

**Interactive Model Evaluation and Diagnostics System**

**V3.1 User Guide**

**October 27th, 2010**

**Eve-Marie Devaliere, MS**

University of North Carolina – Institute of Marine Sciences

US Army Corps of Engineers – Field Research Facility

**Jeffrey L. Hanson, PhD**

US Army Corps of Engineers - Field Research Facility

# Table of Contents

[Table of Contents 2](#_Toc278203133)

[I. IMEDS Description 5](#_Toc278203134)

[II. IMEDS Requirements 6](#_Toc278203135)

[III. Starting IMEDS 7](#_Toc278203136)

[IV. Preprocess Data Module 8](#_Toc278203137)

[4.1 Observed Data Pre-processors 10](#_Toc278203142)

[CHL NDBC Formatters (spe1D or spe2e and onln) 10](#_Toc278203143)

[NODC formatter 10](#_Toc278203144)

[NOS 10](#_Toc278203145)

[IMEDS Generic Format 11](#_Toc278203146)

[Already Pre-Processed 11](#_Toc278203147)

[4.2 Model Data Pre-processors 12](#_Toc278203148)

[WAVEWATCH III (WW3) Formatter 12](#_Toc278203149)

[SWAN Formatter 12](#_Toc278203150)

[Oceanweather Formatter (Wave) 12](#_Toc278203151)

[Oceanweather Formatter (Wind) 13](#_Toc278203152)

[ADCIRC Water-Level Formatter (fort.61) 13](#_Toc278203153)

[ADCIRC Wind Formatter (fort.72) 13](#_Toc278203154)

[IMEDS Generic Format 14](#_Toc278203155)

[Extract from SURA OpenDAP Server 14](#_Toc278203156)

[Extract from local NetCDF file 15](#_Toc278203157)

[4.3 Auto-Filling of the IMEDS Setup Form 16](#_Toc278203158)

[V.](#_Toc278203159) [Evaluate Data Modules 17](#_Toc278203160)

[5.1 Wind & Wave Run Setup 17](#_Toc278203166)

[5.2 Storm Surge Run Setup 25](#_Toc278203167)

[5.3 High Water Marks Module 27](#_Toc278203168)

[Setup 27](#_Toc278203169)

[Results 28](#_Toc278203170)

[VI. Display Modules 30](#_Toc278203171)

[6.1 Overview of the IMEDS Display Module and Performance Sub-Module 30](#_Toc278203172)

[6.2 Stations Statistics Sub-Module 32](#_Toc278203173)

[6.3 Monthly Statistics Sub-Module 33](#_Toc278203174)

[6.4 Peak Analysis Sub-Module 33](#_Toc278203175)

[6.5 Extremes Analysis Sub-Module 37](#_Toc278203176)

[Appendix 1: IMEDS Analyses Types 40](#_Toc278203177)

[Temporal Correlation (TC) Analysis 40](#_Toc278203178)

[Quantile-Quantile (QQ) Analysis 40](#_Toc278203179)

[Peak Event (PE) Analysis 40](#_Toc278203180)

[Extremes Analysis (EA) 40](#_Toc278203181)

[Performance Scores 41](#_Toc278203182)

[Appendix 2: Wavedat Format Description 42](#_Toc278203183)

[Appendix 3: IMEDS Generic Format Description 44](#_Toc278203184)

[Description 44](#_Toc278203185)

[Example 46](#_Toc278203186)

[Reference 48](#_Toc278203187)

# I. IMEDS Description

The Interactive Model Evaluation and Diagnostics System (IMEDS) is a custom GUI-driven toolbox for MATLABTM to assess coastal process model performance using a variety of temporal and spatial metrics. A significant challenge in evaluating large temporal- or spatial-scale simulations is the need to statistically reduce millions of model estimates to a meaningful measure of prediction skill, yet retain sufficient level of detail to identify model strengths and deficiencies. This challenge is now met with IMEDS.

IMEDS requires both an observation data set as ground truth, and a model or test data set to evaluate. The data are composed of time series output at specific geographic stations (i.e., NDBC buoys). The observation data are decomposed into a series of components, such as wind-sea, young swell and mature swell for wave spectrum data. These are further broken down into specific data attributes, such as the height, period and direction in the case of wave analysis. More information on this particular technique can be found in Hanson et al. (2009). A variety of error metrics (such as root-mean-square error, bias, scatter index) are calculated for the model predictions at each station, by component and attribute. A performance score is then calculated from the errors. The performances are then folded through space and time giving the user an assessment of the total model performance for each attribute. As a diagnostic tool the user can then explore model errors and performance as a function of many variables (station, time, components, etc). Further details about the various IMEDS analysis types are available in Appendix 1.

In addition to evaluating model data, IMEDS can be used to compare 2 observation stations or model runs together. The user simply identifies one as being the ground truth data and the other one the test data.

IMEDS is fully functional with wind, wave, and water-level data.

It is recommended to run IMEDS on a windows machine as the GUI looks much better on Windows. It is however possible to run it on a linux machine as well. There hasn’t been any test on a Mac.

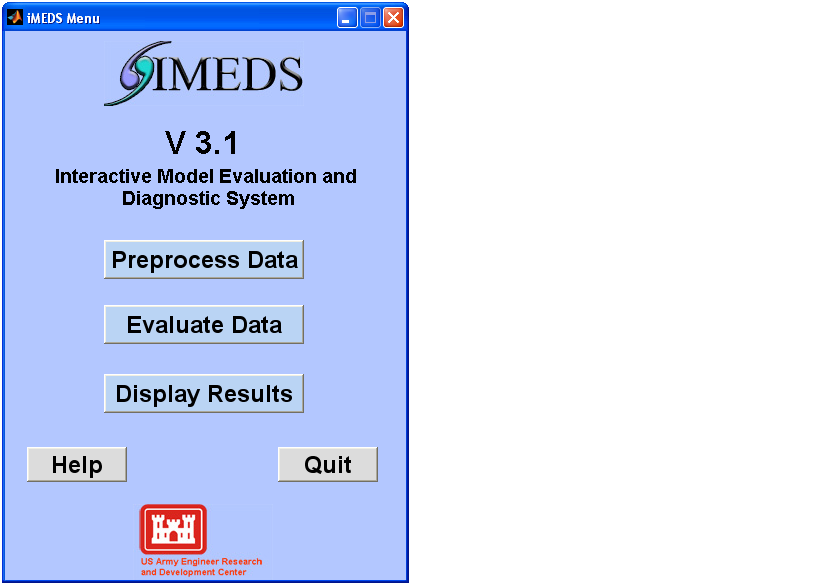
# II. IMEDS Requirements

In order to run IMEDS successfully a few conventions have to be used. Those are listed below:

* Models (or test) and observation data have to be in a ‘wavedat’ format. For a description of this format, please see Appendix 2. Since IMEDS v2.2 you can use a data preprocessing module if your data is not in wavedat format. Please see the IMEDS pre-processor section.
* In most cases, the wavedat file has to hold one month or less of data. If validation is needed over a longer period of time, the data needs to be broken up into monthly files. More details as to when this applies are given in the pre-processors section.
* Filename conventions:
  + The station ID (as the user wants it to appear in the plots) needs to be in the filename for both ground-truth and test data. This is how IMEDS knows which files to pair up. The user has 2 choices to organize the data:
    - Group it in folders holding the station ID with the station ID figuring anywhere in the filenames.
    - Not organize it in folders but have the station ID as the beginning of the filename up to the underscore.
  + If the run spans many months the model and observation files need to specify the month and year in the filename using 3 letter or 2 digits month and 4 digit year as in 41025\_NDBC\_Jun2000.mat. (the order of the code, month and year doesn’t matter as long as they are all included in the filename)
  + As of version 3.1, only one file with the station code can be in the directory that iMEDs is looking in for single month runs.
* If the user is using the preprocessors to format the data, the filename convention will be taken care of. However, other conventions need to be followed for the use of the preprocessors. Please refer to the preprocessors section for further information.
* Reserved Words: Certain words are not supported by MATLABTM because they conflicting with MATLABTM commands. Do not use the followings in any edit field:
  + Default

# III. Starting IMEDS

* From the MATLABTM command window, set the current directory to /*imedsRoot*/IMEDS/src/process where *imedsRoot* is the path to the directory IMEDS had been placed at.
* Type ‘imeds’ to open the imeds Module Selection Window (Figure 1)
* Select between the 3 IMEDS modules
  + Preprocess Data: Format data. If the data is not in wavedat format, a finite set of preprocessors are available for formatting the user's data.
  + Evaluate Data: Initiate the IMEDS run set up process with the ‘Run Setup’ window. Once the IMEDS run is completed it automatically starts the results display.
  + Display Results: View results from previously analyzed data



*Figure 1: IMEDS Module Selection Window*

Each of these modules is described below.

# IV. Preprocess Data Module

The preprocessor module shown in Figure 2 allows one to pre-process both model and/or observed data. If one doesn’t want to process both model and observed data at the same time (e.g. second run of a model, the observed data would have been already processed) one can skip the selection of the directory for the given data. In this case one has to be careful to select which event this run is related to in the event pull-down menu so that the observed data can be retrieved accordingly (details follow).

When browsing to the model or observed data, IMEDS pre-processor automatically points to /*imedsRoot*/IMEDS/data/raw. It is suggested that the user put the raw data in sub-directories of this one, but one could browse anywhere.

Each observed data is associated to an event (e.g. Isabel, year 2007…). The observed data produced by the preprocessor will be in:

*/imedsRoot*/IMEDS/data/obs/event\_name/ (if the event and run name contains spaces they will be replaced by underscore to avoid bugs on Linux systems ; *imedsRoot* is the directory in which IMEDS has been unzipped to)

Each model data set is associated to an event and a run (model run) since a model can be run many times with different settings for the same event. The data produced by the preprocessor will be in (The space rule mentioned above is also valid here for event and run names): */imedsRoot*/IMEDS/data/obs/event\_name/run\_name

Instructions for specific Preprocessor inputs follows.

* **Single Event Checkbox**

If the raw data covers many months, the formatters will automatically split the data into monthly files. However, it may be that the user doesn’t want monthly statistics (e.g.: the studied hurricane event is occurring at the end of a month, trailing into the other month). In this case, the user can force the preprocessors to handle the multi-month event into a single month event, by checking the ‘Single Event’ checkbox.

* **Add to Existing Run Checkbox**

In some cases, one will need to either reprocess data and/or add some data to an already setup run. Alternatively, the user might not want IMEDS to set up the directory structure as previously stated. Either way, the ‘Add to Existing Run’ checkbox allows the user to specify the directory where he wishes the pre-processed data to be stored at.

