```
% file SIOC 221A HW 9
%
% author Julia Dohner
%
% due date December 13, 2017

% analysis of weekly flask-collected co2 data at Mauna Loa and La↙
Jolla
% stations

clear all; close all;

%% load CO2 data

dataMLO = fopen('monthly_flask_co2_mlo_JLD.txt');
dataLJO = fopen('monthly_flask_co2_ljo_JLD.txt');
dataSPO = fopen('monthly_flask_co2_spo_JLD.txt');

valsMLO = textscan(dataMLO, '%f %f', ...
    'delimiter','\t');
valsLJO = textscan(dataLJO, '%f %f', ...
    'delimiter','\t');
valsSPO = textscan(dataSPO, '%f %f', ...
    'delimiter','\t');

fclose(dataMLO);
fclose(dataLJO);
fclose(dataSPO);

% format of .txt files is year, co2 value
LJOyear = valsLJO{1};
LJOco2 = valsLJO{2};

MLOyear = valsMLO{1};
MLOco2 = valsMLO{2};

SPOyear = valsSPO{1};
SPOco2 = valsSPO{2};

% remove flagged data
for i = 1:length(MLOco2)
    if MLOco2(i) == -99.99
```

```
        MLOco2(i) = nan;
    end
end
for i = 1:length(LJOco2)
    if LJOco2(i) == -99.99
        LJOco2(i) = nan;
    end
end
for i = 1:length(SPOco2)
    if SPOco2(i) == -99.99
        SPOco2(i) = nan;
    end
end

% remove nan's
% TODO: if time, go back and change this to a linear interpolation
addpath↙
('/Users/juliadohner/Documents/MATLAB/SIOC_221A/HW9/Inpaint_nans/Inpa↙
int_nans');
MLOco2 = inpaint_nans(MLOco2);
LJOco2 = inpaint_nans(LJOco2);
SPOco2 = inpaint_nans(SPOco2);


%% inspect spacing of data

% inspect time spacing between measurements
MLO_t_diff = diff(MLOyear);
LJO_t_diff = diff(LJOyear);
SPO_t_diff = diff(SPOyear);

% figure
% plot(1:length(MLOyear)-1,MLO_t_diff);
% figure
% plot(1:length(LJOyear)-1,LJO_t_diff);
% figure
% plot(1:length(SPOyear)-1,SPO_t_diff);

minDiffLJO = min(LJO_t_diff);
maxDiffLJO = max(LJO_t_diff);
minDiffMLO = min(MLO_t_diff);
maxDiffMLO = max(MLO_t_diff);
```

```
minDiffSPO = min(SPO_t_diff);
maxDiffSPO = max(SPO_t_diff);

% too big differences in time increments of daily LJO flask sampling,↙
will try
% monthly data instead

% update: time differences range between 0.0767 and 0.0850 years for↙
both
% MLO and LJO records (difference of 0.0083) or ~3 days, deemed small
% enough difference in time increments
% read: counts in my book as even spacing

%% shorten data to same lengths

startYear = LJOyear(1,1); % 1.969041100000000e+03 (latest start)
endYear = SPOyear(length(SPOyear),1); % earliest end
startIndex_LJO = find(LJOyear == startYear);
startIndex_MLO = find(MLOyear == startYear);
startIndex_SPO = find(SPOyear == startYear);
endIndex_LJO = find(LJOyear == endYear);
endIndex_MLO = find(MLOyear == endYear);
endIndex_SPO = find(SPOyear == endYear);

% create new vectors
LJOco2_2 = LJOco2(startIndex_LJO:endIndex_LJO);
LJOyear_2 = LJOyear(startIndex_LJO:endIndex_LJO);
MLOco2_2 = MLOco2(startIndex_MLO:endIndex_MLO);
MLOyear_2 = MLOyear(startIndex_MLO:endIndex_MLO);
SPOco2_2 = SPOco2(startIndex_SPO:endIndex_SPO);
SPOyear_2 = SPOyear(startIndex_SPO:endIndex_SPO);


% plot timeseries
figure('name','Atmospheric CO2 Timeseries');
plot(LJOyear_2,LJOco2_2, '-r', MLOyear_2,MLOco2_2,'-b',SPOyear_2,↙
SPOco2_2,'-g');
xlabel('\fontsize{14}year')
ylabel('\fontsize{14}ppm')
title('\fontsize{16}Atmospheric CO2 Records')
legend('\fontsize{12}La Jolla Station','\fontsize{12}Mauna Loa',↙
'\fontsize{12}South Pole','Location','northwest');
```

```
%% compute spectra (from in-class coherence example)

N = length(LJOco2_2);
segment_length = N/2; % length of each chunk of data (aka segment↙
length)
M = segment_length/2;
Nseg = N/segment_length; % number of segments splitting data into

% using three segments with 50% overlap (concatenating in lines↙
below)
LJO_use=[reshape(LJOco2_2,segment_length,Nseg) reshape(LJOco2_2(M+1:↙
end-M),segment_length,Nseg-1)];
MLO_use=[reshape(MLOco2_2,segment_length,Nseg) reshape(MLOco2_2(M+1:↙
end-M),segment_length,Nseg-1)];
SPO_use=[reshape(SPOco2_2,segment_length,Nseg) reshape(SPOco2_2(M+1:↙
end-M),segment_length,Nseg-1)];

LJO_ft=fft(detrend(LJO_use).*(hann(segment_length)*ones(1,3)));
MLO_ft=fft(detrend(MLO_use).*(hann(segment_length)*ones(1,3)));
SPO_ft=fft(detrend(SPO_use).*(hann(segment_length)*ones(1,3)));

% question: mean is still in here. Do I want it?
