Initialization and Setup of the Coastal Model Test Bed: STWAVE

by A. Spicer Bak, Tyler Hesser, Jane Smith, and Mary Bryant

PURPOSE: The purpose of this Coastal and Hydraulics Engineering Technical Note (CHETN) is to introduce the Coastal Model Test Bed (CMTB) and the application of STWAVE. Numerical models are used by engineers, scientists, and coastal planners in U.S. Army Corps of Engineers Districts, U.S. Army Engineer Research and Development Center (ERDC), academia, and other public and private agencies to simulate complex scenarios at a reasonable cost. The CMTB has been initiated to evaluate and understand the strengths and weaknesses of numerical hydrodynamic and morphologic models using high-resolution temporal and spatial measurements at the Coastal and Hydraulics Laboratory (CHL) Field Research Facility (FRF) in Duck, NC. The improved evaluation methodology will promote rapid enhancement of model capability and focus research efforts in the coastal environment. The CMTB concept and initial setup methodology are described in this document.

INTRODUCTION: Coastal processes research aims to understand the physics of waves, circulation, sediment transport, and morphological change. These physics are often developed into predictive numerical models, which help scientists and engineers understand complex natural systems and solve real-world problems. The CHL FRF has focused efforts on field data collection in and around the nearshore for 35 years. This extensive dataset has helped advance the physics-based understanding of coastal processes by providing critical observations in the surf-zone that have been used to develop and assess a wide range of coastal numerical models. Pertinent data types, including waves, water levels, nearshore currents, bathymetry, and meteorological measurements, are collected in high spatial and temporal resolution. The measurements at the FRF provide a unique coastal dataset on an open coast that spans a broad range of meteorological, oceanographic, and morphologic conditions, resulting in an optimal location to establish the CMTB.

The CMTB provides a plug-and-play Testing and Evaluation platform for numerical models that are used by the districts and ERDC. The CMTB is executed by operating nearshore hydrodynamic and morphologic models in a real-time, data-rich environment, such as the FRF, with dedicated computational resources. By running models in a data-rich environment, high-resolution measurements are available for model verification, restricting the user-required adjustment of coefficients and parameters (tuning knobs). Running the models in real-time not only allows for greater user feedback in smaller temporal segments but also lends itself to evaluation of various data assimilation techniques. The CMTB includes dedicated high-performance computational resources, which allow for new spin-off simulations, such as testing alternative scenarios with the same input data (e.g., running the same simulations at different grid resolutions or friction coefficients, spatially variable wind models). The CMTB is developed using Python 2.7, an open-source language. Developing in an open-source architecture allows for the framework to exist in multiple locations without requiring licenses, thus lowering costs. The
ability to run models in near real-time, and the analysis that follows, improves quantification of both the errors associated with model results and the uncertainties associated with each model’s assumptions. The first model incorporated into the CMTB is the nearshore wave transformation and generation model STWAVE (Massey et al. 2011).

This document begins with a brief description of the FRF available data, including the cross-shore array of wave gauges and the historical bathymetric record, allowing the reader to understand the unique position of the FRF. Next, the STWAVE model is briefly explained. Finally, the setup of STWAVE into the CMTB framework is described in detail, including specifics on the processes developed for incorporation into the CMTB.

FIELD RESEARCH FACILITY (FRF) DATA COLLECTION: The FRF is a coastal field data collection center for ERDC-CHL, making it an ideal place to set up the CMTB. Data collection ranges from various oceanographic, bathymetric, and meteorological data. Data types include wind, barometric pressure, infrasound, Argus imagery (Holman and Stanley 2007), lidar (Brodie et al. 2015), wave (Hanson et al. 2009), ocean currents, and water level data measured in various locations in and approaching the surf zone. Monthly bathymetric surveys are collected by Lighter, Amphibious Resupply, Cargo (LARC) or Coastal Research Amphibious Buggy (CRAB) (Birkemeier and Mason 1984) to accompany these data.

As part of the FRF Data Integration Framework effort, all of the locally collected meteorological and oceanographic data are loaded onto a publically accessible Thematic Real-time Environmental Distributed Data Services (THREDDS) server. Geospatial survey data are stored on a separate ARCServer. All data are also accessible through a data portal, accessible through the FRF website ([www.frf.usace.army.mil](http://www.frf.usace.army.mil)). The CMTB algorithms pull data from these servers and use these data to both prepare the forcing files as well as the validation data.

