Clim. Past Discuss.,
doi:10.5194/cp-2016-78-SC3, 2016

Climate of the Past Discussions

Interactive comment on "The 1816 'year without a summer' in an atmospheric reanalysis" *by* Philip Brohan et al.

P. Brohan

philip@brohan.org

Received and published: 6 September 2016

The air temperature anomaly correlations for the 23 stations are 0.49 0.50 0.61 0.53 0.46 0.50 0.62 0.66 0.62 0.71 0.77 0.76 0.70 0.60 0.49 0.53 0.63 0.73 0.49 0.74 0.57 0.75 0.72 (without volcanic forcing) and 0.56 0.50 0.67 0.54 0.53 0.57 0.64 0.72 0.67 0.75 0.80 0.79 0.71 0.62 0.61 0.55 0.66 0.78 0.56 0.78 0.58 0.79 0.76 (with forcing)

If we take a null hypothesis that these are independent identically distributed (IID) samples with the same mean, then we could do a t-test. The t.test function in R gives:

t = -1.3992, df = 43.903, p-value = 0.1688 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.10186122 0.01838296 sample estimates: mean of x mean of y 0.6165217 0.6582609

A p-value of 0.16 would not allow us to conclude with confidence that the mean had changed, which is (I think) the point you are making here.

However, the data are NOT IID. Most obviously they are not independent - they are paired samples, so we should use the paired-sample t-test. Repeating t.test with paired=TRUE gives:

t = -7.2588, df = 22, p-value = 2.851e-07 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.05366419 -0.02981407 sample estimates: mean of the differences -0.04173913

This is very much a significant result, reflecting the low probability that all stations would increase in correlation (as seen in the data) if the data were samples from the same distribution.

This is better, but it is still not correct, as the correlations are not samples from the same distribution (not ID): the correlation at a station depends on the precision of the reanalysis, the nature of the weather, and the quality of the observations, all of which differ between stations.

It's hard to say in this case what an appropriate null hypothesis would be - what sort of differences should we expect if the forcing really made no difference. The fact that the correlation increases for all the stations is a strong signal of improvement, but we can't do a formal significance test.

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2016-78, 2016.