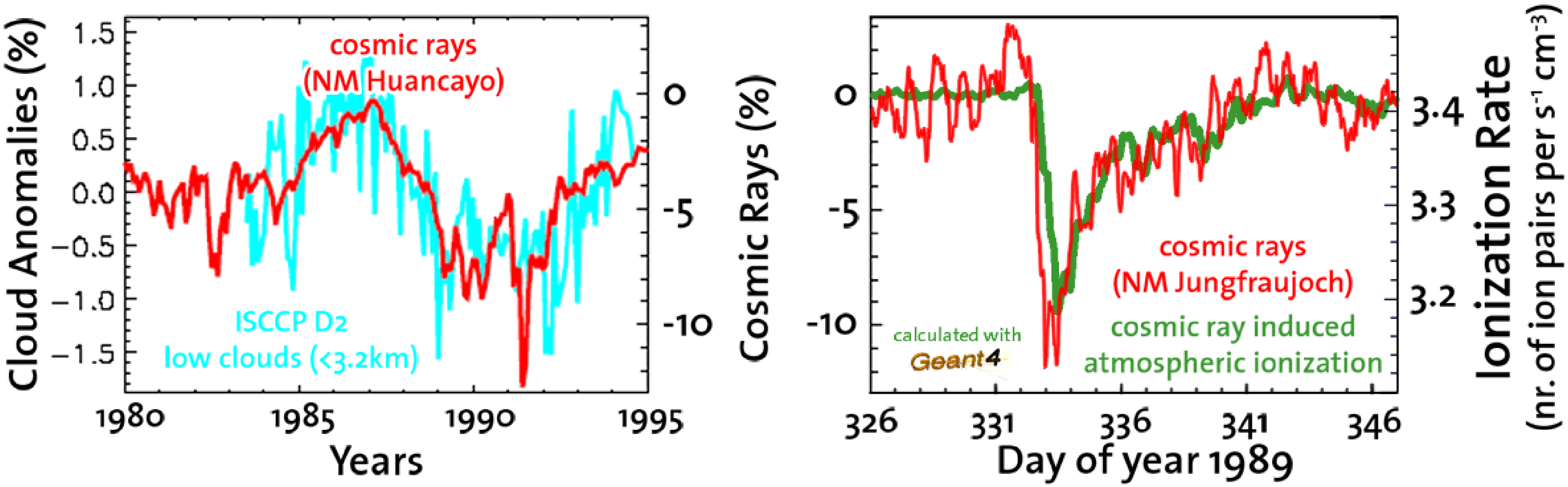
Correlation b/w low and high cloud cover in ISCCP: r = -0.79



Claims of observed CR – cloud links are made for low level clouds: <u>these data are not real!</u>

Short-term studies give a further opportunity to test GCR-cloud hypothesis

• Short-term changes in cosmic rays (e.g. Forbush decreases) are comparable in mag. to variations during an 11yr solar cycle.