* **Organized by Station Checkbox**

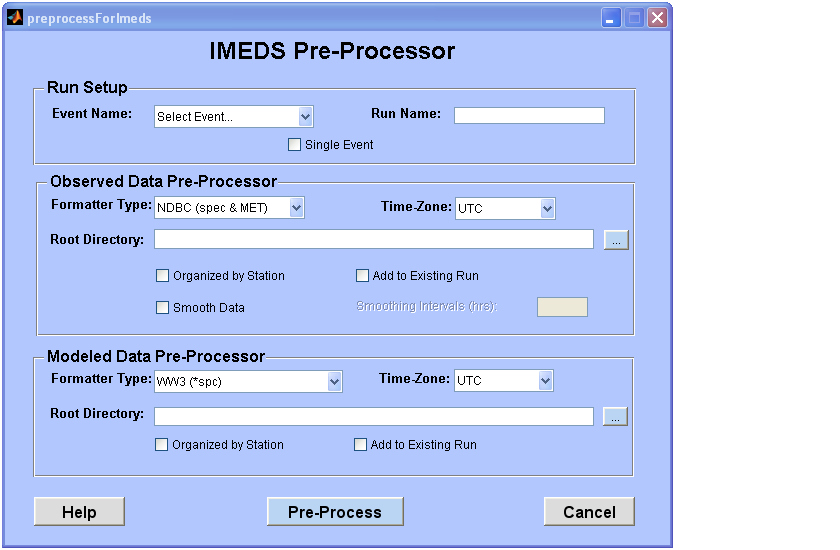
If ‘Organized by Station’ is checked that means the files (model or observed) are organized by folders whose name is the station name inside the root directory that the user selected. If it is not checked the program would be looking for the data files in the root directory selected by the user.

* **Time-Zone Pull-Down Menu**

The user needs to specify which time-zone the results are in so that IMEDS can correctly match the observed data and model results.

* **Smooth Data Checkbox**

For observed data, the user can also choose to smooth the data by checking the ‘Smooth data’ checkbox and entering the number of hours he or she desires the data to be smoothed over in the ‘Smoothing Interval’ edit area. . This data will be saved smoothed. If the user chooses to try different smoothing values, it would be advised not to smooth the data here but use the smoothing option in the Run Setup. This later option will not affect the data saved on the drive, but only the data used in a particular run.



*Figure 2: IMEDS Pre-Processor*



## 4.1 Observed Data Pre-processors

### CHL NDBC Formatters (spe1D or spe2e and onln)

* This is a custom format used by USACE-CHL.
* The IMEDS preprocessor selects the NDBC processor to use depending on the extension of the spectra file (spe1D or spe2D). If the files are organized by station the preprocessor is called once for all the files in the directory. If not it is called file by file since it has to pick which processor to use for each file.
* It will give an error if a spectra file can't be found.
* As for the WW3 formatter if the data is organized by directory the station name (which needs to be the same than for the model data) is assumed to be the name of the directory. If not, it will take the 2nd character of the filename up to the first underscore. (e.g. for n46041\_2000\_11.onln, 46041 is the station name)

### NODC formatter

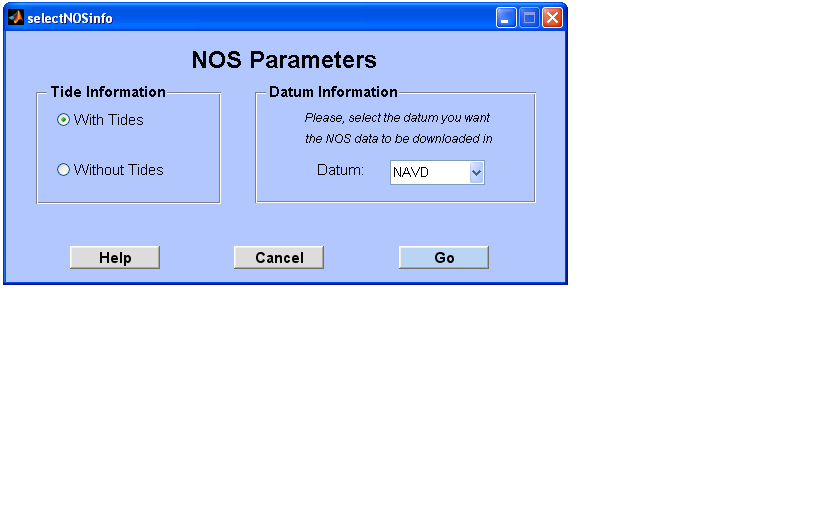
* This formatter operates on wind and wave spectra from the standard NODC File Type 291. NODC data can be downloaded at

http://www.nodc.noaa.gov/BUOY/buoy.html

* Data can be organized by station folders or can be in the same folder. As for the other formatters if the data is organized by folder, the station name is assumed to be the folder name. If it is not, we use NODC conventions as in 44009\_200902 with stationName\_YearMonth used to identify the station name and dates. The year and month will be used as described above in both cases.
* The data needs to be unzipped before running the preprocessor as MATLABTM unzip function doesn't support the .Z extension used by NODC.
* Since NODC files don't have any extension, IMEDS will consider each file found in the specified directory to be a valid NODC format.

### NOS

* This formatter/downloader gets stations and time information from the model water-level preprocessors (including ‘Already Pre-Processed) and automatically downloads matching NOS data using NOS web-services and formats it for IMEDS. It thus needs to be run simultaneously with a model preprocessor.
* When the NOS formatter is selected, a new window, presented in Figure 3, opens for the user to specify tides and datum information.
* The user has to specify if the data needs to be detided, in which case, the data is downloaded in the STND datum which is the station datum (all stations have data in STND). If this is not a detided run, the user needs to choose which datum he needs the data in. Please just be aware that all stations are not serviced in all datums.
* Each download takes some time and because we are remotely logging to a server many different errors can occur. Unless the error message is to warn the user that the data is not available or that the station is not a NOS station, IMEDS will try to download the file once again. If one gets numerous error messages from the NOS server, we would suggest trying again.



*Figure 3: NOS Parameters Window*

### IMEDS Generic Format

* A generic ASCII format was built to be able to ingest any wind, wave, water-level data into IMEDS. This format has been thought of to also allow other data types as future versions of IMEDS arise. The format description is presented in Appendix 3.
* Once this is selected, the user is asked to browse to a filename (instead of the root directory).
* There is so naming convention for the filename
* As of version 3.1, this formatter doesn’t have the ability to decompose the data set in monthly files if the event covers more than one month.

### Already Pre-Processed

This is not a formatter but rather indicates that the user has already formatted the observed data. Some model formatters (such as the ADCIRC Wind preprocessor or if the user wants to extract data from the SURA OpenDAP server) use already pre-processed observation files to figure out where (latitude and longitude location) they need to extract data from the model. In that case, the user will be asked to select an observation folder pointing to the files with the points that need to be extracted.

## 4.2 Model Data Pre-processors

### WAVEWATCH III (WW3) Formatter

* As of release 3.0, one unique WW3 formatter handling single month/station and multi months/station exists.
* The only filename convention is that the filename ends with ‘spc’.
* The station name is obtained from the file. The station name will be 46042 if the header looks like:
  + '224 46042' 37.00 237.50 988.6 10.73 318.1 0.00 270.0 or
  + '46042 ' 37.00 237.50 988.6 10.73 318.1 0.00 270.0.
* If your header line looks any different, you may encounter problem. Please let us know if that’s the case.
* The data can be organized by stations or all the stations can be in the same folders
* Each file can have as many records as desired. IMEDS will take care of dividing it into monthly files if necessary.

### SWAN Formatter

* The SWAN formatter works on single station SWAN output. It needs the spectra file and optionally the TAB file for wind and depth information. If the TAB file is missing, one can still evaluate the wave data but not the wind data for a particular station.If the files are organized by folder, the formatter gets the station name from the folder name. If they are not, it will get the station name from the spectra file, after ‘spec2d.out.’. (e.g. 45032 will be the station name from the file spec2d.out.45032)
* The station name also needs to be included in the TAB file as in 45032\_TAB.
* Each station file can have as many records as desired. IMEDS will take care of dividing it into monthly files if necessary.

### Oceanweather Formatter (Wave)

* This formatter is developed for Oceanweather, Inc formatted wave spectra files.
* The data can be organized by stations or all the stations can be in the same folders
* The filename convention is buoyName.asc (as in 47012.asc). This can be modified if necessary.
* Each station file can have as many records as desired. IMEDS will take care of dividing it into monthly files if necessary.

### Oceanweather Formatter (Wind)

* This formatter is developed for Oceanweather, Inc formatted wind file.
* No naming convention is required but only one file (with the full OWI grid) should be in the specified folder.
* !!! As of version 3.0, this formatter needs to be run in parallel with the NODC formatter to know which data points to extract

As of version 3.0, the OWI wind formatter doesn’t separate a given file in monthly files if it spans more than a month.

### ADCIRC Water-Level Formatter (fort.61)

* This formatter is developed to read the fort.61 file from ADCIRC. This file contains water elevation at specific locations.
* Both the fort.61 (water elevation),and fort.15 (stations list) need to be in the directory. The fort.221 (input wind file) is an optional file. This is needed to know the model start time. If one knows the model start time and doesn’t want to download the quite large fort.221 file, IMEDS gives the option of entering the model start time if the fort.221 is missing.
* The formatter will extract the water elevation data from the fort.61 file accordingly to the list of stations found in the fort.15 file and will align the model times to the observed time with the date found in the fort.221 (or to the date entered).
* The files need to be named fort.61, fort.15 and fort.221.
* IMEDS is keying on ‘TOTAL NUMBER OF ELEVATION RECORDING STATIONS’ to find the beginning of the stations list. This comment needs to be in the fort.15 file.
* As of version 3.1, the ADCIRC water-level formatter doesn’t separate a given file in monthly files if it spans more than a month.

### ADCIRC Wind Formatter (fort.72)

* This formatter is developed to read the fort.72 file from ADCIRC. This file contains wind data at specific locations.
* Both the fort.72 (wind), and fort.15 (stations list) need to be in the directory. The fort.221 (input wind file) is an optional file. This is needed to know the model start time. If one knows the model start time and doesn't want to download the quite large fort.221 file, IMEDS gives the option of entering the model start time if the fort.221 is missing.
* The formatter will extract the wind data from the fort.72 file accordingly to the list of stations found in the fort.15 file and will align the model times to the observed time with the date found in the fort.221.
* The files need to be named fort.72, fort.15 and fort.221 (or to the date entered).
* IMEDS is keying on 'NSTAM - NUMBER OF MET RECORDING STATIONS' to find the beginning of the stations list. This comment needs to be in the fort.15 file.
* As of version 3.1, the ADCIRC wind formatter doesn't separate a given file in monthly files if it spans more than a month.

