An analysis of the spatial mean and dispersion of surface temperatures over the last 1200 years
Martin Tingley (mingley@fas.harvard.edu) and Peter Huybers, Harvard University

Abstract
The time evolutions of the spatial mean and variance of the Northern Hemisphere temperature distribution are described, and confidence intervals estimated. Considerable effort has been applied to the problem of estimating the mean, and the accompanying debate over the relative warmth of present and past warmings. Here we explore the time evolution of the variance about the mean. Results from the instrumental data indicate that periods of increased mean surface temperature can be associated with increased spatial variance. There is a weak indication that this is also true for the proxy data.

1. Introduction
It is of fundamental importance to understand the evolution of the Northern Hemisphere (NH) mean temperature, and a great deal of work has been done on this problem. This work complements that body of research by quantifying the spatial dispersion about the mean value, information important for several reasons.

• The greater the spatial variance, the greater the uncertainty in the mean temperature estimate.

• The variance signal associated with past warm events can potentially be used to characterize them, and determine if, for example, the Medieval Warm Period was more or less of a regional phenomenon than the current warming.

• The temporal evolution of the spatial variance is interesting in and of itself.

Crawley and Lowery (2000) argue that there are significant disparities in the timings of warm periods in the 1000-2000AD interval at different locations in the Northern Hemisphere. This scenario should leave a large dispersion signature, which is found by the analysis presented here primarily in the 1050-1150 interval.

The method employed here is complimentary to the threshold based approach utilized by Osborn and Briffa (2006). Whereas their approach is primarily sensitive to changes in the mean, the method used here distinguishes changes in the mean (first moment) from the variance of temperature anomalies around the mean (second moment).

One issue facing this analysis is the trade off between temporal resolution and uncertainty: by considering a large window of years, the confidence interval widths for the mean and dispersion decrease, at the expense of smoothing out some of the structures identified in Crawley and Lowery (2000).

2. Data
Instrumental: The locations of the 250 Climate Research Unit (CRU) Northern Hemisphere temperature records with at least 8 monthly observations a year over the 100 year period 1901-2005 are shown in Figure 1. Each record was standardized by removing the mean over the full 100 year interval, standardizing over a shorter interval results in biased estimates of the second moment.

Proxy: The 14 proxy records from Osborn and Briffa (2006) and two from Moberg et al (2005) used in the study are shown in Figure 2, and the locations in Figure 3. The proxy records are smoothed to reduce variations with timescales shorter than 20 years. This is likely not the optimal method of preprocessing the data, and the impact of this smoothing on the results must be further investigated. In the future it will be of interest to work with unsmoothed proxy datasets.

3. Methods
The mean and the Interquartile Range (IQR), are used to measure the first and second moments, respectively. The IQR is defined as the distance between the 25 and 75 percentiles of a distribution, and is a robust measure of spread. To increase the sample size and reduce uncertainty, observations from 10 (CRU) or 100 (Proxy) year intervals are pooled together. This pooling biases the IQR estimate high in intervals that contain large changes in the mean value. To correct for this bias the IQR is calculated for shorter time intervals in each record.

Confidence intervals (CIs) for the mean and IQR-Adj are calculated using the Bias Corrected (BC) bootstrap method of Efron and Tibshirani (1986) with 2000 bootstrap samples. The bootstrap methodology assumes independent samples, whereas the temperature data is assumed to be correlated in space and time. To correct for this correlation, a number $N_x < N$ of the date points are resampled each year (see next section).

4. Sample Size
Figures 3 and 4 show the dependency of the 90% BC-Bootstrapped assumed value of $N_x$ for annual (CRU and Proxy), decadal (CRU) and centennial (Proxy) binning intervals. Estimation of $N_x$ requires a bias increase in sample size, while temporal autocorrelation complicates the estimation. The interpretation of the IQR (dispersion in time and space).

5. Results
Main results are shown in the Figures 5-8. Panels (a, b) of Figures 5&7 show the mean and IQR-Adj for the CRU and proxy data. Light grey: statistic value at each year; black: value for a 10 (CRU) or 100 (Proxy) year sliding window.