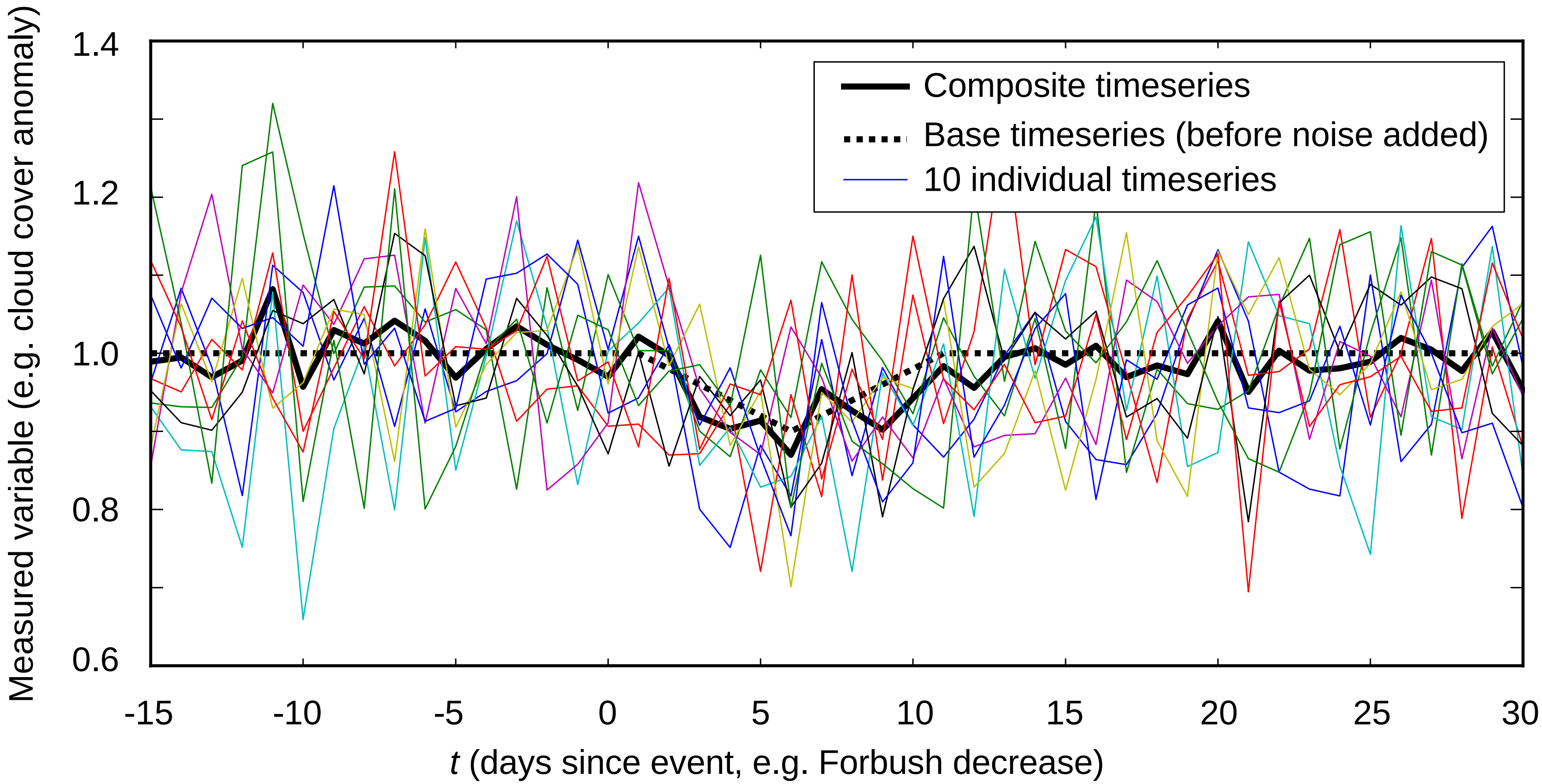


Čalogović et al. 2010, GRL

• If a response time in clouds is at most one week, it should be fast to observe changes in clouds on short timescales (Arnold, 2007).

• Problem: meteorological variability in clouds has to be **reduced** to be able to detect the signal.

What is a composite and why is it useful?



- Successive averaging of events (in time or space)
- Used to increase signal-to-noise ratio (SNR)
- Enable detection of small amplitude signal against large variability

Composite studies of Fd events and cloud properties show conflicting results:

positive correlations…

Tinsley & Deen, 1991; Pudovkin & Vertenenko, 1995; Todd & Kniveton, 2001; 2004; Kniveton, 2004; Harrison & Stephenson, 2006; Svensmark *et al.*, 2009; Solovyev & Kozlov, 2009; Harrison & Ambaum, 2010; Harrison et al. 2011; Okike & Collier, 2011; Dragic et al. 2011; 2013; Svensmark et al., 2012

…no correlations or inconclusive results…

Pallé & Butler, 2001; Lam & Rodger, 2002 ; Kristjánsson *et al.*, 2008 ; Sloan & Wolfendale, 2008; Laken *et al.*, 2009; Čalogović *et al.*, 2010; Laken & Kniveton 2011; Laken et al., 2012

… and even negative correlations

Wang *et al.*, 2006; Troshichev *et al.*, 2008.

What are potential causes of this conflict?

1. A [CR–cloud] signal exists above detection threshold, and some studies have found it.
2. A signal does not exist above detection thresholds, and positive results were false-positives.
3. A signal exists sometimes above detection thresholds, and thus some results find it and others don't given different criteria. **

** e.g. true if signal only detectable for high mag. events., or if CR–cloud link a secondary phenomena.

How to proceed?

To narrow these possibilities, we should find the optimum procedure for generating and assessing composites:

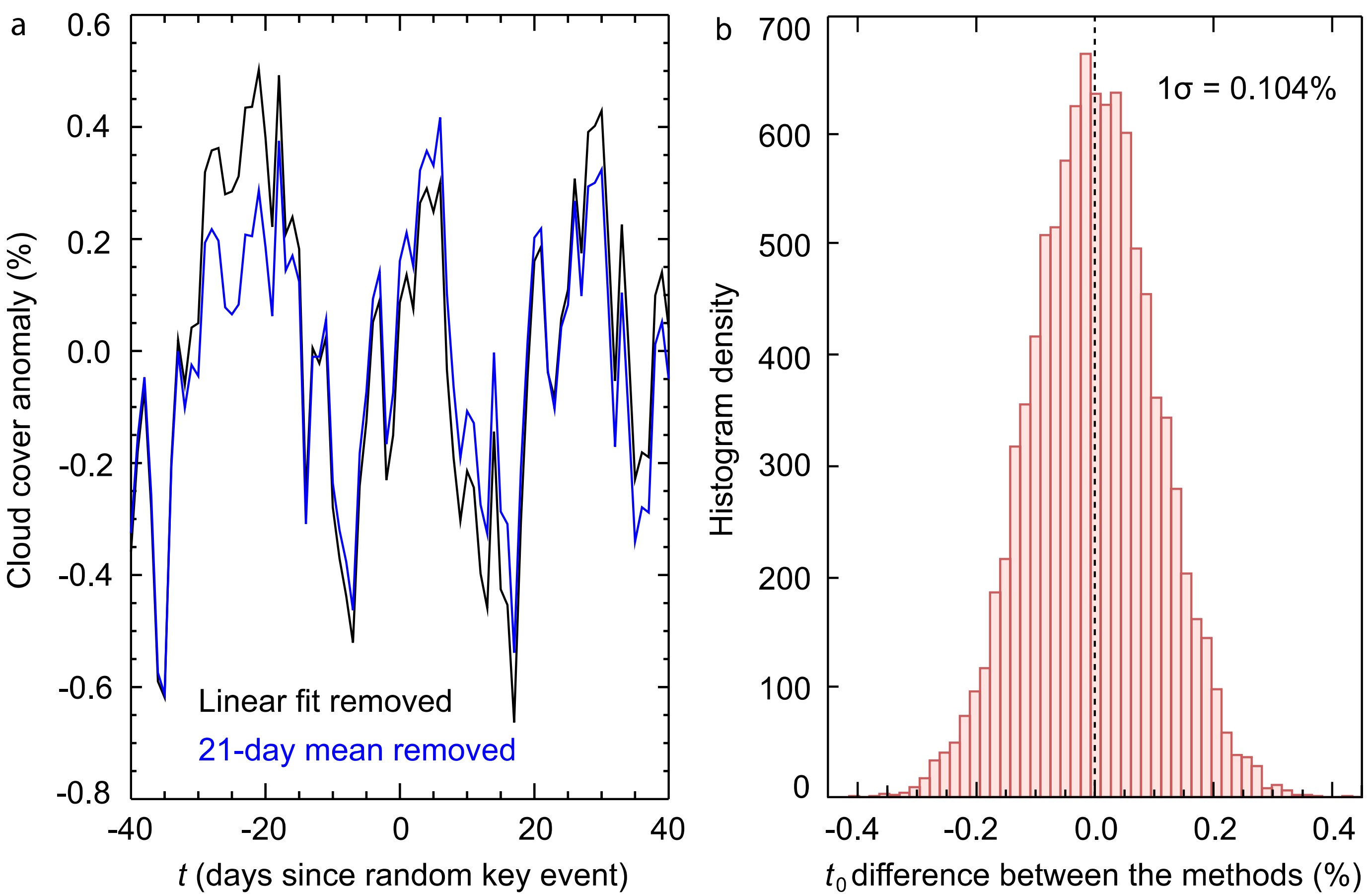
1. Creating anomalies for testing
2. Accurately identifying significance (p-values)

Composites should be made with anomalies rather than raw data, as unrelated variations can be reduced

- **minimize variations in data unconnected with hypothesis testing**

Two methods of removing variations compared for random composite (n=20): linear trends removed (black), and a 21-day running mean (only) removed

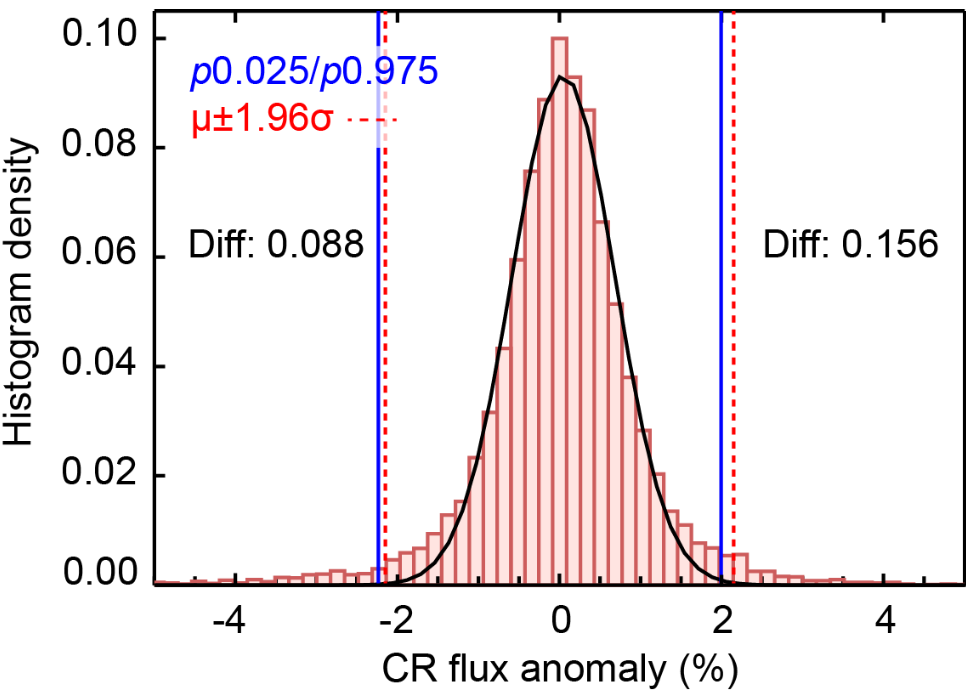
a) one random composite over $t_{\pm 40}$ *b*) 100,000 composites at t_0



