1 Introduction

The long term history of Earth's climate has a large potential to inform our knowledge about the frequency of ice ages, which is an area of strong interest to geologists and paleoclimatologists. Ice age frequency is closely related to a set of phenomena referred to as *Milankovitch cycles*, which encompass variations in Earth's orbit that have periods much longer than one year. The detailed dynamics governing each Milankovitch cycle is complex and beyond the scope of this project, but we can highlight the three major cycles of interest that may influence temperatures. In this section, we use NASA's explanation of the Milankovitch cycles [1], as well as a 1980 paper by John Imbrie and John Z. Imbrie [2] (which used mathematical modelling to find more precise values for the periods of each cycle) as sources for each period. Below, we summarize the three main cycles:

- Eccentricity Variations: since the Earth-Sun system is subject to additional gravitational perturbations by other planets (namely Jupiter), the Earth's orbit eccentricity is not constant, and varies in a cycle with period of approximately 100,000 years. It is suggested by Imbrie and Imbrie that the eccentricity cycle has its strongest components with periods of 95.0 ka and 125.0 ka, but with many other components in a spectral spread between 95.0 – 136 ka.

- Obliquity Variations: the angle made between Earth's axis of rotation and its orbital axis is also not constant, and varies in periodically. Note that obliquity plays a strong role in determining the relative intensity of the seasons experienced in each hemisphere. Imbrie and Imbrie again suggest that the obliquity variation is the simplest of the three, and has period of 41.0 ka with relatively low spectral spread.

- Axial Precession: Combined tidal forces from the Sun and the Moon have the effect of gradually rotating the orientation of Earth's axis of rotation. Imbrie and Imbrie suggest that the dominant components of the precession signal have periods of 23.0 ka and 19.0 ka.

Unfortunately, prior to the beginning of the global temperature record (in 1850), there is essentially no *direct* temperature information that can be used to reconstruct Earth's historical climate, and thus, finding direct Earth temperature data over a time range large enough to analyze these cycles is impossible. However, paleoclimatologists frequently use *climate proxies*, one of which are foraminifera. Benthic (bottom-of-ocean) foraminifera are small marine organisms that exist at the bottom of the ocean, and whose shells are often fossilized. Each shell contains a certain ratio between ^{18}O and ^{16}O, and due to ^{18}O-enriched water's slightly different evaporation rate, this ratio is heavily influenced by the temperature of the surrounding water at the time of its death (and subsequent fossilization) [3].

This ratio is measured using a standard variable $\delta\ ^{18}$O, given by the following definition:

$$\delta^{18}\text{O} = \left(\frac{R_{sample}}{R_{standard}} - 1\right) \times 1000 \tag{1}$$

where R_{sample} is the ratio between [^{18}O] and [^{16}O] in the sample, and $R_{standard}$ is the corresponding ratio of a constant standard sample, independent of the measurement [4]. It was also shown by Epstein et al. in [4] that the temperature (T, in $^{\text{o}}$C) of the foraminifera is empirically related to δ^{18}O through a quadratic best fit:

$$T = 16.5 - 4.3\left(\delta^{18}\text{O}\right) + 0.41\left(\delta^{18}\text{O}\right)^2 \tag{2}$$

Finally, in this project, we use a deep-ocean foraminifera time-series called the LR04 Stack, from Lisiecki et al. [5]. The stack was assimilated from 57 different δ ^{18}O time series from globally distributed locations into one time-series spanning 5.3 million years. Since the time-series is from deep-ocean foraminifera, vertical mixing is relatively small and thus temperatures are fairly constant on short time scales – this lends the dataset well for use in analyzing long term temperature behaviour.

Given this temperature proxy, the questions that I attempt to address are as follows:

1. Can we extract significant $125, 95.0, 41.0, 23.0$, and 19.0 ka signals from from the temperature time series (corresponding to frequencies of $0.008, 0.0105, 0.0244, 0.0435$, and 0.0526 ka^{-1})? How does the sampling rate effect how well we can extract these signals?

