

MarMOT experiment configuration: an illustrative example

J. C. P. Hemmings and P. G. Challenor

June 16, 2011

This supplementary material includes all input and output files for a simple set of MarMOT experiments. The experiments are designed to evaluate the misfit of 4 different simulations to synthetic observation data used in the twin experiments described in Section 3 of “Addressing the impact of environmental uncertainty in plankton model calibration with a dedicated software system: the Marine Model Optimization Testbed (MarMOT)”. The experiments are defined by 4 single-record experiment control tables in the files **exTT**, **exT0**, **exOT** and **ex00**. These are 3-site simulations for (a) the true parameter vector in the true environment, (b) the true parameter vector in a realization of the optimization environment (Environment 1 in Table 4: Cost minimization), (c) the optimal parameter vector associated with this optimization environment in the true environment and (d) the same optimal parameter vector in the optimization environment. Optimal parameter vectors are those determined using the Experiment 3 cost function weightings.

The names of the input files used in each experiment are listed in the experiment’s control table as values of variables having names of the form *tabletypef*. The files defining the input item tables are located in the directory **in** and are the same for all experiments. There is one case table for each experiment. These are in the files with names of the form **case_***. The files **outvar_*** give lists of variables selected for output.

The output file **gfan.log** together with the output table files in the directory **out** was produced by running all 4 experiments in sequence with the command

```
marmot exTT exT0 exOT ex00
```

Two output tables were produced in each GFAn experiment: the first gives daily mean scalar output, relating to the surface layer, including forcing variables,

state variables and tracer rates of change due to perturbation; the second gives misfit information at each observation point.

After giving some general information about the MarMOT Model Evaluator application, the log file `gfan.log` shows the configuration for ‘GFAn Experiment 1’ as a formatted version of the control file `exTT`. This indicates each input item to be read plus the case table, followed by the output tables to be written and any files specified for sub-selecting from the lists of output table variables. Log information follows for each input table loaded. GFAn then reports the cost function value for the experiment as the ‘objective function value’. The sequence is repeated for the remaining 3 experiments. Only the case table is changed. Item tables are not re-loaded.

The production of log file output for the input item and case tables is controlled by optional prefix characters added to the file name given for each in the experiment control file. The log file here includes a formatted version of all data loaded. The ‘+’ character preceding each domain item file names indicates that the data table should be logged in addition to the metadata. Formatted records are numbered in the first column. Long records span multiple lines each starting with the same record number. A linked version of the case table produced by the GFAn cross-referencer follows, showing the data item record numbers applied in each simulation case.

GFAn automatically generates templates for each recognized type of input file on request. The templates for the development version of MarMOT used here (MME version 1.1 alpha - 15 Jun 2011) can be found in the `templates` directory. These have file names of the form `gfan.template.tabletypef` for item tables or output tables, `gfan.template.tabletypedataf` for domain item data tables and `gfan.template.tabletypevarf` for output table variable selection files. Parameter initial value tables for the optimizer initialization are parameter-set specific and the corresponding template file names have an additional suffix indicating which parameter set item they are associated with. Full descriptions of each variable are given in the parameter set item and domain item data table templates. For descriptions of output variables refer to the templates for the output table variables files or to header comments in the output files.

Standard GFAn input and output files consist of variables in space-separated columns with one row per data record and variable names in the first row. An alternative, easily-editable form is allowed (and used here) for tables with only one record. Introduced by the word `NAME` in row 1, this form has 2 columns with variable names in column 1 and values in column 2. Text on any line following a `#` character is treated as a comment.