Differences (*b*), indicates remaining non-linear variations in composite from synoptic scale variability. If not removed, **this will bias results of composites**.

Calculate thresholds for statistical significance Monte Carlo approach

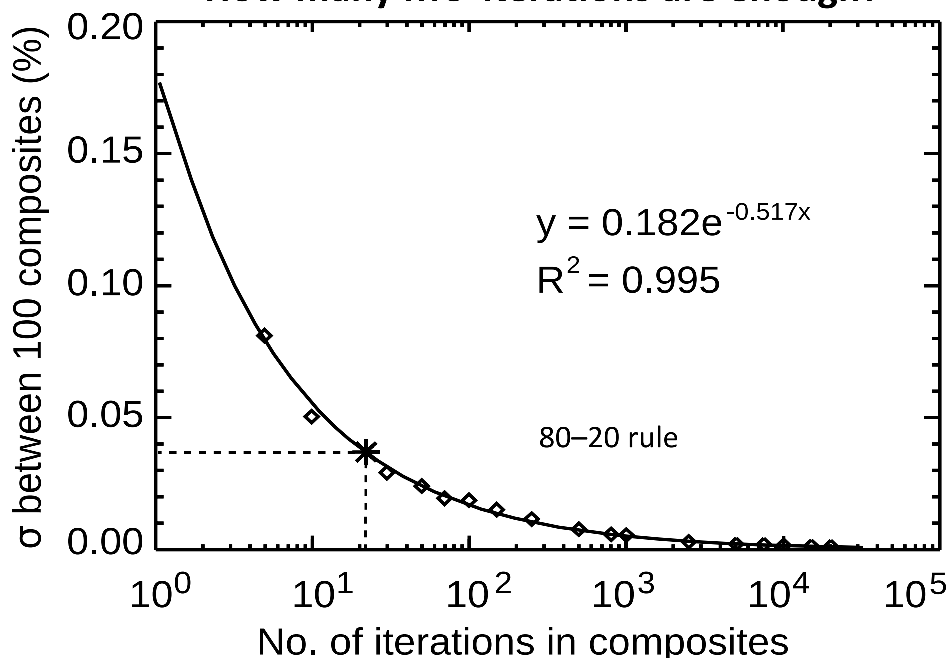
Distribution of daily anomalies



By generating large populations of random events identical in design and parent data to a real composite, the probability (p) of obtaining a given value in a real composite by chance can be accurately known.

This has advantages over traditional tests (e.g. T/U tests), as it <u>requires no minimum sample size or specific distribution, and doesn't need adjustment for autocorrelation</u>.

How many MC iterations are enough?



Wait wait wait, rewind…

I have skipped a step…

I explained how to test significance of a value in a composite **before** I discussed a crucial step of how to design a composite.

Why?

Because awareness of the significance thresholds should be built in to the design of composites!

