Daniel Otto
23/03/10

Generating synthetic climate data for glacier modeling

Introduction

In my research I am working on simulating the probability density of glacier volume in 1850 for a set of well-observed glaciers in North America and Europe. By simulating each glacier's fluctuations from 850-1850 for an ensemble of global climate model simulations, I will be able to estimate an uncertainty for the glacier's equilibrium state in 1850 and the envelope of natural variability. The length of the industrial era is short relative to glacier timescales. Glaciers are losing mass, but detecting whether it is attributable to natural climate variability or anthropogenic forcing requires understanding glacier variability over time. Glaciers have a response time on the order of decades, resulting in memory of 1-2 centuries for medium-sized mountain glaciers and much longer for large ones. Thus, knowing whether a glacier was growing or shrinking in the decades leading up to 1850 influences its response to industrial-era warming. By controlling for the effect of inertia at the beginning of the industrial-era, I can assess whether present-day observed glacier fluctuations exceed the range of natural variability.

To model a glacier's variability on the landscape, I first need to calibrate the glacier model to have a length approximately equal to the observed pre-industrial length and a mass balance variability that matches the observed record. This ensures the glacier will be approximately the correct size at the end of the simulation and has the correct sensitivity to climate variability. Thus, under different realizations of past climate, any deviation from the known pre-industrial extent will be a result of the glacier's lagged response. To perform the calibration, I need to force the glacier with a synthetic climate.

This problem presented in this project became apparent after simulating glaciers with recently available data from the Last Millennium Reanalysis (LMR; Tardif, et al., 2019), which derives climate anomalies using paleoclimate proxies. While I had calibrated the glacier models with white noise, forcing them with the LMR did not produce expected results, hinting that it was statistically distinct from white noise. Where previous work has used white noise for glacier calibration, climate variability is generally assumed to be red noise, which is reflected in the LMR. Roe and Baker (2016) show that glaciers are very sensitive to persistence in climate variability, where even small amounts of persistence generates a much larger envelope of natural variability.

To correctly calibrate my glacier models, I assess the spectral characteristics in surface temperature and precipitation for global climate models (GCMs) and the LMR and replicate them in a synthetic dataset. I use CMIP5 simulations of the past 1,000 years from multiple modeling groups. I compare these datasets for two glaciers: Argentiere glacier and Wolverine glacier. Shown in Figure 1, Argentiere glacier is located in the French Alps and Wolverine glacier is in the Chugach range of Southcentral Alaska.

Figure 1. Argentiere glacier (left); Wolverine glacier (right).

Spectral analysis

Last Millennium Reanalysis (LMR)

The LMR is primarily red noise, which can be seen from the distribution of its power spectral density (hereafter power spectrum) in Figure 2. This is an asset to the LMR, as we expect climate variability to be primarily red noise.[1] While precipitation is not as red as temperature in the LMR, it is certainly not white noise. Both become white noise for short periods (high frequencies) under approximately 10 years.

The colored lines in Figure 2 show the change in fitted red-noise shape if an increasing number of high frequencies are left out, corresponding to the part of the curve which most resembles white noise. These show that the low frequencies are redder than the high frequencies, as the curve becomes steeper with more frequencies removed. It also demonstrates the resiliency of the red-noise fit. Where there is little spread, the autocovariance has a similar shape (relative to a power-law distribution) at both low frequencies and high frequencies.

The other notable contrast is between Wolverine and Argentiere. Wolverine is in a sub-arctic maritime climate, compared to Argentiere's relatively continental climate of the Alps. Additionally, climate variability in Europe and Alaska are driven by different (though not unrelated) climate dynamics. Precipitation spectra are similar, with Argentiere's high frequencies being slightly redder on account of the smaller spread between fitted-noise shapes. In temperature, Wolverine exhibits markedly different characteristics by both very closely fitting the red-noise shape and having little spread between fitted lines, indicating equal redness at all frequencies. The difference in redness spread between temperature and precipitation for Wolverine is large, potentially reflecting the character of its maritime climate.

[1] It is also plausible to imagine that the paleoclimate proxies the LMR is derived from could contribute some color, but I have not investigated this.