### IMEDS Generic Format

Please refer to the description of the IMEDS Generic format in the above ‘Observed Data Pre-processors’ section

### Extract from SURA OpenDAP Server

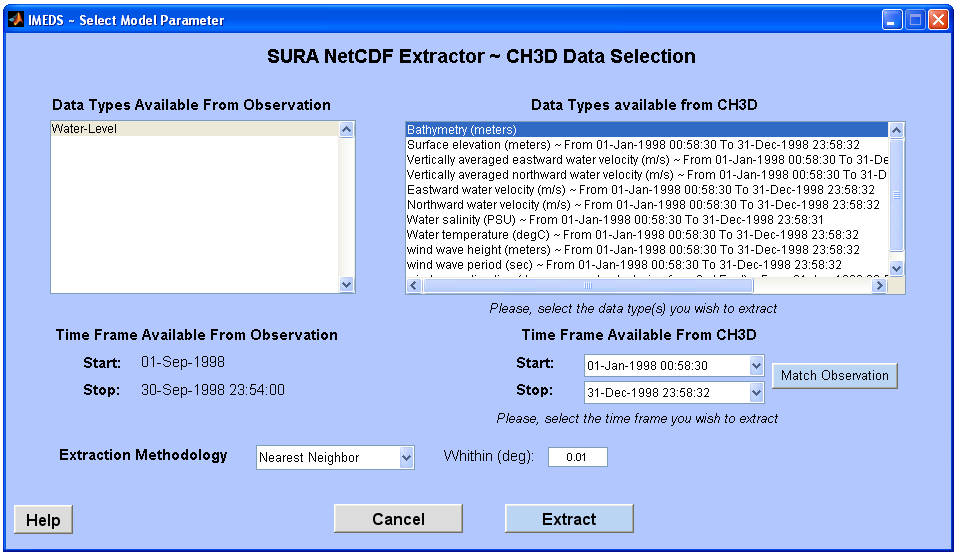
* This option has been added in version 3.1 to support the SURA project. More information about this project can be found here:

<http://www.sura.org/programs/coastal.html>

* This needs to be selected at the same time than an observation preprocessor, so IMEDS knows where to extract data.
* When selected, another window, presented in Figure 4, will open to let the user choose which model he or she wants to use from the list of models available from the SURA OpenDAP server.
* The user will then be directed to another module, presented in Figure 5. Information about the data found in the observation file is displayed on the left hand side. Information about the data available from the chosen model is available on the right hand side. A pull down menu lets the user picks the start and end date of the data type(s) he wishes to extract. In order to make it easier on the user, a push button ‘Match Observation’ selects the date of the model to surround the observation dates.
* Finally, the user needs to select the extraction methodology he desires. Bilinear interpolation will be an option in the future, but as of version 3.1, the nearest neighbors’ methodology is the only one available. For this method, the user is asked to input a distance threshold (in degrees) to find a point to extract within a certain range of the observation point location.
* Messages warn the user if something he is trying to do is impossible, such as if the variable he/she is trying to download is not (yet) handled by IMEDS, if a match for an observation station couldn't be found or if a limitation of version 3.1 has been reached.

## choose mod.PNG

*Figure 4: IMEDS SURA Model Selection*

**

*Figure 5: IMEDS SURA NetCDF Extractor*

### Extract from local NetCDF file

* The local netCDF extractor works the same way than the SURA OpenDAP server extractor but on a local file. The user thus needs to browse to a local file after choosing this formatter instead of choosing the model from a list.
* Please refer to the above section ‘Extract from SURA OpenDAP Server’ for further details on the extractor.
* This also needs to be run simultaneously with an observation preprocessor so IMEDS knows where to extract the data.

### Already Pre-Processed

This is not a formatter but rather indicates that the user has already formatted the model data. The NOS downloader/formatter can use already pre-processed observation files to figure out where (latitude and longitude location) it needs to download data. In that case, the user will be asked to select a model folder pointing to the files with the points that need to be downloaded.

## 4.3 Auto-Filling of the IMEDS Setup Form

* Once the pre-processor is done preprocessing the input data (and if no problem was encountered) the IMEDS setup GUI is launched and auto-filled with the information from the preprocessor. More details are given in the following section about Run Setup Windows, but IMEDS picks the appropriate Run Setup Window depending on the preprocessed data.
* If one only decides to preprocess model data, the observed data directory in the IMEDS setup would be automatically selected as being

/imedsRoot/IMEDS/data/obs/event\_name/. A warning is displayed to tell the user that this information needs to be checked.

* If one only decides to preprocess observed data, the model data directory would be automatically selected as being:

/imedsRoot/IMEDS/data/obs/event\_name/run\_name if the run name has been specified. A warning will also be displayed. If the run name has not been specified in the preprocessor, the user will have to browse to the model data directory in the IMEDS setup screen.

* In both above cases the ‘Organized by station’ is checked by default since it seems to be the most likely. Please, make sure this is correct before proceeding.
* The output directory is automatically selected as being

/imedsRoot/IMEDS/data/processed/event\_name/run\_name but the user can browse to another directory if he desires. Again, if the run name hasn't been specified the user will have to browse to the output data directory in the IMEDS setup screen. This directory will then be created while IMEDS is running. If the user tries to change this directory name, the browsing will take him or her to an upper level directory since the directory would not exist yet.

* The data type is automatically selected.
* Temporal correlation is run by default and no other analysis type is selected by default. No output (display or save) is selected.
* The stations list is automatically populated.

*All those settings can be modified at the user discretion.*

### 

# V.

# 

# 

# Evaluate Data Modules

As of IMEDS 3.1, two Run Setup Windows and a High Water Marks processing are available.

* Wind & Wave Run Setup: Here the user can evaluate wind and wave data together or separately.
* Storm Surge Run Setup: Here the user can evaluate water-level data.
* High Water Marks: Here the user can evaluate High Water Mark (HWM) data.

The High Water Mark module is a separate module and most setup options are similar in both Wind & Wave and Storm Surge Run Setup Windows. This section will start by describing the Wind & Wave Run Setup Window, and will then describe any Storm-Surge related options. The description of the High Water Marks processing will follow.



## 5.1 Wind & Wave Run Setup

The Wind & Wave Run Setup Window, depicted in Figure 6, allows the user to setup all necessary information for IMEDS to process both observed and modeled wind & wave data. Below is a description of each feature available in the Wind & Wave Run Setup Window.

* **Preset Event and Preset Run**

User-defined run setups can be saved for future use. In order to simplify the access of a run, different runs are associated with an event. The user can thus run different versions (runs) of Katrina (an example event). Selecting a given event in the *Preset Event* would give the user a list of associated runs in the *Preset Run* pull-down menu. If this is a new event and/or a new run, both pull-down menus have an ‘Enter New…’ option that will guide the user into entering a new event and run name if chosen from the *Preset Event* pull down menu or a new run only if chosen from the *Preset Run* pull-down menu. Figure 7 shows the popup window the user needs to fill when choosing a new event.

If the user selects a preset event and a preset run then the form is filled in automatically.

* **Test Directory, Ground Truth Directory and Output Directory**

This is where IMEDS looks for the test (model) data, ground truth data and where to output the results. The user can type the name of the required directories or browse to them by clicking the *Browse* buttons. The *Organized by Station* checkbox needs to be selected if the data is organized by station. In this case the folder name needs to be the station code. (E.g.: It needs to be checked if the model data file is:

/*imedsRoot*/IMEDS/data/mod/myTest/41025/41025\_SWAN.mat and unchecked if it is :

/*imedsRoot*/IMEDS/data/mod/myTest/41025\_SWAN.mat (/*imedsRoot*/IMEDS/data/mod/myTest/ being the *Test Directory* entry))

* **Test Name**

This can be a model name (e.g.: SWAN) or a test name. It needs to be what was referred to as test in the IMEDS requirements paragraph. If the name the user is looking for is not in the list, one can be entered by selecting 'Enter New'. The user would then be guided into entering a test name.

* **Multi-Months Selection**

A basic IMEDS run is based on a monthly file. However, the IMEDS run can be performed over more than one month of data. If the user chooses to do so, the time selection check box needs to be checked and the time span of the run declared using the pull down menus and edit boxes situated below the time selection check box.

* **Data Selection**

The user has the choice between wind, wave and water-level analysis. Wind and Wave can be done at the same time as long as both data type are included in the wavedat file.

* **Extra Analysis Type**

Select types of analysis desired: Quantile-Quantile (QQ) or Peak Analysis (PK). Temporal Correlation (TC) is executed by default. If PK is selected, the Peak Analysis setup window will pop up.

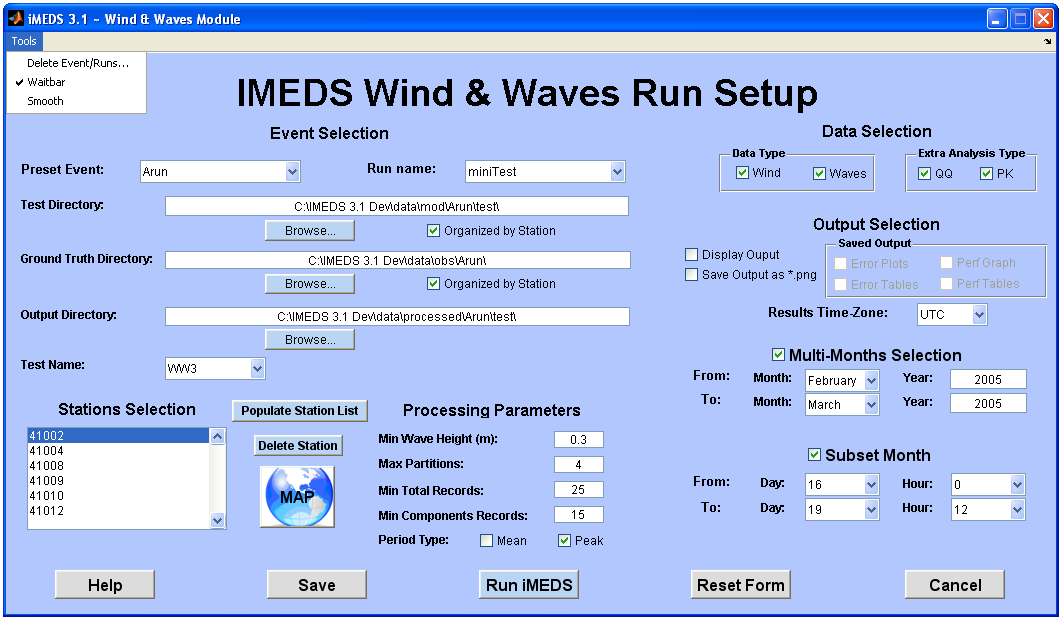
* **Peak Analysis Setup**

As of version 3.1, two types of Peak Analysis are possible using a fixed elevation or using standard deviation. The peak event analysis will occur for each data type selected in the Data Selection area in the run setup window. The other ones will be grayed out.

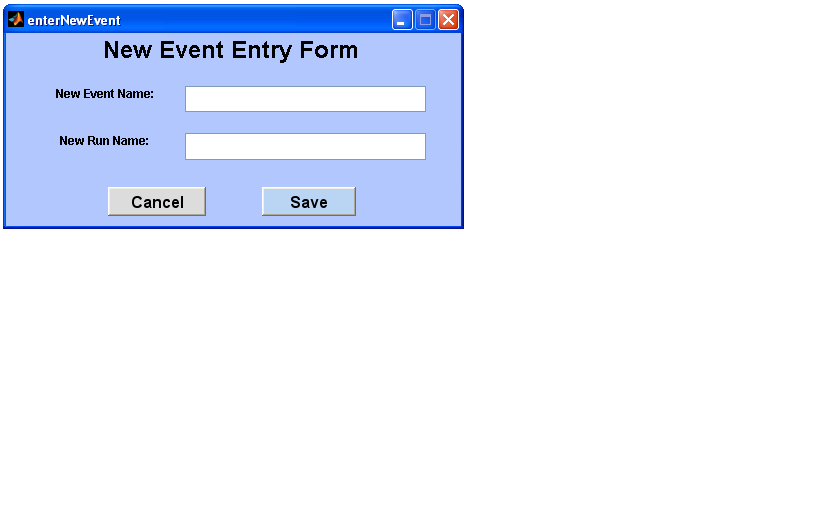
If the user select the fixed elevation methodology, thresholds have to be entered in terms of wave height, wind speed and/or water-level amplitude.