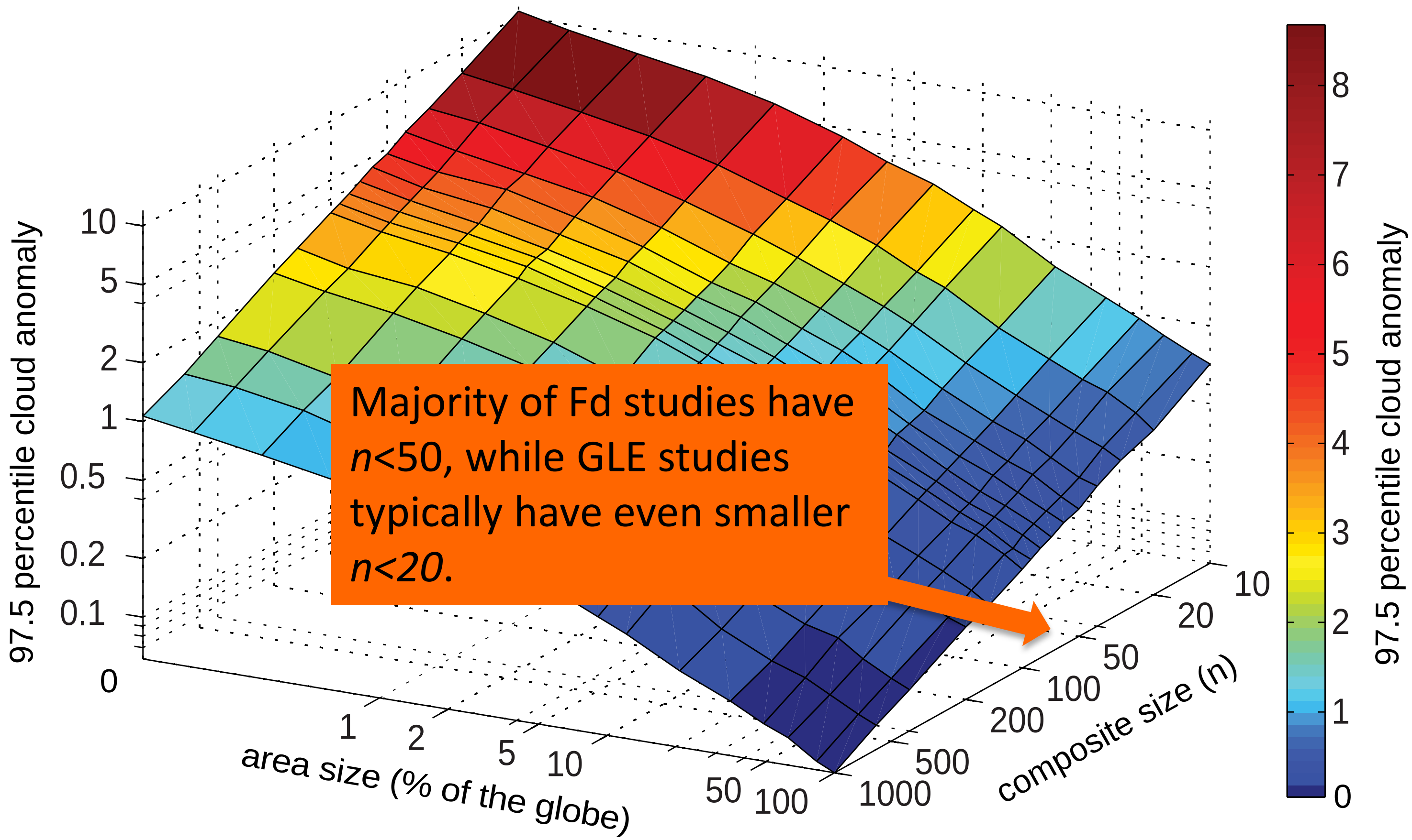
Designing a composite

How big for the sample area and how many events?

Noise levels of data govern detectability of a signal. In composites, the noise varies with both the spatial area (*a*) considered by the data, and the number of composite events (*n*).