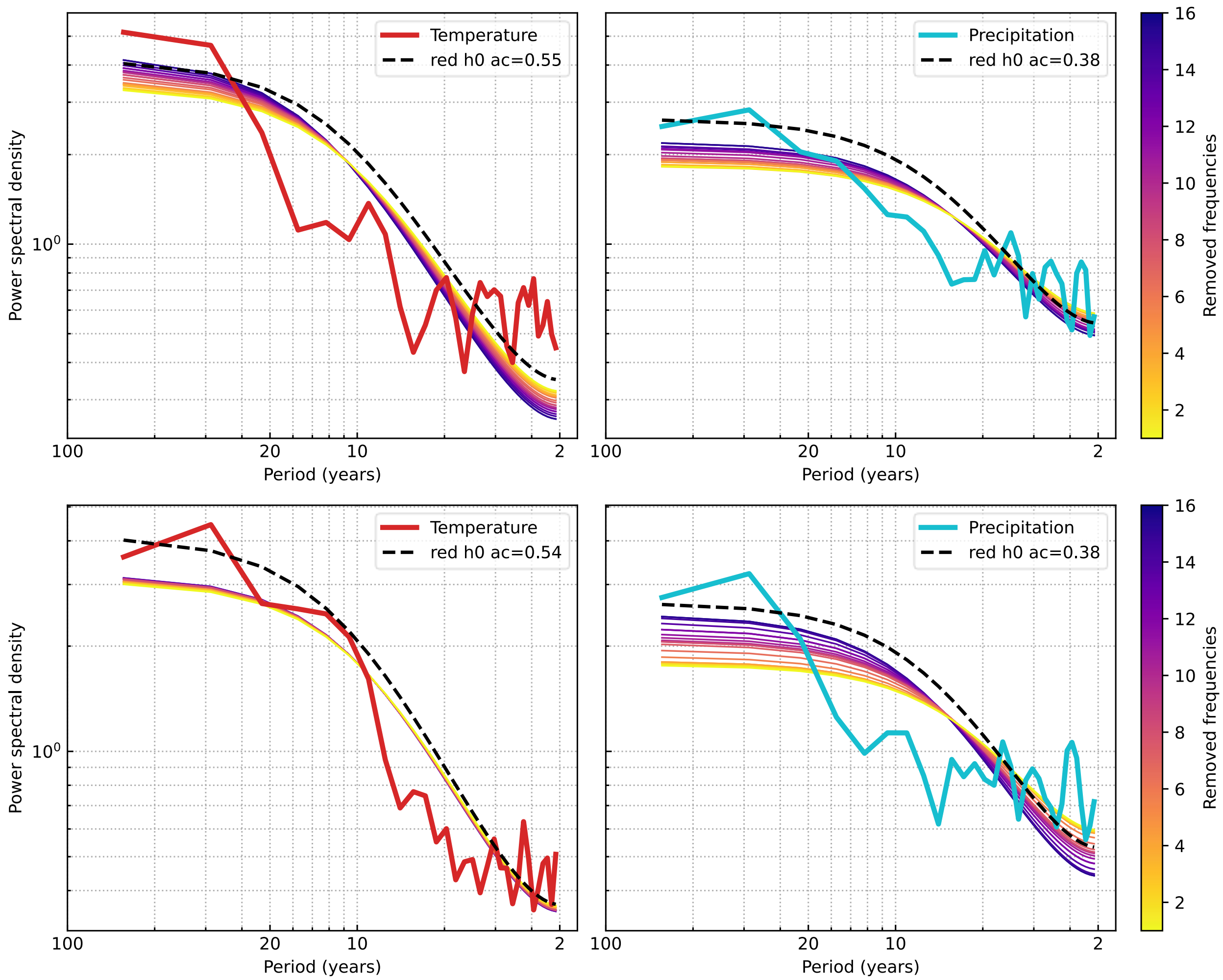

Figure 2. Argentiere glacier (top); Wolverine glacier (bottom). Chunk length = 64. The black dashed line represents the null hypothesis red-noise shape for the given autocorrelation. The colored lines represent fitted red-noise shapes after removing the nth-highest frequencies.

Global Circulation Models (GCMs)

The spectra for GCMs are much whiter than the LMR and are largely similar in distribution of noise between models. Shown in Figure 3 are individual model power spectra, the multi-model grand mean, and standard deviation of model spectra. Temperature is predictably redder than precipitation, though both are closer to white noise than expected. There is a notable spread between models regarding power in particular frequencies, but the overall shape is similar. Past 1,000 simulations were used specifically for their inclusion of volcanoes, and variable model response to aerosol forcing may explain some of the larger deviations. There is much less difference between the two glaciers in GCMs than in the LMR. The spectra from GCMs generally seem less realistic if our baseline expectation is red noise.