If the standard deviation is chosen, the user needs to specify the number of standard deviation above the mean for each data type. An example is shown in Figure 8.

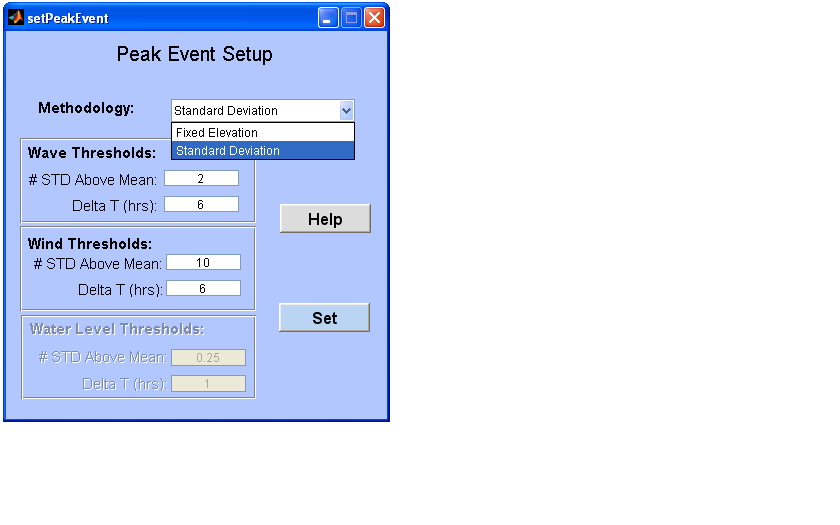
Whichever methodology is used a delta T in hours needs to be specified for any given data type. If 6 hours is selected the peak analysis will look for the model peak 6 hours behind and ahead of the observed peak.



*Figure 6: IMEDS Wind And Waves Run Setup Window*



*Figure 7: IMEDS New Event Entry Form*



*Figure 8: IMEDS Peak Event Setup for Standard Deviation for Wind and Wave*

* **Subset Month**
  + If a single-month run is being setup, this section would subset the monthly file from starting day and hour to ending day and hour. The user is warned if the time-range specified is out of the file time-range. Model outputs have usually larger time-steps than observed data. If a user requests a subset from the 5th at 7 am to the 15th at 8am and the output is every 6 hours (00,06,12 and 18) the file would be subset from the 5th at 12 am to the 15th at 6am. One thus needs to be careful by requesting a subset large enough to include both modeled (or tested) and observed data. The time-zone used here is the time-zone selected for the output.
  + If a multi-month run is being setup, the first month will be subset from the starting day and time to the end of the month and the last month will be subset from the beginning of the month to the ending day and time. The same rules as the single month subset apply here.
* **Output Selection**

After IMEDS is done running the display module opens and all plots, graphs and tables are then available to the user. However, the user has 2 options concerning the output while IMEDS is running.

* Display Output: the figures produced by IMEDS will be visible and pop up to the screen while IMEDS is running.
* Save Output as \*.png: The figures are save as a \*.png file while IMEDS is running. (This is particularly useful if the user intends to use all or most of the figures into a presentation) .The user has the possibility to choose what type of output he wants to be save. The error plots and tables are associated to each station statistics and performances, while the performances tables and graphs are associated to the summary of the model run (when the stations statistics are folded up).
* **Time-Zone**

This is the time-zone the results are going to be displayed in. Once the wavedat files are loaded in IMEDS, the data is switched to this time-zone for the processing and results display.

* **Station Selection**

The user needs to specify which stations the IMEDS run need to consider. This is done in the Station Selection part of the IMEDS Run Setup GUI.

Five parameters common to all stations for a given run need to be specified.

* + Min Wave Height: the minimum height a wave partition needs to be to be considered in the performance calculations. Default to 0.3 m.
  + Max Partition: The maximum number of partitions to consider at a given point in space and time. Default to 4.
  + Min total records: The minimum number of observed records to be available for IMEDS to do any processing. If a file has a number of records below this threshold IMEDS would behave as if this file does not exist. Default to 25.
  + Min components records: The minimum number of record for a component to be included in some statistics and any performance calculations. A data set needs to have at least this number of wind-sea for example for the wind-sea performances to be calculated. This number thus corresponds on the number of matching records between observed and tested data. (generally less than the number of observed records). Residuals, scatters and bias are still calculated if the number of components records fall below this threshold. Default to 15.
  + Period type: If the wave data holds wave spectrum information (and not bulk statistics only), the user can choose between using the peak or the mean period for the wave analysis. Default is peak period. (Please note that even if mean period is chosen, the tables will still say Tp in version 3.1)
* **Populate Station List**

At the push of this button, the stations list is populated from files IMEDS could find in the observation and tests directories. For more information on how IMEDS finds those station IDs, please refer to the IMEDS Requirements section at the beginning of the User Guide. If the user chooses to do so, a run can be done over a subset of those stations. One needs to select the station he doesn’t want to process from the station list and push the ‘Delete Station’ push button.

* **Tools Menu**

Three options are available from the ‘Tool Menu’:

* + Delete Event/Runs

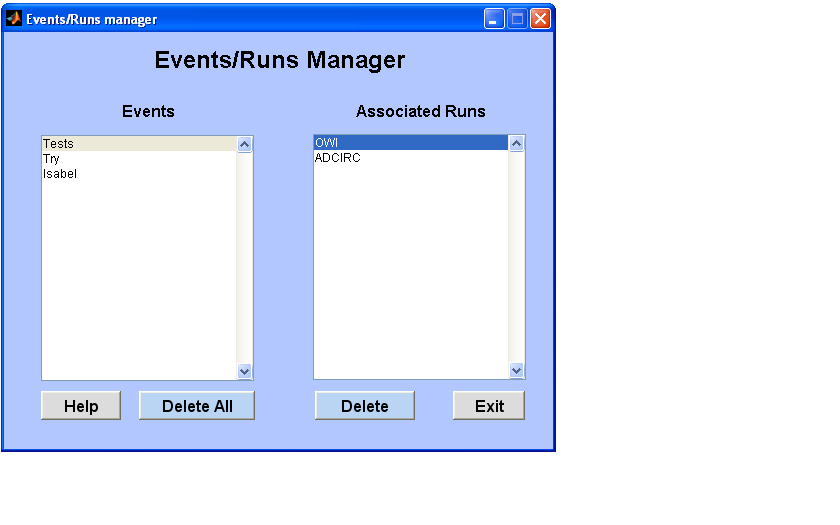
This option will open the ‘Events/Runs Manager’ as shown in Figure 9. This allows this user to clear some of his old events/runs. This will delete events and/or runs from the Run Setup GUI lists but won’t delete any processed data associated to a particular run.

If a run is selected in the ‘Associated Runs’ list, hitting the ‘Delete’ button will remove this run.

In order to remove all the runs from a particular event, a faster option is to select the event itself from the ‘Events’ list and hit ‘Delete All’.

The ‘Help’ button gives a reminder on how to use this Events/Runs Manager.

The ‘Exit’ button closes the window.



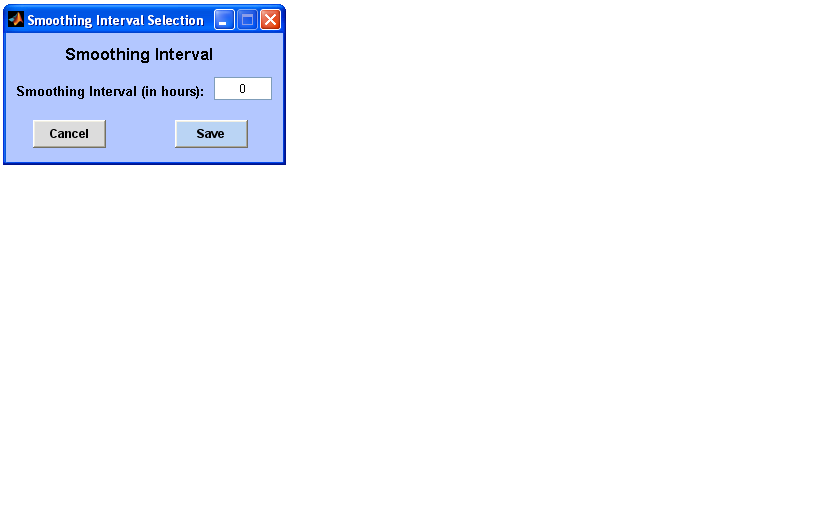
*Figure 9: IMEDS Events/Runs Manager GUI*

* + Waitbar

Waitbars are automatically turned on during IMEDS processing. The user can turn them off by checking this option off in the Tools menu.

* + Smooth

This option enables data smoothing for the observed dataset over a specific interval. The Smoothing Setup window, shown in Figure 10, pops up when selecting the ‘Smooth’ option. The user can then enter the smoothing interval in hours that he wishes for. The smoothed data will be used for the run but won’t be saved on the hard drive. The user can thus try different smoothing values without preprocessing his data again. This can be particularly useful for the Peak Analysis.



*Figure 10: Smoothing Setup Window*

* **Save Setup Button**

This button allows the user to save the particular setup without running IMEDS

* **Run IMEDS Button**

*Run IMEDS Button* first asks the user if he wishes to save the setup and then run IMEDS. Once IMEDS is done running the Display Module will get launched.

* **Reset Form**

The Reset Form button resets the Run Setup form as it was found when starting the GUI

* **Cancel Button**

This closes IMEDS.

* **Help Button**

This button opens the help document

## 5.2 Storm Surge Run Setup

As previously mentioned most options from the Wind & Wave Run Setup are available in the Storm Surge Run Setup, presented in Figure 11. This section is describing the few options that are specific to the Storm Surge Run Setup

* **Data Type**

The only data-type currently available in the Storm Surge Run Setup is water-level. It is expected that in a future version harmonics analysis will also be available.

* **Extra Analysis Types**

Extremes Analysis (EA) is available for water-level processing. EA is similar to the Peak Analysis but for the fact that it is not only finding the high peaks but the low peaks as well. This is particularly useful in tidal analyses. Similarly to the PK analysis, 2 different options are available. The user can choose between a fixed elevation and a certain number of standard deviation above the means with the pull-down menu. The user then has to respectively enter the amplitude or the number of standard deviation. Finally a delta T in hours needs to be specified to know how far in time to look for a model extreme matching an observed extreme.

* **Processing Parameters**
  + Time-step (hrs): The observed and model data will be matched to the same times, using the time-step specified here.
  + Obs Smoothing Interval (hrs): The smoothing interval for water-level observed data.
  + Mod Smoothing Interval (hrs): The smoothing interval for water-level model data.

## stormSurge.PNG

*Figure 11: IMEDS Storm-Surge Run Setup Module*

## 5.3 High Water Marks Module

This module allows the user to assess High Water Marks statistics. We apologize for the roughness of this module but this has been added at the last minute.