Cross-shore Array at the FRF. The FRF cross-shore array of wave gauges (Hanson et al. 2009) is a series of wave gauges starting at the continental shelf and continuing to the swash zone. These wave measurements are collected utilizing various types of established technologies, including surface buoys, acoustic profilers, and pressure sensors. Waves are measured over a 34 min record (sample frequency defined in Table 2), and statistics ($H_m$, $T_p$, $T_m$, directions, etc.) are reported hourly. The two Datawell Directional Waverider buoys (17 m and 26 m sites, NDBC 44056 and 44100) report at 30 min intervals. The cross-shore array components are described in Table 1 and locations in modeling domain are shown in Figure 1.

<table>
<thead>
<tr>
<th>Gauge Type</th>
<th>Nominal Depth</th>
<th>Measurement Type</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Sample Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBC 44014</td>
<td>47.6 m</td>
<td>ARES 4.4 payload</td>
<td>36.611</td>
<td>-74.842</td>
<td>1.71 Hz</td>
</tr>
<tr>
<td>Waverider</td>
<td>26 m</td>
<td>Directional wave</td>
<td>36.25867</td>
<td>-75.59217</td>
<td>1.28 Hz</td>
</tr>
<tr>
<td>Waverider</td>
<td>17 m</td>
<td>Directional wave</td>
<td>36.20017</td>
<td>-75.71533</td>
<td>1.28 Hz</td>
</tr>
<tr>
<td>AWAC</td>
<td>11 m</td>
<td>Directional wave and current</td>
<td>36.18961</td>
<td>-75.73940</td>
<td>2 Hz</td>
</tr>
<tr>
<td>AWAC *</td>
<td>8 m</td>
<td>Directional wave and current</td>
<td>36.18818</td>
<td>-75.74323</td>
<td>2 Hz</td>
</tr>
<tr>
<td>AWAC</td>
<td>6 m</td>
<td>Directional wave and current</td>
<td>36.18733</td>
<td>-75.74654</td>
<td>2 Hz</td>
</tr>
<tr>
<td>AWAC</td>
<td>4.5 m</td>
<td>Directional wave and current</td>
<td>36.18677</td>
<td>-75.74871</td>
<td>2 Hz</td>
</tr>
<tr>
<td>Gauge Type</td>
<td>Nominal Depth</td>
<td>Measurement Type</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Sample Frequency</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------</td>
<td>------------------------------------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Aquadopp</td>
<td>3.5 m</td>
<td>Directional wave, current, and bottom elevation</td>
<td>36.18670</td>
<td>-75.74898</td>
<td>2 Hz</td>
</tr>
<tr>
<td>Paros pressure</td>
<td>2.8 m **</td>
<td>Point wave and bottom elevation</td>
<td>36.18621</td>
<td>-75.75082</td>
<td>2 Hz</td>
</tr>
<tr>
<td>Paros pressure</td>
<td>2.1 m **</td>
<td>Point wave and bottom elevation</td>
<td>36.18607</td>
<td>-75.75135</td>
<td>2 Hz</td>
</tr>
<tr>
<td>Paros pressure</td>
<td>1.9 m **</td>
<td>Point wave</td>
<td>36.18599</td>
<td>-75.75162</td>
<td>2 Hz</td>
</tr>
<tr>
<td>Paros pressure</td>
<td>0.75 m **</td>
<td>Point wave</td>
<td>36.18593</td>
<td>-75.75187</td>
<td>2 Hz</td>
</tr>
</tbody>
</table>

* to be deployed in 2016; ** the seafloor depth varies greatly with the nearshore pressure gauges.

Figure 1. The regional parent STWAVE simulation domain is shown with the cross-shore array shown (black dots) with the nested simulation area (boxed area nearshore) and the nested simulation forcing (red triangles) with regional background bathymetry rendered in the Surface-water Modeling System (SMS). The nearshore portion of the cross-shore array is shown in greater detail in Figure 3.
In the CMTB, the cross-shore array gauges are used as primary evaluation sites. In the nearshore, the cross-shore array is accompanied by a tower-mounted lidar system above the dune that is capable of measuring both waves and beach topography. Using lidar, data measuring the sea surface, nearshore, remotely sensed wave gauges (Brodie et al. 2015) are created and co-located with those of the nearshore pressure gauges. These “synthetic” gauges are used as a secondary evaluation.

