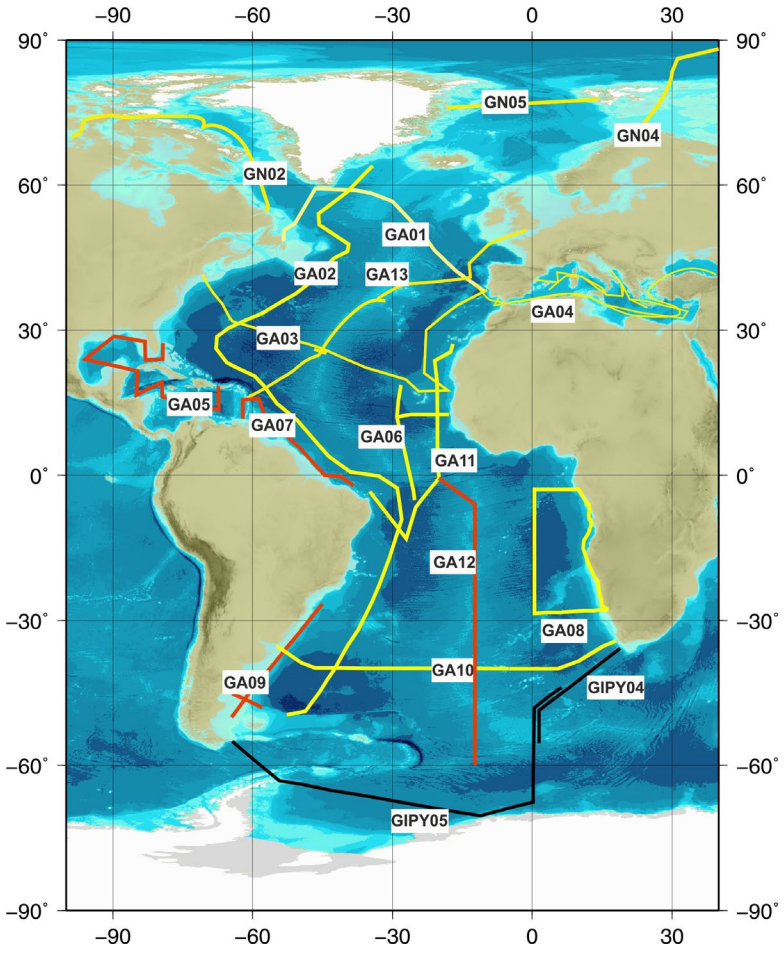
2021



Source: www.geotraces.org

FeMIP – Iron model intercomparison project- SCOR WG151

FeMIPeval

Iron model evaluation tool with GEOTRACES IDP2017

MATLAB version

Jonathan Rogerson
Marcello Vichi
University of Cape Town, South Africa

FeMIPeval – Quick Start Guide

For those who can't wait to read the documentation and have already an overall idea of the scope of this tool, here is a minimal quick start guide:

1. What do you need?

A netCDF version of the 2017 GEOTRACES Intermediate Data Product
CDO and NCO installed
Matlab, including m_map
Your model output as a netCDF (ideally CMOR-ised)

2. What is the workflow?

(A) Extract the GEOTRACES data for the section and variable of interest using *GEOTRACES_section.m* matlab script
(B) Regrid your model output using *FEMIP_regrid* shell script
(C) Extract the model data for the GEOTRACES section using *model_section.m* matlab script
(D) Construct the skill assessment using *modelplot.m* matlab script

Contents

1 INTRODUCTION

FeMIPeval consists of a MATLAB and shell library that will allow a user to quickly conduct model-observation comparison and analysis. The package is intended to compare coupled biogeochemical model outputs with observational data from the GEOTRACES Intermediate Data Product 2017 (IDP2017). The manual is divided into three parts; with the first summarizing the various univariate statistics that FeMIPeval employs while the second details the installation process as well as downloading the IDP2017 and other subsidiary packages. The final section will explain the code structure and usage of FeMIPeval with the aid of example model outputs. In comparing model-observational data, FeMIPeval uses the approach of Vichi & Masina (2009), combining qualitative plots and quantitative statistics for a thorough skill evaluation. The coding package was born out of the SCOR working group 151 Iron model intercomparison project (FeMIP) and the naming convention for some of the scripts is based on the work presented in Tagliabue *et al.* (2016).

