VTide Executable

# Introduction:

VTide is a re-implementation of Mike Foreman’s Versatile Tidal Analysis program. Much of the astronomical and matrix assembly code is borrowed verbatim, although the solver and inputs are changed and we have added a prediction.

A journal reference for this work that describes this program in contrast to the previous state of the art is:

Foreman, M. G. G., J. Y. Cherniawsky, V. A. Ballantyne, 2009: Versatile Harmonic Tidal Analysis: Improvements and Applications. *J. Atmos. Oceanic Technol.*, **26**, 806–817.

This documentation is a manual to the software inputs, with minimal discussion of how to select constituents. It is suggested that you get

# Analysis: vtide\_analyze

## Basic usage:

vtide\_analyze analysis\_main\_filename

## Inputs

### Main input file:

The main input file is a short file listing names of other inputs and output files. The extensions can be anything you like. Here “.inp” is used for the main input and “.dat” for input data and “.out” for output.

>>>>>>>>>>Sample analysis\_main.inp with annotations

sanfrancisco\_analysis\_2010\_07.dat # analysis request file

tide5n.dat # constituent database

run66b\_2010\_07\_obs.dat # input observations file

run66b\_2010\_07\_out.out # analysis output file

run66b\_2010\_07\_constituent.out # coefficients for predication

run66b\_2010\_07\_fitted.out # fitted values

>>>>>>>>>>> End main\_input.txt

### Analysis request file:

The analysis request file contains data specific to the station and constituents to be fitted. The contents are as follows:

1. One line containing 4 integers:
   1. The number of main frequencies
   2. Number of components (dimensions) being analyzed (1 for elevations and 2 for velocities)
   3. Whether to include a trend (0 for no, 1 for yes)
   4. Number of locations to be analyzed
2. One line for each constituent to be included in the analysis along with its frequency in cycles/h. The frequency is for reference only – the analysis uses exact astronomical arguments.
3. One line for start date of analysis in format yyyy-mm-dd
4. One line for end date in format yyyy-mm-dd
5. A line with station info:
   1. station\_id (integer)
   2. station\_name (20 characters or fewer, quotes need to be used if there are spaces)
   3. time\_zone (3 characters, e.g. “PST”)
   4. latitude degrees
   5. latitude minutes
   6. longitude degrees
   7. longitude minutes
6. For each main frequency associated with inferences
   1. One line identifying the main frequency:
      1. The name of the main frequency (e.g. “K1”)
      2. its frequency in cycles/hr
      3. number of frequencies inferred from this main frequency
   2. For each inference one line giving:
      1. The name of the inferred constituent
      2. Frequency in cycles/hr
      3. Amplitude ratio (inferred constituent amplitude divided by analyzed constituent amplitude)
      4. Phase lag of the inferred constituent subtracted from the phase of the analyzed constituent in degrees.

### Constituent database (tide5n.dat)

This file contains astronomical arguments and descriptions of each constituent such as Doodson numbers. This file will seldom change. You will typically want to use the default file tide5N.dat.

### Input observation file:

The format for observation is to list the date, then the data for each station. In the case of velocity, the u and v components are grouped together for each location. Here is an example line for July 23, 2010 at 21:00 and two stations (or one station with u and v):

2010-07-23 21:00 2.0970 1.9568

or

2010-07-23T21:00 2.0970 1.9568

The use of 24 hour time is required and you must use the time “00:00”, not “24:00” as would be the custom with military formats such as HEC-DSS. The separators between the year month and day are arbitrary, as is the following character (the letter “T” is a sort of truncated ISO formatting, but there are still no seconds). The program calculates phase in whatever time zone you give it. So if you give it GMT it will give Greenwich-based phase. If you give it a local standard time, it will use that.

### Analysis results output file

The analysis results file gives is the most comprehensive output file. It includes constituent fits in the order in which the stations were given as well as significance statistics that can be interpreted by reference to the Foreman article.

### Constituent coefficient output file

The constituent coefficient output file is an abbreviated form of output that focuses on the harmonic constants. It is used to load harmonic constants for prediction.

### Fitted values output file

After the fit is performed, it is evaluated immediately at the same time points as given in the input. The purpose of the fitted values file is to give the user or client code access to these values for residual analysis and to test the prediction program for correctness.

# Prediction: vtide\_predict

## Basic usage:

vtide\_predict predict\_main\_filename

## Inputs

The prediction executable uses many of the same files as analysis program.

### Main input file

The main input file contains the names of a few other input/output files from the analysis as well as the dates and time step of the prediction. You can predict any contiguous period you like, although extrapolation of the trend a long distance may produce unusual results.

>>>>>>>>>>Sample prediction\_main.inp with annotations

run66b\_2010\_07\_constituent.out # coefficient file from analysis

tide5n.dat # constituent database

run66b\_2010\_07\_pred.out # output file for fitted values

2009-03-12 # prediction start date

2010-12-31 # end date (hour/minute ignored)

10 # interval in minutes

>>>>>>>>>>> End main\_input.txt

## Choosing constituents

Choosing the right number and selection of constituents is a central issue of harmonic analysis. One thing you will certainly need to know is what is available. Please see the constituent\_menu.txt file for a list of names and frequencies.

Once you have the list of possibilities, the key issues are these:

1. Culling the list of main constituents to ones that are well separated given the data on hand
2. Choosing shallow water frequencies that are well suited to the circumstances
3. Selecting appropriate inferences.

For task 1, a good place to start is the Rayleigh pairings charts for each species (diurnal, semi-diurnal, etc) that was originally distributed with Foreman’s F77 tide program and originates with Godin. The rows of the chart describe increasing dataset lengths (in hours). As the data get longer, arrows describe good additions to make to the constituent list based on frequency separation and importance of the constituent based on the tide potential. The Rayleigh criterion has been criticized as being too conservative in terms of frequency differences, so it may be reasonable to err on the side of more frequencies by including items a few rows lower on the chart and then cull the list based on statistical significance. At the same time, some constituents may not be easy to identify statistically or may be very small.

The analysis produces several significance tests. The easiest to understand are the amplitude standard deviation and t-test values; as usual, a t-test statistic of 2.0 represents significance at the 95% value. There is some question whether the assumptions of the test are met – see Foreman (2009). The other statistics come from Equation (6) of Cherniawsky (2001) Ocean tides from TOPEX/Poseidon sea level.

Shallow water is a more site-specific subject. The Rayleigh pairings chart includes some shallow water frequencies that are often included in analyses. We have also included a chart of frequently used shallow water constituents as well as a chart of shallow water constituents and main constituents that may be easily confused. These may become distinguishable based on node factors if the record is long enough (say 8 years). Otherwise, the composition of nearby stations and likelihood of shallow water constituent development should be considered.

Inference should be used if a significant constituent lies too close to a main constituent to separate them based on the data in hand. A good example is P1, which is pretty big according to the tide potential but not distinguishable from K1 for a record of a month. Inference is carried out by assuming you know something about the relationship of the main (K1) and inferred (P1) constituent, and this is most easily achieved if you have a reference station with a long tidal history (e.g. San Francisco). In this case you will need to assess the amplitude ratio and relative phase of the two constituents by assuming they are the same as at the reference station.