Setting a New Standard for Metocean Databases
R. de Graaff, Deltares; H. Awda, D. Majumdar, M. Satya, ADMA OPCO; S. Caires, Z. Peng, E. Moerman, Deltares

Abstract
A highly accurate and detailed database of operational and extreme metocean parameters was recently developed for the offshore waters of Abu Dhabi, U.A.E. The metocean database provides base information for planning of offshore installations and operations, as well as input for design of offshore platforms, pipelines, artificial islands or other structures. The development of the metocean database was based on the latest available topographic, bathymetric and environmental information and state-of-the-art numerical models, statistical data pre- and post-processing and presentation tools. Recent marine surveys provided good quality metocean measurements, with which the applied models and the resulting database contents could be calibrated and validated.

Generally, metocean studies of an offshore region include pre-formatted output of metocean parameters, such as joint-occurrence tables and roses, at a limited number of output locations having fixed spacing. Thanks to parallel computing of numerical wave, tide and storm surge models on a high-speed Linux cluster and data storage with high-speed interrogation of NetCDF data files, it was possible to derive multi-year hindcast time-series and extreme statistics of a large set of relevant metocean parameters data at any location in a large offshore area. With a purpose-built graphical user-interface, the user can perform additional data analysis, for example assessing down-time statistics at a planned offshore loading facility, from the retrieved hindcast time-series. The new database with large, consistent and accurate datasets, will eventually result in safer and cost-optimal offshore operations and designs. The challenge for the developer of the database is to safeguard high quality of all data analyses, hydrodynamic modelling, data processing and consequently the database contents.

This paper describes the state-of-the-art shallow-water wave, water level and current modelling and statistical analyses that are behind the dataset along with the calibration and validation of the data. Also the graphical user interface, developed to retrieve and present data from the database, is presented.

Introduction
For design, installation and maintenance of offshore structures, reliable metocean information is required in various formats. These include for instance wind, wave or current climates in the form of joint occurrence tables or roses, tables of individual wave heights, persistence tables, spectral diagrams or tables with extreme value estimates. These metocean information is based on analyses of hindcast datasets and, in shallow-water cases, on offshore to nearshore wave, tide and surge modelling. Typically, the metocean data are presented for a number of agreed output locations (typically 5 – 20) and presented in a metocean report.

Operating a large offshore area, as in the case of an oil company or wind farm operator, will require huge effort and costs to have metocean data at all locations within the operating area. In those cases, a metocean database with a dedicated user interface will be preferred, as it allows for deriving consistent sets of metocean data at any location within the database area. This will obviously require some capital investment, but in the end will become highly useful and profitable.

Recently a metocean database was developed covering all oil fields operated by the Abu Dhabi Marine Operating Company (ADMA OPCO), an area of approximately 200 x 120 km² offshore Abu Dhabi (Figure 1). Instead of filling
the database with pre-processed output at fixed locations, it was decided to store multi-year hindcast time-series and extreme statistics of a large set of relevant metocean parameters at about 13,900 locations within the database area. The creation of such database requires huge computational and data storage capacity, which was feasible thanks to parallel computing of numerical wave, tide and storm surge models on a high-speed Linux cluster, and data storage with high-speed interrogation of NetCDF data files. With a purpose-built graphical user-interface, metocean output can be derived at any location in the database area by means of interpolation, and additional data analysis can be performed on various metocean parameters, such as for instance the assessment of down-time statistics at a planned offshore loading facility.

**Data collection and review**

In spite of all offshore activities since the early 1960’s, e.g. over 300 platforms and numerous subsea pipelines and cables, only limited long-term metocean data have been collected in the southern part of the Arabian Gulf. The only known long record of wind and wave measurements were in the Abu al Bukhoosh (ABK) field, consisting of 3-hourly visual observations of wind and wave parameters from 1989 until 2001, and hourly observations of wave height and period (no directions) at Zerkou Repeater Tower from 1984 until 1993. In recent years, additional metocean data have been collected near Das Island and in Zakum, Umm Shaif, SARB and Umm Lulu fields (Fugro, 2005, 2009a,b). These data provided a reliable source of data for the purpose of model validation in the shallow-waters of Abu Dhabi. The observation locations are presented in Figure 2.

