MarMOT 1.1 Changes with Respect to MarMOT 1.0

John Hemmings 19 Nov 2013

MarMOT 1.1 is a significantly improved product in terms of functionality, useability and numerical accuracy. In the absence of comprehensive documentation for MarMOT 1.1, the MarMOT 1.0 description and user guide (Hemmings, 2009) serves as the basic user guide. It should be read in conjunction with the following notes to get the best out of the new system.

This distribution includes the Oschlies and Garçon (1999) NPZD model that was supported in MarMOT 1.0 and the MEDUSA 1.0 model of Yool et al. (2011). Support is retained for the Met Office HadOCC model and extended to include support for data assimilation. However, the distribution of the model code is restricted and it is not available in the public release. Model interface routines for each model define external model parameter names for use by MarMOT. To relate these to internal parameters in the model codes, refer to template files gfan.template.model<n>f (produced on request using the command 'marmot' without arguments). Internal names for the NPZD model differ from those in Hemmings, 2009, Table 6.1.

1. MarMOT Model Evaluator 1.1: New Features

1.1. GFAn Application Environment Changes

1.1.1 New Input Items

model3: model parameter set item for MEDUSA
assim2: data assimilation parameter set item for Hemmings et al. (2008) chlorophyll assimilation scheme used in HadOCC.
pertseed: perturbation seed parameter set item for use with random tracer perturbations
ft2 .. ft5: additional scalar forcing items to allow different grids to be used for different input variables
fkt3 .. fkt5: additional full-depth profile forcing items to allow different grids to be used for different input variables

1.1.2. New Output Tables

outoption: options table indicating active options including defaults
outtmth: monthly mean surface & scalar variables table
outktmt: monthly mean full-depth profile table
outtyr: annual mean surface & scalar variables table
outktyr: annual mean full-depth profile table
outtassim: surface & scalar data assimilation variables table
outktassim: full-depth data assimilation profiles table
outcost: cost-table giving mean weighted misfit over all observations

1.2. Initialization Changes: New State Variables

The list of initializable variables given in Hemmings, 2009, Section 3.4 has been extended with the addition of support for the MEDUSA 1.0 model (Yool et al., 2011). The new variables are: pndia (non-diatom phytoplankton), pdiat (diatoms), zmi (microzooplankton), zme (mesozooplankton), dsi (dissolved silicon or silicic acid), dfe

(dissolved iron), rcchlndia (non-diatom C:chlorophyll ratio), rcchldiat (diatom C:chlorophyll ratio), rcnpndia (non-diatom C:N ratio), rcnpdiat (diatom C:N ratio), rcnzmi (microzooplankton C:N ratio), rcnzme (mesozooplankton C:N ratio), rsinpdiat (diatom Si:N ratio), rfen (particulate organic Fe:N ratio).

When MEDUSA is selected, the phytoplankton, zooplankton and chlorophyll variables and associated ratios are not initialized from the 'init' item table but are diagnosed from the MEDUSA state variables by aggregation. Note that the version of MEDUSA included with this distribution overrides any initialized C:N ratios.

1.3. Changes to Environmental Forcing

1.3.1. New & Renamed Forcing Variables

New forcing variables have been introduced for dust deposition and to support new MarMOT functionality for tracer perturbation and relaxation (modifying the list given in Hemmings, 2009, Section 3.5). Further variables are included for the chlorophyll data assimilation scheme and a number of spare 'auxilliary' variables are provided to allow for extraction of additional fields from NetCDF data sets. Reference tracer concentrations used for relaxation have been renamed from <tracer_name>ref to ref<tracer_name> for ease of identification.