### Setup

In this module’s interface, presented in Figure 12, the user first needs to specify the project name (for the plots titles) and the output directory (default in data/processed). The user next needs to browse to the observed and model High Water Marks files (the default browsing directory is data/raw). The files can either be in the IMEDS Generic format for time-series data, or in tabular format for HWM data only. The user needs to specify the file format in the models and observed ‘File Format’ sections. Both formats can be used within one run. If the file is in IMEDS Generic format, the maximum elevation is extracted from the time-series. The tabular format is defined as lat lon hwm (in meters) , such as:

37.178333 -76.396944 1.7221

37.108611 -76.393056 1.9489

37.110556 -76.319167 1.9324

37.141944 -76.375833 1.8227

37.492778 -76.310000 1.3167

36.906667 -76.088333 1.8349

37.177222 -76.805833 1.7861

If the user so chooses, comments can be added to the file but they need to start with ‘%’ (without quotes).

In order to match the model HWM with the correct observed HWM and leave some freedom to the modelers, the HWM don’t need to be in the same order in both files but the lat/lon values from the model need to be within a ‘degree distance’ from the observed lat/lon. This distance needs to be specified by the user in the ‘HWM Match’ area. The default value is 0.

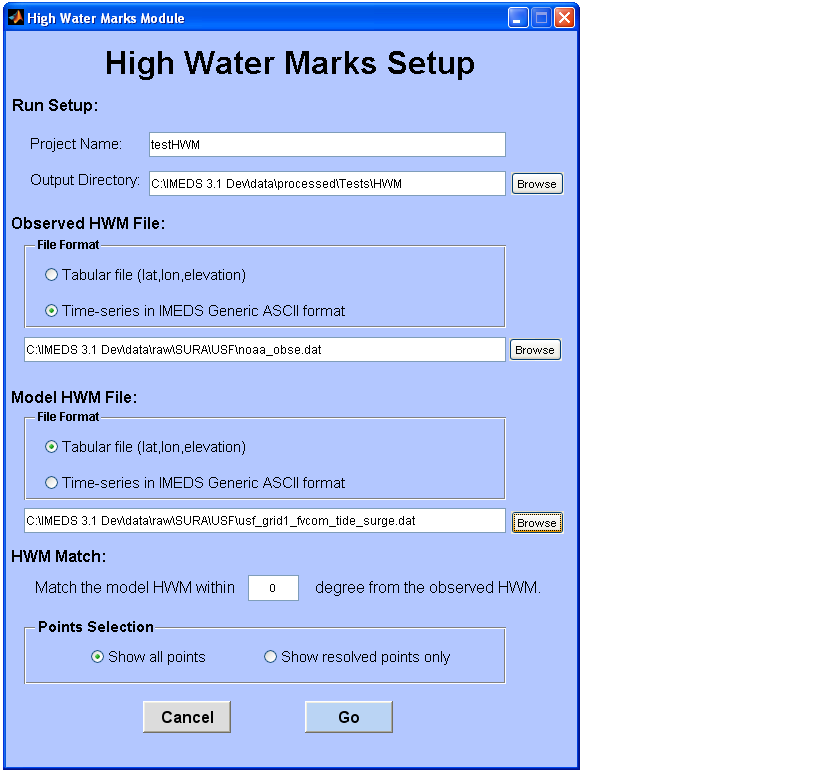
IMEDS will warn when an observed HWM can’t be match with any model HWM.

In order to avoid a warning, modelers can also stipulate an unresolved HWM by a ‘NaN’ (without quotes)

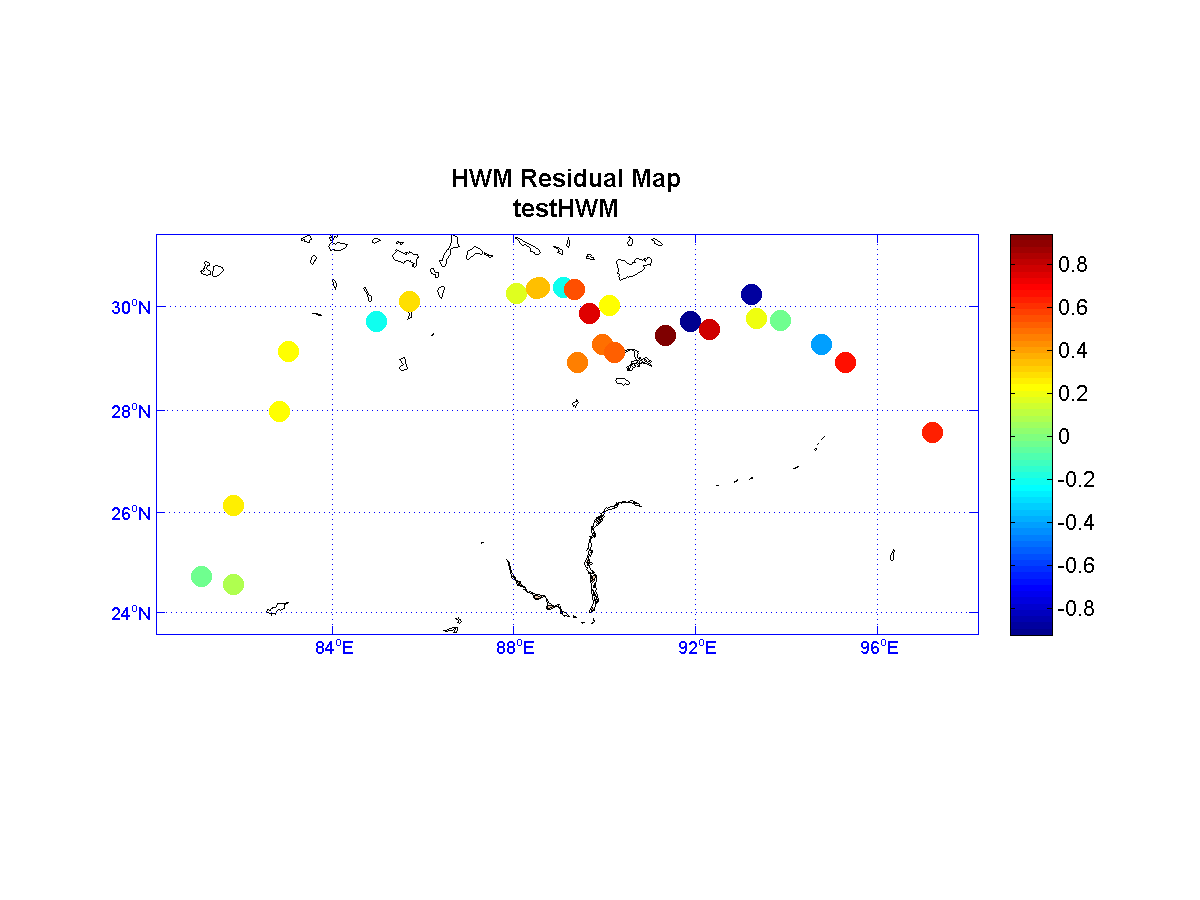
The user also can choose between getting statistics on all the observed HWM or only on the ones that the model has resolved. This option is available in the ‘Points Selection’ section. Default is on all the points.

### Results

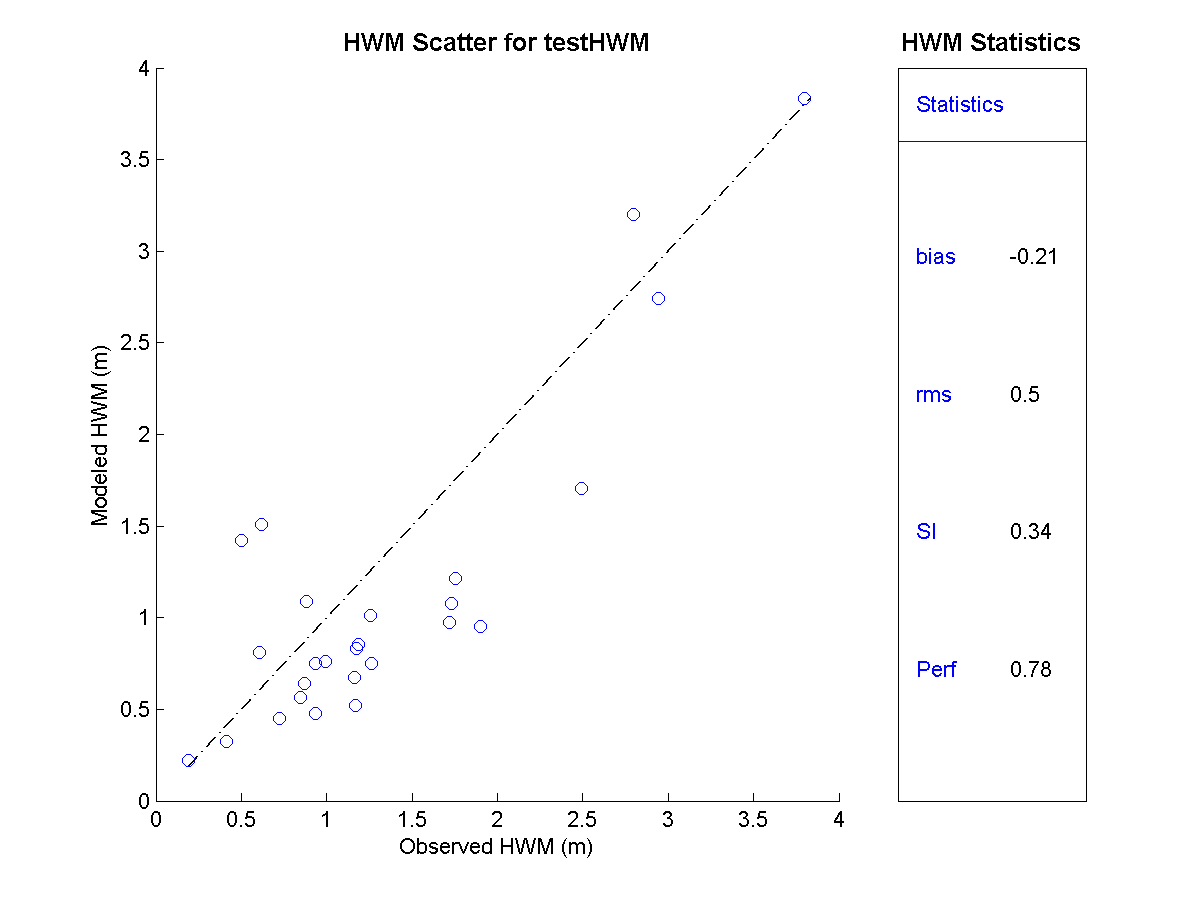
Two different plots are produced for the HWM analysis: a residual HWM map and a HWM scatter plot. The residual map, presented in Figure 13, shows color coded residual (observation - model) for each station. The scatter plot, depicted in Figure 14, presents a scatter plot of observed versus model, as well as a statistic table relating the bias, RMS error, scatter index and performance associated to the High Water Marks.



*Figure 12: High Water Marks Setup Interface*



*Figure 13: HWM Residual Map*

**

*Figure 14: HWM Scatter Plot and Statistic Table*

# VI. Display Modules

As of version 3.1, IMEDS has 3 different display modules: one for waves, one for wind, and one for water-level. The Display Modules open once an IMEDS run is over or can be called directly from the Module Selection Window for a previous run. When it opens, it automatically opens the performance module for all the stations and month (if more than one month was selected) for each component. The Display Module is interactive and greys out options that are not available in particular circumstances (e.g. No direction data for 1D buoys). It also warns the user if any data is missing (e.g.: some stations are out of service over several months in the course of a year) or if some statistics haven’t been computed because of too few data points. Temporal correlation is available for all data-types, quantile-quantile analysis and peak analysis are available for wind and waves, and extremes analysis is available for water-levels, as long as they have been included in the analysis. The displays for temporal correlation and quantile-quantile analysis are very similar and described in the three following sections. The peak analysis and extremes analysis displays are slightly different and are presented separately.