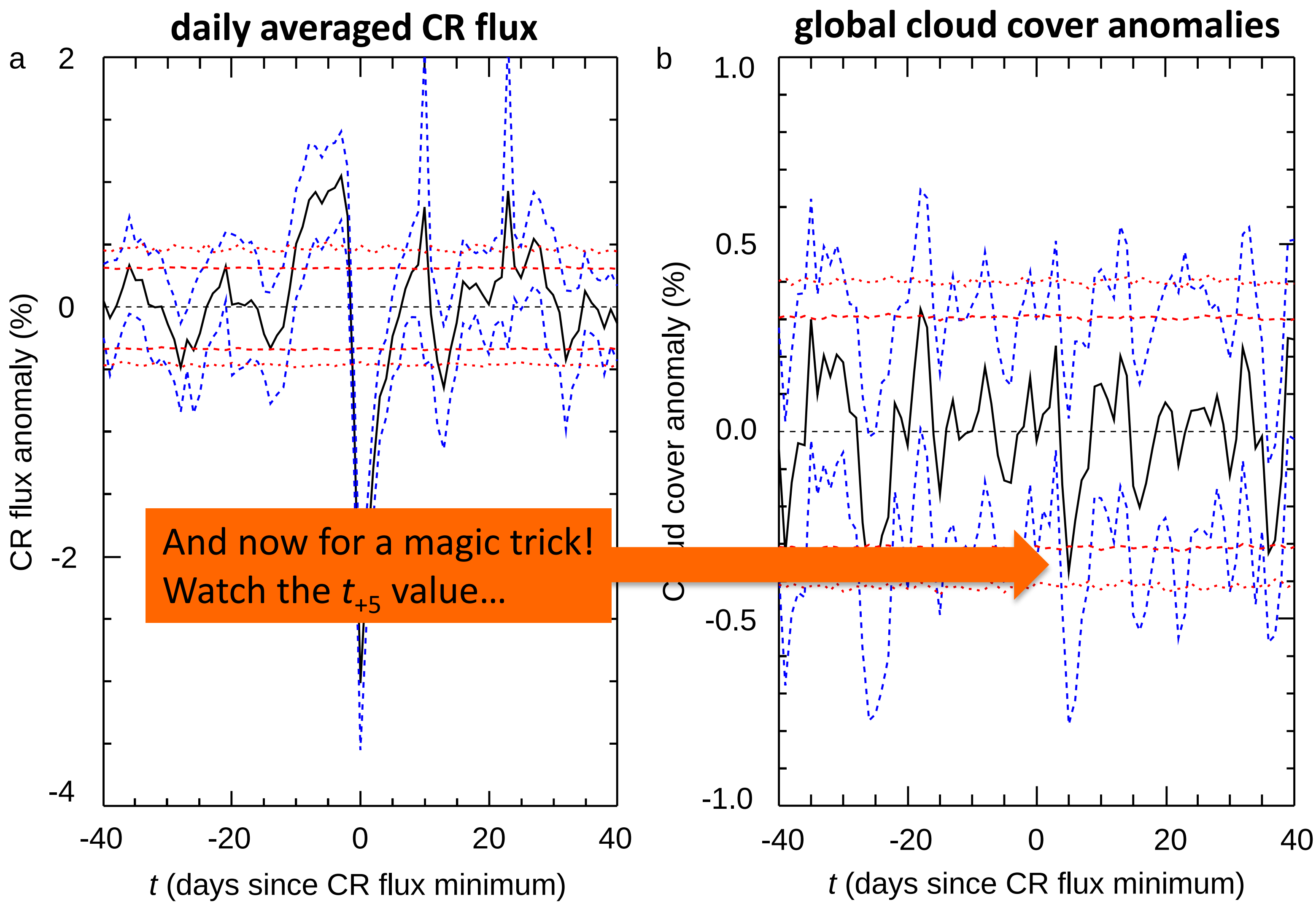
'*Noise*' indicated by 97.5th percentile values from 10,000 random composites of varying *a* and *n* size.

Each point of grid represents another independent set of 10,000 MC simulations (as in previous slide)



Given possible upper limit effects, and a consideration of noise, it is possible to see how large *a* and *n* would need to be at minimum to see a hypothesized effect.

Composite examples



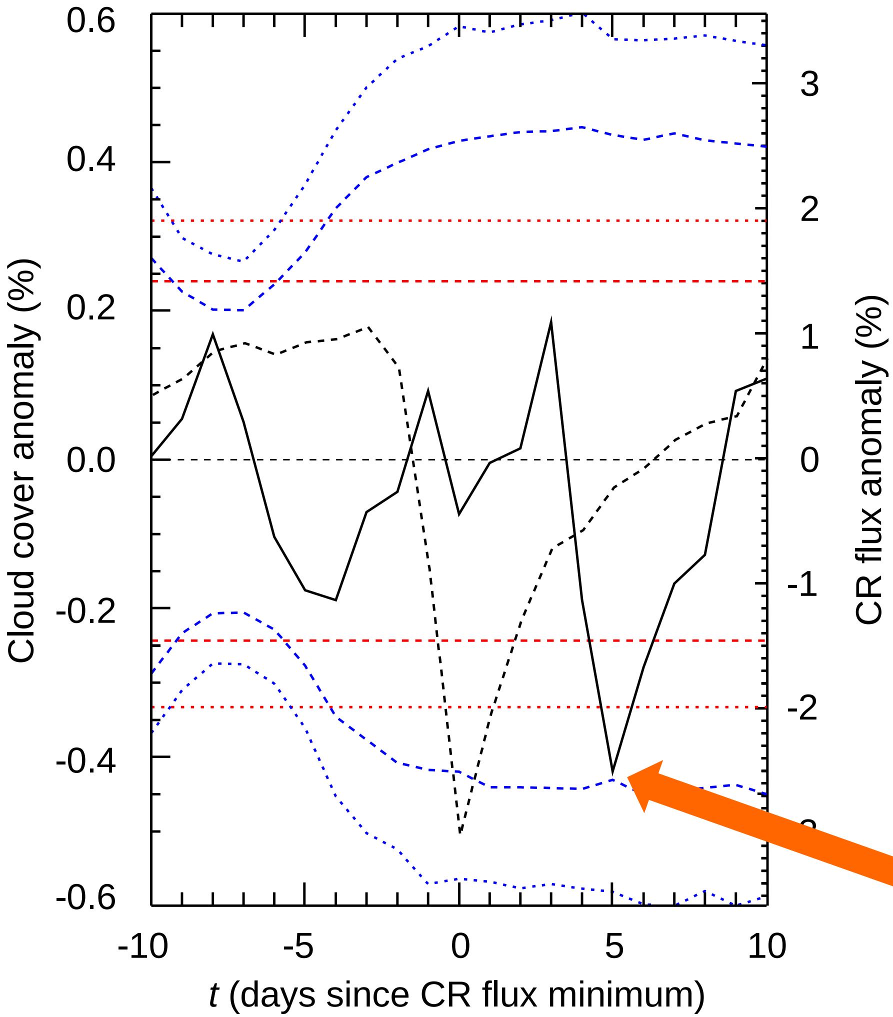
Fd events, adjusted so t_0 aligned to largest CR deviation (not Fd onset) n = 44

