Estimating multi-decadal trends in ENSO: A case study using Darwin SLP

A.M. Chiodi and D.E. Harrison

Motivation

- Revisit the issue of whether or not there is evidence of long-term change in the El Nino-Southern Oscillation (ENSO): Are multi-decadal and longer-term perspectives consistent?

- Use a high quality, long-term (136yr) record of marine climate to examine issues faced in trend studies.

Outline

- Background/Motivation
- Results

  -Sliding-window examination of trends in Darwin SLPA and their statistical significance

  -Relationship between trends over full-record (136 years) and shorter decadal/multi-decadal segments

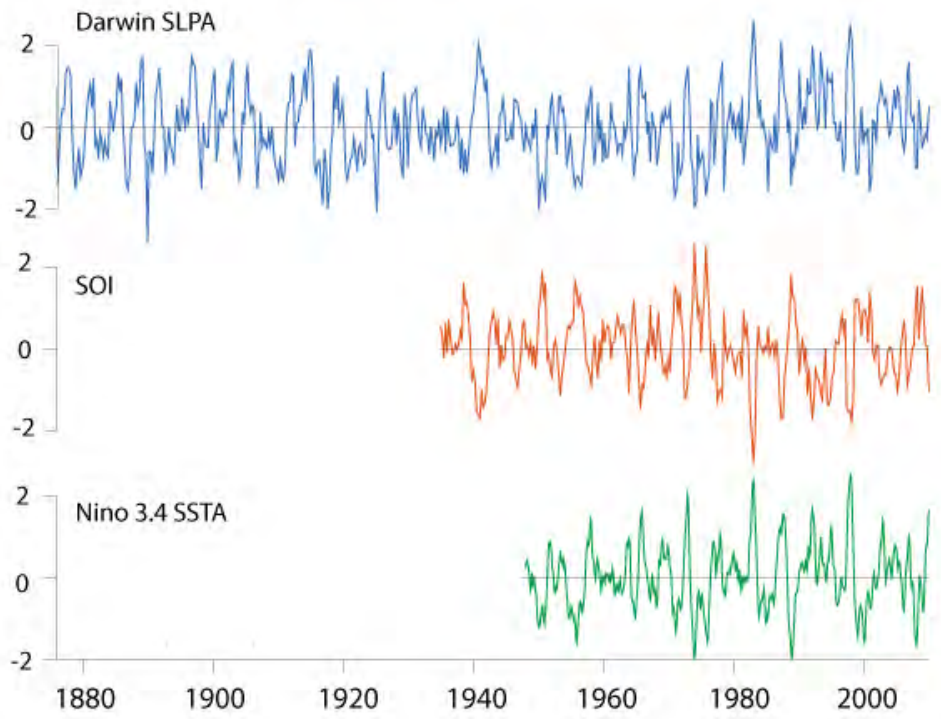
  -Difficulties in using statistical tests on “red” time series
- Summary/Conclusions

ENSO and Darwin

- Darwin SLPA is a good proxy for ENSO state

  Darwin SLPA is highly correlated with other ENSO records at interannual and longer timescales (r=0.9 and 0.8, for SOI and Nino 3.4, when smoothed with a 13-month Parzen filter)

- Previous studies suggested evidence of climate change in Darwin record (e.g. Trenberth & Hoar, 1996, c.f. Harrison and Larkin, 1998).
- IPCC model projections of future ENSO trends differ (Vecchi et al., 2008)
- Nicholls (2008) found a 50yr (1958-2007) trend in SOI when considering just some seasons, but not others.



Determining very long-term change with the available historical records

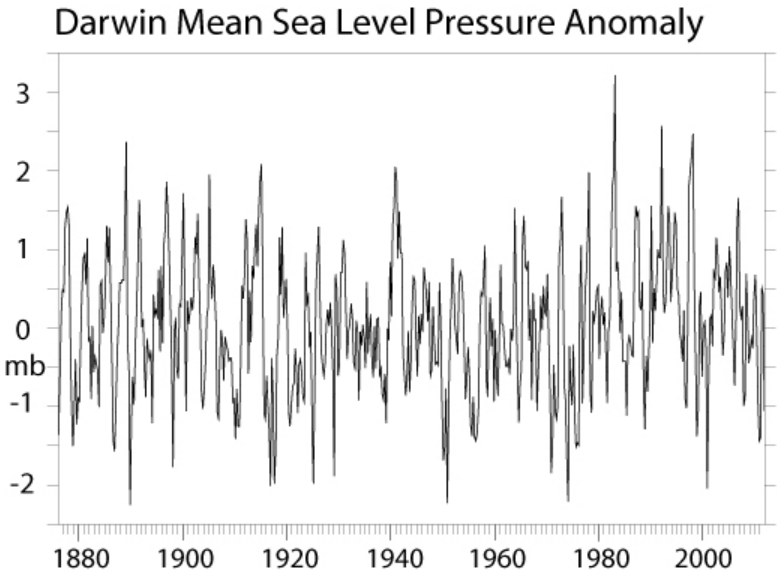
- Climate variability has been well sampled only recently
- Interest in determining which aspects of climate are undergoing long term change plus limits imposed by the historical availability of climate-quality records mean we often look for trends in climate variables over the full-length of the available historical records, yet are ultimately interested in trends over longer periods
- What are the risks of extrapolating trends in this situation?

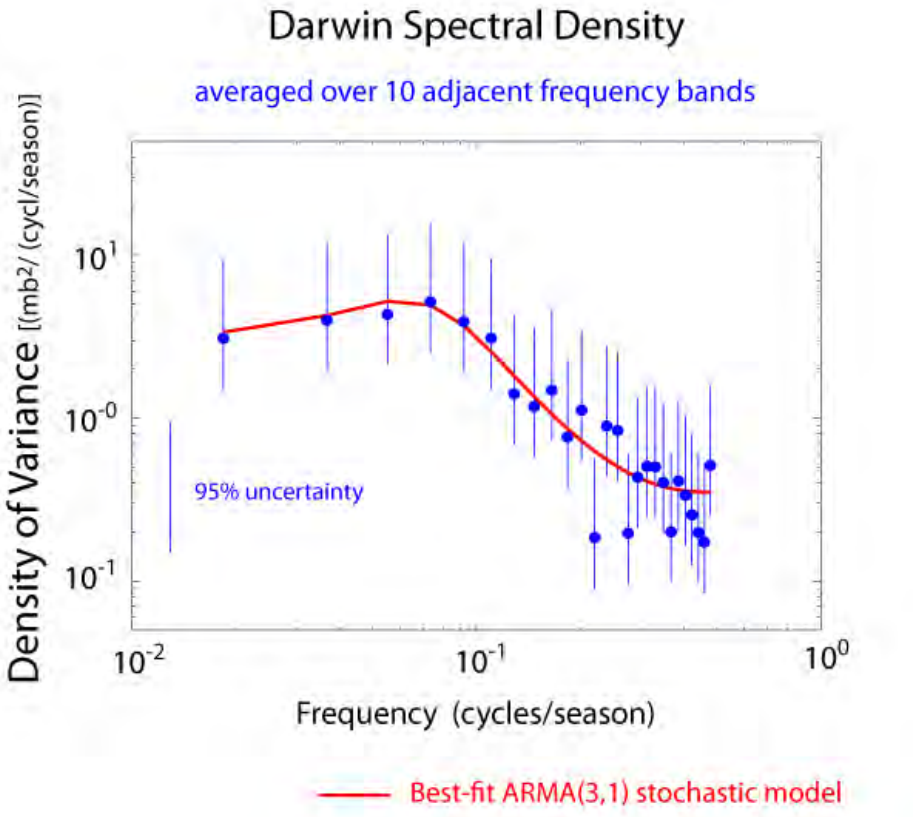
The role of tests for trend statistical significance