## 6.1 Overview of the IMEDS Display Module and Performance Sub-Module

Figure 15 illustrates the options available in the Display Module and some options from the Performance Sub-Module.

The features are listed from top to bottom and left to right.

* **Toolbar**

The toolbar options allow the user to explore the data in more depth by (in order) zooming in, zooming out, pan, obtain the data for a particular point (graphs only), print and save.

* **Analysis Push Buttons** (Temporal Correlation, Quantile-Quantile Analysis, Peak Event Analysis)

These push buttons switch between the different analysis types. If a display type is chosen that is not available in one particular analysis type, the user will be automatically redirected to the performance or statistics table. A message will alert the user if this happens. Extreme Analysis is also accessible here in the water-level display module.

* **Module Options**

This radio-button group selects one of the different IMEDS display modules: global performance scores, statistics grouped by stations or by months. The global performance scores is selected in Figure 13.

* **Display Options**

This radio-button group changes the IMEDS display type. In the performance module, the user can review summarized performances in a table or in a bar graph.

* **Data Options**

In the performance module, the user can select to view performances grouped by stations or by components (e.g: wind-sea, swells for wave data..)

* **Map**

The map button, in the lower-left corner, represented by a globe, plots a map of the different stations used in the given run.

* **Help**

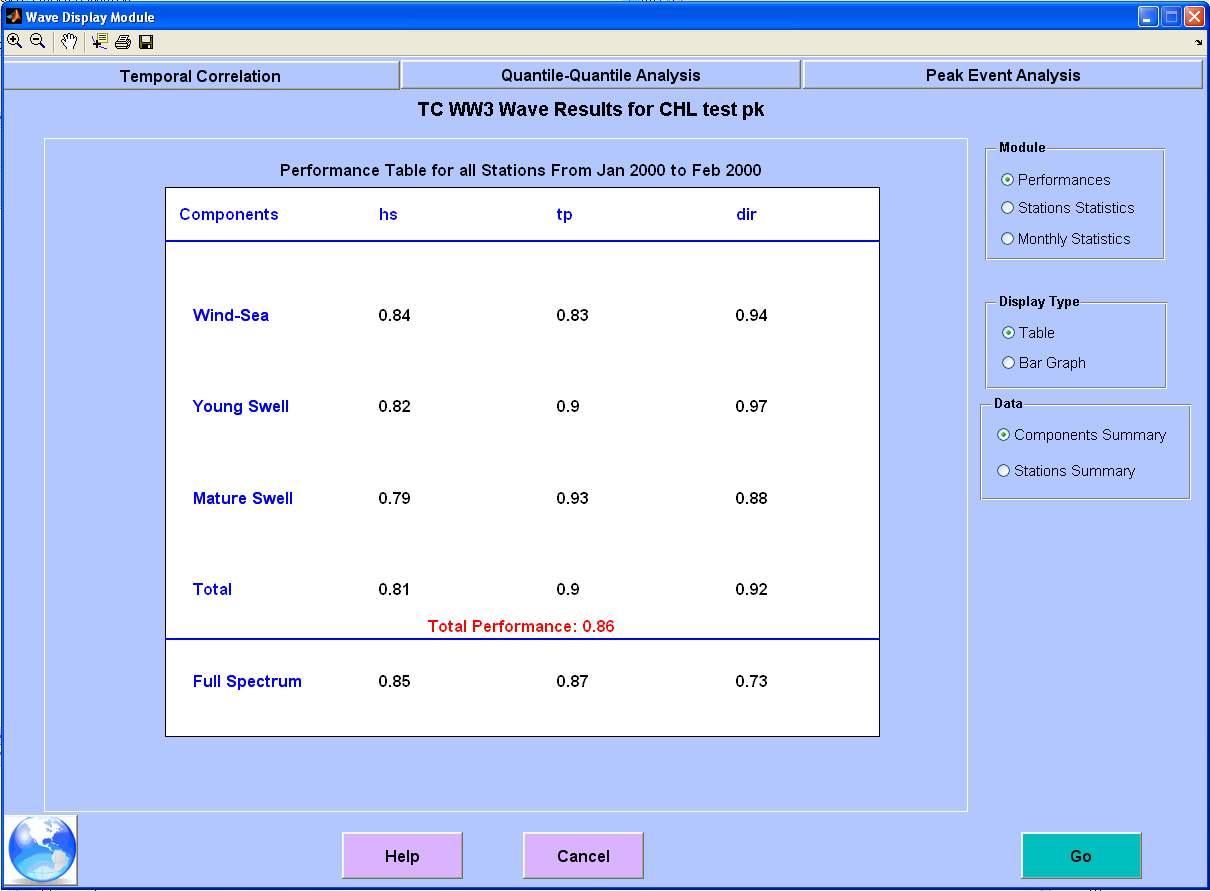
The help button opens the wave display help.

* **Cancel**

Cancel closes the wave display window.

* **Go**

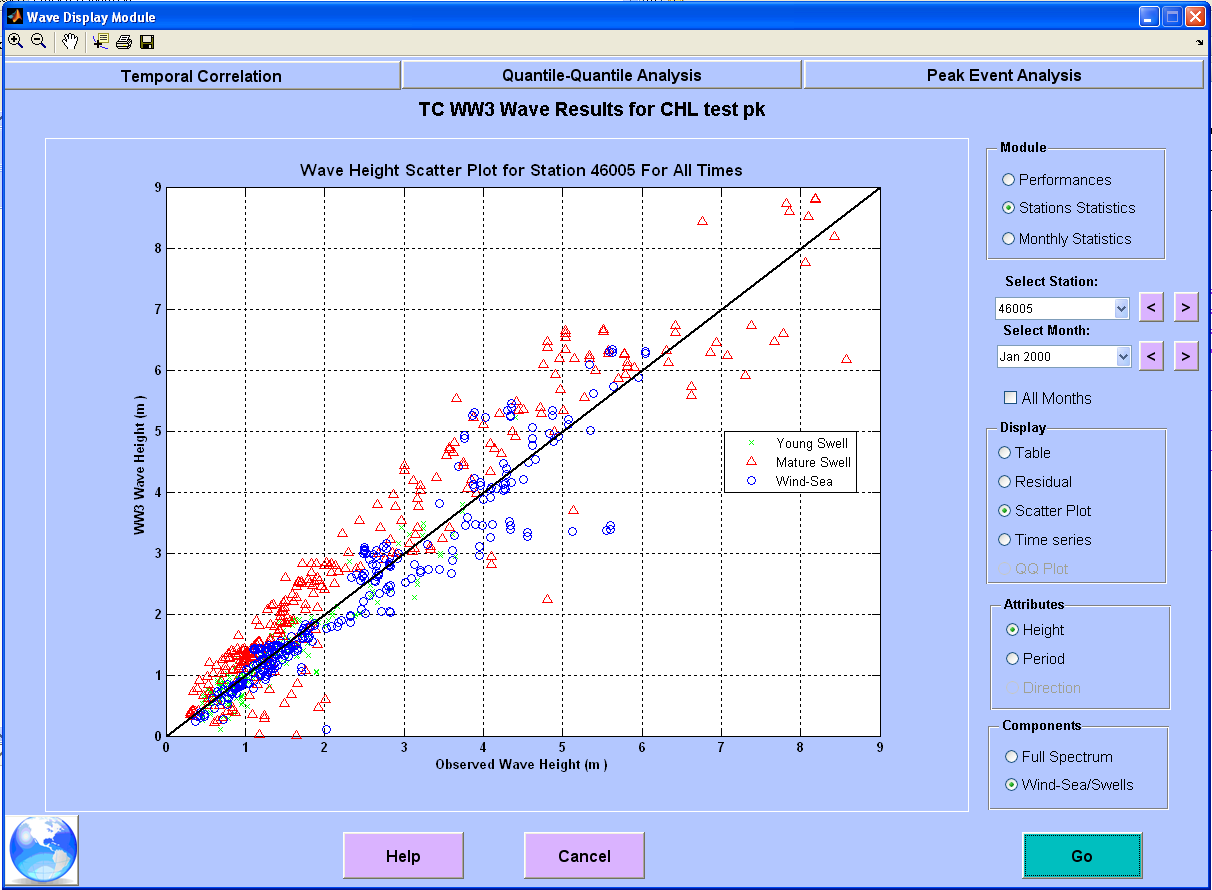
Once the selection of the different options in the right hand-side of the Display Module is done, hitting the Go button displays the results.



*Figure 15: Display Module and Performances Sub-Module Features*

## 6.2 Stations Statistics Sub-Module

The Stations Statistics sub-module gets deeper into the IMEDS results and gives information related to each station for each given month. Figure 16 illustrates the features available in the Stations Statistics Sub-Module. As previously, the features are described from top to bottom. Only the ones that differ from the Performance Sub-Module are stated.



*Figure 16: Stations Statistics Sub-Module Features*

* **Station and Month Selections**

The user can select the station and the month he wants the statistics for using the stations and month pull-down menus. The left and write arrows make it easier to navigate through stations and months once a particular display type is chosen (no need to push ‘Go’). For a multi-month run, an extra check-box ‘All Months’ (below the months selection) is available in the Stations Statistics Sub-Module for a given station statistics for all months.

* **Display Options**

The user can review statistics in a table, a residual graph, a scatter plot or time-series for TC analysis, and a table or QQ plot for QQ analysis. Residuals, scatters and time-series can also be plotted for all months if the ‘All Months’ check box is selected.

* **Components Selection**

In the Stations Statistics module, for all display types but the table, the graphs can be plotted using the full spectrum information of the different wave components.

* **Attributes Selection**

The wave attribute can be selected using radio buttons.