2 STATISTICS

Coupled biogeochemical models of marine ecosystems are utilized in the scientific community for: numerical experiments, hypothesis testing and the synthesis of information (Stow *et al.* 2009). Furthermore, coupled models are used as forecasting tools, most notably being projections of change relating to rising atmospheric CO_2 concentrations. As numerical models become increasingly used in science and policy, the importance of understanding the limitations and usefulness of a numerical model becomes paramount in evaluating the outputs. Typically, comparisons between coupled biogeochemical model outputs and observational data is accomplished through visual (qualitative) means, most often in the form of data maps or section plots. However, univariate statistics may provide useful insight on model performance when the outputs bisect observational data. In assessing the accuracy or skill of a model, FeMIPeval utilizes the work of Stow *et al.* (2009) and Vichi & Masina (2009), who suggest several simple quantitative metrics for comparing model outputs with observational data:

r – the Spearman correlation coefficient

$$r = \frac{\frac{1}{N}\sum_{i=1}^{N}(O_i - \bar{O})(P_i - \bar{P})}{\sigma_O \sigma_P}$$

or the Pearson correlation coefficient if the data are not normally distributed and cannot be normalized

σ_O – the standard deviation of observations

$$\sigma_O = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - \bar{O})^2}$$

σ_P – the standard deviation of predicted values

$$\sigma_P = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - \bar{P})^2}$$

AAE – the absolute average error

$$AAE = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i|$$

RMSD – the root mean square difference

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$$

$RMSD_{CP}$ – the centered (unbiased) root mean square difference

$$RMSD_{CP} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(P_i - \bar{P}) - (O_i - \bar{O})]^2}$$

B – the bias

$$B = \frac{1}{N}\sum_{i=1}^{N}P_i - \frac{1}{N}\sum_{i=1}^{N}O_i = \bar{P} - \bar{O}$$

MEF - the Nash-Sutcliffe Model Efficiency

$$MEF = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2}$$

RI – the reliability index

$$RI = \exp\sqrt{\frac{1}{N}\sum_{i=1}^{N}(log\frac{O_i}{P_i})^2}$$

Where N = the number of observations, O_i = the ith of N observations, P_i = the ith of N predictions, and $\bar{O}$ and $\bar{P}$ are the observation and prediction averages, respectively. The above list of equations is not meant to be exhaustive; however, they were chosen because they have favorable qualities for model analysis.

The correlation coefficient (r) measures the tendency of predicted and observed values to vary together. It has a range of -1 to 1, where negative values indicate an inverse relationship and ideally, the score should be as close to 1 as possible, indicating a direct relationship between observed and predicted values. However, Taylor (2001) notes that the correlation coefficient alone is not enough to determine whether two data sets have the same amplitude of variation. Therefore, to establish the degree of difference in two data sets, the unbiased $RMSD_{CP}$ is used. Collectively, the RMSD, bias and average absolute error are all measures of the size of the discrepancies between predicted and observed values and a score near zero indicates a close match.

The $RMSD_{CP}$ differs from the RMSD because any difference in the means of the observed and predicted values are first removed. As the RMSD approaches zero, two patterns are thus more alike; however, Taylor (2001) shows that for a given $RMSD_{CP}$ value it is subsequently impossible to determine how much of the

error is attributed to a difference in structure and phase and how much is as a result of the difference in the amplitude of the variations. Thus, the $RMSD_{CP}$ and correlation coefficient are complimentary statistics and to understand where the difference in the two data sets occurs, the standard deviation for both the predicted and observational data sets can be used. The RI quantifies the factor by which model predictions differ from observations and the RI is the exponentiated RMSD and ideally should be as close to 1 as possible. Nonetheless, the RI does not distinguish whether the multiplicative factor is related to an over or underestimation; therefore, it requires the concurrent analysis of the bias. The MEF (Nash & Sutcliffe 1970) measures how effectively a model predicts relative to the mean of the observations. Ideally, the score should be close to 1, indicating a close match between predicted and observed values. Nash & Sutcliffe (1970) categorize the performance levels as follows: > 0.65 excellent, 0.65-0.5 very good, 0.5-0.2 good, < 0.2 poor and if the index is less than zero, then the model is a worse predictor than the mean of the observations.