- We often look to tests for trend statistical significance to guide our interpretation of a given trend.

  "significant" = indicative of change in underlying processes
- But, reliably determining trend statistical significance can be discouragingly difficult for many geophysical time series, which have spectra that characteristically are "red".
- As the "redness" of a process increases, so does the range of trends seen in finite segments, even when there is no very long term trend.

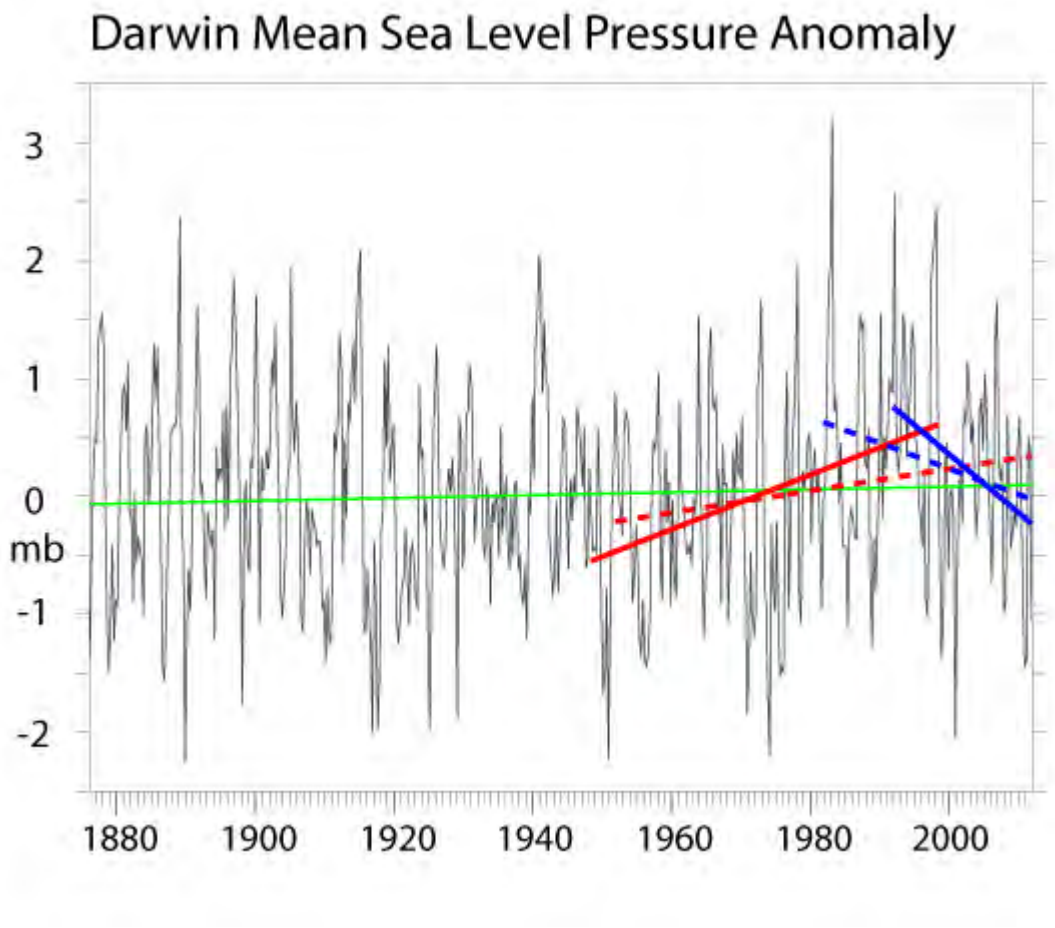
Darwin SLP Anomaly





- Like many geophysical records, Darwin's spectrum is "red", or contains more variance at lower than higher frequency.
- The reliability of tests for trend statistical significance depends on their ability to accurately estimate the amounts of stationary (no-trend) low frequency variability present

Multi-decadal trends: the importance of start/end points



Trends over different recent multi-decadal periods and the record-length (136 year) trend have different character

Last 20 years : significant/ negative (99%)