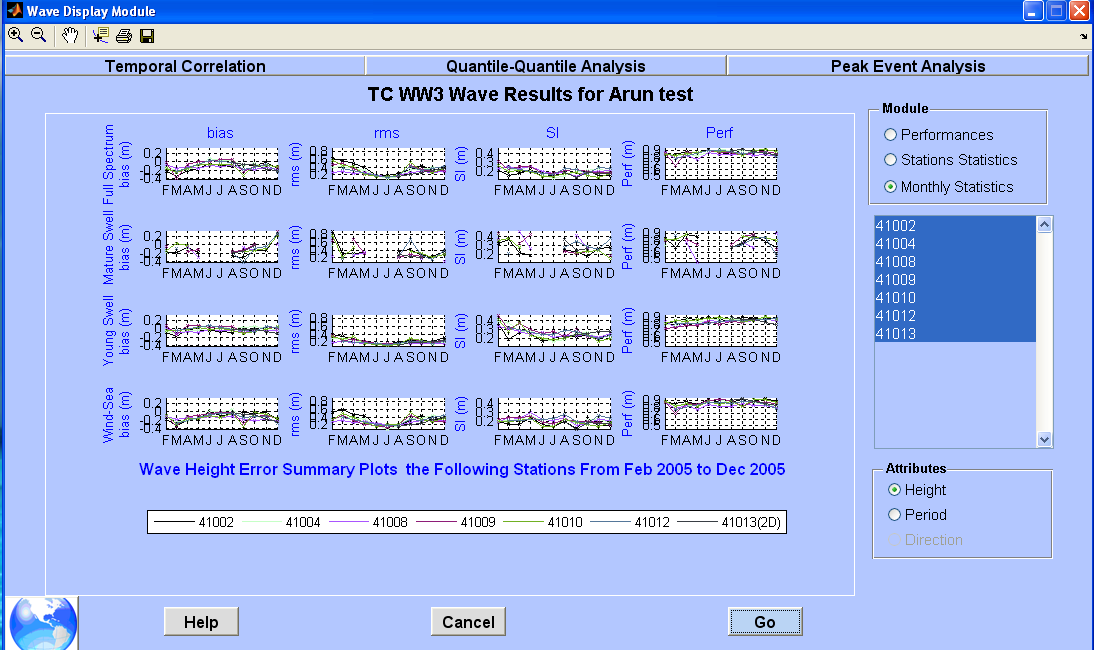
## 6.3 Monthly Statistics Sub-Module

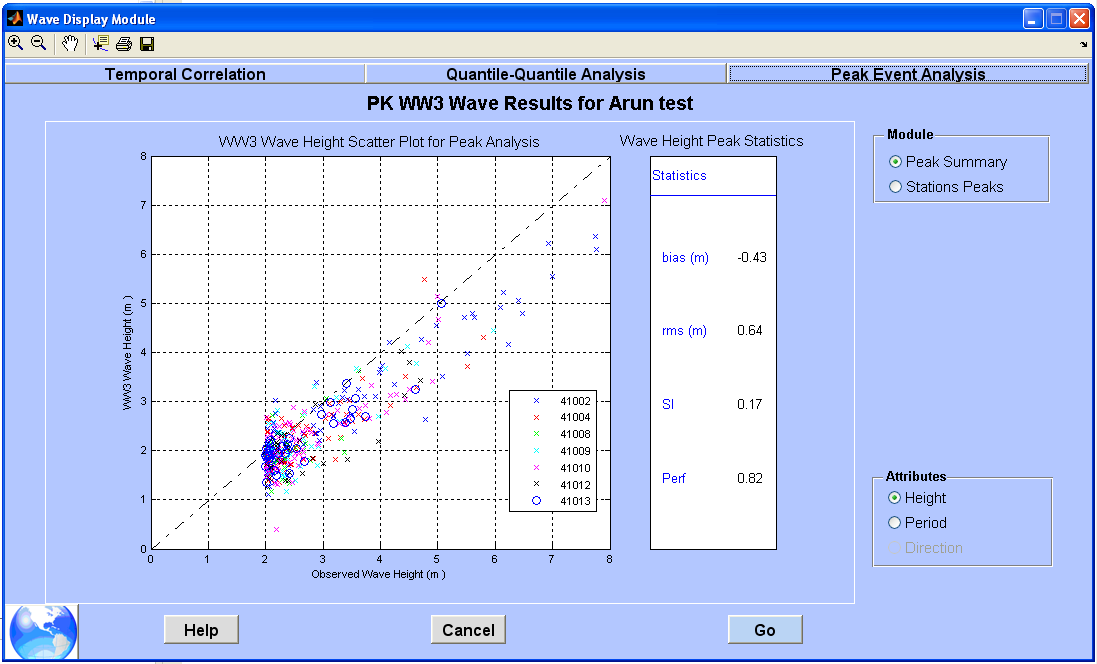
The Monthly Statistics Sub-Module is only available for a multi-month run. The features of the Monthly Statistics Sub-Module are presented in Figure 17. The stations list on the right-hand side allows the user to choose which stations and which wave attribute he wants to plot. It then plots time series of the bias, RMS Error, Scatter Index and Performances for each selected station for the full spectrum and each wave component in the example of wave analysis.

## 6.4 Peak Analysis Sub-Module

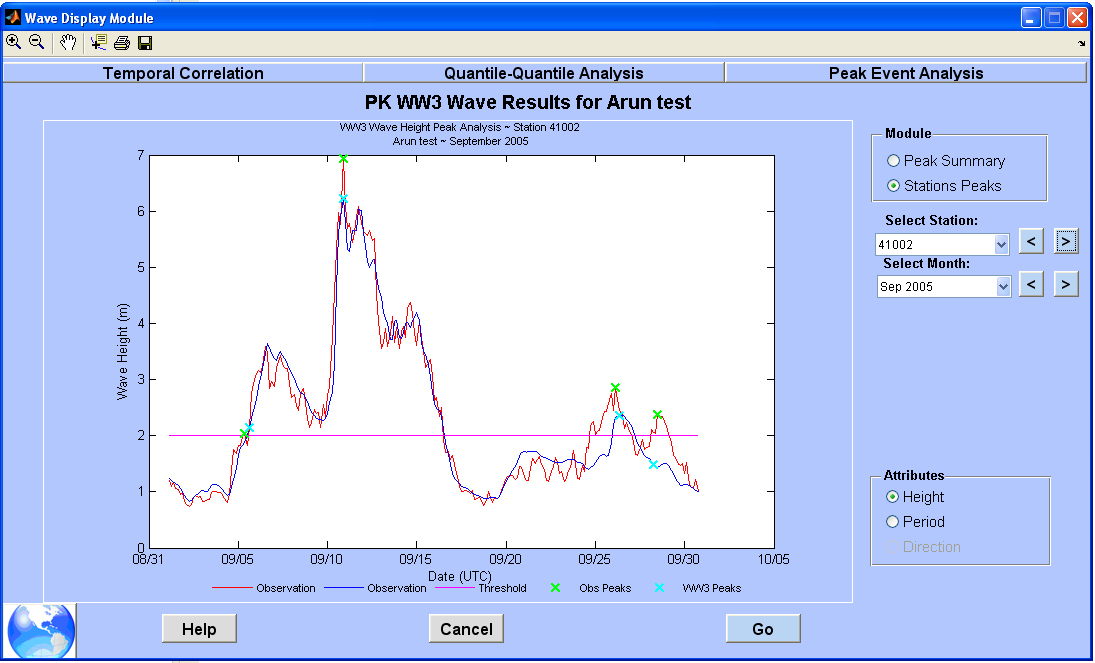
When selecting the Peak Event Analysis, the Peak Summary is automatically loaded. This summary, shown in Figure 18, includes a scatter plot with the peaks color and symbol coded by stations. Many peaks through the course of a month or many months are plotted with the same symbol and color for each station. A peak statistics table is also included in this summary. The table gives the bias, RMS error, scatter index and performance score for the overall peak analysis. The summary can be obtained for any wind or wave attributes, using the ‘Attributes’ radio-button group.

The time-series for each station and month, illustrating the different peaks used for the analysis can be obtained by selecting the ‘Station Peaks’ radio-button. Similarly to the Peak Summary, each attribute can be selected. Note that the threshold line would only appear for the wave height or wind speed. An example can be seen in Figure 19. The observed data is plotted in red while the model data is plotted in blue. If available, the peak threshold is represented as a horizontal magenta line. The model peaks are represented with a cyan cross while the observed peaks are shown with a green cross.

*Figure 17: Monthly Statistics Display*



*Figure 18: Peak Analysis Summary Display*

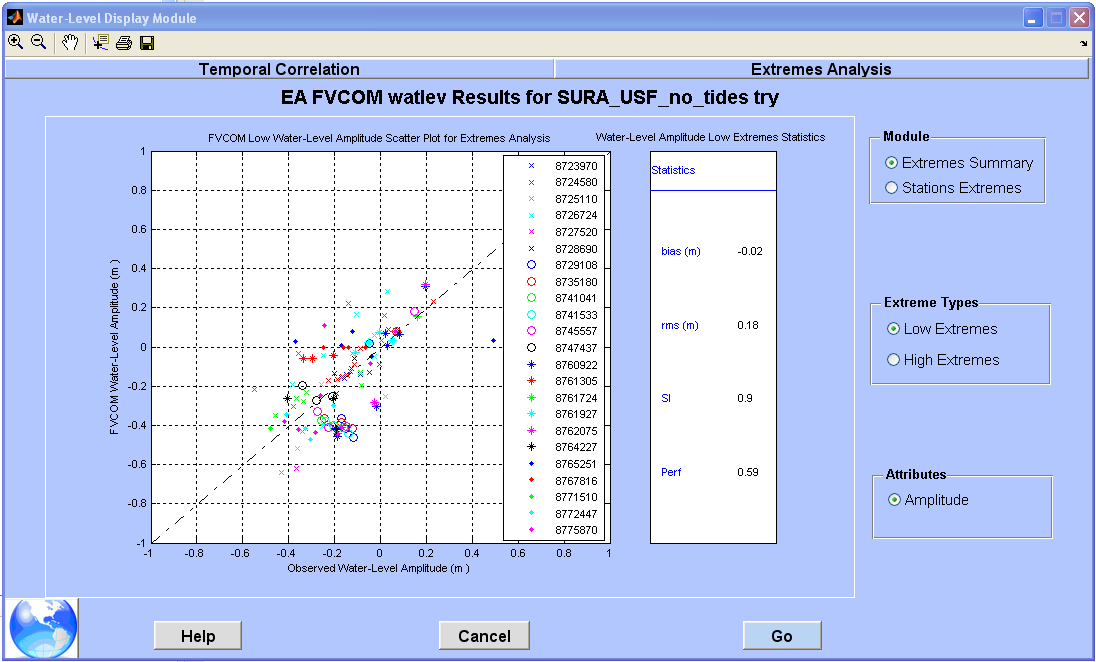


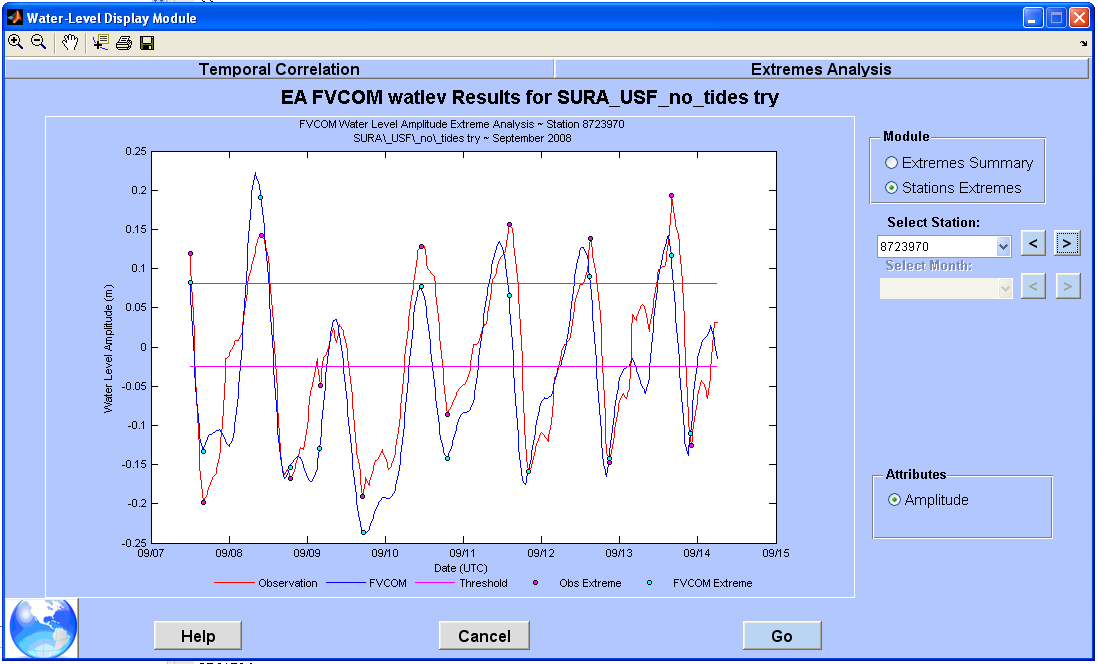
*Figure 19: Peak Analysis ~ Station Peaks Display*

## 6.5 Extremes Analysis Sub-Module

The Extreme Analysis Sub-Module is very similar to the Peak Analysis Sub-Module. Figure 20 presents the low extremes summary display. A similar display is available for high extremes, when the high extremes radio button is selected. As for the peak analysis, summary includes a scatter plot with the peaks color and symbol coded by stations. The table gives the bias, RMS error, scatter index and performance score for the high or low extreme analysis.

