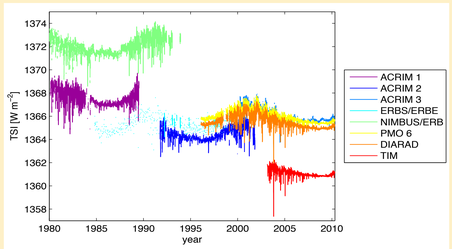
How to make composites out of multiple observations ?

Thierry Dudok de Wit[1], Greg Kopp[2]

[1]LPC2E, CNRS/University of Orléans, [2]LASP, Boulder

No composites without data stitching…

Total Solar Irradiance (TSI) records overlap but disagree in their absolute level. Merging them into one single composite is a crucial but also very challenging task. This has also become a topic of considerable debate [Krivova et al., Fröhlich & Lockwood, Willson & Scafetta, etc.].



Some of the TSI observations made since 1980.

What is the best strategy ?

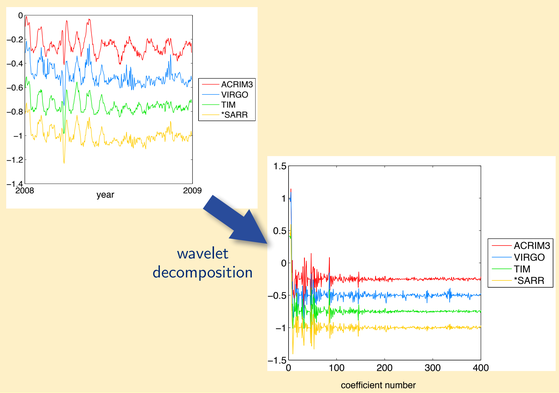
- ☐ take a weighted average
- ☐ for each day, select the least noisy instrument
- ☐ for each day, select your favorite instrument
- ☐ take a simple average
- ☐ none of the above

The approach we advocate

- work in wavelet domain rather than in the time domain: compute the wavelet transform to convert each record into a series of wavelet coefficients
- for each day, merge the wavelet coefficients by using a **Bayesian approach** = use all the available information - don't discard any data
- do an inverse wavelet transform to get back in the time domain. End result is the Bayes composite and its confidence interval (posterior probability distribution)

Multiscale analysis is needed

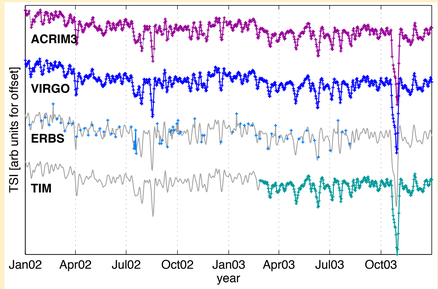
Using a discrete wavelet transform, we convert each record into a series of time-dependent wavelet coefficients (time-scale decomposition).



How to handle data gaps ?

Data gaps are a problem when computing the wavelet transform. However, they can be easily filled in by expectation-maximization [Dudok de Wit, A&A 533 (2011)]. We flag them, so that they do not affect the final outcome.

Rationale: all TSI records emanate from observations of the same Sun with instrument measurement noise added.



Example of observed TSI records (color) and their interpolation (grey), based on the expectation-maximization technique.

Conclusions

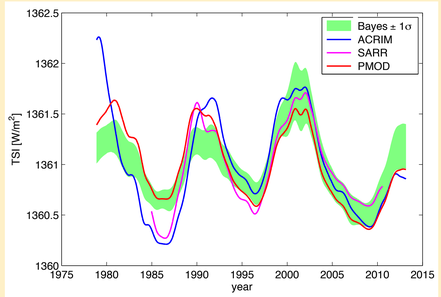
- The Bayesian approach has been successfully used before for paleoclimatic reconstructions [Tingley et al., J. Climate 23 (2010)]. Here we advocate the same approach for merging TSI records because it is rigorous and **naturally incorporates all available information**.
- By merging of the observations in wavelet space, we overcome problems caused by discontinuities in time & are able to pick out at each scale those records which are most relevant.
- The Bayes composite we obtain (still preliminary) indeed shows **better noise properties** than other composites.
- This approach is now being considered by an ISSI team (lead by Greg Kopp) that will deliver a new TSI composite. It will also be used for merging SSI records in the SOLID project (see talks by Margit Haberreiter & Micha Schöll).

Still in progress

- A fully Bayesian scheme requires a lot of computation & mathematics. We're still working on that… and this will take time
- This method requires **realistic confidence intervals for the observations** (time-dependent or not), which hardly exist. We're now developing our own (empirical) scheme for estimating such confidence intervals.

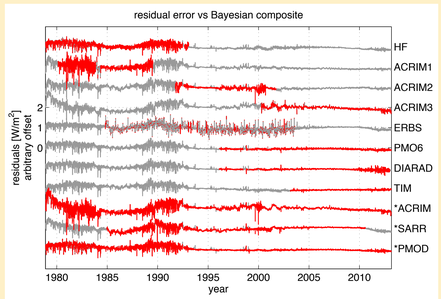
The TSI composite

Next, to build the composite:
1) Do a weighted average of the wavelet coefficients by using a Bayes scheme to determine the optimal weights (based on the uncertainties given for each record).
2) Inverting the wavelet transform gives the Bayes composite and its confidence interval.



Bayes composite (green) and 3 common composites

How does each TSI record compare with our composite ?

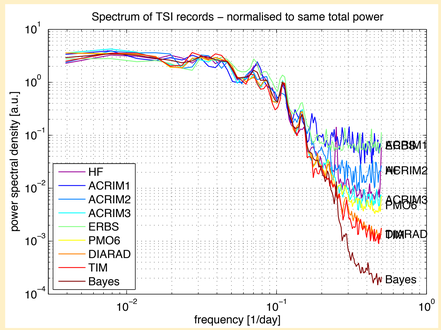


Th residuals = difference between observations and the Bayes composite.

red = observed
grey = interpolated

Before 1990 the VIRGO and ACRIM composites don't describe the most probable value of the TSI but rather the upper and lower edges of the distribution. The agreement is better for the last solar cycle, except for the SARR composite.

The Bayes composite has lower high-frequency noise than each individual record



Power spectral densities of all records (estimated by windowed Fourier transform).

The Bayes composite exhibits a lower noise floor at high frequency (daily-weekly variations).