New scalar variables:

```
dust        #  aeolian dust deposition (mg/m^2/d)
fedust      #  aeolian input of soluble iron (mmolFe/m^2/d)
rcchlda     #  surface C:Chl forcing for data assimilation scheme (gC/gChl)
chlda       #  surface chlorophyll forcing for data assimilation scheme (mg/m^3)
pgrowda     #  phytoplankton specific growth rate forcing for DA scheme (/d)
plossda     #  phytoplankton specific loss rate forcing for DA scheme (/d)
```

New full-depth variables:

```
pertdin     #  perturbation rate for tracer (transformed tracer units /d)
pertnh4     #  perturbation rate for tracer (transformed tracer units /d)
pertphy     #  perturbation rate for tracer (transformed tracer units /d)
pertpndia   #  perturbation rate for tracer (transformed tracer units /d)
pertpdiat   #  perturbation rate for tracer (transformed tracer units /d)
pertzoo     #  perturbation rate for tracer (transformed tracer units /d)
pertzmi     #  perturbation rate for tracer (transformed tracer units /d)
pertzme     #  perturbation rate for tracer (transformed tracer units /d)
pertdet     #  perturbation rate for tracer (transformed tracer units /d)
pertdic     #  perturbation rate for tracer (transformed tracer units /d)
pertalk     #  perturbation rate for tracer (transformed tracer units /d)
pertdsi     #  perturbation rate for tracer (transformed tracer units /d)
pertdfe     #  perturbation rate for tracer (transformed tracer units /d)
pertsddin   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdnh4   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdphy   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdpndia #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdpdiat #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdzoo   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdzmi   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdzme   #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsddet   #  perturbation rate s.d. for tracer (transformed tracer units /d)
```

```
pertsddic     #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsdalk     #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsddsi     #  perturbation rate s.d. for tracer (transformed tracer units /d)
pertsddfe     #  perturbation rate s.d. for tracer (transformed tracer units /d)
rlxdin        #  relaxation rate for tracer (/d)
rlxnh4        #  relaxation rate for tracer (/d)
rlxphy        #  relaxation rate for tracer (/d)
rlxpndia      #  relaxation rate for tracer (/d)
rlxpdiat      #  relaxation rate for tracer (/d)
rlxzoo        #  relaxation rate for tracer (/d)
lxzmi         #  relaxation rate for tracer (/d)
rlxzme        #  relaxation rate for tracer (/d)
rlxdet        #  relaxation rate for tracer (/d)
rlxdic        #  relaxation rate for tracer (/d)
rlxalk        #  relaxation rate for tracer (/d)
rlxdsi        #  relaxation rate for tracer (/d)
rlxdfe        #  relaxation rate for tracer (/d)
refdin        #  dissolved inorganic nitrogen (mmolN/m^3)
refnh4        #  ammonium (mmolN/m^3)
refphy        #  phytoplankton nitrogen (mmolN/m^3)
refpndia      #  non-diatom phytoplankton nitrogen (mmolN/m^3)
refpdiat      #  diatom nitrogen (mmolN/m^3)
refzoo        #  zooplankton nitrogen (mmolN/m^3)
refzmi        #  microzooplankton nitrogen (mmolN/m^3)
refzme        #  mesozooplankton nitrogen (mmolN/m^3)
refdet        #  detrital nitrogen (mmolN/m^3)
refdic        #  dissolved inorganic carbon (mmolC/m^3)
refalk        #  alkalinity (meq/m^3)
refdsi        #  dissolved silicon (silicic acid) (mmolSi/m^3)
refdfe        #  dissolved iron (mmolFe/m^3)
auxvar1       #  unprocessed auxiliary variable
auxvar2       #  unprocessed auxiliary variable
auxvar3       #  unprocessed auxiliary variable
auxvar4       #  unprocessed auxiliary variable
auxvar5       #  unprocessed auxiliary variable
```

1.3.2. Transport Processes

The advection scheme options have changed. New option codes (replacing previous codes 1-4) are:

1 = Basic Upstream Scheme (Donor Cell Scheme)
2 = Monotone Upstream Scheme for Conservative Laws (MUSCL; Van Leer, 1977; Hourdin & Armengaud, 1999)

The implementation of the MUSCL scheme was derived from NEMO 3.4 code (Madec et al., 2008). Its performance for biogeochemical tracers was evaluated by Levy et al. 2001. The diffusion scheme in MarMOT 1.1 is an implicit scheme likewise derived from NEMO 3.4. The centred difference advection schemes included in MarMOT 1.0 have been removed.