Two offshore buoys, Datawell Waveriders in depths of 26 m and 17 m, calculate directional waves using an accelerometer package to track the water surface and calculate the wave parameters. Farther inshore of these buoys, the cross-shore array utilizes Nortek Acoustic Wave and Current (AWAC) instruments. These devices use calculated phase shifts between acoustic backscatter signals to calculate wave parameters and currents. In a nominal 3.5 m depth of water, a Nortek Aquadopp is used. These are acoustic instruments similar to the AWACs, though smaller and less intrusive in shallow water. Starting at the location of the Aquadopp and moving inshore, each gauge is accompanied by an altimeter. The altimeters are acoustic instruments that provide backscatter values to “locate” the seafloor, allowing depth measurements at each gauge. In the shallowest regions, wave parameters are measured using buried Paroscientific pressure sensors (Raubenheimer et al. 1998), which account for signal attenuation due to water depth and burial. These gauges are placed at FRF cross-shore coordinates of 100, 125, 150, and 200 m, which approximately correspond to swash zone across the sand bar to approximately 2.75 m of depth.

**Bathymetry.** One of the most important aspects of nearshore wave modeling is utilizing accurate bathymetry. The FRF has a long, continual record of bathymetric surveys. The record, started in 1980, has a total of almost 900 bathymetric surveys collected to date. A historical snapshot of the available bathymetric data is shown in Figure 2. The data are collected with the LARC and/or CRAB and processed into transect data (Birkemeier and Mason 1984). These data are then translated into a local gridded data product.

One of the main advantages of establishing the CMTB at the FRF is the availability of high-quality survey data in the nearshore. With these extensive bathymetric surveys, the model bathymetry is updated automatically to match the simulation timeframe, whether it is a historical or live simulation. This is one of the unique features available by using the CMTB.
INITIAL SETUP METHODOLOGY: As previously mentioned, the first model tested in the CMTB is STWAVE. An initial modular structure was developed to allow for plug-and-play capability and the addition of other wave models in the future. While each new model has different setup (input/output), the modular design allows for multiple instances of the same model to evaluate different model options or inputs (under the same forcing conditions) and the addition of new models (e.g., wave, circulation, or morphological models) into the CMTB.

Model Description. The STWAVE model provides a simple-to-use, robust numerical model for simulating nearshore wind-wave growth, propagation, and transformation (Smith 2001; Smith 2007). STWAVE is a phase-averaged, spectral nearshore wave model that incorporates wind-wave generation, wave shoaling, refraction, and breaking (Massey et al. 2011). The model has two modes of operation: half-plane and full-plane. The half-plane mode operates with energy only propagating towards shore to save computational resources while the full-plane model allows for propagation of energy in any direction. Currently, the CMTB is running STWAVE in both half-plane and full-plane modes. (For more information on the STWAVE model, please see Massey et al. [2011]).

The STWAVE model uses an orthogonal, Cartesian grid defined in a local coordinate system, where the FRF pier is aligned with “X-direction” of the grid. All input/output data operate on this shore-perpendicular Cartesian grid convention with angles measured counterclockwise from the x-axis, with zero degrees indicating a wave moving toward the shore. This is standard in mathematical operations, while the geographical convention holds true north as zero degrees and utilizes a positive clockwise convention. This convention requires the FRF-measured wind and wave directions and wave spectra (measured in the geographical convention) to be rotated from true north to shore perpendicular for data comparison. For half-plane model input, the measured offshore propagating energy must be removed from the spectra, and the wave directions are linearly interpolated to 5° direction bands.

STWAVE has an option to create nested simulations, in which a larger domain outputs wave spectra at select locations to drive a higher-resolution nearshore model. The “nested” simulation allows for a finer resolution in the area of greatest need or interest, in this case, the nearshore area at the FRF. The nested output spectra of the larger grid serve as the input spectra for the nested simulation (Figure 1). These files are automatically named and output by the parent simulation, allowing for seamless integration into the CMTB work flow.

COASTAL MODEL TEST BED (CMTB) OPERATIONS: The development language for the CMTB is based in Python 2.7, a general-purpose, high-level programing language. The design philosophy of Python emphasizes readability while allowing compact code design. The open source nature of Python allows for the developed system to be used without the need for licenses. One of the benefits of Python is the use of the classes for data handling, as in object-oriented programing. Oftentimes in the CMTB, data are passed between functions in neatly named packages called dictionaries, similar to MATLAB structure.

Setup. STWAVE is deployed at the FRF at two grid scales, run sequentially, a regional simulation beginning at approximately 26 m depth and a nested model beginning at approximately 8 m depth. The regional, parent model is run at a lower resolution and provides transformed wave spectra to the boundary of the nested, high-resolution nearshore simulation.
Because the CMTB is comprised of a coarse parent simulation and a high-resolution nested simulation, two input bathymetries are required. The background bathymetry data for the parent simulation was obtained from the Renaissance Computing Institute (RENCI) for the Federal Emergency Management Agency Region 4 (Blanton 2008) study. This regional digital elevation model (DEM), with a cell size of 10 m, was generated from numerous datasets collected at different times. The STWAVE parent simulation’s bathymetry is interpolated from these data to a 50 m resolution using an inverse distance weighted algorithm (Figure 1). The grid spans 17,250 m in the cross-shore direction and 38,650 m in the alongshore direction. Both the nested and the regional grids are updated as the most recent bathymetry becomes available, typically at a monthly interval.