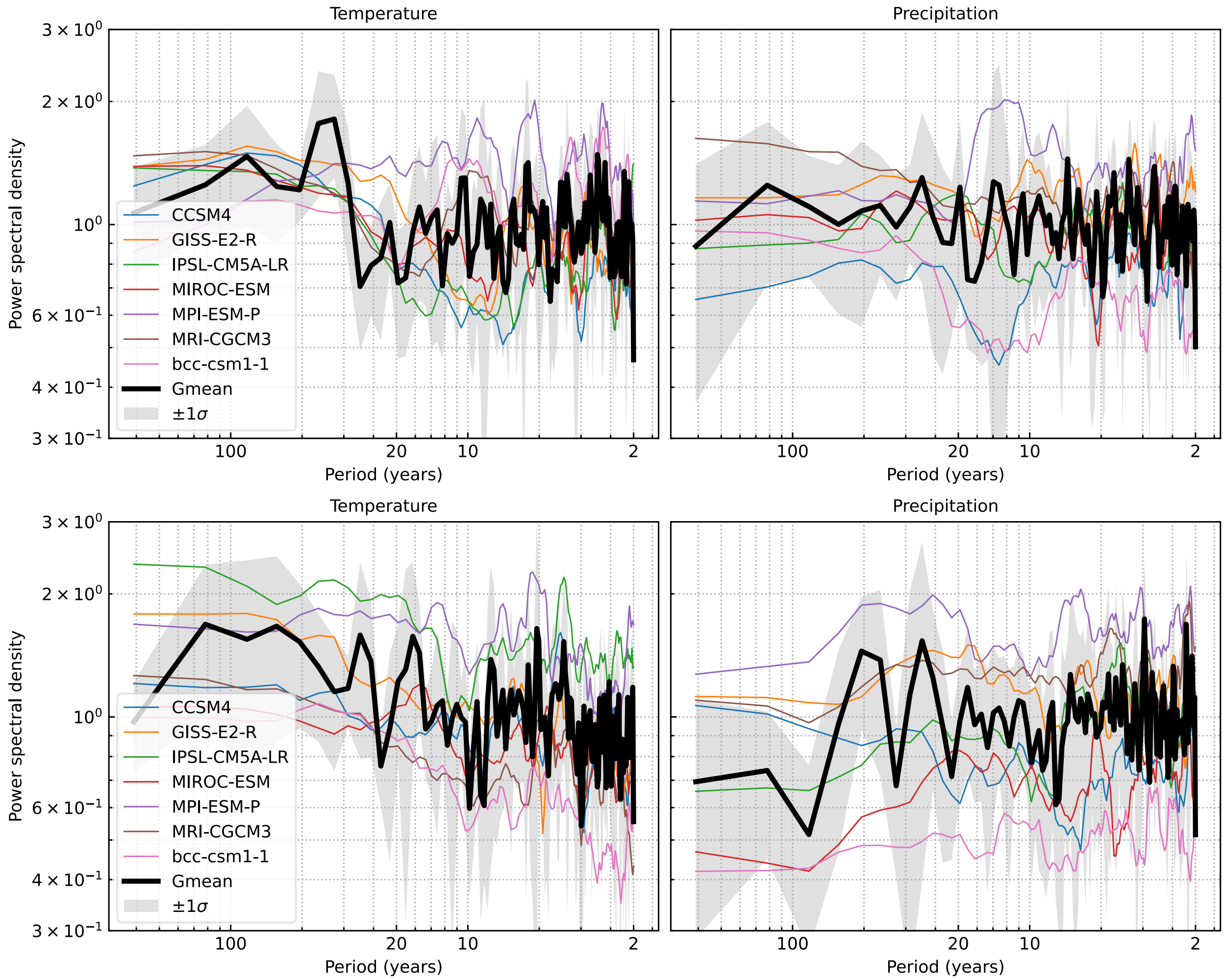

Figure 3. Argentiere glacier (top); Wolverine glacier (bottom). Colored lines are individual models. The black line is the multi-model grand mean. GCM power spectra are smoothed for clarity. The mean spectrum (black) is unsmoothed. Chunk length = 256. A larger chunk length was used to help determine whether low-frequency variability was present.