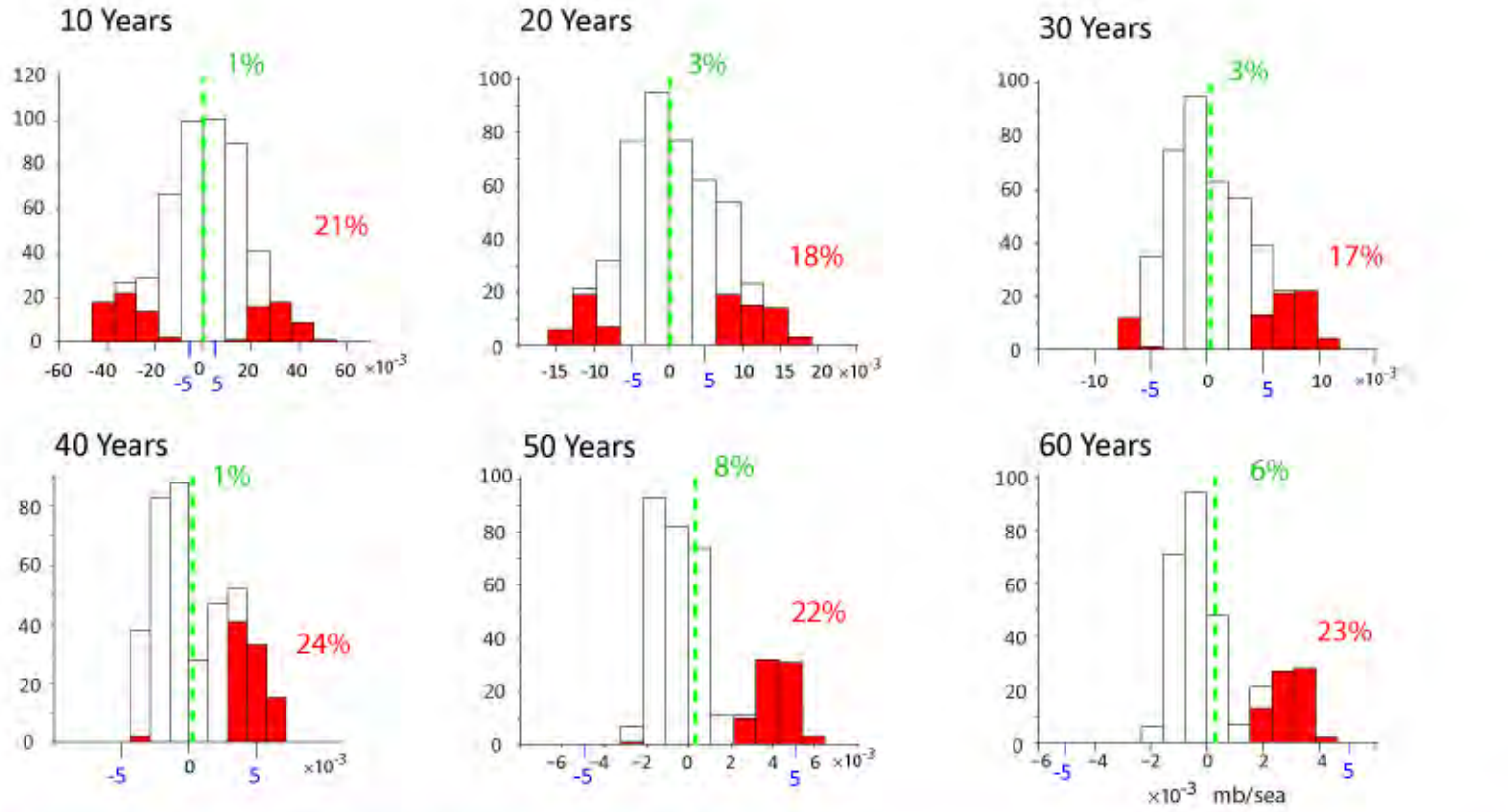
Last 30 years : nearly significant/negative (86%)

Last 60 years : significant/positive (95%)

1949-1998 : significant/positive (>99.9%)

1876-2011 : not significant (<66%)

Distribution of sliding-window multi-decadal trends in Darwin

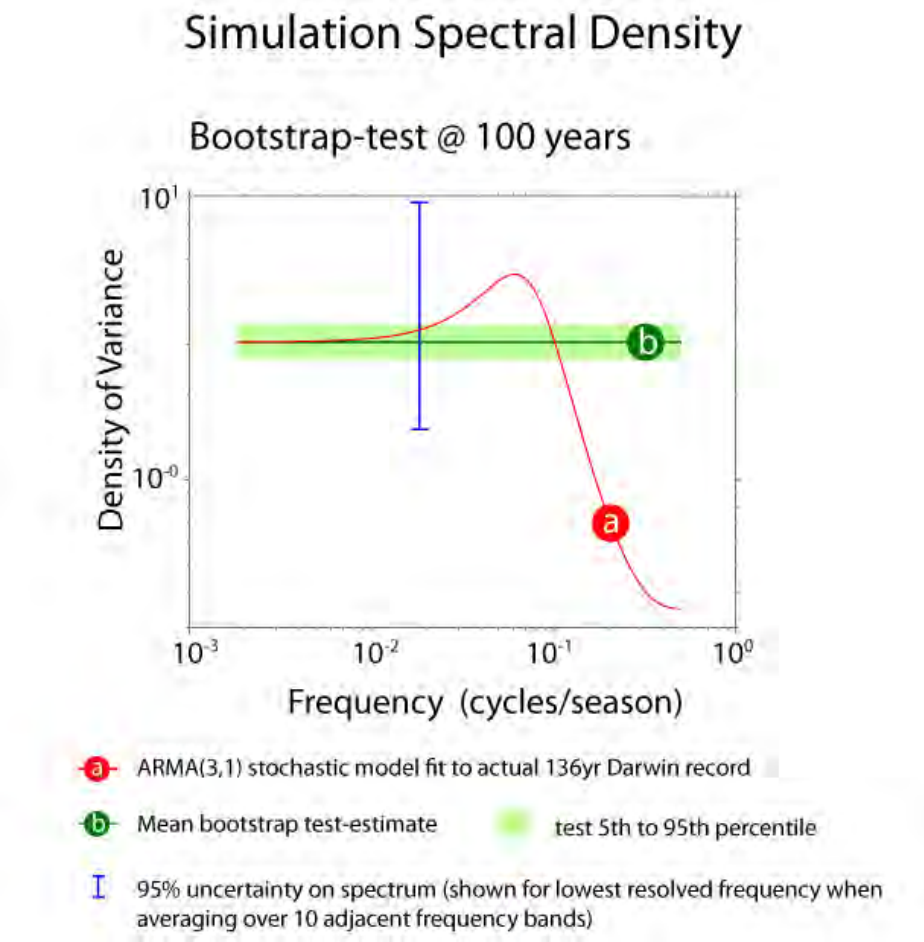


Multi-decadal trends do not provide a reliable estimate of the longer term (136 yr) behavior in this case.

Test for trend statistical significance

Test procedure:

1) Select record (or segment)
2) Determine trend (least squares fit)
3) Remove trend
4) Estimate amounts of non-secular variability present
5) Randomly generate many synthetic series based on results from step 4 to determine likelihood that the trend seen (step 2) is an artifact of this type (non-trend) of behavior.



We have confirmed that in the scenario that Darwin is well-represented by the stochastic model shown above (curve a), the version of the bootstrap test used here performs well when 100 years (400 seasons) of "data" are provided

Summary

- All tests considered find that the 136 year trend in Darwin is not statistically significant
- There is a rich variety of behavior seen at multi-decadal time scales, with relatively equal numbers of positive and negative trends
- Despite this, a disproportionally large number of multi-decadal sliding-window segments yield statistically significant trends based on tests that showed good performance in a best-fit scenario