In assessing model performance, univariate statistics can complement qualitative tests in providing a thorough skill evaluation. Furthermore, using several metrics simultaneously has two advantages: the first being different aspects of model performance can be evaluated and the second that certain statistical tests can be complementary to each other and further aid in the interpretation of the score. Consequently, FeMIPeval uses a combination of univariate statistics and qualitative plots to conduct model-observation comparisons (refer to Sec. 4).

3 DOWNLOADS

The primary function of FeMIPeval is to allow a user to quickly compare coupled biogeochemical model outputs with observational data from the GEOTRACES IDP2017. This chapter consists of three sections, with the first detailing the installation of FeMIPeval. The second describes the necessary system requirements and steps to download the various data sets and packages required by FeMIPeval while the final section gives an example directory structure for a user.

3.1 INSTALLATION

FeMIPeval is free to use and is available via GitHub for download (github). Follow the link provided to download the zip file. Once unzipped, the following MATLAB and shell scripts should be available in your FeMIPeval directory:

```
FeMIP_regrid
FeMIP_makegrid
GEOTRACES_section.m
model_section.m
modelplot.m
mll2grid.m
skillplot.m
rmsd.m
pvalues.m
relindex.m
nashsutcliffe.m
```

3.2 SYSTEM REQUIREMENTS

FeMIPeval can be run on any UNIX-based architecture and has been tested on Linux, Windows Subsystem for Linux (WSL) and Linux on a virtual machine.

The software requirements are:

- MATLAB (version 2015b and above)
- The freely-available mapping toolbox `M_Map`, which can be obtained from eoas. It must be installed separately.
- Climate Data Operators (CDO) (Schulzweida 2020)
  Install with: `sudo apt-get install -y cdo`
- netCDF Operators (NCO) (Zender 2021)
  Install with: `sudo apt-get install -y nco`

The data packages required:

- GEOTRACES IDP2017 available at: https://www.bodc.ac.uk/geotraces/data/idp2017/. Download file “Discrete Sample Data — NetCDF format (10.7 MB) IDP2017 digital dataset”. You will need to login or register before downloading the file.

When downloading and unzipping the GEOTRACES IDP2017 data, the netCDF file containing the observational data is named: “GEOTRACES_IDP2017_v2_Discrete_Sample_Data.nc” and its default

location is: GEOTRACES_IDP2017_v2/discrete_sample_data/netcdf/. This file is very large (4.2 GB), so please make sure you have enough space before unzipping. Please see "nc_info.txt" for an overview of structure and content of the IDP2017 v2 netCDF file. See "nc_variables.txt" for a list of netCDF variable names and their associated GEOTRACES variable names. This manual and the scripts use the default netCDF variable names (e.g. var73 corresponds to the GEOTRACES variable Fe_D_CONC_BOTTLE, which is the concentration of dissolved Fe in nmol/kg).

3.3 EXAMPLE DIRECTORY STRUCTURE

There is no fixed directory structure required for FeMIPeval, but the following is suggested for ease of use (Fig. 1). Note that the following directory structure is used in Sec. 4 when describing the functioning of FeMIPeval.

The user should first create a main directory in which all the data, outputs and source code will be stored. An example would be WORK, but the user can choose any name:

- FeMIPeval: containing the MATLAB and shell scripts for processing
- DATA: has the GEOTRACES IDP2017 data and the M_Map mapping toolbox (these data files can be anywhere, as long as they are included in the search path)
- OUTPUT: directory where FeMIPeval outputs will be stored
- MODEL_DATA: the original model data in netCDF format

This is the default structure unless the user specifies its own names with the optional parameters (see below).