In addition to the collected measurement data, environmental hindcast data from the PERGOS database were applied. The PERGOS database was developed by Oceanweather Inc., and includes hindcasts of wind, wave, water level and current parameters for a continuous period of 27 years (1983 – 2009). A total of three data blocks were used for the definition of deep-water boundary conditions and shallow-water statistics of extreme surge heights and currents.

All collected data were analysed with the objectives to (1) review the consistency and quality of the collected datasets, (2) determine correlations between datasets, (3) increase understanding of the meteo- and hydrodynamic processes and (4) to derive a sound basis for the validation and calibration of modelling results and definition of model boundary and input conditions. For analyses of the various datasets, the data analysis and processing tool ORCA (Van Os et al., 2011) was used. The analyses included the assessment of normal conditions (wind, waves and currents), average and maximum tide levels and surge heights and extreme value estimates (based on PERGOS hindcasts). The results of the analyses are presented for instance as joint occurrence tables and wind, wave and current roses of which examples are given in figures 3 and 4.

**Tide and surge modelling**

**Model setup**

Normal and extreme water level and current conditions in the metocean database were based on extensive hydrodynamic modelling using Delft3D-FLOW (Deltares, 2009). Delft3D-FLOW solves the three-dimensional shallow-water equations for given boundary conditions. The equations are solved by an implicit finite difference method (ADI) on a staggered (spherical or orthogonal curvilinear) grid. For the present application, Delft3D-FLOW was used in the one layer (depth-averaged) mode. To enable modelling at various spatial scales within one modelling environment, Delft3D-FLOW applies the domain decomposition technique. Domain decomposition is the dynamic (two-way) coupling of model grids in one simulation. This allows a model to consist of more than one grid with different resolutions and using internal coupling boundaries at the interface.

A calibrated model covering the southern part of the Arabian Gulf with a horizontal resolution of approximately 1.6km formed the basis for the water level and current modelling, see Figure 5. A series of detailed fine grid models were nested in the regional model by means of the domain decomposition technique in areas where high resolution was required, such as around shoals and islands. The detailed domains are shown in Figure 6. Around Das Island and Zerkou Island a minimum grid resolution of about 60m was used in order to properly schematise the local bed gradients in the model. The depth schematisations of all domains were based on recent bathymetric surveys and digitised nautical charts, see Figure 7.

The tidal dynamics in the model are driven by the prescription of tidal water level variations along the open sea boundary of the regional model with astronomic constituents. A set of 54 tidal constituents were defined at a total of 46 boundary segments, including the main constituents Q1, O1, K1, P1, N2, NU2, M2, L2, S2, K2 and the seasonal constituents SA and SSA. These constituents were obtained by nesting in the overall Arabian Gulf Model (Delft Hydraulics, 1994) and subsequent calibration by comparison with observations.
Spatial- and time-varying fields (based on the PERGOS data) were used as forcing in the Delft3D model domains. Surge related water levels, those resulting from meteorological variations (e.g. wind and air pressure), were imposed on the open boundary of the model as time series, separately from the tide variations, which are imposed as astronomic constituents. Following a sensitivity study, it was concluded that by forcing time series of surge heights (e.g. resulting from the harmonic analysis of the PERGOS data points at the boundary of the regional model) at the open model boundary in combination with simultaneous time series of wind speed and direction, would accurately represent the water level variation during storm events in the database area.

Calibration of the tide and surge model
The regional model was calibrated against observed water levels in the frequency domain, i.e. by comparing computed amplitudes and phases of the main astronomic constituents with observed sets. Main focus of calibrating the vertical tide was to position the M2 amphydrionic point at the correct location, see Figure 8. Error statistics were derived by analysing the computed (c) and observed (o) amplitudes (H) and phases (G) for a total of 30 stations in the regional model domain for 10 astronomic constituents having the largest amplitudes. The statistics are summarised in Table 1 and show that except for the P1 constituent all amplitudes are within about half a centimetre and the phases within 10° of the observed values.