Anomalies (black lines) = daily mean – 21-day running mean

±1.96 Standard error of mean (SEM) (dashed blue lines)

Two-tailed p0.05 (dashed red lines), two-tailed p0.01 (dotted red lines)

How to obtain a false positive



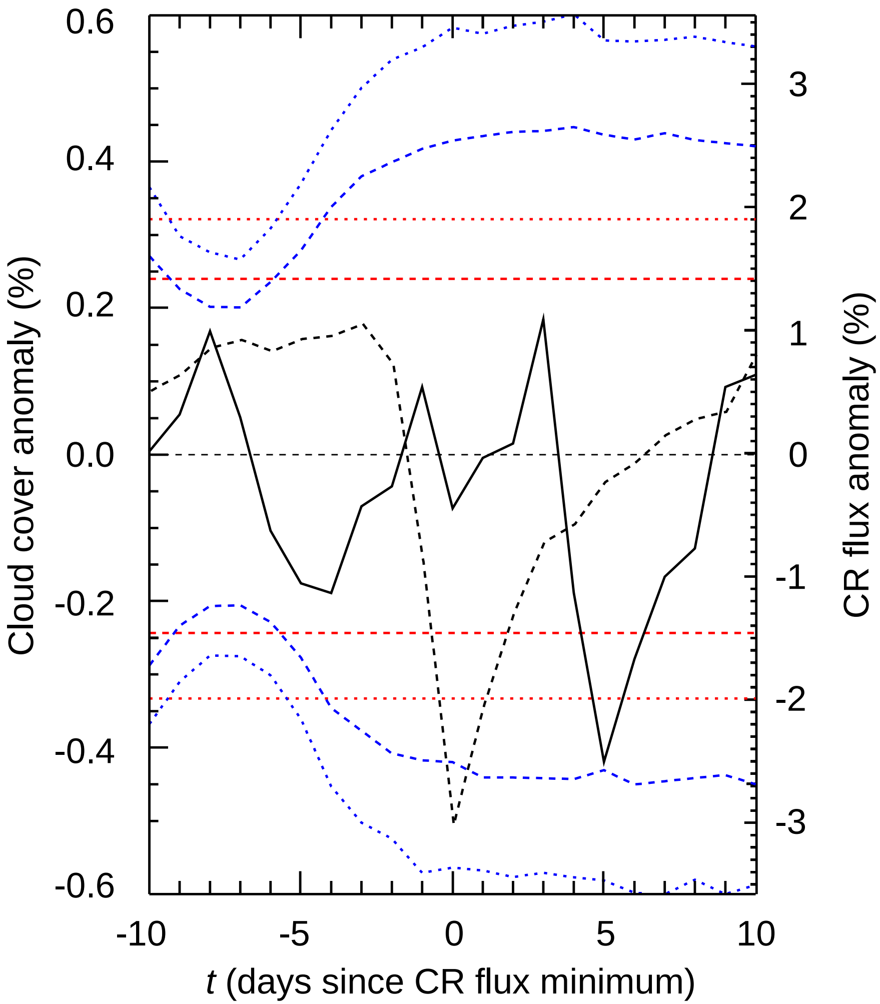
The new t_{+5}!

Cookbook…

1. Identify a base or 'undisturbed' period before the key events, that represent 'normal conditions' (e.g. figure uses t_{-10} t_{-5})
2. Calculate deviations against this 'undisturbed' period (i.e. subtract every *t* point from mean of 'normal conditions')
3. Statistically compare the data anywhere to the 'undisturbed' period (e.g. T-test, or even MC from the base period [red lines p0.05 p0.01])

Normalization to base period reduces population variability towards base period, narrowing confidence intervals.

How to ~~obtain~~ avoid a false positive



Overcoming bias with Monte Carlo (MC):

Use MC-calculated thresholds which are independent for each *t* point - autocorrelation is automatically accounted for, even if a base-period normalization is applied! (expansion of confidence intervals)

Autocorrelation effects are automatically taken in to account - random samples in the MC all treated with an identical approach to the analyzed composite

Comparing the blue lines to the red, indicates how tests that can't account for autocorrelation would view the data as being from different populations.