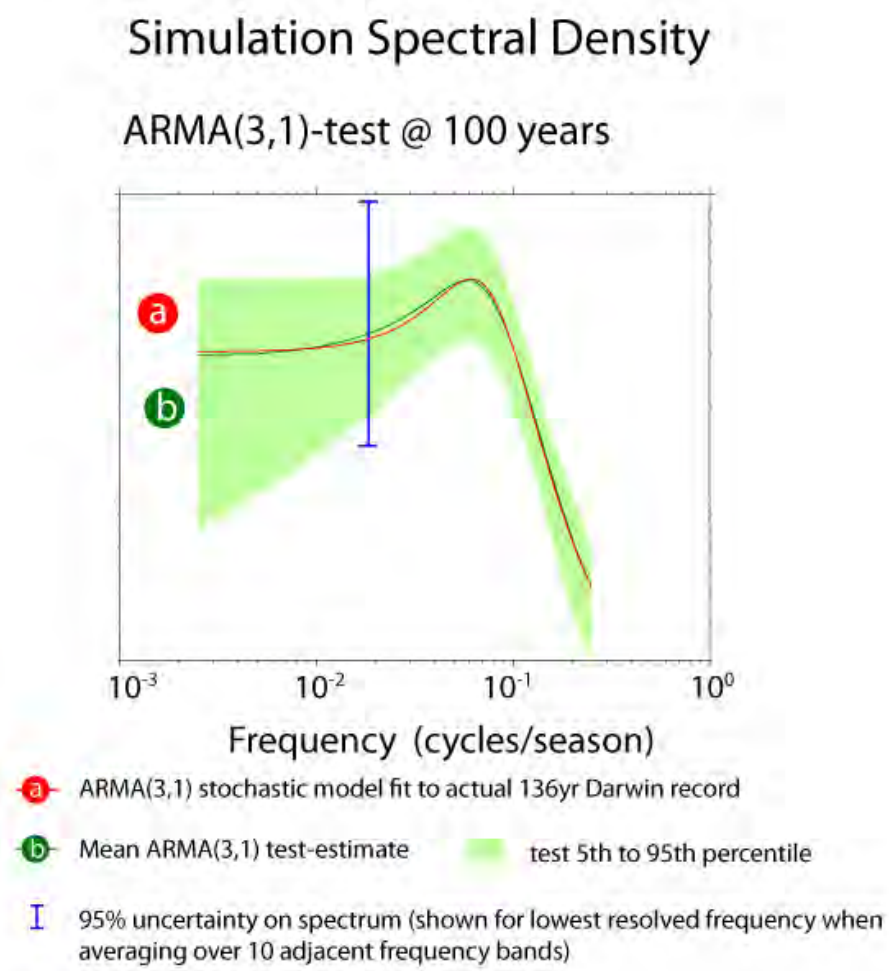
Conclusions

- We suggest that the best answer is that the null hypothesis that there is no very long-term trend in Darwin (ENSO) is true
- Focusing just on a particular multi-decadal segment could easily lead to a different story
- For Darwin, multi-decadal records are too short to be useful for inferring longer term trends. We suggest this is a feature of “red” time series
- Caution is needed when using statistical tests to judge the “significance” or likely persistence of a record-length trends
- Understanding the physical processes responsible for the observed multi-decadal variability and possible longer term change is key.

Test for trend statistical significance

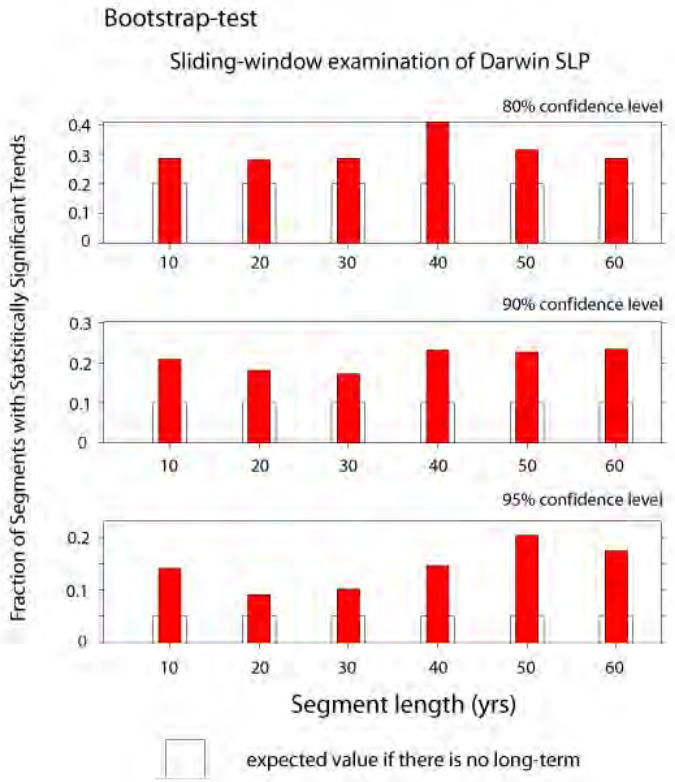
ARMA-simulation = parameters from actual 136yr Darwin record

ARMA-test = parameters from simulated (or sliding-window) segment



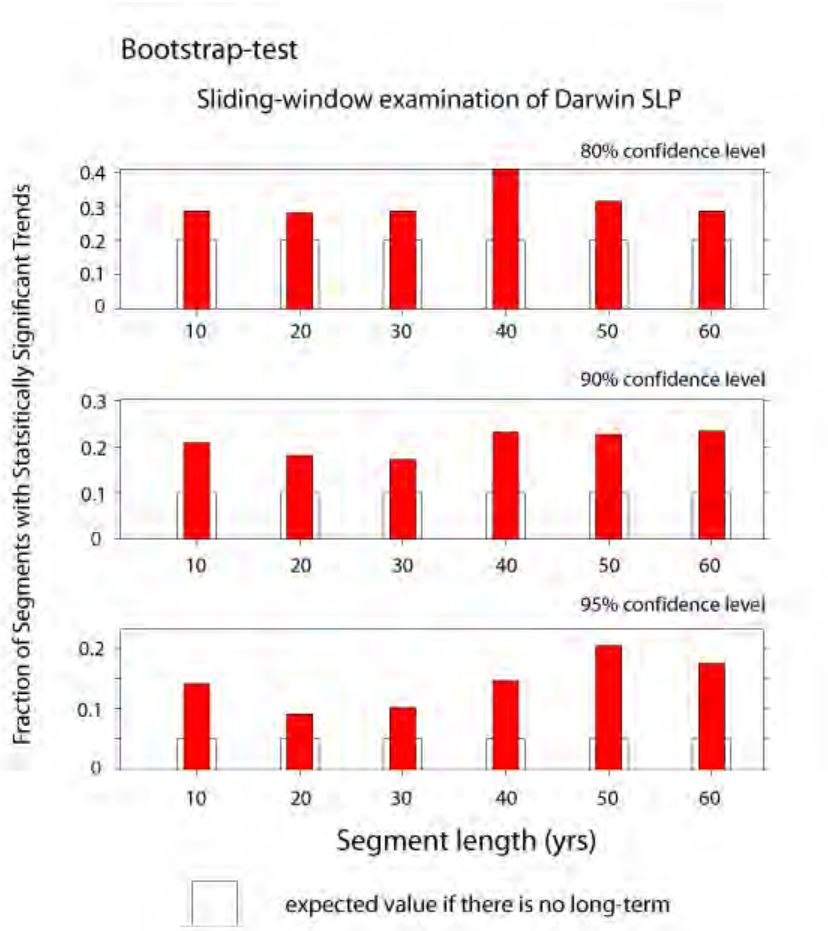
An ARMA-based test can also be used (c.f. Trenberth and Hoar 1996). In this case a larger range of test-estimated low frequency behavior is seen in the simulation experiment and results are less stable than in the bootstrap-case.