<table>
<thead>
<tr>
<th>Tidal constituent</th>
<th>mean Hc-Ho [m]</th>
<th>mean Hc/Ho [-]</th>
<th>mean Gc-Go ['']</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>-0.001</td>
<td>1.16</td>
<td>-8.8</td>
</tr>
<tr>
<td>O1</td>
<td>-0.006</td>
<td>0.98</td>
<td>-3.5</td>
</tr>
<tr>
<td>K1</td>
<td>-0.006</td>
<td>0.99</td>
<td>-0.6</td>
</tr>
<tr>
<td>P1</td>
<td>-0.015</td>
<td>0.88</td>
<td>-7.7</td>
</tr>
<tr>
<td>N2</td>
<td>-0.001</td>
<td>1.00</td>
<td>-1.1</td>
</tr>
<tr>
<td>NU2</td>
<td>0.001</td>
<td>1.07</td>
<td>-2.9</td>
</tr>
<tr>
<td>M2</td>
<td>-0.002</td>
<td>1.01</td>
<td>2.1</td>
</tr>
<tr>
<td>L2</td>
<td>-0.001</td>
<td>1.11</td>
<td>2.0</td>
</tr>
<tr>
<td>S2</td>
<td>0.004</td>
<td>1.12</td>
<td>6.8</td>
</tr>
<tr>
<td>K2</td>
<td>0.002</td>
<td>1.20</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Table 1 Error statistics between observed (o) and computed (c) of amplitude (H) and phases (G) of the main 10 tidal constituents for 30 stations in the regional model domain for 10 astronomic constituents

Modelling of normal tide and surge conditions
Tide and surge modelling was carried out for a 10 year hindcast period (2000 - 2009) using the calibrated Delft3D-FLOW models covering the database area. The models were forced by (1) vertical water level variations at 46 segments along the open boundary of the regional model including the seasonal variation, (2) time series of surge heights obtained from the harmonic analysis of a PERGOS data point located at the boundary of the regional model and (3) time series of wind speed and directions from a PERGOS data point located at the boundary of the regional model. Validation showed that this type of model forcing resulted in reliable hindcast water levels and currents in the database area.

A total of 13,900 output locations were selected in the database area, see Figure 9. The selection of these output locations was a balanced choice between obtaining sufficient spatial detail in the output to allow interpolation of conditions between points (by the user-interface) on the one hand, and limiting the size of the output files on the other.

The tide and surge models were validated against the measurements at SARB and Umm Lulu. Figure 10 presents density scatter diagrams of water levels and currents for the complete measurement period (Aug. 2008 – Sep. 2009) for SARB and Umm Lulu. These figures show that the tide and (wind-induced) surge can be accurately simulated using the calibrated Delft3D-FLOW models covering the database area.

After completion of the hindcast tide and surge modelling, the model results were checked for consistency. Next, the computed water levels, depth-averaged current magnitudes and current directions were retrieved using automated (MATLAB®) scripts and were stored in the database.

Modelling of extreme tide and surge conditions
Extreme still water levels were derived for return periods of 1, 10, 100, 1,000 and 10,000 years. The methodology included combining the statistical distribution of the tidal high waters with the extreme statistics of the surge heights
in the database area. The latter component was derived by computing a spatial-varying surge factor field, based on computed surge heights from an extreme storm simulation, and multiplying the factors with extreme surge height statistics from a nearshore PERGOS point. The Highest and Lowest Astronomic Tide levels were based on the respectively maximum and minimum water levels in the database area that occurred in 2007. The extreme still water and tide levels were checked for consistency and stored in the database.

Extreme current statistics were derived from the 10-year hindcast time-series at each output location in the database. The extreme value analysis was based on the peak-over-threshold method, assuming an exponential distribution.