Emulating power spectra

Method

The rest of the analysis was done using the LMR because it has more notable structure to its power spectra than GCMs, as well as being the original motivation for the work. For simplicity, the following figures show only Argentiere glacier. Following Owens (1978), the procedure to generate a timeseries of noise for a given power spectrum is to invert a Fourier series that was "randomly sampled" from the power spectrum. However, power spectra lack phase information, and cannot be inverted to reproduce their original time series. Because I seek a synthetic representation of the original, phase is randomly added to the power spectrum, sampling a uniform distribution (0, 2π) repeatedly until the desired length is reached. The inverse Fourier transform is then taken, resulting in a time series that has been "reconstructed" from multiple realizations of the original power spectrum. I will refer to this as the reconstructed power spectrum.[2]

[2] This is probably a misnomer, but I already put it on the figure labels.

Effect of filtering

Relative to centennial climate variability, the 1,000 years of the LMR is relatively short, and can be thought of as a sample of the climate system's true variability. For a sufficiently long time series of past climate, the reconstructed power spectrum will converge towards the true long-term climate variability and true power spectrum. In this regard, the length of the LMR is probably not sufficient to produce a robust estimate of the true distribution of spectral density. Assuming a uniform sampling bias at all frequencies, I filtered the power spectrum using a Savitzky-Golay filter to estimate the true power spectrum. While a Savitzky-Golay filter seems a better choice than other filters for approximating the mean distribution, it still likely mangled the shape of the low-frequency spectrum due to the small number of samples.

Figure 4 shows the power spectrum of the resulting reconstructed time series. Two issues became apparent. First, as mentioned, is that the LMR itself is a sample of climate variability, and estimating a representative power spectrum at low frequencies is impossible for the given number of samples. By increasing the chunk length, more low-frequency variability can be resolved, and a more representative fit can be obtained. The Savitzky-Golay filter is, in a way, combining spectral estimates to produce a more robust result, and benefits from its input signal having higher resolution. This is evident from comparison in the time domain of reconstructions with different chunk lengths, which is discussed in the next section. Second, a 1,000-year time series with random phase will not always produce the same century-scale variability as the original. The dashed lines in Figure 4 show that the power spectrum of the reconstructed time series converges on the original power spectrum as length increases. However, any 1000-year segment of a longer series is still subject to the same issue of sampling variability.