The time-series for each station and month, illustrating the different high and low extremes used for the analysis can be obtained by selecting the 'Station Extremes' radio-button. An example can be seen in Figure 21, where the model data is plotted in blue, the observed data in red, and the extremes thresholds are presented with 2 horizontal magenta lines. The observed extremes are shown in magenta, while the model matching extremes are displayed in green.

*Figure 20: Extreme Analysis ~ Low Extremes Summary Display*

*Figure 21: Extremes Analysis ~ Station 8723970 Extremes Display*

# VII. Multiple Model Compare

## 7.1 Overview of the IMEDS Multi Model Compare

## 7.2 Overview of the IMEDS Multi Model Compare

## 7.3 Overview of the IMEDS Multi Model Compare

Appendix 1: IMEDS Analyses Types

The four statistical approaches available in IMEDS are Temporal Correlations (TC), Quantile-Quantile (QQ), Peak Event (PE) and Extreme Analysis (EA). The results of each of these are folded into a Performance Score computation. Each of these computations is briefly described below.

## Temporal Correlation (TC) Analysis

The TC analysis is a direct comparison of time-paired data attributes. The TC analysis provides an indication of how well the hindcast quantities match the observed quantities in absolute time. The following metrics are used to quantify the TC errors:

For non-directional data (speed, time, height and period) the error metrics are:

* RMS error
* Bias
* Scatter index.

For directional data the error metrics are:

* Circular correlation
* Circular bias

## Quantile-Quantile (QQ) Analysis

The QQ analysis provides information on how the distribution of data attributes compare between observation and model results. Quantile-Quantile distributions computed for both data sets (observed and modeled) are statistically compared to compute a set non-directional error metrics (see list above).

## Peak Event (PE) Analysis

The PE analysis extends the IMEDS capability by isolating and computing statistics on event peak data. User-provided thresholds (constants or Standard Deviation multipliers) are used to identify event peaks in wind and wave records. A standard up-crossing analysis is used to isolate time series segments that contain relevant data. Corresponding peaks are identified in the test data using a user-provided time search threshold. Only the bulk (full-spectrum) statistics are used for this analysis. The attributes extracted from each peak event form data pairs that are then used to compute the standard non-directional error metrics (see list above).

## Extremes Analysis (EA)

Similar to the IMEDS Peak Event analysis for winds and waves, the water-level Extreme Analysis (EA) identifies peaks (and lows) and computes error statistics on the extreme highs and lows during the tidal cycle and/or over the passage of a storm.

### Performance Scores

This above analyses result in a set of error metrics that quantify the hindcast skill in reproducing the physical attributes at each observation station. For a 1-year wave study with 6 stations this can result in a database of 3,500 independent measures of model skill for each hindcast run. A performance scoring method was developed to reduce the error metric database into a small set of performance indicators for overall skill assessment. Performance scores are computed by normalizing the wave component metrics to mean quantities and averaging them across metrics, months and stations with contributions weighted by sample size. The resulting non-dimensional performance scores range from 0 (uncorrelated) to 1 (perfect correlation) and relate to the fraction of the mean that is not impacted by error. So a performance score of 0.8 can be interpreted that error levels are within approximately (1-0.8)\*100 = 20% of the attribute means.

Appendix 2: Wavedat Format Description

Data processed by IMEDS are stored in special structure called *wavedat*. The various preprocessors modify the given raw format (e.g.: WW3 or NDBC) into this wavedat format. A summary of the required *wavedat* input fields follows:

* file: (string) Processing History
* time: [1x m double] Observation Date and Time
* espt: {1x m cell} Directional Wave Spectra (m2/(hz x deg)
* dwfhz: [1x64 double] Frequency Bins (Hz)
* dwdeg: [1x24 double] Angle Bins (deg)
* dwAvv: {1x m cell} 1D Energy-Frequency Spectga
* hs: [1x m double] Significant Wave Height (m)
* fp: [1x m double] Peak Frequency (Hz)
* thetap: [1x m double] Peak Direction (Hz)
* winddir: [1x m double] Wind Direction (Deg From True North)
* windspeed: [1x m double] Wind Speed (m/s)
* name: (string) Data Set Name
* size: [1x1] Number of Records
* type: '2D' or 1D, Specifies Directional or Non-Directional Data
* lat: [1x1] Latitude (decimal degrees +N/-S)
* lon: [1x1] Longitude (decimal degrees +E/-W)
* depth: [1x1] Water Depth (m)

Partitioning the data stored in wavedat (automatically done by IMEDS while running) adds an ‘out’ field. The fields in wavedat.out contain information on the individual wave components (spectral partitions):

* time: [1x m double] Date and Time
* rms: [1x m double] RMS Wave Height (m)
* azimuth: [1x m double] Mean Direction (Deg from True North)
* freq: [1x m double] Peak Frequency (Hz)
* grp: [1x m double] Wave System Number
* sys: [1x m double] Swell Event Number
* totnrg: [1x m double] Total Energy (m2)
* sighight: [1x m double] Significant Wave Height (m)
* dirspread: [1x m double] Directional Spread
* par: [1x m double] Partition Number
* wforce: [1x m double] Wave Force Ratio

# Appendix 3: IMEDS Generic Format Description

### Description

The file follows the following format:

% IMEDS generic format version 1.0 - data types list

% year month day hour min sec dataType (units)

% Source Timezone datum

Station\_1\_ID Station\_1\_lat Station\_1\_lon

Year\_1 Month\_1 Day\_1 Hour\_1 Min\_1 dataType\_1 ...

Year\_i Month\_i Day\_i Hour\_i Min\_i dataType\_i ...

Year\_n Month\_n Day\_n Hour\_n Min\_n dataType \_n

Station\_i\_ID Station\_i\_lat Station\_i\_lon

Year\_1 Month\_1 Day\_1 Hour\_1 Min\_1 dataType\_1 ...

Year\_i Month\_i Day\_i Hour\_i Min\_i dataType\_i ...

Year\_n Month\_n Day\_n Hour\_n Min\_n dataType\_n

Station\_n\_ID Station\_n\_lat Station\_n\_lon

Year\_1 Month\_1 Day\_1 Hour\_1 Min\_1 dataType\_1 ...

Year\_i Month\_i Day\_i Hour\_i Min\_i dataType\_i ...

Year\_n Month\_n Day\_n Hour\_n Min\_n dataType\_n

Where:

* The first line holds a description of the file, along with a version number and a data types list.
  + data types list : example water-level and wind - This is more for the reader to know what data type is in the file - IMEDS won't read this
  + This can also hold whatever the modeler would like to add such as their run name/number/version
* The second line holds a description of the data available in the file. Keywords are very important here as IMEDS will parse this to know what to expect. The units come in parenthesis, after each data type. The keywords are:
  + year
  + month
  + day
  + hour
  + min (optional)
  + sec (optional)
  + watlev - water-level
  + hwm - high water marks
  + hs - wave height
  + tp - wave-period
  + dir - wave-direction
  + wndSpd - wind-speed
  + wndDir - wind-dir
* The units abbreviations are the usual convention , such as 'm' for meters, 'ft' for feet, 'm/s' for meter/second .
  + Default units in IMEDS are metric units (MKS system).
  + Communication will have to occur between the modelers and IMEDS developer(s) for any other data-types that are not yet available for processing in IMEDS
* source: The source/model of the dataset, such as 'ADCIRC', 'NOS', ....
* Timezone: The time-zone, such as UTC, EST, etc
* datum: Optional - The datum such as NAVD88... For now IMEDS doesn't switch datum automatically but the datum information is hold in our internal structure...

With getting our observed data from the NOS SOAP services we can also gather datum information and could potentially have the datum switched on the fly - the problem with that is that not all datum information is available for all stations

* Station\_i\_ID: The station ID, such as the NOS ID (8635753)
* Station\_i\_lat : The station latitude
* Station\_i\_lon : The station longitude
* Year\_i: The year for record i
* Month\_i: The month for record i
* Day\_i: The day for record i
* Hour\_i: The hour for record i
* Min\_i: The minutes for record i
* dataType\_i: The data for record I, non available values have to be marked as -999

### Example

% IMEDS generic format version 1.0 - water-elevation

% year month day hour min watlev (m)

NOS UTC NAVD

8635750 37.9950 76.4650

2009 11 9 0 0 0.2440

2009 11 9 1 0 0.1580

2009 11 9 2 0 0.0530

2009 11 9 3 0 -0.0310

2009 11 9 4 0 -0.1000

2009 11 9 5 0 -0.1460

2009 11 9 6 0 -0.1520

2009 11 9 7 0 -0.0900

2009 11 9 8 0 0.0140

2009 11 9 9 0 0.1000

8636580 37.6150 76.2900

2009 11 9 0 0 -0.0330

2009 11 9 1 0 -0.1110

2009 11 9 2 0 -0.1630

2009 11 9 3 0 -0.1950

2009 11 9 4 0 -0.1770

2009 11 9 5 0 -0.1210

2009 11 9 6 0 -0.0230

2009 11 9 7 0 0.0710

2009 11 9 8 0 0.1050

2009 11 9 9 0 0.0910

8638610 36.9470 76.3300

2009 11 9 0 0 -0.3470

2009 11 9 1 0 -0.3630

2009 11 9 2 0 -0.3110

2009 11 9 3 0 -0.1880

2009 11 9 4 0 -0.0180

2009 11 9 5 0 0.1720

2009 11 9 6 0 0.2770

2009 11 9 7 0 0.2940

2009 11 9 8 0 0.2290

2009 11 9 9 0 0.1030

8639348 36.7780 76.3020

2009 11 9 0 0 -0.3740

2009 11 9 1 0 -0.4120

2009 11 9 2 0 -0.3480

2009 11 9 3 0 -0.2420

2009 11 9 4 0 -0.0720

2009 11 9 5 0 0.1540

2009 11 9 6 0 0.3360

2009 11 9 7 0 0.3810

2009 11 9 8 0 0.2940

2009 11 9 9 0 0.1330Reference

Hanson, J.L., B. Tracy, H. Tolman and R. Scott, 2009. Pacific hindcast performance of three numerical wave models, *J. Atmos. Oceanic Technol*., **26**, pp. 1614-1633.