Wave modelling

Model setup
The normal and extreme wave conditions in the ADMA OPCO database area were determined by means of shallow-water wave modelling using the state-of-the-art 3rd generation phase-averaged wave model SWAN (Booij et al., 1999). The wave modelling was based on the same model schematisations as used for the water level and current modelling (Figure 6). The directional space in SWAN was defined by 36 directional bins. The spectral resolution covered a frequency range from 0.03 Hz to 2.5 Hz, which corresponds to wave periods between 0.4 and 33.3 seconds. The number of frequency bins was set to 43, resulting in a relative resolution (∆f/f) of approximately 0.1. SWAN was run in 3rd generation mode for wind input, quadruplet interactions and white-capping. The white-capping formulation of Van der Westhuysen et al. (2007) was used. Depth-induced breaking was activated with a constant breaker parameter and default coefficients. Bottom friction was enabled using the semi-empirical expression derived from the JONSWAP results for bottom friction dissipation (Hasselmann et al., 1973) and triad wave-wave interactions were also activated. For the non-stationary simulations of the normal wave conditions, the BSBT numerical scheme was used. Sensitivity runs showed that an integration time step of 1 hour could be applied.

Validation wave model
The SWAN wave model results were validated using all the available measurements in the database area. A comparison between SWAN model results and 1-hourly measurements of H_s at SARB and Umm Lulu (see locations in Figure 2) for the period August 2008 until October 2009 are presented in Figure 11. Overall results are presented in Table 2.

<table>
<thead>
<tr>
<th>Station (Figure 3)</th>
<th>H_s Bias</th>
<th>T_p Bias</th>
<th>T_m02 Bias</th>
<th>MWD Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Das Island</td>
<td>0.07 m</td>
<td>0.13 s</td>
<td>-0.29 s</td>
<td>31°</td>
</tr>
<tr>
<td>Umm Shaif</td>
<td>-0.05 m</td>
<td>-0.27 s</td>
<td>-0.47 s</td>
<td>9°</td>
</tr>
<tr>
<td>Zakum</td>
<td>0.09 m</td>
<td>0.07 s</td>
<td>-0.30 s</td>
<td>33°</td>
</tr>
<tr>
<td>SARB</td>
<td>-0.03 m</td>
<td>0.16 s</td>
<td>-1.20 s</td>
<td>26°</td>
</tr>
<tr>
<td>Umm Lulu</td>
<td>-0.00 m</td>
<td>0.15 s</td>
<td>-1.21 s</td>
<td>25°</td>
</tr>
</tbody>
</table>

Table 2: Error statistics for five locations in the SWAN wave model

The model results (summarised in Table 2) were shown to be in good agreement with measurements of H_s, T_p, and MWD. This conclusion is supported by the scatter density plots (e.g. Figure 11) and statistical analysis, which show that the scatter is small and correlations between model results and measurements are high. Measured storm peaks at different locations and different periods are successfully captured by the SWAN wave model (Figure 12). Only the T_m02 was shown to be underestimated by on-average 1s, which was related to the application of the white-capping formulation of Van der Westhuysen et al. (2007). It was therefore decided to add 1s to all computed mean zero-up-crossing wave periods (T_m02) in the database area.

Modelling of normal wave conditions
To define the normal wave conditions in the database area, the SWAN domains were run in non-stationary mode for a 10-year period (2000 – 2009). The boundary conditions consisted of hindcast wind, waves and water levels from the PERGOS data points located at the relatively deep-water northern boundary of the regional model (Figure 5). Boundary wave conditions for SWAN were prescribed by the following integral wave parameters: H_s, T_m01, MWD and angular spreading (D_angle).

Modelling of extreme wave conditions
For the extreme wave modelling, the input wave conditions at the open boundary of the regional model were based
on the results of extreme value analysis, which are based on the 27-year hindcast time series of the PERGOS data points. Extreme wave runs were made for a matrix of five return periods, for omni-directional and 12 directional sectors and for lower, point and upper extreme estimates of $H_s$. Associated extreme wind speeds, spatial varying still water levels, mean wave periods ($T_{m01}$), directional spreading ($D_{spr}$) and mean wave direction (MWD) were based on relationships found at the offshore PERGOS points, and were applied in the extreme wave runs. For the extreme wave modelling, 0.3m was added to the water level fields to take into account an additional clearance for expected sea level rise over a period of 50 years due to global warming. An example of a computed $H_s$ field for an extreme 1 in 100 year event is given in Figure 13.