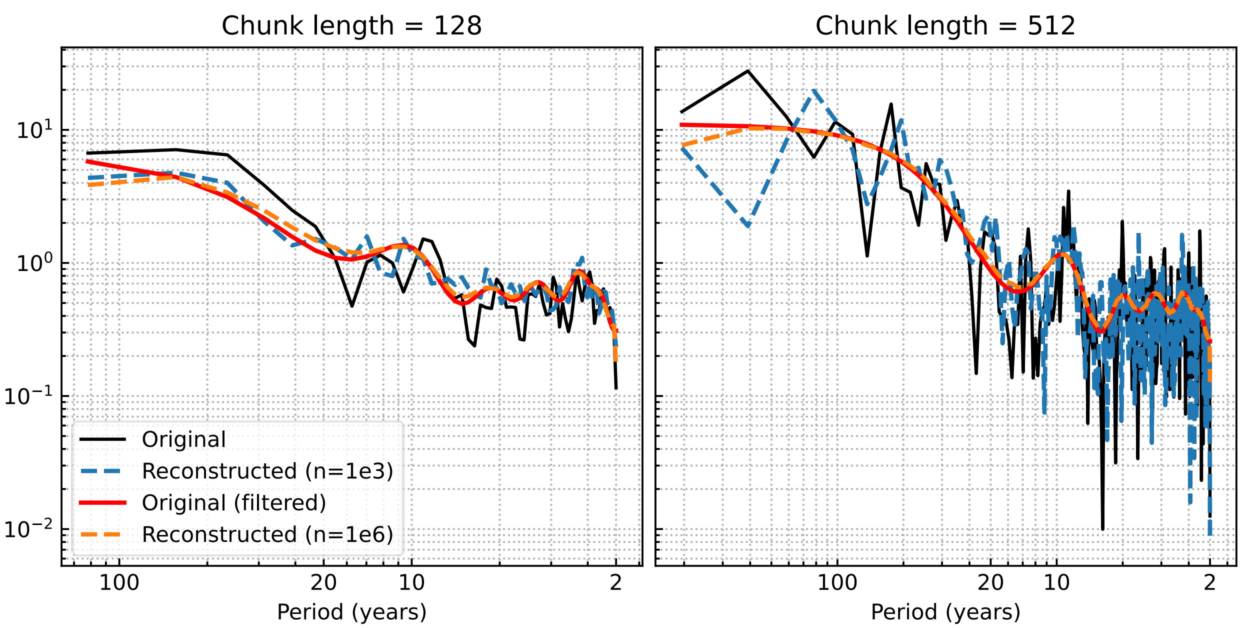

Figure 4. Demonstration of the process to replicate a power spectrum shown for two chunk lengths. The original power spectrum is shown (black) with the filtered power spectrum (red). The additional resolved frequencies for the larger chunk length results in a shape more resembling red noise. The dashed lines are the power spectra of two reconstructed time series. As the number of points increases, the reconstructed power spectrum approaches the filtered spectrum. The Savitzky-Golay filter used had an order = 4 with a window = ¼ * number of frequencies. The filtering was applied 4 times to approximate the shape more generally. The total power was adjusted to match the original spectrum.

Welch's method vs. periodogram

If we accept that the quality (exactness) of the reconstruction increases with power spectrum resolution and the number of resolved low frequencies, the logical conclusion to this optimization is to replace Welch's method with a periodogram, where the resolution is only limited by the sampling resolution and sample length. Using a periodogram results in a near-perfect match of the input power spectrum when the length of the generated time series equals or exceeds the length of the original. The tradeoff is amplification of noise in the reconstructed time series and the risk of a biased original sampling distribution being reproduced in the synthetic data.

Shown in Figure 5 are a comparison of the methods in the time domain. While differences are not immediately apparent, they become visible with further inspection (and a magnified view). Note the number and size of high-magnitude excursions for each method relative to the original data. 5a shows chunk size = 64, which most clearly misrepresents the original variability. No extreme peaks in the reconstruction reach the magnitude of the original data, and they occur less frequently. Additionally, there is some symmetry in the shape of adjacent peaks. An example of this can be found at 250 years, where there is a sequence of approximately 100-year "double peaks" that are more frequent in the reconstruction than the original data. The same can be found for troughs. 5b,c show chunk size = 512, with and without filtering. Both have more realistic amplitude than 5a and more detail at high frequencies. There may be more low-frequency variability in the unfiltered version, peaks appear larger, and there is more persistence at decadal scales. Overall, the differences from filtering are small (itself an interesting result). Finally, 5d is the periodogram-generated time series. Again, differences are small, though the location of peaks does not correspond as well with the preceding plots. Numerically, this is the best match to the original data. Subjectively however, I struggle to call it better than 5b,c, and I harbor some suspicion about the frequency of 50-year peaks relative to the original data.