2. If we can extract these trends, what period does the temperature data suggest that each has, and is the discrepancy large compared to the findings of Imbrie and Imbrie [2]?

2 Analysis and Results

2.1 Data Analysis Procedure

In this section, we present the data analysis procedure used to analyze the LR04 stack. Our primary goal is to create a spectral decomposition of the temperature signal. Note the LR04 stack is not sampled on equally-spaced increments; the sampling increment varies from 1 ka to 5 ka, thus, we must sub-sample (in step 2) by an appropriate rate. We choose to use the sampling increment of $\Delta t = 5ka$ (or $f_s = 0.2ka^{-1}$), in order to see how sampling frequency affects the DFT amplitude spectrum. The procedure is outlined below.

1. Determine the temperature signal, T_n, from the $\delta^{18}\text{O}_n$ signal, from equation 2.

2. Subsample the temperature signal T_n, such that it is sampled on an equal time increment $\Delta t_1 = 5$ ka (corresponding to $f_s = 0.2$ ka^{-1}). Note that this is justified by the Nyquist sampling theorem, since we are only interested in resolving frequencies with $f < 0.5 f_s = 0.1$ ka^{-1}.

3. De-trend the temperature signal (removing the linear trend), hence reducing the amplitude of the zero-frequency component in frequency space.

4. Apply a bandpass filter to isolate the signals with periods in the range of interest. The high frequency cutoff of the bandpass filter should be set to the Nyquist frequency in order to prevent aliasing, hence rejecting components of the signal with frequency greater than $0.5 f_s$.

5. Compute the Fourier transform of the temperature signal, $\mathcal{F}[T_n]$.

6. Smooth the temperature signal in frequency space, by convolution with a suitable window function (Hann window). This is done in order to get a more accurate value for the peak of the DFT amplitude, as the signal is fairly noisy in the frequency domain. Compute the frequencies corresponding to the peak values of the smoothed amplitude spectrum.

2.2 Results and Discussion

Below, we show the unfiltered temperature time series T_n computed directly from the data. We note that the temperature scale is roughly what is expected in the deep ocean (between 0ºC and 4ºC) [6]. Note that we see that the temperature value occasionally dips below 0 ºC – this is likely a combined effect from the empirical relationship (given by equation 2) being only approximate, and also from the assumption that the δ^{18}O signal is solely due to temperature variation. In reality, the δ^{18}O is also contributed to partially by effects such as water salinity [5]. Regardless, our main interest is in the periodic trends in the data, rather than specific temperature values.