LJO_spec=sum(abs(LJO_ft(1:M+1,:)).^2,2)/N; % sum over all spectra↙
(2nd dim)
LJO_spec(2:end)=LJO_spec(2:end)*2; % multiply by 2 to make up for↙
lost energy

MLO_spec=sum(abs(MLO_ft(1:M+1,:)).^2,2)/N;
MLO_spec(2:end)=MLO_spec(2:end)*2;

SPO_spec=sum(abs(SPO_ft(1:M+1,:)).^2,2)/N;
SPO_spec(2:end)=SPO_spec(2:end)*2;

%% uncertainty estimate on spectra

% question: is this nu calculation right??
nu = 1.9*2*Nseg; % DOF = 2*number of segments*1.9 (1.9 for the↙
Hanning)
err_high = nu/chi2inv(0.05/2,nu);
err_low = nu/chi2inv(1-0.05/2,nu);
```

```
ratio_chi2 = err_high/err_low;

%% plot spectra

% Question: I added the +1 to the M here to make it match the dims of↙
the specs
% (145x1) because means are still in there
frequency=(1:M+1)/(segment_length/12);
frequency = frequency';

figure('name','Power Spectra of CO2 Records');
% TODO: fix the locations of the uncertainty estimates
loglog(frequency,LJO_spec, '-r', [.2 .2],[err_low err_high]*LJO_spec↙
(100), ...
    frequency, MLO_spec, '-b',[.1 .1],[err_low err_high]*MLO_spec↙
(100), ...
    frequency, SPO_spec, '-g',[.3 .3],[err_low err_high]*SPO_spec↙
(100))
xlabel('\fontsize{14}cycles per year')
ylabel('\fontsize{14}ppm^2/cpy')
title('\fontsize{16}Power Spectra of CO2 Records')
legend('\fontsize{12}La Jolla Station','\fontsize{12}Mauna Loa↙
Station', '\fontsize{12}South Pole','Location','northeast');

%% compute coherence

% compute cross covariance of LJO/MLO
ccLM=sum(LJO_ft(1:M+1,:).*conj(MLO_ft(1:M+1,:)),2)/N;
ccLM(2:end)=ccLM(2:end)*2;
% compute coherence
C_LM=abs(ccLM)./sqrt(LJO_spec.*MLO_spec);
phase_LM = atan2(-imag(ccLM),real(ccLM));
deltaPhase_LM = sqrt((1-C_LM.^2)./(abs(C_LM).^2*2*Nseg));

% compute cross covariance of MLO/SPO
ccMS=sum(MLO_ft(1:M+1,:).*conj(SPO_ft(1:M+1,:)),2)/N;
ccMS(2:end)=ccMS(2:end)*2;
% compute coherence
C_MS=abs(ccMS)./sqrt(MLO_spec.*SPO_spec);
phase_MS = atan2(-imag(ccMS),real(ccMS));
deltaPhase_MS = sqrt((1-C_MS.^2)./(abs(C_MS).^2*2*Nseg));
```

```
%% uncertainty estimate

alpha = 0.05; % 95% confidence level
gamma_threshold= sqrt(1-alpha^(1/(Nseg-1)));


%% plot the coherence

figure('name','Coherence Plots of CO2 Records');
subplot(2,1,1)
plot(frequency, C_LM,[frequency(1) frequency(end)],[gamma_threshold↙
gamma_threshold]);
axis tight;
xlabel('\fontsize{14}cycles per year')
ylabel('\fontsize{14}coherence')
title('\fontsize{16}Coherence of CO2 Records (LJO and MLO)')
legend('\fontsize{12}La Jolla vs. Mauna Loa','\fontsize{12}↙
Uncertainty Threshold','Location','southeast');

subplot(2,1,2)
plot(frequency, C_MS,[frequency(1) frequency(end)],[gamma_threshold↙
gamma_threshold]);
axis tight
xlabel('\fontsize{14}cycles per year')
ylabel('\fontsize{14}coherence')
title('\fontsize{16}Coherence of CO2 Records (MLO and SPO)')
legend('\fontsize{12}Mauna Loa vs. South Pole','\fontsize{12}↙
Uncertainty Threshold','Location','southeast');

% Question: is this threshold right? seems really high... but I↙
suppose it's that 5%
% of the data will be above the threshold just due to random chance,↙
so it should be
% high?
% "set a threshold for evaluating whether a calculated coherence↙
exceeds
% what we might expect from random white noise"

%% plot the phase

% Question: I'm having a hard time understanding this plot (read Lec↙
15
```

```
% Notes)
% Question: what are the units on the y axis? (again, read lec 15↙
notes)
figure('name','Phase Plot for Mauna Loa and South Pole Coherence');
semilogx(frequency,[phase_MS phase_MS+deltaPhase_MS phase_MS-↙
deltaPhase_MS]);
xlabel('\fontsize{14}cycles per year')
ylabel('\fontsize{14}phase units')
title('\fontsize{16}Phase Plot for Mauna Loa and South Pole↙
Coherence')
legend('\fontsize{12}phase', 'phase+delta phase','phase-delta↙
phase','Location','northeast');
axis tight
```