As in MarMOT 1.0, the advection schemes can handle additional ‘active’ tracer velocities due to sinking, floating or swimming. MarMOT 1.1 supports vertically varying active tracer velocities (although this feature is not used in the present model versions).

1.3.3. Perturbation-rate Forcing

Optional perturbation of biogeochemical tracers as described by Hemmings & Challenor (2012) is included. Each primary tracer may be perturbed at each time step by prescribed perturbations and/or stochastic perturbations. Derived tracers are also modified by the perturbations such that composition ratios are preserved. Perturbation rate forcing can be time- and/or depth-dependent. The feature is designed primarily to support parameterization of horizontal flux divergences of the tracers associated with un-modelled lateral advection. Perturbation rates above the upper mixed layer depth, specified by the 'mld' forcing variable, can be averaged over the boundary layer at run-time if required to ensure homogeneity.

Prescribed perturbations are introduced by specification of non-zero values for the pert<tracer> forcing variables. Stochastic perturbations are introduced as a first order auto-regressive process specified by a 24 hour auto-correlation coefficient and a standard deviation. (See Hemmings & Challenor, 2012 for details). Seeds for each tracer are specified in the 'pertseed' item table by variables seed<tracer_name>. Perturbation for a tracer is activated if perturbation rate forcing data are found or if stochastic forcing is specified by presence of a seed and forcing data for the perturbation standard deviation. The following parameters in the 'option' item table control the application of the perturbation forcing data.

perttr<tracer_name>	#	transformation type for tracer perturbation (0-2)
pertautocorr	#	autocorrelation of random tracer perturb'n rates after 24 h
pertmixopt	#	perturbation b.l. mixing option (0-2)

See file gfan.template.README (produced with other templates on request using the command 'marmot' without arguments) or GFAn log output for meaning of option codes.

1.3.4. Tracer Relaxation

Tracer relaxation rates can now be time- and depth-dependent. As an alternative to relaxation towards reference fields, individual tracers can be completely replaced by the reference fields at each time step allowing detailed analysis of process dynamics. This override function is specified by prescribing relaxation rates above an upper limit of 1000 tracer units per day. (It is indicated in the output tables by the value 9999.) At times and locations where override is indicated, the end-of-time step reference tracer concentration is used, whereas normally the start-of-time step value is used to ensure stability. Although tracer relaxation forcing is directly applicable to primary tracers only, the derived tracers are also modified such that composition ratios are preserved. Reference concentrations and relaxation rates above the upper mixed layer depth, specified by the 'mld' forcing variable, can be averaged over the boundary layer at run-time if required to ensure homogeneity.

Relaxation rates can still be set as 'option' item parameters but they can now also be provided as forcing data (i.e. in 'environ' or forcing items). The option item parameters serve as constant default values in the absence of forcing data.

It is possible to control maximum tracer relaxation rates using the perturbation rate standard deviations as described by Hemmings & Challenor (2012; Appendix B2), although the details of the scheme differ. In the scheme described in the paper, a uniform relaxation rate was applied to all relaxed tracers to ensure preservation of relationships between different tracers in the reference state. In the present scheme,

a uniform relaxation constraint factor is applied to all tracers as a multiple of their individual relaxation rates. This allows greater flexibility. However, preservation of relationships is no longer assured unless their prescribed relaxation rates are identical.