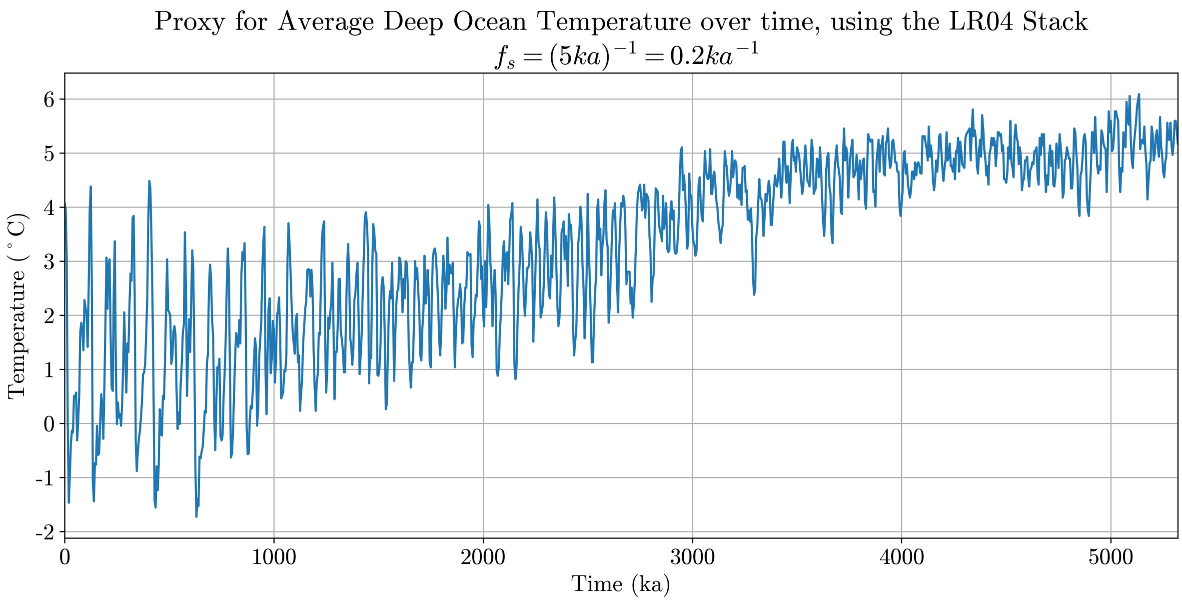
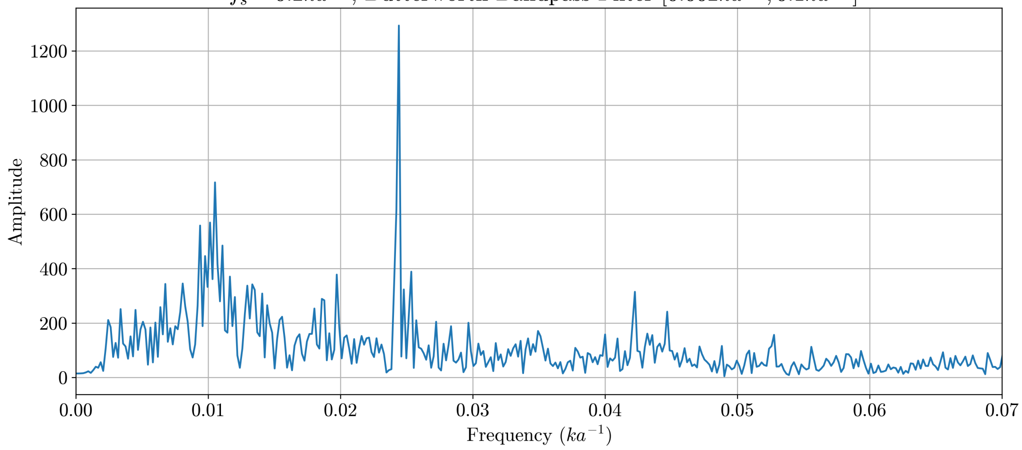


Figure 1: Temperature time series at a sampling increment of $\Delta t_1 = 5$ ka

After de-trending, bandpass filtering, and computing the fourier transform of the data, we show below both the un-smoothed and smoothed amplitude of the DFT, in the frequency domain (we only show the positive frequencies, since the amplitude spectrum is symmetric). Figure 2 shows the smoothed and un-smoothed frequency-domain signal, after the application of a 5^{th} order Butterworth bandpass filter over the frequency interval of $[0.002\ ka^{-1}, 0.1ka^{-1}]$. We can see that there is a fairly strong signal around the frequencies of $0.01ka^{-1}$ (corresponding to a period of around 100 ka), and around $0.025ka^{-1}$ (corresponding to a period of around 40 ka). The smoothed plot in figure 2 also shows a smaller frequency peak at around $0.0435ka^{-1}$, which is one we would expect to see from the axial precession signal. Beyond that, the amplitude doesn't show notable peaks in either the raw or smoothed spectra.

