PHY 477/577: Obs. Methods & Data Analysis in Phys. Ocn. Derek J. Grimes Assignment 1: Plotting & Evaluating Periodic Fit to Data grimesdj@uncw.edu

Goal: develop familiarity with plotting, and computing/evaluating statistics from coastal data.

Task: Astronomical tides are an excellent example of a complex periodic signal. However, water-levels are also influenced by a myriad other processes. The dominant tidal constituents for most sub-tropical regions are the the principal lunar tidal period is roughly Tl ≈ 12h 25min and principal solar tide with a period Ts ≈ 12h. We will determine what fraction of water-level variability is described by these two constituents using a 1-month tide gauge record (choose your birthday month). First, navigate to [https: //tidesandcurrents.noaa.gov/](https://tidesandcurrents.noaa.gov/) and choose a station somewhere you’d someday like to visit.

1. Download an ASCII water-level file (*e.g.,* .csv) of the 6-minute verified observations. You can either do this manually, or develop code to download data directly from the NOAA API. Here is example code to download water-level data from a NOAA tide gauge, just modify the variables *gaugeID, startDate, endDate* to match your desired query (note: the % symbol tells matlab that text on the remainder of the line is a comment not code):

1. % In order to r e t r e i v e water − l e v e l data using NOAA’ s API we need :
2. % 1) the t i d e gauge ’ s i d e n t i f i e r

# 3 gaugeID = ’ 8658163 ’ ;

1. % 2) the s t a r t / end date for the data − set
2. s t a r t D a t e = ’ 20230101 ’ ; 6 endDate = ’ 20230201 ’ ;

|  |  |  |
| --- | --- | --- |
| 7 | % 3) | the desired v a r i a b l e / i n t e r v a l / u n i t s / formatting , |
| 8 | % | see s t r i n g a f t e r ” product ” below . See also : |
| 9 | % | h t t p s : / / api . t i d e s a n d c u r r e n t s . noaa . gov / api / prod / |
| 10 | % 4) | the URL syntax to queiry data using t h e i r API ( a p p l i c a t i o n |
| 11 | % | programming i n t e r f a c e )−−a method for two computers to communicate , |
| 12 | % | here it ’ s your computer t a l k i n g to the f i l e −system housing the data . |

# 13 url = [ ’ h t t p s : / / api . t i d e s a n d c u r r e n t s . noaa . gov / api / prod / d a t a g e t t e r ? begin date = ’ , startDate , ’&end date= ’ , endDate , ’&s t a t i o n = ’ , gaugeID , ’&product= water level&i n t e r v a l =6&datum=NAVD&time zone=gmt&u n i t s =metric&format=csv ’

] ;

1. %
2. % 5) create a filename to save the data on hard − drive :

# 16 fileName = [ ’ / path / to / data / data / w a t e r l e v e l s ’ , gaugeID , ’ ’ , startDate , ’ ’ , endDate , ’ . csv ’ ] ;

1. websave ( fileName , url ) ;
2. %
3. Read in the data using matlab’s built in routines:
   1. the function *readtable*() generates a data-table class variable.
   2. the function *table2array*() will extract row/column data from the table.
   3. the function *convertTo*() can be used to convert to ‘datenum’ format which is a decimal Julian day. To learn more execute:*help datenum*.

19 % 6) Now you ’ re ready to open \& read the data in MATLAB.

|  |  |
| --- | --- |
| 20 data = | readtable (????) ; |
| 21 time = | t a b l e 2 a r r a y ( data (????) ) ; |
| 22 waterLevel = | t a b l e 2 a r r a y ( data (????) ) ; |
| 23 waterSigma = | t a b l e 2 a r r a y ( data (????) ) ; |

1. What are the relevant dependent and independent variables and what are their units?

The dependent variable is Water Level (measured in feet) which is the calculated average water level for the six-minute interval. The independent variable is time (measured in 6-minute intervals for a day).

1. Is there a quality control flag? Or are there any NaNs? If so, write code to either mask bad data or find all NaN’d indices.

There are predicted water levels and verified water levels. In the data table, if the predicted water level matches exactly with the verified water level it is marked with I, if not it is marked with 0 (none of the predicted values match exactly with the verified water level). There is also a quality column labeled ‘Quality’ with the character ‘v’ used throughout the entire column. No information is given as to what the character ‘v’ means, however I assume it is another way of saying a water level was verified. There are not any NaNs in the downloaded data. There is also the water sigma column, which is the sd of the one-minute interval measurements in relation to the six-minute interval average. This information can be used to potentially identify “bad data” or data that has a large sd.

1. Based on the principal lunar and solar periods, what is the fundamental period (in days) of these constituents? What astronomical phenomena prevent periodicity on this time scale?

“  A minimum of 30 days of data is used in order to observe the majority of the lunar and solar cycles. A year of data is necessary to directly observe all of the 37 tidal harmonic constituents.”

“M2 – The largest Lunar constituent. The Earth rotates on its axis every 24 hours, but the Moon is orbiting in the same direction as the Earth’s rotation. It takes a location on the Earth an additional 50 minutes to “catch up” to the Moon. This results in a tidal signal (M2) which has 2 peaks every 24 hours and 50 minutes.” Thus the fundamental period .510

“S2 – The largest solar constituent – is related to the direct gravitational effect of the Sun on the tides. The Earth rotates on its axis every 24 hours. This results in a tidal signal (S2) which has 2 peaks every 24 hours.” Thus the fundamental period .5

1. What is the sample mean *η*¯, standard deviation *s*, and skewness *γ* of the water level record *η*(*t*)? Discuss.

Water Level Sample Mean (η̄): 4.018675

Water Level Standard Deviation (s): 1.8543221695138978

Water Level Skewness (γ): -0.05443303652523521

Sigma Sample Mean (η̄): 0.02188125

Sigma Standard Deviation (s): 0.010849148899206823

Sigma Skewness (γ): 0.38876873932244593

The water level sample mean was 4.019 ft with a standard deviation of 1.854. Seeing that the Water level skewness is negative this suggests that the left tail of the distribution is likely to be longer and the values are likely to be below the mean.

1. Included with the data is the standard deviation *si* of the raw 1-minute interval measurements used to generate the 6-minute mean record *ηi* (variables *waterSigma* and *waterLevel*, respectively). Use this time series of standard deviation to estimate the uncertainty of each 6-minute mean.
2. Plot the water-level record *η*(*t*) and uncertainty on separate axes. Label your axes, include proper units, and a legend, and make a figure caption to accompany it (see Lecture 2 for examples).

A graph of water level record and uncertainty

Description automatically generated

1. Describe the observations. For example,

-Are there any anomalies in the tidal record? I would say there are no anomalies in the tidal record, the values are for the most part consistent with the general tidal patterns.

-What is the relative magnitude of the uncertainty to the tidal amplitude? – For the most part, the magnitude of the uncertainty correlates with the magnitude of the tidal amplitude. It certainly is not an exact fit, however, seems to match the general pattern.

-Are the uncertainties constant? Or do they vary? Why? The uncertainties are not constant, rather they vary. As demonstrated by the plot, the uncertainties seem to increase as the water level increases (in ft) and decreases as the water level decreases. A possible reason for this could be the type of measuring equipment used or even where the equipment was placed to measure the water levels. Perhaps the equipment used is more accurate at measuring lower water levels than higher levels, or perhaps the equipment was placed in a spot that made it more difficult to accurately measure high water levels.

2