The nearshore nested simulation domain measures 1,800 m by 1,000 m in the alongshore by cross-shore, respectively, and is shown in Figure 3. It extends just seaward of the 8 m array of pressure gauges (Long and Oltman-Shay 1991) in the cross-shore and spans the length of the FRF property with approximately an additional 540 m north of the property to reduce boundary effects approaching the cross-shore array from the north. The construction of the modeled bathymetry begins by using a portion of the background DEM dataset (Blanton 2008). The most recent gridded FRF survey data, measured along the property, are then placed into the background DEM, and the edges smoothed to remove any discontinuities that might exist between the two bathymetries.
Model parameters, as well as spatially constant wind and water levels, are passed to the model via a simulation file, and boundary wave spectra are passed through the spectral file. The wind data are measured at the end of the pier using a combination of four impeller-type anemometers while the water level data come from the National Oceanic and Atmospheric Administration gauge (DUKN7) at the end of the pier.

**CMTB Process Roles.** The development of the CMTB code is designed to be flexible and modular to allow for easy plug-and-play incorporation of new models or ensemble modeling. The design of the algorithms is done with primary focus on live data; however, with a few minor alterations, the system is also capable of running historical data from the FRF. This is important flexibility, allowing for updates to the setup or processing algorithms. The development of the CMTB code has four basic roles listed below. For the remainder of the document these roles will be referred to by numbered place in the workflow.

1. **Retrieve data**
   - The relevant data are retrieved from various data servers and packaged into Python dictionaries.

2. **Translate model input data**
   - This code is responsible for processing, preparing, and translating the data from the #1 code into the format required by the individual model. Wave spectra are rotated and truncated to half-plane requirements (if required); forcing data are averaged and time matched for the required simulation time-step, and the model input files are created.

3. **Translate model output data**
   - This code translates the model output data. In the case of STWAVE, the model output are first parsed and packaged into dictionaries. The model output are then processed, rotated back to true north, and time matched to available live data. Model outputs are also sent to a THREDDS server to be stored for further, more detailed analysis.

4. **Analyze data**
   - Depending on the data output from the model, 1-1 scatter plots and time-history plots are created for each daily simulation that compare model output and live data. Statistical analyses are performed on time-paired, observation-model data to initially evaluate the model performance. More thorough analysis must be done over longer timescales after a significant amount of data are processed.

All of these processes run autonomously in the workflow detailed in Figure 4.
Retrieve Data. The *GetData* class is responsible for retrieving the data and is the backbone upon which the CMTB is built. This module is responsible for retrieving the data autonomously from the previously mentioned servers that store data collected by the FRF and packaging the data for the “translate data in” (#2) code. With given dates, the code will search through the various data servers holding relevant information and return the data in the form of Python dictionaries. A Python dictionary is a data structure that can hold variables containing data while allowing for addition or deletion of variables internal to the dictionary.

The expandable *GetData* class has five basic data retrieval mechanisms including *getbathy*, *getwavespec*, *getwind*, *getWL*, and *getcurrents*. Each of these functions returns a dictionary of available data during the requested time period. Each function is capable of retrieving data from any gauge at the FRF, with the appropriate gauge ID, and within the selected category of data.

Translate Input. The *STprepdata* holds the second role described in the CMTB Process Roles section. This process takes the data output from the *GetData* class and prepares the STWAVE input files for the parent simulation and the nested simulation. During this preparation of data, missing and questionable data are interpolated and flagged accordingly. Here, data are averaged, rotated, and written to files.

STWAVE is run on a shore-perpendicular grid with the x-axis aligned parallel to the FRF pier (angle of 71.8° measured clockwise from true north). This results in an STWAVE grid azimuth of 198.2° (measured counter-clockwise from east). The regional STWAVE simulation is forced with a two-dimensional spectra estimated from the Fourier directional coefficient measurements and the Maximum Likelihood Estimator method from the 26 m Waverider buoy. The input
spectra are rotated to align with the grid orientation and truncated for half-plane simulations (Figure 5). For a live simulation, the bathymetry is updated to the most recently collected. With all of the appropriate input files created, the simulations are run automatically.