Computing the exact local maxima of the smoothed signal resulted in period values for each of the resolvable cycles, which are displayed in table 1. Note that by far, the 96.8 ka and 40.6 ka had the largest amplitudes, and the higher frequency signals had a relatively small deviation from the background of the spectrum. Also note that the lowest frequency eccentricity signal (corresponding to a 125 ka period), as well as the highest frequency precession signal (corresponding to a 19.0 ka period), were unable to be resolved. There were also other frequency peaks in the amplitude spectra that didn't correspond precisely to one of the three Milankovitch cycles (for example, there is a the peak at around $0.019ka^{-1}$).

Since this doesn't correspond to one of the predicted cycle periods identified by Imbrie and Imbrie, this peak is not identified in this report. However, identification of this peak (and others) could possibly be an extension of this analysis.



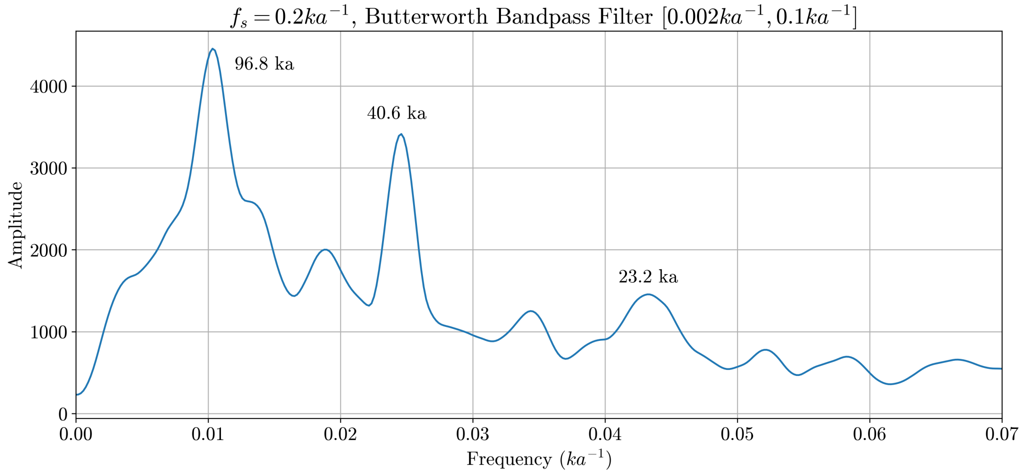


Figure 2: DFT Amplitude Results with $f_s = 0.2\ ka^{-1}$. The four identified cycle periods in 1 are annotated directly on the bottom subfigure.

	Cycle Periods (ka)				
Theoretical [2]	125.0	95.0	41.0	23.0	19.0
LR04 Stack	N/A	96.8	40.6	23.2	N/A

Table 1: Values obtained for the Milankovitch cycle periods from the LR04 Stack with $f_s = (5\ ka)^{-1}$

3 Conclusion and Extensions

In summary, we were able to resolve signals with periods 96.8 ka, 40.6 ka, and 23.2 ka using the LR04 Stack. These are very closely comparable to the theoretical values found by Imbrie and Imbrie in [2] of 95.0 ka, 41.0 ka, and 23.0 ka, corresponding to eccentricity, obliquity, and precession cycles, respectively. The three observed cycle periods have relative error of $1.9\%, 1.0\%$, and 0.9%, respectively. The 40.6 ka signal was found to have much lower spread in the frequency domain relative to the 96.8 ka and 23.2 ka signals (see the top plot of figure 2). This agrees with the predictions of Imbrie and Imbrie. Thus, using the benthic δ^{18}O data, we have successfully resolved three Milankovitch cycle frequencies, one from each cycle (eccentricity, obliquity, and precession). Other signal components, such as 125 ka and 19.0 ka signals, were too weak to be properly resolved with the given signal to noise ratio of the LR04 Stack.

Notable caveats of this analysis includes the underlying assumption that the δ ^{18}O signal is solely a function of temperature – in reality, water salinity and density also influences the signal [5]. Additionally, due to the varied sampling rate of the data, there is a significant fraction of the data that is left unused. Finally, we note that a rigorous analysis of the uncertainty in each of the observed cycle periods was left as beyond the scope of this report. Nonetheless, an extension of this analysis could be to perform a thorough statistical analysis, to find some measure of uncertainty on the observed cycle periods.