**Database filling and additional parameters**

Upon completion of all wave, water level and flow modelling, the relevant parameters were retrieved for all 13,900 selected output locations from the computational result files using dedicated MATLAB® scripts. Automated quality checks were included to identify and remove erroneous data. After quality checks, the time series (normal conditions) and singular values (extreme conditions) at all output locations were imported in a set of NetCDF files. NetCDF (Network Common Data Form) is a set of interfaces for array-oriented data access and a freely-distributed collection of data access libraries to support the creation, access, and sharing of scientific data (www.unidata.ucar.edu). The total size of the NetCDF files forming the database was about 7 Gigabytes.

As part of the post-processing of the computational results, orbital wave velocities were computed from the $H_s$, $T_{m01}$ and local water depth, taking into account the distribution of wave energy over frequencies and directions assuming a Jonswap spectrum. The directional distribution was applied by means of a cosine-$m$ function with a fixed value for $m$ ($m = 2$). The transformation of the wave spectrum to the spectrum of orbital velocities was carried out for 3 water depths: 1m above the seabed, mid-depth and at the surface.

The maximal wave height $H_{\text{max}}$ is defined as the largest wave height in 1000 waves during a certain sea state ($H_{0.1\%}$). These were determined using the Composite Weibull distribution developed by Battjes and Groenendijk (2000). This distribution is preferred to the Rayleigh distribution in regions where the highest waves in the sea state are depth-limited, which is the case in the southern part of the Arabian Gulf. The Composite Weibull distribution estimates of the maximal wave height already account for eventual depth-induced wave breaking. When waves are not depth-limited, the Composite Weibull distribution estimates are the same as those from the Rayleigh distribution.

Based on an analysis of a large number of measurements, Goda (1978) has shown that the most likely wave period associated with the highest waves in a sea state is closely related to the peak wave period $T_p$. According to Goda, this wave period is 0.9 to 1.0 times $T_p$. Standard practice is to take the wave period associated with the maximum wave height $T_{H_{\text{max}}} = 1.0 T_p$.

Extreme current velocity profiles were obtained (as part of a workflow in the metocean database) from the depth-averaged currents, applying the vertical profile model of Zitman (1992). This model accounts for the effect of the wind on the vertical current profile.

**User Interface**

A purpose-built graphical user-interface (GUI) was developed in MATLAB® to select, view and export all metocean parameters stored in the database at any location in the offshore waters of Abu Dhabi. The GUI allows for (1) different modes of data selection (point, polygon, transect), (2) retrieval of data from the binary NetCDF database files, (3) execution of various post-processing scripts that are part of the ORCA toolbox (Van Os et al., 2011), (4) presentation of selected metocean output formats, (5) exporting figures (pdf, png, jpg, etc.), MS Excel summary sheets or data (e.g. ASCII, MS Excel, ArcView shape files) and (6) loading and saving of work sessions. The various communication interfaces of the GUI are schematised in Figure 14. A screenshot of the main window of the GUI is presented in Figure 15.

When selecting a location in the database and specifying what type of output is required (e.g. wind rose or joint-occurrence table of $H_s$ versus $T_p$), the GUI will retrieve the requested datasets from the three surrounding datapoints stored in the database. The datasets can consist of either 10-year hindcasts of a specific metocean parameter or singular values in case of extreme values. The retrieved datasets are then transformed to the selected location by means of triangular interpolation. Finally, the datasets are processed into the requested presentation format such as a wind rose, joint-occurrence table or a table with extreme estimates. Other types of presentation formats are maps of an extreme parameter (e.g. maximum 100 year $H_s$), 2-dimensional wave spectra, vertical current profiles and time-series of extreme storm events. All these output formats can be individually requested using the GUI. A summary of normal and extreme wind, wave, water level and current output tables can
be obtained in the form of a MS Excel summary sheet. The contents of the summary sheet have been set-up specifically for the purposes of structural design.