Statistically significant multi-decadal trends in Darwin



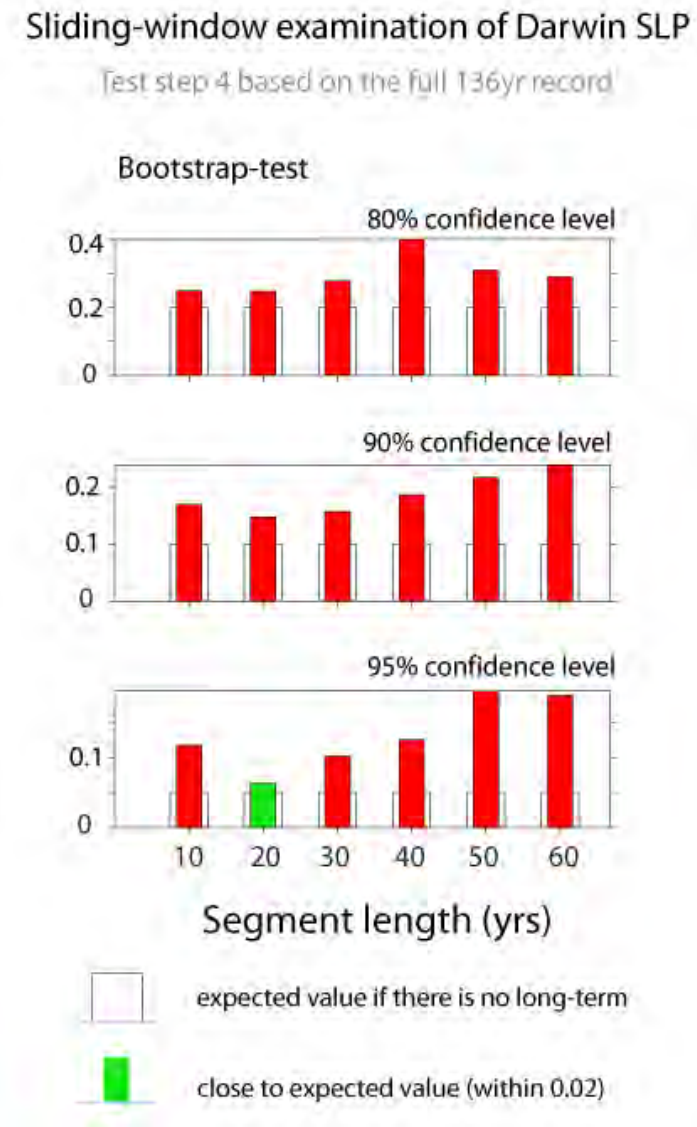
More sliding-window segments (by a factor of 2 to 5) reach trend statistical significance than are consistent with the null hypothesis that there is no very long-term trend in Darwin, being true.

Statistically significant multi-decadal trends in Darwin



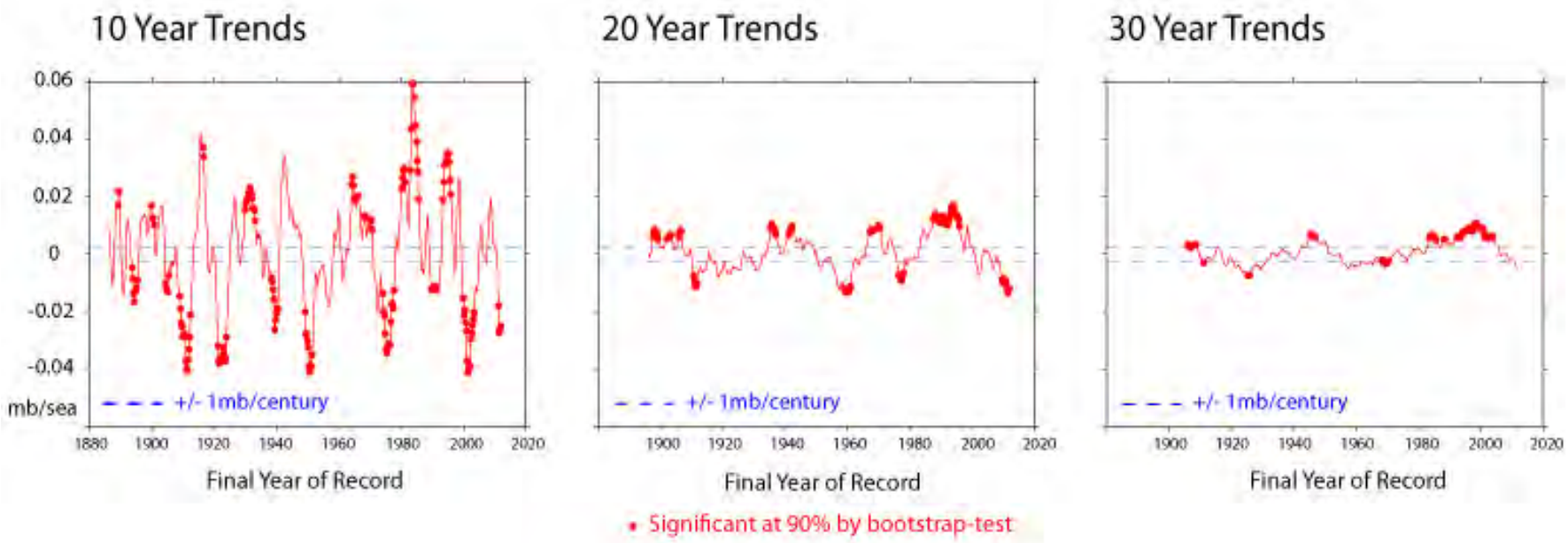
More multi-decadal sliding-window segments (by a factor of 2 to 5) reach trend statistical significance than are consistent with the null hypothesis that there is no very long-term trend in Darwin, being true.