Other extension of this analysis could include identification of other peaks in the amplitude spectrum, specifically, the peak located at approximately $0.019ka^{-1}$. As mentioned above, this doesn't correspond directly to one of the frequency components identified by Imbrie and Imbrie [2]; thus, a possible theoretical basis for a cycle of frequency $0.019ka^{-1}$ could be explored. Another posssible extension of this analysis could be using other long-term climate proxy data (e.g. ice core data) to try to resolve a stronger signal for the 125 ka and 19.0 ka period cycles.

References

[1] Milankovitch (orbital) cycles and their role in earth's climate – climate change: Vital signs of the planet. *NASA*, Mar 2020.

[2] John Imbrie and John Z Imbrie. Modeling the climatic response to orbital variations. *Science*, 207:943–953, Feb 1980.

[3] Monica Bruckner. Paleoclimatology. *Paleoclimatology: How Can We Infer Past Climates?*, Mar 2020.

[4] S Epstein, R Buchsbaum, H A Lowenstam, and H C Urey. Revised carbonate-water isotropic temperature scale. *Bulletin of the Geological Society of America*, 64:1315–1326, Nov 1953.

[5] Lorraine E. Lisiecki and Maureen E. Raymo. A pliocene-pleistocene stack of 57 globally distributed benthic 18o records. *Paleoceanography*, 20(1), 2005. Note: Paper and Data available online at `https://lorraine-lisiecki.com/stack.html`.

[6] Jennifer Bergman. Temperature of ocean water. *Temperature of Ocean Water - Windows to the Universe*, Feb 2011.