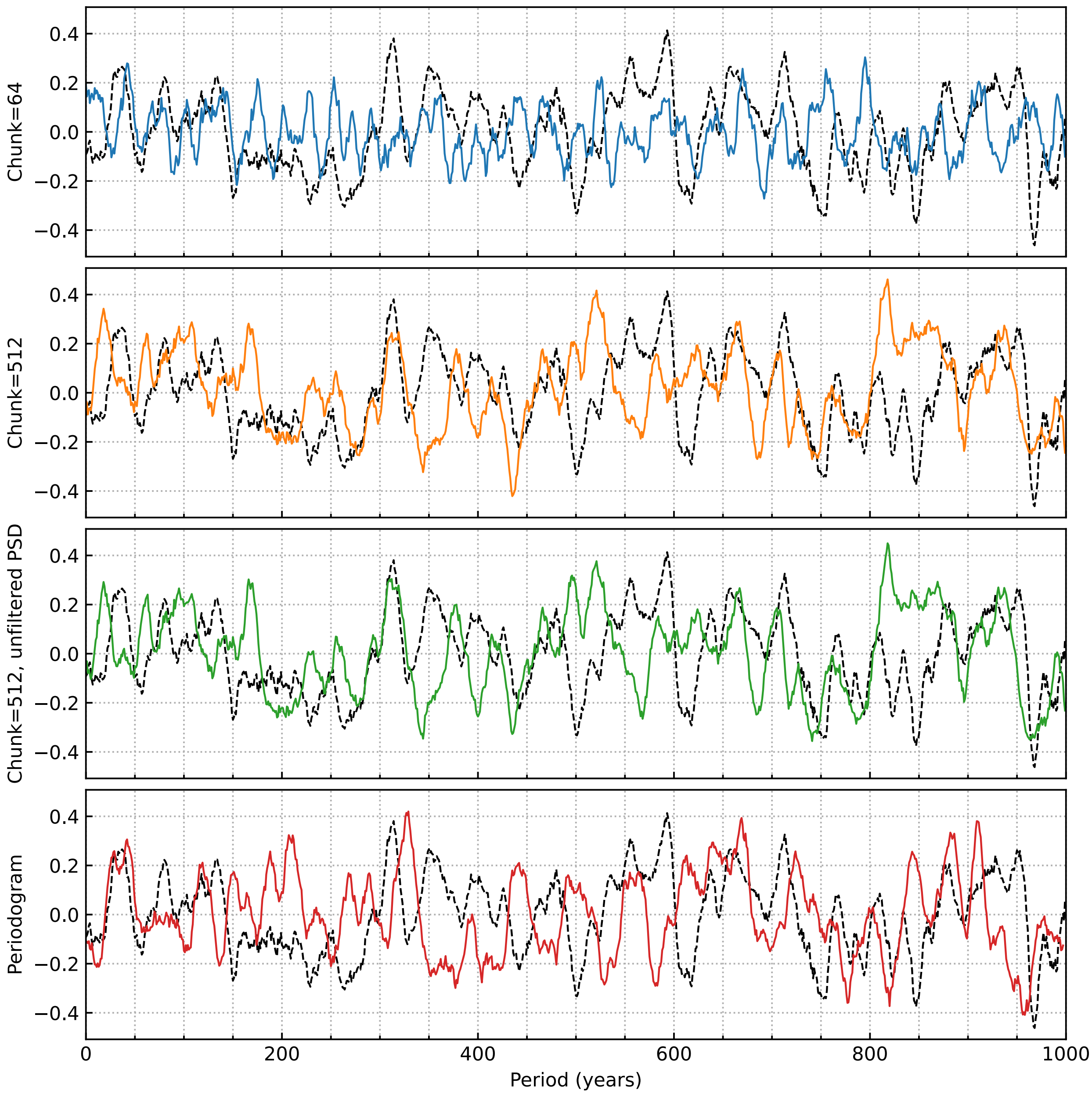

Figure 5. Colored lines are synthetic noise generated with various methods. The dashed black line is the original temperature data from the LMR for Argentiere glacier. The methods shown are: a) filtered Welch's with chunk size = 64 (blue), b) filtered Welch's with chunk size = 512 (orange), c) unfiltered Welch's with chunk size = 512 (green), d) and a periodogram-based reconstruction (red). Note the number and size of high-magnitude excursions for each method relative to the original data. The only difference between b) and c) is whether the original power spectrum was filtered. The same random seed was used for each reconstruction, so all reconstructions have the same phase at each time. 10-year smoothing for clarity.

However, in the case of a short sample period relative to the climate variability we wish to reproduce (such as with the LMR), aliasing low-frequency variability in the original data (that may or may not be truly representative) with Welch's method might be less desirable for the final application than mistakenly portraying a coincidental 1-in-10,000 year event as a 1-in-1,000 year event using a periodogram. For modeling industrial-era glacier change and glacier response to transient 19th-century cooling, the null hypothesis is that glacier change does not exceed the envelope of natural variability. Therefore, assuming natural climate variability (and thus glacier variability) is well represented by the fluctuations of the past 1,000 years is the conservative choice resulting in the most generous envelope of natural variability. The alternative situation, where there is a 1-in-1,000 year event in the LMR that is not reproduced in the synthetic reconstruction, could generate an artificially-small estimate of glacier variability. Even moderately sized mountain glaciers such as Argentiere and Wolverine have a response time of approximately 30 years, resulting in memory on the 1-2 century scale, increasing the importance correctly representing low-frequency variability. While there may be technical aspects I have overlooked, both the filtered Welch’s method with a large chunk size and the periodogram, depending on the application, seem to be viable methods for generating synthetic climate variability.

Future work

I plan to go forward with this method in my research. The only challenge which remains is the covariance between temperature and precipitation in climate variability. Especially considering surface energy budgets, any covariance temperature and precipitation will affect the magnitude of surface energy variability. Remaining work is to investigate the co-spectral characteristics between temperature and precipitation for the LMR and GCMs, particularly phase. Using the same random seed, I think it will be possible to use the co-spectral phase as a PDF to randomly lag the phase of precipitation Fourier series relative to temperature before they are inverted. The effect on glacier variability to the uncoupled and hypothetical coupled synthetic time series remains to be seen.