Statistically significant multi-decadal trends in Darwin



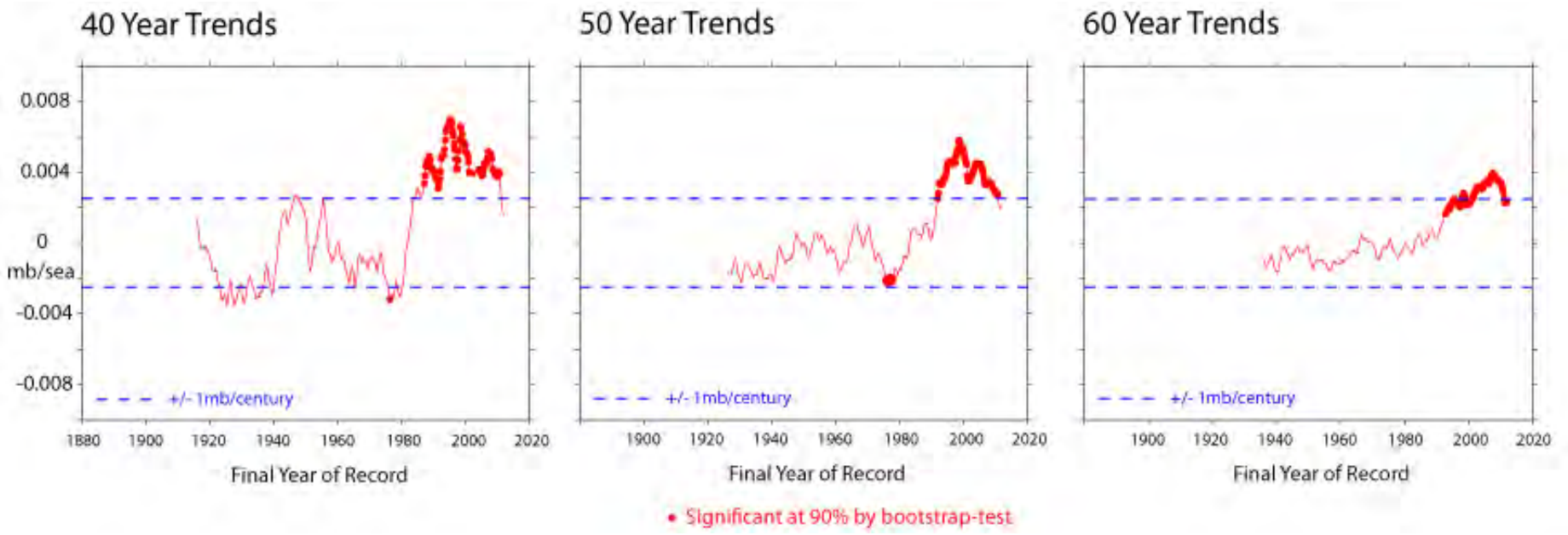
This inconsistency remains even when the amounts of non-trend variability present are estimated from the full record.

Sliding-window trends in Darwin:
10 to 30 years



- Similar numbers of positive and negative trends are seen
- Many of each reach statistical significance

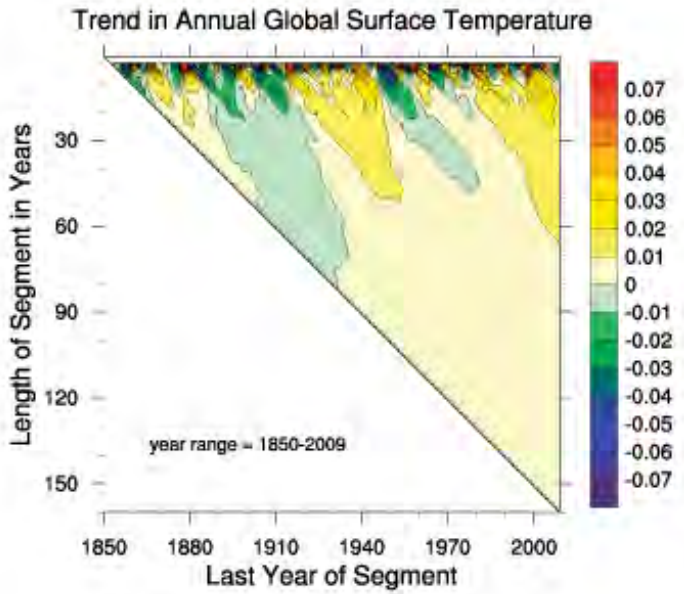
Sliding-window trends in Darwin:
40 to 60 years



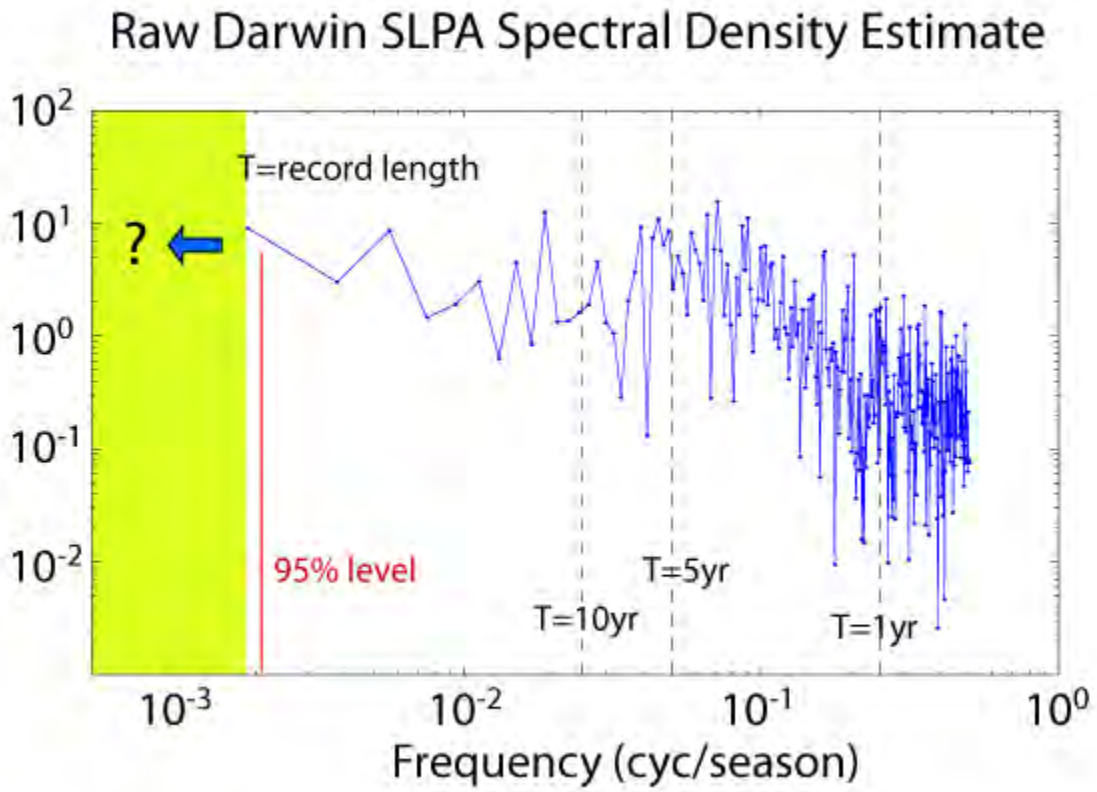
- Still similar numbers of positive and negative trends (up to 13% more negative)
- Positive trends ending in late 1990s/early 2000s are the largest/most statistically significant