**Testing and validation**

During and after developing the database and user-interface, various tests were carried out to ensure that the metocean database performed properly. The tests included:

- **Speed tests**: functional tests focusing on time required for data retrieval from the database, data processing and data presentation;
- **Platform tests**: final tests to make sure the metocean database runs on any Windows computer without installation or licensing checks;
- **Accuracy tests**: tests to ensure that the data is stored correctly in the database and that the correct data is retrieved when selecting a certain location and data type.

As final check, the summed totals of the joint-occurrence tables derived from the database at SARB and Umm Lulu were compared with measured currents ($u_{ave}$) and $H_s$ and MWD. These comparisons are presented in figures 16 to 19, and show that the database provides reliable metocean datasets in the relatively shallow waters offshore Abu Dhabi.

**Discussion**

The main advantage of developing a metocean database of a large offshore area is that it provides a consistent set of metocean data that can be used by multiple users. The present setup allows the user to define numerous output formats for a wide range of purposes and at any location in the database area. Having one source of metocean data instead of various independent sources with varying quality, will increase the coherence of designs in a large offshore area, eventually resulting in safer and cost-optimal offshore operations and designs. In case significant changes to the system are made (e.g. large scale reclamations or dredging), the database will need to be re-evaluated.

The challenge for the developer of the database is to safeguard high quality of all data analyses, hydrodynamic modelling, data processing and consequently the database contents. It is impossible to manually inspect all data stored in the database. However, by using properly calibrated wave, tide and surge models and running and processing the model results in a consistent way using well structured Matlab® scripts, will minimise the possibility of errors in the database. This requires validated sets of measurements for calibration and in-depth knowledge of the (shallow-water) physics in the database area.

For the assessment of extreme wind, wave, water level and current parameters, the exponential distribution was applied. Assuming the exponential distribution instead of the more general Generalized Pareto Distribution leads to comparable return value estimates, but narrower confidence intervals. The assumption was therefore made mostly to not end up with over-conservative lower and upper bounds of the 95% confidence intervals. In specific for the mature fields of Abu Dhabi where about half of the offshore structures have passed their original design life, reassessment of the designs on the basis of accurate rather than conservative estimates of extremes are preferred by the operating companies. However, for some cases it may be found that the amplitudes of the 95% confidence intervals (derived on the basis of the exponential distribution) are underestimated. When using extreme estimates from the metocean database for the purpose of design, the design engineers should therefore evaluate whether or not to apply an additional safety margin.

**References**


Figure 1  Abu Dhabi territorial waters, oil fields and database area (shaded blue)

Figure 2  Location of observation data
Figure 3 Joint-occurrence table and rose of hourly wind speed versus direction based on observations at SARB (Aug. 2008 – Sep 2009).

Figure 4 Overview of wind (top left), wave (top right) and current (bottom left) climate roses on the basis of observation data and PERGOS data points.
Figure 5 Computational grid of the regional south Arabian Gulf model

Figure 6 Detailed domains in database area
Figure 7 Depth schematisation of the regional model and all detailed domains

Figure 8 Co-tidal chart of M2 as computed by the regional model
Figure 9  Selected output locations
Figure 10  Density scatter between model results and measurements of water levels (top), depth-averaged current magnitudes (mid) and directions (bottom) for SARB (left panel) and Umm Lulu (right panel) for the period 1 Sep. 2008 till 1 Sep. 2009

Figure 11  Density scatter between model results and measurements of $H_s$ at SARB (left) and Umm Lulu (right)
Figure 12 Comparisons between SWAN model results and measurements at Umm Lulu during a storm period in February/March 2009

Figure 13 Computed 100-year extreme $H_s$
Figure 14  Communication interfaces of Graphical user Interface

Figure 15  Screenshot of the GUI

Figure 16  Comparison between %-exceedance of $u_{meas}$ (left) and $u_{dir}$ (right) from observations and from the database for location SARB
Figure 17  Comparison between %-exceedance of $u_{\text{ave}}$ (left) and $u_{\text{dir}}$ (right) from observations and from the database for location Umm Lulu.

Figure 18  Comparison between %-exceedance of $H_s$ (left) and MWD (right) from observations and from the database for location SARB.

Figure 19  Comparison between %-exceedance of $H_s$ (left) and MWD (right) from observations and from the database for location Umm Lulu.