As in the original scheme, the maximum permitted magnitude for the rate of change due to relaxation is determined by the product of the relaxation control factor (rlxctrlfact) and the perturbation standard deviation (pertsd<tracername>). The relaxation constraint factor is first determined for each relaxed tracer at each depth level. It is the ratio of the maximum permitted rate of concentration change to the potential rate of change. The potential rate of change is defined here as the rate of change associated with complete relaxation in 1 time step (i.e. for a relaxation rate equal to the reciprocal of the time step, this being the maximum relaxation rate without overshoot in the absence of other tendencies). The minimum relaxation constraint factor over all tracers and depth levels is used. If relaxation rates prescribed are smaller than the reciprocal of the time step then the limit will not be reached. If greater, it may be exceeded.

The following new parameters in the 'option' item table affect the application of the relaxation forcing data.

```
rlxctrlfact     #  multiple of perturbation s.d. for limiting relaxation rate
rlxmixopt       #  relaxation reference profile b.l. mixing option (0-2)
```

[Erratum: Hemmings, 2009, Section 3.5.2. Equation 3.7 R.H.S. has incorrect sign.]

1.4. Photosynthesis Sub-model Change

The point-in-time photosynthesis model (photmodel = 2) now provides level integrated ratio of photosynthesis to $\alpha_d PAR_d$ to support model parameterizations of photo-adaptation such as that in the MEDUSA model.

1.5. Improved Time-stepping Numerics

The implementation of the time-stepping schemes has been greatly improved over that in MarMOT 1.0 (Hemmings, 2009, Section 3.7.2). Partial time steps are no longer used. Instead, tendencies for each process are calculated using the system state at the appropriate time level. Users requiring more detail should refer to the source code (modsock.f90).

1.6. Simulation Output Changes: Time of Validity & New Diagnostics

Diagnostics now have a ToV at the mid-point of the time step (or double time step in leapfrog mode), irrespective of the selected time-stepping options. Linear interpolation is used where necessary. The new rules for ToV are summarized here:

1) ToV for all tracer variables and composition ratios, including reference tracer values used as forcing data, is end of time step.
2) Concentrations are assumed to vary linearly during the time step.
3) ToV for all forcing variables specific to a data assimilation analysis is also end of time step. (These variables are only defined at the end of the assimilation time step.)
4) Other forcing variables and non-tracer diagnostics (including those like primary production that are calculated by MarMOT as a function of tracer values and other variables) are treated as constant for the time step and ToV is mid-step.

5) If leapfrog time-stepping scheme is used, the ToV for non-tracer forcing and diagnostics becomes the middle of a double time step which is the end of a standard time step and these values are treated as varying linearly between valid times for averaging purposes.

New diagnostics have been added and the output table variables have been re-arranged with respect to the list in Hemmings, 2009, Section 3.8.1. The new list of variable groups is:

Case variables
Time variables
Depth variables
Scalar forcing variables (in scalar output only)
Profile forcing variables
Primary tracers
Composition ratios
Derived tracers (including element totals)
Physical process diagnostic
Production, scavenging & sinking fluxes
Profile diagnostics for photosynthesis sub-model
Sinking flux diagnostics
Tracer tendencies due to implicit horizontal flow
Tracer tendencies due to perturbation
Tracer tendencies due to relaxation
Scalar diagnostics for photosynthesis sub-model (in scalar output only)
Surface & bottom flux diagnostics (in scalar output only)
Depth integrals of tracers (in scalar output only)
Depth integrals of production and scavenging (in scalar output only)
Depth integrals of element tendencies due to implicit horizontal flow (in scalar output only)
Depth integrals of element tendencies due to perturbation (in scalar output only)
Depth integrals of element tendencies due to relaxation (in scalar output only)

2. GFAn 1.1: New Features

The MarMOT 1.1 code includes a new version of the host Generic Function Analyzer (GFAn) program providing extended services as follows.

2.1. New Data Output Options

Double precision binary option for output tables (as an alternative to ASCII). Table toolbox binary versions are produced by prefixing output file names with ‘=’ in the control table.

Output of internal item table data to external files is now possible after the data are loaded. This is achieved by prefixing input file names with ‘!’. Use ‘!=’ to write binary files. This allows data extracted from larger files or NetCDF data sets to be saved, avoiding the need for re-extraction.