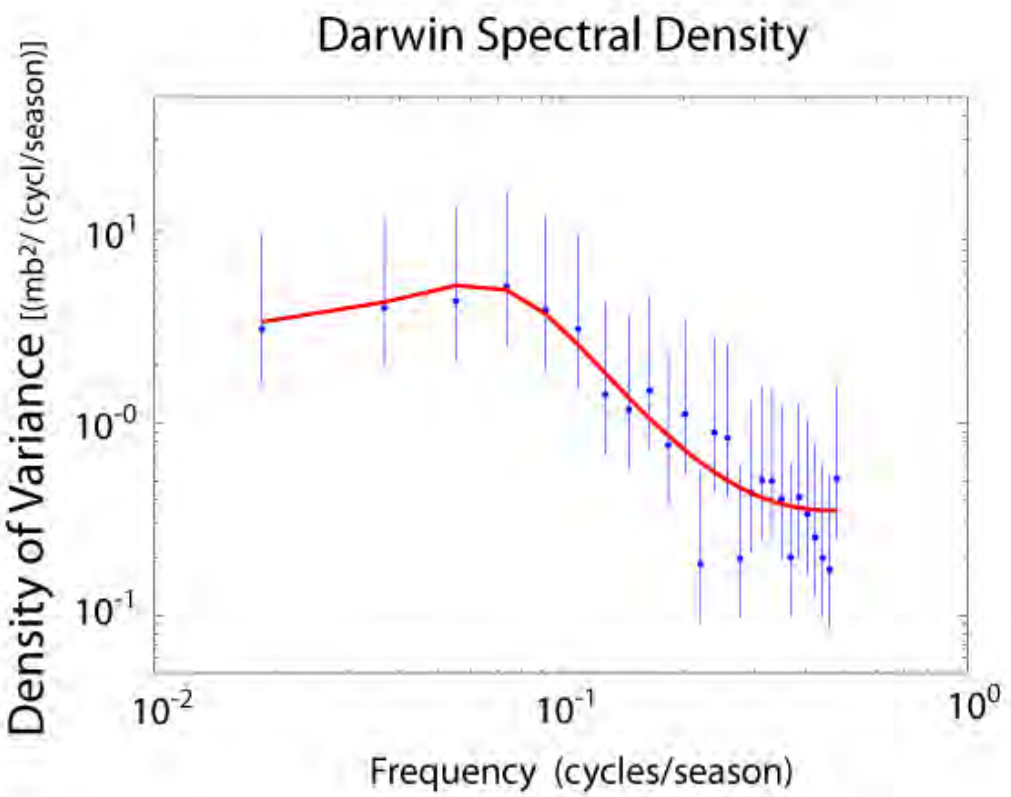
Multi-decadal and century-length trends

- Need for understanding relationship between relatively short and long term trends has been recognized
- Visual tool suggested by Leibmann et al. (2010)
- Interest in our work is in seeing how this plays out for a different type of variable – trend significance methods from Liebmann et al. do not apply to Darwin SLPA.

