Environmental Modelling in Industry Study Group,
21-24 September 2015, Isaac Newton Institute, Cambridge

**Organisers:** The EPSRC *Living with Environmental Change - Maths Foresees* Network and the NERC Probability, Uncertainty and Risk in the Environment (PURE) Network.

**Problem Statement to be Presented by Fugro GEOS**

# Statistical Framework for Utilisation of Modelled Data for Tropical Cyclones

## Summary

We are increasingly asked to produce wind and wave design criteria e.g. 100-year or 10,000-year return period values for tropical cyclones in regions where such events occur relatively infrequently. We are able to model individual storms, although getting both the storm track and intensity to match historical data is not easy. Multiple model runs with small differences in the input parameters and model physics can produce a collection of plausible storm events, and a large enough data sample from which to predict extremes.

However, we would like to explore the statistical validity of this approach and to see whether there is a rigorous way of deciding how many of the multiple model runs for an individual storm can be included in the database.

## Details

As part of its work, Fugro GEOS undertakes risk analysis for offshore sites, and this includes estimating, for instance, the 1-in-100-year maximum wave height at specified locations.

In regions where tropical cyclones or hurricanes are frequent we typically use existing hindcast databases, such as those available for North-West Australia and the Gulf of Mexico. The Gulf of Mexico hindcast includes all 201 hurricanes between 1950 and 2008 and the North-West Australia hindcast models 111 tropical cyclones between 1970 and 2010. Pooling of data from neighbouring grid points is used to maximise the number of events included in the extreme value analysis. Until recently, pooling of data was based on recommendations made in API (American Petroleum Institute, 2007. API Bulletin 2INT-MET: Interim Guidance on Hurricane Conditions in the Gulf of Mexico). Now, measures of extremal dependence (S. Coles, J. E. Heffernan and J. A. Tawn (1999), Dependence measures for extreme values analyses. Extremes, 2, 339 – 365) are used to determine the geographical extent of the data pool. Recent projects in the South China Sea, which used the extremal dependence test, yielded in the order of 100 to 500 events pooled from 6 or 7 grid points with which to perform EVA.

In regions where tropical cyclones are relatively infrequent, there are no existing hindcast databases available. We have the capability to generate atmospheric models of individual cyclones and to use these models to drive local wave models and determine cyclonic wave heights at a site of interest. However, we need a methodology to produce sufficient events to enable statistically valid extreme values to be predicted. In addition to the pooling technique, where wave heights at a number of different locations are combined, we can also include multiple attempts at modelling one individual cyclone. Small differences in input parameters and model physics can change both the track and intensity of the modelled storm, any of which could be considered as a possible representation of a cyclone in the region.

Currently our method for deriving extreme wave criteria in regions where cyclones are infrequent is:

1. Gathering the cyclone data for the region in question over the last 45 years. There is good satellite data since 1970, and there may have been, say, 6—10 severe storms in the area of interest during that period.
2. For each of those storms, running simulations of it with numbers of combinations of different microphysics models, different boundary layer models, different sea surface models etc. The simulations of the storm area take external conditions from a large-scale atmospheric model, interpolated to give values on the grid size of the storm simulation. The microphysics models are different ways of representing the homogenised effects of the physics at sub-grid scale. The boundary layer models are, similarly, different ways of representing the effects of the atmospheric boundary layers. They include, for instance, models of enthalpy transfer across the air-sea interface.
3. Using each of those simulated storms to predict wave height at the specified location. This is done using a standard wind-wave model.
4. Using that set of data to estimate the 1-in-100-year maximum wave height at the location.

Further detail can be provided of each of these steps, but the particular question for the Study Group is how to choose the number and type of different simulations to carry out in step 2.

The method is, in effect, assuming that the climate conditions are stable (at least as far as estimating the 1-in-100-year maximum wave height is concerned) over the last 45 years and over the future period for which an estimate is required. The aim of using the various models in step 2 is to simulate what might have happened in a slightly different realisation of that particular storm - might it have been more intense, less intense, might it have taken a different track, and so on.

The ultimate aim is still to produce a single number, the estimate of the 1-in-100-year maximum wave height.

We wish to know what confidence can be placed in the results of this analysis and how that confidence level depends on:

1. Number of variations used in step 2;
2. Type of variation used in step 2;
3. Other factors within our control; and
4. Factors beyond our control.

An additional question that Fugro Geos are interested in is: *What additional considerations enter the picture when we have to estimate the 1-in-10,000-year maximum wave height instead of 1-in-100?*

**<u>Data</u>**

Data can be made available, consisting of tables showing maximum predicted wave height at a given location indexed by storm number, and which combination of submodels was used in the simulation.

Uploaded to the TGM website 10/09/15