Appendix

The complete code is included below. Note that the data and its source referenced in [5].

```
import numpy as np
import matplotlib as mpl
import matplotlib.font_manager as font_manager
from matplotlib import pyplot as plt
from scipy import signal
EPS = 10**(-5)
mpl.rcParams['font.family']='serif'
cmfont = font_manager.FontProperties(fname=mpl.get_data_path()\
                                      + '/fonts/ttf/cmr10.ttf')
mpl.rcParams['font.serif']=cmfont.get_name()
mpl.rcParams['font.size']=16
mpl.rcParams['mathtext.fontset']='cm'
mpl.rcParams['axes.unicode_minus']=False

# define plotting subroutine
def makePlot(filename,x,y,xlab,ylab,title,legend=None,xlim=None,\
             ylim=None,figsize=(14,6),semilog=False,annotations=None):
    fig = plt.figure(figsize=figsize)
    if not semilog:
        plt.plot(x,y)
    else:
        plt.semilogy(x,y)
    plt.xlabel(xlab)
    plt.ylabel(ylab)
    plt.title(title)
    plt.grid()
    if legend:
        plt.legend(loc='best')
    if xlim:
        plt.xlim(xlim)
    if ylim:
        plt.ylim(ylim)
    if annotations:
        for note in annotations:
            plt.annotate(note[0],xy=note[1])
    plt.show()
    fig.savefig(filename+".png", dpi=300)
    plt.close()

# define bandpass filtering subroutine
def butterworthFilter(x, lowf, highf, fs, order=5):
    low = lowf/(0.5*fs)
```

```
    high = highf/(0.5*fs)
    b,a = signal.butter(order, [low, high], btype='band')
    return signal.lfilter(b, a, x)

# load data
temp = np.loadtxt("benthic_data.txt",skiprows=1)
times = temp[:,0]
d18O = temp[:,1]
sterr = temp[:,2]

# subsample at period of 5 ka
times_sub = []
d18O_sub = []
for i,time in enumerate(times):
    # note that due to the variation in sampling frequency, between
    # indices of 600 and 1051, the signal was sampled at increment of 2 ka,
    # so we must average adjacent signal values in this range.
    if (time % 5.000 == 0.0): # all other sampling increments divide 5 ka evenly.
        times_sub.append(time)
        d18O_sub.append(d18O[i])
    elif ((int(time) % 10 == 4) and (i > 600) and (i < 1051)):
        times_sub.append(0.5*(time+times[i+1]))
        d18O_sub.append(0.5*(d18O[i] + d18O[i+1]))
dt = times_sub[1]-times_sub[0]
times_sub = np.asarray(times_sub)
d18O_sub = np.asarray(d18O_sub)

# compute temperature
temp = 16.5 - 4.3*d18O_sub + 0.14*(d18O_sub**2)
# plot raw data
makePlot("temp_5ka_samp",times_sub,temp,"Time (ka)","Temperature ( C )",\
        "Proxy for Average Deep Ocean Temperature over time, using the LR04"+\
        " Stack\n $f_s = (5 ka)^{-1} = 0.2 ka^{-1}$",\
        xlim=(times_sub[0],times_sub[-1]))

# detrend data
p = np.polyfit(times_sub, temp, 1)
trend = np.polyval(p, times_sub)
temp_detrended = temp - trend
# filter data
lowcut = 1./500.
highcut = 1./10.005
fs = 1/dt
temp_filt = butterworthFilter(temp_detrended, lowcut, highcut, fs)
# plot filtered and detrended temperature
makePlot("temp_filtered_5ka_samp",times_sub,temp_filt,"Time (ka)","Detrended "+\
        "Temperature ($\Delta$ C)",\
```

```
         "Detrended and Filtered Temperature over time, using the LR04 Stack"+\
         "\n $f_s = 0.2 ka^{-1}$, Butterworth Bandpass Filter"+\
         " $[0.002 ka^{-1}, 0.1 ka^{-1}]$",\
         xlim=(times_sub[0],times_sub[-1]))

# compute DFT of filtered and detrended temperature
temp_filt_f_sh = np.fft.fftshift(np.fft.fft(temp_filt*dt))
freq_ax = np.fft.fftshift(np.fft.fftfreq(len(temp_filt_f_sh), dt))
# plot DFT amplitude of filtered and detrended temperature
makePlot("temp_fourier_5ka_samp_amp",freq_ax,np.abs(temp_filt_f_sh),\
         "Frequency ($ka^{-1}$)","Amplitude",\
         "DFT Amplitude of Filtered Temperature"+\
         "\n $f_s = 0.2 ka^{-1}$, Butterworth Bandpass Filter "+\
         "$[0.002 ka^{-1}, 0.1 ka^{-1}]$",
        xlim=(0,0.07))

# define convolution window
N = 25
hann = np.hanning(N)
# plot convolved DFT amplitude spectra
makePlot("temp_fourier_5ka_samp_amp_conv",freq_ax,\
        np.convolve(hann,np.abs(temp_filt_f_sh),'same'),\
         "Frequency ($ka^{-1}$)","Amplitude",\
         "DFT Amplitude of Smoothed, Filtered Temperature"+\
         "\n $f_s = 0.2 ka^{-1}$, Butterworth Bandpass Filter "+\
         "$[0.002 ka^{-1}, 0.1 ka^{-1}]$",
        xlim=(0,0.070),annotations=[["96.8 ka",(0.012,4200)],\
                                    ["40.6 ka",(0.022,3600)],\
                                    ["23.2 ka",(0.041,1600)]])

# finding the 41, 100, and 23 year periods
convolved = np.convolve(hann,np.abs(temp_filt_f_sh),'same')
from scipy.signal import argrelextrema
where = argrelextrema(convolved, np.greater) # compute local maxima
print("Printing amplitude values, frequencies, and period values at peaks:")
for wh in where[0]:
    if freq_ax[wh] >= 0.0:
        print("Amplitude: {:3f} \tFrequency: {:5f} \tPeriod: {:5f}".\
          format(convolved[wh],freq_ax[wh],1/freq_ax[wh]))
```