![Input Spectra from FRF 26m Datawell Waverider Buoy at 2016-04-13 15:30:00](image)

**Figure 5.** The figure shows the measured spectrum (from 2016 April 13 at 1530 UTC) – left – being rotated and converted to the STWAVE convention – center – centered and truncated to 5° bins – right. The white dotted section is the shoreward energy, which is removed for half-plane simulations.

Time gaps and errors in the live data can occur if the gauges are not reporting or have quality control (QC) issues. In these cases, model input spectra are linearly interpolated between the previous and next-available spectra in order to keep uniformity in the simulation steps.

Available wind (10 min records) and water level (6 min records) data are vector and scalar averaged, respectively, to match the time-step of the model run (30 min records). If subsequent data are still unavailable or have a bad QC flag, these data are omitted, and model input or validation use linearly interpolated values. The time-step and the specific data are marked with a flag for any instance when interpolated values are used during analysis.

For any initialization data (wave – at boundary, wind, or water level), if there is a gap in the data greater than 6 hr, the model run is aborted, and the user is prompted to create new run times that have more available data. For the instances in which the model input data is interpolated or questionable data, a flag record is created, indicating that these values are marked for consideration during further processing.

**Translate Output and Analysis.** Once both of the simulations, parent and nested, have completed, the translation and analysis begins. The model output data, both parent and nested
simulations, are parsed and distributed by the translate output routine (#3). This routine opens and reads the text output from the STWAVE model and puts the data into Python dictionaries that are used in the analysis process (#4). Angles are rotated back to true north and model outputs are time matched to observed data. Simulated and measured parameters are compared with scatter plots and time-history plots. Examples of these are shown in the Figures 6–9 for a 31-day run in covering the month of February 2016. The first comparisons, shown in Figure 6 and Figure 7, are to the 17 m Waverider buoy.

The time histories (Figure 7) show when the model performed well while the scatter plots, shown in Figure 6, show if there is bias in the model output relative to the measurements. These plots are constructed for every location in the cross-shore array. As an example, another comparison, at the 3 m Aquadopp (reference Figure 3 for location), is shown in Figure 8 and Figure 9. The plots show that for smaller wave heights, the model data compare well to observations; however, the model is biased high for the largest wave event during the modeled period (February 2016) at the 3 m
station. This underprediction of wave heights during some wave conditions by the model is an example of a potential research question that could be addressed using the CMTB. The last portion of the script involves translating the model output and passing it to a local THREDDS server for easy data recall for more sophisticated analysis. Bias, root-mean-square error, correlation coefficients, and other statistics can be calculated between the observed data and the model output for any duration of time using the archived netCDF files stored on the THREDDS server.

Figure 8. 1-1 plots for the 3 m Aquadopp from the cross shore array.

Figure 9. Time-series plot of the 3 m Aquadopp from for the month of February 2016.

**Summary of Work Flow.** The work flow, shown in Figure 4, is fairly simple and is controlled by a shell script that is run as a scheduled job on a Centos 7 Linux machine. Currently, the STWAVE model is run over a 24 hr period starting at 0000 UTC through 2330 at 30 min increments. The time frame allows for a sizable portion of data to be processed while still maintaining a high update rate. Data from the latest 24 hr prior to simulation initiation are gathered using `retrieve data (#1)` then passed to `translate input (#2)` where model input files are created. The parent simulation is run, data are output directly to formats that are used as inputs for the nested simulation file, and the nested simulation is run. Once both models are complete, the `translate out` script (#3) selects the appropriate observation locations and returns the
appropriate data in the form of dictionaries that are passed to the analysis routines (#4). These routines output cursory comparisons between modeled and observed data before passing the data to a local THREDDS server for storage and more thorough analysis at another time, allowing for more data to be gathered to perform a more complete evaluation of STWAVE over a variety of forcing conditions.

**CONCLUSION:** The CMTB is initially tested and implemented with the nearshore, phase-averaged STWAVE model. The CMTB is built in a generic and flexible form that allows easy integration of new models. As the project proceeds, new metrics or processes will be developed and integrated into the present methodology of model evaluation. Once models are evaluated under a broad range of conditions for extended periods of time, areas for improvement will be identified, and research on both the numerical modeling processes and the underlying physical processes can be better directed.

**ADDITIONAL INFORMATION:** For additional information, contact Spicer Bak, Coastal Observation and Analysis Branch, Coastal and Hydraulics Laboratory, 1261 Duck Road, Duck, NC 27949 at 252-261-6840 x 230 or email: Spicer.Bak@usace.army.mil.

This CHETN should be cited as follows:


**REFERENCES**


NOTE: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.