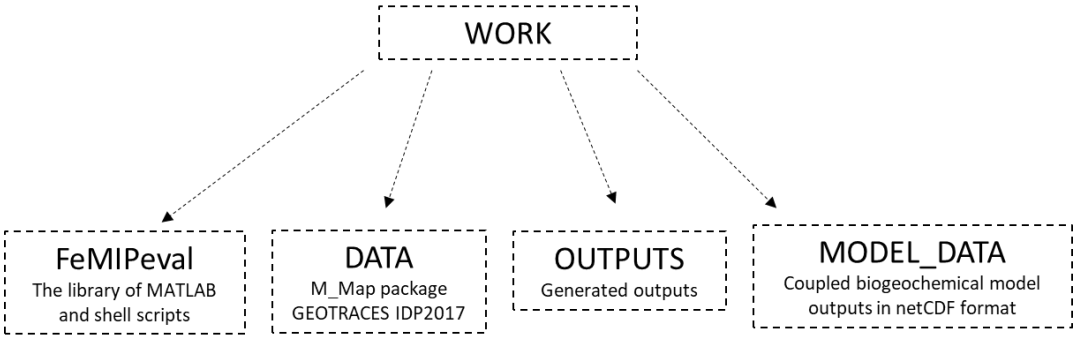


Figure 1: Flow diagram of example directory structure

4 CODE DESCRIPTION

FeMIPeval has been tested with model outputs made available by the SCOR working group participants. The following coupled biogeochemical models have been used: PISCES, GFDL, ReCOM, UKSEM and NorESM2. The script is expected to work for any ocean model output that is compliant with the Climate Model Output Rewriter (CMOR, https://cmor.llnl.gov/) developed for CMIP, although most of the netCDF written according to the CF (v1.8 or above) convention is expected to work.

This chapter is divided into two sections, the first details the functionality of the various MATLAB and shell scripts while the second showcases an example MATLAB script for running FeMIPeval.

4.1 Overview of the scripts

FeMIPeval uses a combination of MATLAB and shell scripts to conduct model-observation comparisons (refer to Sec. 3.1). The shell scripts are intended to preprocess the coupled biogeochemical model outputs, after which, the MATLAB scripts conduct the statistical analysis and figure generation of the processed model outputs with the GEOTRACES IDP2017. Fig. 2 outlines the order of operations when using FeMIPeval and in overviewing the various scripts, focus will be given to the MATLAB and shell scripts that require user inputs.