For the CRU data (Figure 5), the IQR-Adj seems to track the mean: there is a local maxima in the IQR-Adj nearly concurrent with the local maximum in the mean at ~1940, and both have risen over the last few decades. The correlation is 0.77 at the decadal time scale. There is a hint of a similar pattern in the proxy data (Figure 7): local maxima in the IQR-Adj are nearly concurrent with local maxima in the mean at ~1150 and ~1450 (the correlation at the centennial time scale is only 0.26). Crawley and Lowery (2000) argue that different regions of the NH warmed at different times to produce the Medieval Warm Period (MWP) feature. This is not inconsistent with the finding here that the MWP maxima in mean and IQR are not concurrent.

Defining the MWP as the years 956-1151 when the pooled mean estimate is greater than the peak at 1403, the null hypothesis that the mean IQR value is the same within the MWP as it is outside the MWP is rejected using the Mann-Whitney test (Zur, 1999) with $p<0.001$. The test is conducted using annual estimates of the IQR. The null is similarly rejected defining the MWP as the years 1000-1200 ($p=0.001$). The test is less conclusive ($p=0.075$) when the IQR is defined as the years 913-1206, the interval when the pooled estimate of the mean is greater than zero.

Panel (c) of Figures 5&7 is a measure of the spatial homogeneity of the temperature change between intervals: the mean (over locations) of the absolute change in the mean change (at locations) between time intervals. For example, during consecutive time intervals, first Europe is warm and North America cool, then the opposite, the mean and IQR-Adj could be unchanged while the mean deviation is large. The spatial dispersion of the temperature change for the MWP (Figure 5c) generally decreases from 1945: temperature changes at different locations become less variable. In contrast, the mean dispersion in the proxy data (Figure 7c) seems to slowly increase from ~1300.

The maps show temperature patterns for three different intervals. Figure 6: CRU data for (a) 1910 (cool), (b) 1950 (plateau), and (c) 2000 (warm). Figure 8: Proxy data for (a) 1050 (warm), (b) 1650 (cool), and (c) 1950 (warm). Regional scale features dominate the CRU temperature patterns: Eastern North America at 1950, Europe and Japan at 2000; it is difficult to identify similar features in the sparser proxy data.

6. Conclusions and Future Work
A mean and variance analysis of NH surface temperatures is presented. Assuming the CIs are reasonable, the variance (measured by the IQR-Adj) undergoes a significant time evolution in both the CRU data and the proxy records. In the CRU data it appears that the spatial variance tends to rise near or at local maxima in the mean. A similar pattern may be present in the proxy records. In this regard, then, the null hypothesis that the modern warming has the same variance characteristics (as measured by the IQR-Adj) as previous warm events cannot be rejected. It should be noted, however, that the moderate extent of the recent increase in variance, along with the gradual decrease in the mean deviation, do suggest that this modern warming is more uniform for its magnitude than the mid 20th century warm interval.

There is a need to repeat the analysis with more proxy records, which would permit a reduction in the binning length required for significant results. Increasing the number of proxy records would allow for a more meaningful investigation of regional behavior, and a comparison to the spatial patterns seen in the CRU warm periods.

7. Comparison with the Osborn and Briffa (2006) method
Figures 9 and 10 compare the mean and IQR method used here to the approach of Osborn and Briffa (2006). The panels show: (a) a set of surrogate time series with mean and median; (b) the IQR for each time point; and (c) the equivalent IQR as predicted by the Osborn and Briffa (2006) method.

In Figure 9, the mean of the records changes through time, while the variance about the mean remains constant. In Figure 10, the variance of the records increases as the mean deviates from zero. The analysis presented here, based on the mean and IQR, distinguishes these two cases as being qualitatively different.

The Osborn and Briffa (2006) method is mainly sensitive to changes in the first moment: as the mean increases, the proportion of records above some threshold grows.