Two different results for t_{+5} (the above with a mean p<0,05 and the earlier, with a mean 0.01>p<0.05: which is correct? (clue, it's a trick question)

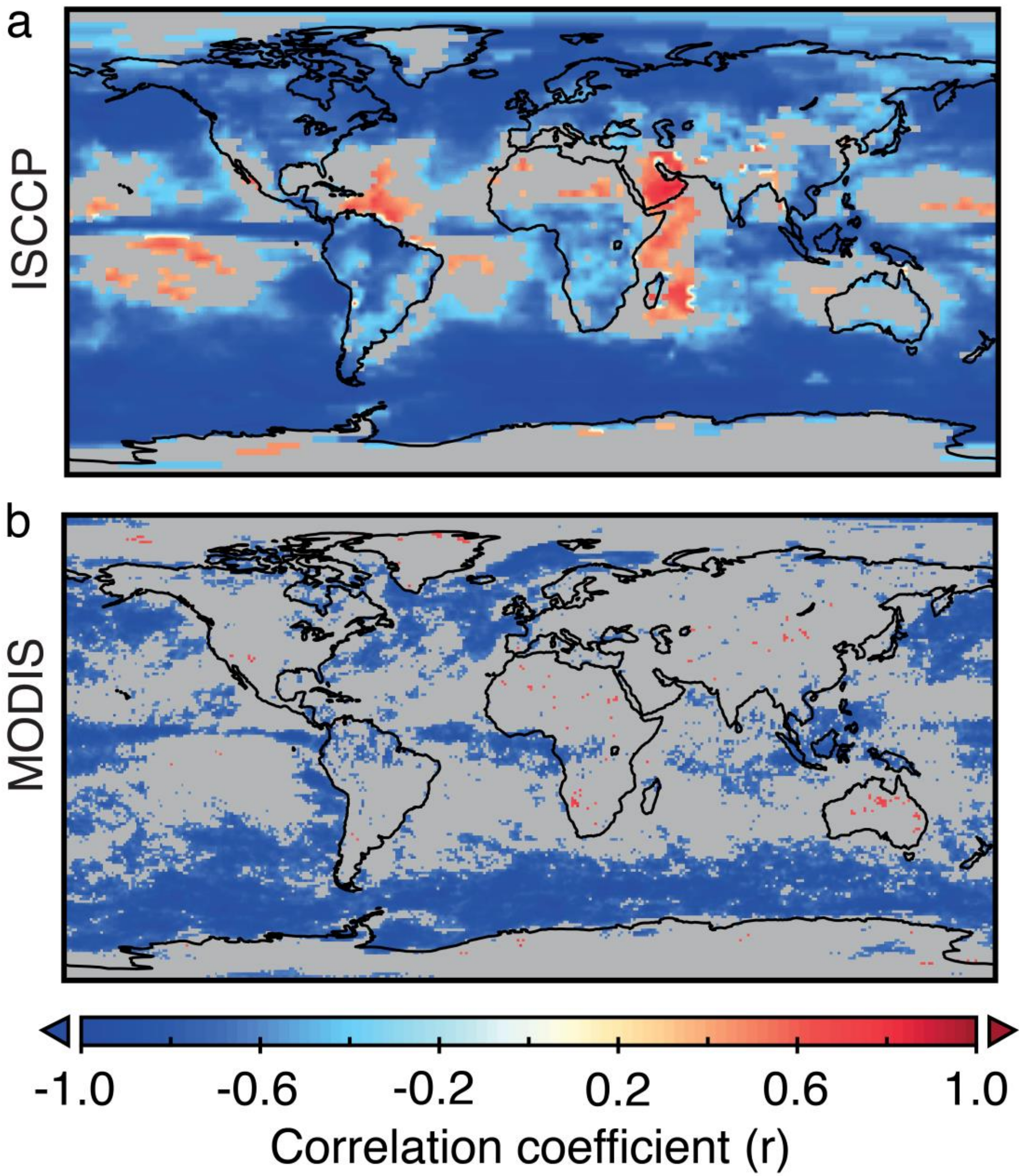
Conclusions

- Minor methodological differences in composite analysis can produce conflicting results.
- These are the likely source of discrepancies between cosmic ray – cloud composite studies.
- Neither short-term or long-term studies of satellite observed cloud support a cosmic ray – cloud link.

You can find more info, publications, and updates at:
www.benlaken.com
Twitter: benlaken

Thanks very much for listening!

Further info on cloud low/mid+high anticorrelation from ISCCP/MODIS:



P<0.05 statistically significant correlation between high+mid and low cloud cover

Degrees of freedom constrained by effective sample size