Data table content written to log output using the ‘+’ prefix are limited to a maximum of NREC_MAX records (defined in toolbox.h, currently 1000000). Use prefix ‘++’ to override this limit and log the full contents.

2.2. Processing Control Options

Options for controlling the GFAn processing sequence to suppress unnecessary steps. The following may be supplied as argument 1 on the command line.

-x to skip cross-referencing & subsequent processing
-c to skip objective function's data checks & subsequent processing
-s to skip parameter optimization & function evaluation
-o to skip parameter optimization
-f to skip function evaluation

3. Table Toolbox Utilities: New Programs

tabacc – cumulative sum
tabbin – convert to table toolbox binary format
tabbody – output data without variables header
tabbound – limits and intervals for plotting
tabcatn – concatenate tables without tabcat checks for matching variables
tabfcat – concatenate fields into a single variable
tabfilt – apply filter
tabhead – output variable names only
tabintrp – linear interpolation
tabisbin – checks if table is binary
tabrange – upper & lower bounds plus mean and standard deviation
tabrms – root mean square
tabssp – sum of squares & products
tabstr – evaluate expressions involving string functions
tabsum – sum
tabtext – convert text file to table toolbox format
tabtsv – create tab-separated variable table
tabtsvin – read tab-separated variable table (converting spaces to underscores)
tabvar – variance & covariance
tabvpos – position of variable in header list
tabxpand – expand variable into separate fields according to a delimiter

IMPORTANT NOTE: Certain toolbox programs (e.g. tabfilt, tabmerge) may fail immediately with memory segmentation faults on some systems because of default limits on the size of the stack used for automatic variables. This can be avoided by increasing the stack size or removing the limit before running the program. The C Shell limit command can be used, eg.

`limit stacksize 30m`	to set to 30 MB or
`limit stacksize unlimited`	to remove.

References

Hemmings, J. C. P., Barciela, R. M. & Bell, M. J., 2008. Ocean colour data assimilation with material conservation for improving model estimates of air-sea CO2 flux, J. Mar. Res. 66, 87-126.

Hemmings, John C.P, 2009. A marine model optimization test-bed for ecosystem model evaluation: MarMOT version 1.0 description and user guide. Southampton UK,

National Oceanography Centre Southampton, 111pp. (National Oceanography Centre Southampton Research and Consultancy Report 67)

Hemmings, J.C.P. & Challenor, P.G. 2012. Addressing the impact of environmental uncertainty in plankton model calibration with a dedicated software system: the Marine Model Optimization Testbed (MarMOT 1.1 alpha), *Geosci. Model Dev.* 5, 471-498. (doi:10.5194/gmd-5-471-2012)

Hourdin F. and A. Armengaud, 1999. The Use of Finite-Volume Methods for Atmospheric Advection of Trace Species. Part I: Test of Various Formulations in a General Circulation Model, Mon. Weather Rev., 127, 822-837.

Lévy, M., Estublier, A. & Madec, G., 2001. Choice of an advection scheme for biogeochemical models. Geophys. Res. Letters 28, 3725-3728.

Madec, G., 2008. "NEMO ocean engine". Note du Pole de modélisation de l'Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619. (Version 3.4: Gurvan Madec & the NEMO team, January 2012)

Oschlies, A. & Garçon, V., 1999, An eddy-permitting coupled physical-biological model of the North Atlantic. 1. Sensitivity to advection numerics and mixed layer physics, Global Biogeochem. Cy. 13, 135-160.

Van Leer, B., 1977. Towards the ultimate conservative difference scheme IV: a new approach to numerical convection, J. Comput. Phys. 23, 276-299.

Yool, A., Popova, E. E. & Anderson, T. R., 2011. MEDUSA 1.0: a new intermediate complexity plankton ecosystem model for the global domain. Geosci. Model Dev. 4, 381-417.