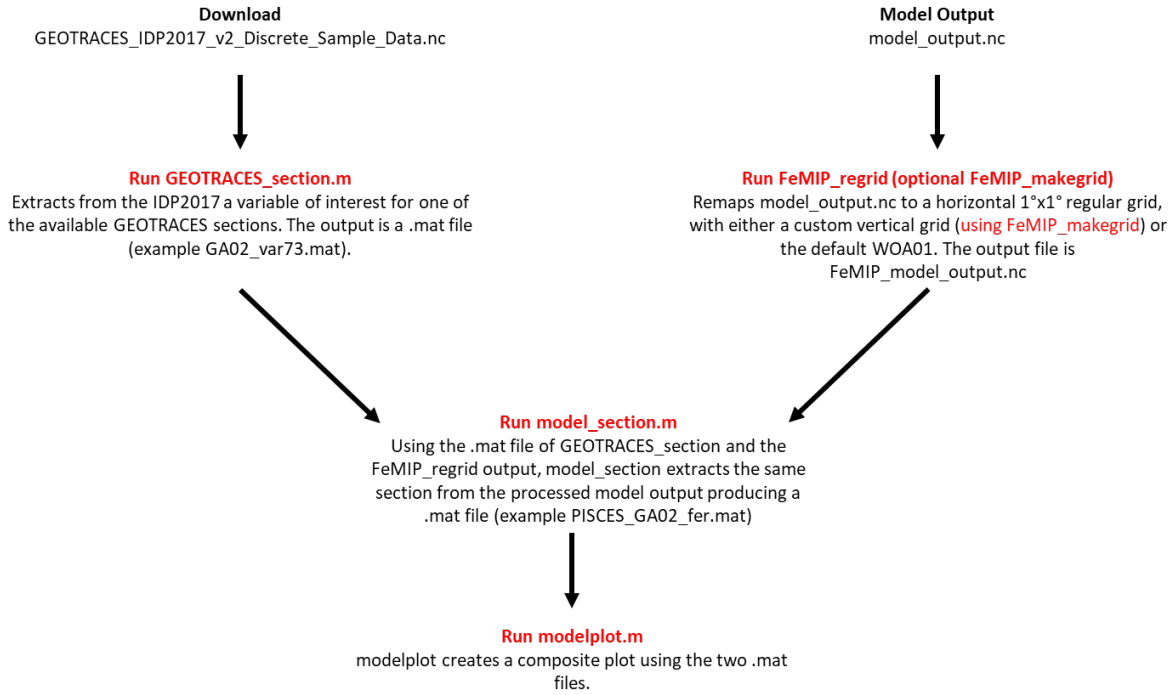


Figure 2: Flow diagram illustrating the order of operations for FeMIPeval

4.1.1 FeMIP_regrid

Biogeochemical model outputs can vary substantially in their resolution as well as grid-structure/design. Consequently, `FeMIP_regrid`'s primary function is to remap a model's coordinates onto a standard model grid. Using the methodology of Tagliabue *et al.* (2016); model outputs are regridded onto a regular 1°x1° horizontal grid with either 33 depth levels corresponding to the World Ocean Atlas 2001 (WOA01) (Conkright *et al.* 2002) as the default option or a custom vertical grid created by the user (refer to Sec. 4.1.2 on FeMIP_makegrid). In doing the horizontal regridding, the nearest-neighbor interpolation scheme is used as this was experimentally determined to be the most conservative mapping method.

Running the command: `./FeMIP_regrid` in terminal will display the user options and arguments for the script:

```
Usage: cmd [-g] [-i INDIR] [-o OUTDIR] [-p PREFIX] [-h help] FILELIST
```

Available user options:

-g: Use the custom vertical grid created with FeMIP_makegrid

-i: Specifies the INDIR, which is the directory where the model files are stored

-o: Specifies the OUTDIR, root output directory where processed model files are stored

-p: PREFIX is the name of the directory created within OUTDIR

-h: Calls a help function

If `FeMIP_regrid` is successfully run, the processed model file(s) will be located in the OUTDIR or in OUTDIR/PREFIX and prefixed with "FeMIP". Also, depending on the options used, several txt files will be in the output directory corresponding to the grid configuration used as well as subsidiary files required by the MATLAB routines.

Example 1

A user can specify the model file(s) (located in the INDIR) to be processed through the FILELIST argument. Running the command: `./FeMIP_regrid FILE.nc` is equivalent to:

```
./FeMIP_regrid -i ../MODEL_DATA -o ../OUTPUTS FILE.nc
```

By default, `FeMIP_regrid`, when called from the FeMIPeval directory, uses the directory structure from Sec. 3.3 and will regrid the model file(s) to the default WOA01 with no additional directory being created in OUTDIR to store the outputs.

Example 2

If a user wished to use a custom vertical grid (created by running `FeMIP_makegrid`, refer to Sec. 4.1.2) on the model file PISCES.nc as well as store the outputs in a new directory called PISCES within OUTDIR:

```
./FeMIP_regrid -g -i ../MODEL_DATA -o ../OUTPUTS -p PISCES PISCES.nc
```

Note, if the `-g` option is used but `FeMIP_makegrid` has not be run, this will generate an error message. For further help, a user can run the command: `./FeMIP_regrid -h` for a full description and usage of the various options.

In using `FeMIP_regrid`, it is important that the model outputs are CMOR-ized to simplify the gridding and comparison or at least the major CF (v1.8) attributes should be added to the netCDF file. For more information, refer to Climate Model Output Rewriter version 3 CMOR. In the `README.txt` file, several

useful NCO commands on editing netCDF attributes are suggested to assist in formatting your model output to make it more compliant.