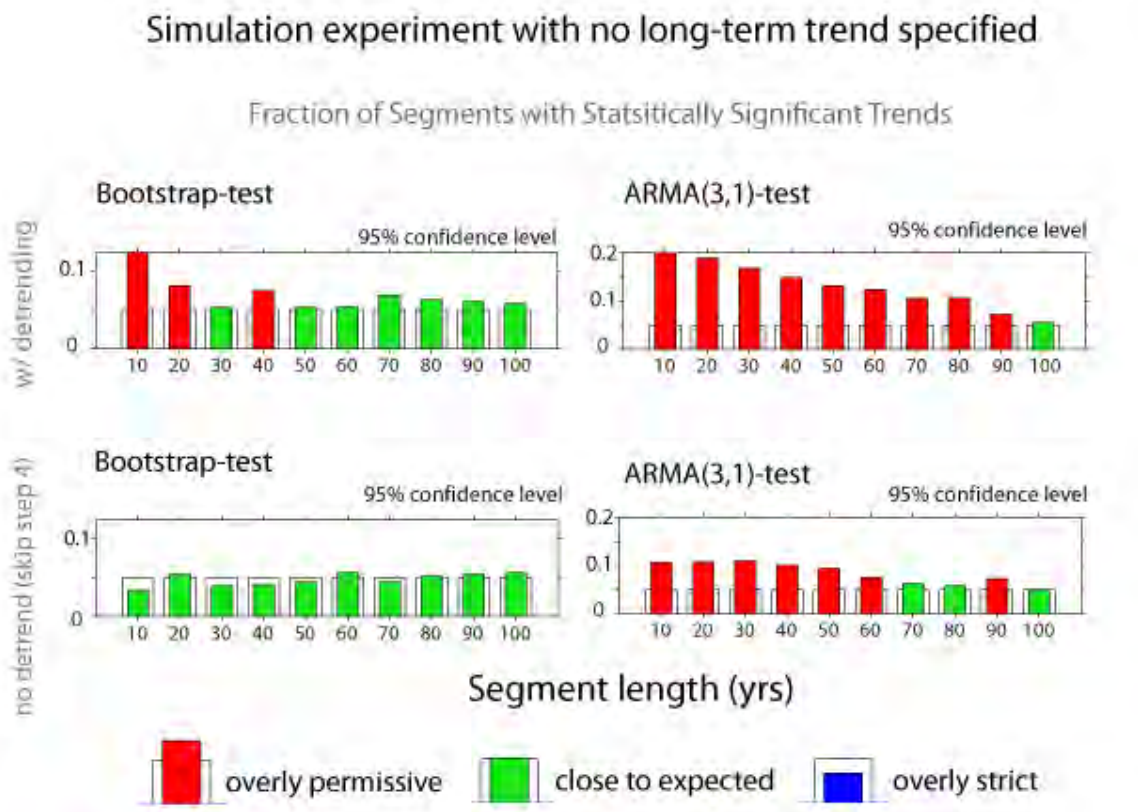


From: Liebmann, et al., BAMS, 2010

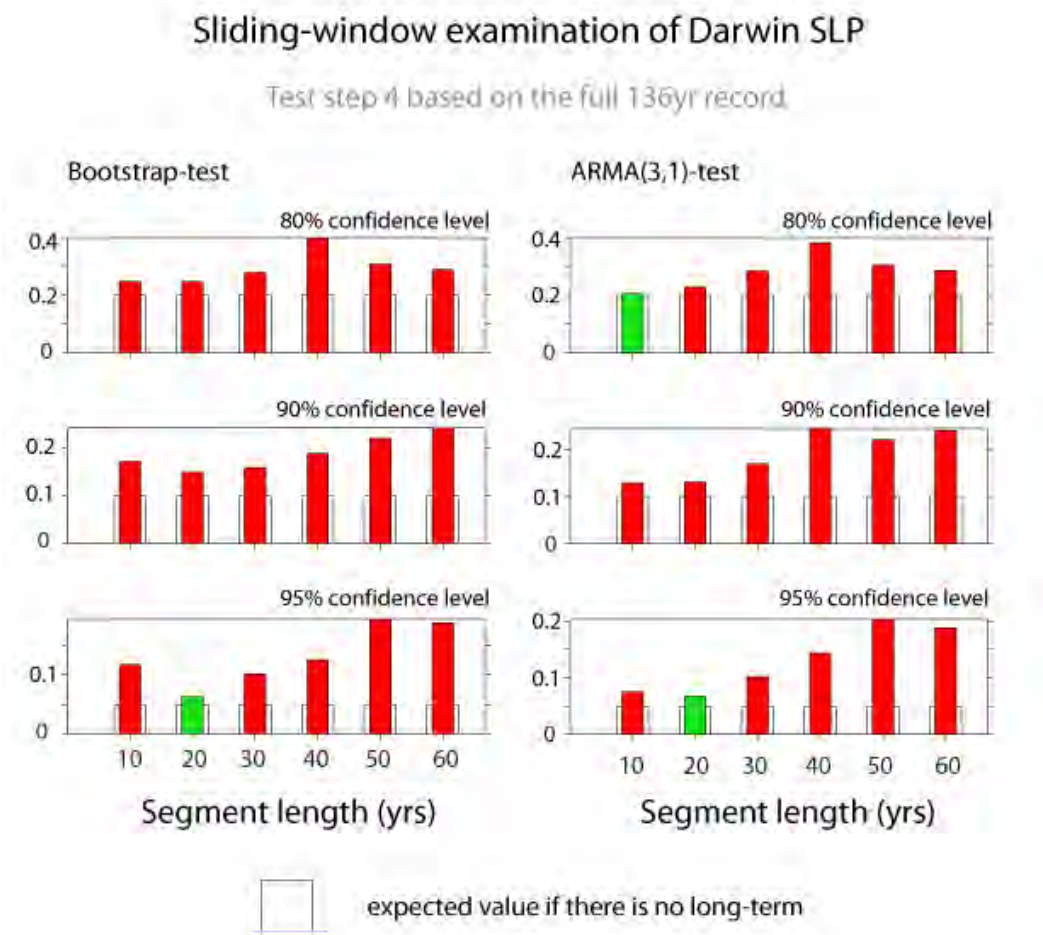

Raw Darwin SLPA Spectral Density Estimate
10^2
10^1
10^0
10^{-1}
10^{-2}
T=record length
?
95% level
T=10yr
T=5yr
T=1yr
10^{-3}
10^{-2}
10^{-1}
10^0
Frequency (cyc/season)


Darwin Spectral Density
Density of Variance [(mb²/ (cycl/season)]
10¹
10⁻⁰
10⁻¹
10⁻²
10⁻¹
10⁰
Frequency (cycles/season)

Simulation experiment results

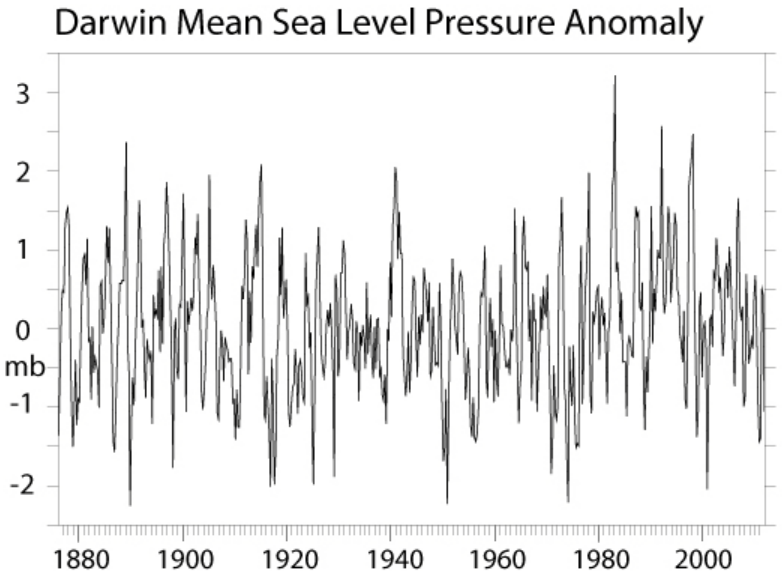


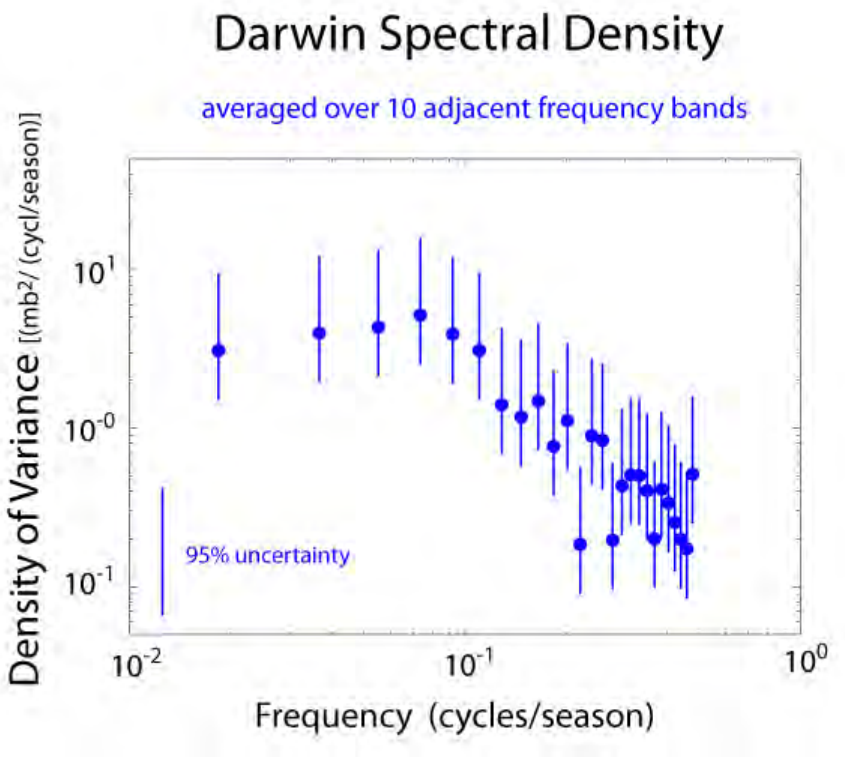
Statistically significant multi-decadal trends in Darwin



This inconsistency remains even when the amounts of non-trend variability present are estimated from the full record (Note that there is some reduction in the fraction of shorter segments that yield statistically significant trends, but not much change at the longer time scales considered).

Darwin SLP Anomaly





- Like many geophysical records, Darwin's spectrum is "red", or contains more variance at lower than higher frequency.
- The reliability of tests for trend statistical significance depends on their ability to accurately estimate the amounts of stationary (no-trend) low frequency variability present