4.1.2 FeMIP_makegrid

In comparing model outputs to observational data, it is necessary for both sets of data to share a similar coordinate system. The default vertical coordinate system is that of the WOA01 and this is used in the FeMIP_regrid script as well as in the MATLAB routines to process the GEOTRACES IDP2017. Consequently, FeMIP_makegrid allows a user to create a custom vertical grid onto which the coupled biogeochemical model outputs will be remapped. The two important txt files produced by `FeMIP_makegrid`, saved in OUTDIR, are: 'mygrid', which contains the grid information required by `FeMIP_regrid` to do the remapping and 'levels', which is used in the MATLAB routine GEOTRACES_section.m (refer to Sec. 4.1.3) to remap the GEOTRACES IDP2017 data section of choice to the same coordinate system. If `FeMIP_makegrid` is not run, the only txt file present in the OUTDIR will be "WOA01", containing the default grid information.

A user must manually edit `FeMIP_makegrid`. Running the script as is will generate the vertical grid of ReCOM and it will be placed in the INDIR (following the directory structure of Sec. 3.3). A user may edit the INDIR appropriately but the corresponding INDIR for `FeMIP_makegrid` must be identical to that then used in `FeMIP_regrid`, else this will generate an error.

```
INDIR="../MODEL_DATA" # Directory where model outputs are stored

shiftdepth="5

# Create the vertical grid file

# When editing the 'mygrid' file, please only edit the 'size' and 'levels'
variables

cat > mygrid << EOR

zaxistype = depth_below_sea

size = 31

units = m

levels = 0 5 15 27.5 45 65 87.5 115 147.5 182.5 220 270 340 430 535 650 775
915 1070 1245 1465 1730 2020 2340 2680 3050 3475 3950 4450 4950 5450

EOR
```

When editing `FeMIP_makegrid`, special consideration must be given to the variable 'shiftdepth'. `FeMIP_regrid` employs a linear interpolation for the vertical grid, therefore, it is important that the value of 'shiftdepth' correspond to the value of the first depth level in the model file. This is necessary, else, the first depth in the processed model file, after running `FeMIP_regrid`, will contain just NaN values. As an example, if the first depth of my desired vertical grid begins at 0 m, but my model grid starts at 5 m, you must set 'shiftdepth' to 5. `FeMIP_regrid` will shift the grid to start at 0 m and not at the depth of 'shiftdepth'.

To create the custom grid, a user need only edit the 'levels' and corresponding 'size' variable appropriately. To run `FeMIP_makegrid`, execute the following in terminal: `./makegrid`.

4.1.3 GEOTRACES_section.m

The IDP2017 is the second publicly available data product of the GEOTRACES program and contains quality-controlled data from the Atlantic, Pacific, Arctic, Southern and Indian ocean (Schlitzer *et al.* 2018). `GEOTRACES_section` extracts from the IDP2017 a variable of interest from a chosen section but only at station locations where full-depth sample profiles were taken (refer to Sec. 3.2 on downloading IDP2017). The output is a .mat file, named after the section and appended with the variable of choice.

```
GEOTRACES_section(section,variable,vgrid)
```

`GEOTRACES_section` requires three inputs:

(1) section: the cruise section of choice. For a list of available sections, the user can refer to Schlitzer *et al.* (2018) or use `help GEOTRACES_section`.
(2) variable: the desired variable to extract. For a list of variables, refer to Schlitzer *et al.* (2018), but note the choice of variables is limited, depending on the chosen section.
(3) vgrid: set to 'F' if you want the section data to be vertically interpolated onto the WOA01 grid (default). Else, setting to 'T' will use the custom grid created with FeMIP_makegrid (Sec. 4.1.2).

IMPORTANT NOTE: The full path to the folder containing the file "GEOTRACES_IDP2017_v2_Discrete_Sample_Data.nc" (refer to Sec. 3.2). needs to be added to the MATLAB search path. In the MATLAB command window, type

```
>> addpath PATH_TO_FOLDER
```

Where PATH_TO_FOLDER is the full path to the folder (e.g. ../DATA or /home/user/GEOTRACES/discrete_sample_data/netcdf)

Example 1

From the GA02 section, extract dissolved iron (var73) and interpolate onto the WOA01 vertical grid.

```
GEOTRACES_section('GA02','var73','F')
```

This will produce the .mat file GAO2_var73.mat.

The file contains the following variables.

```
Name                      Size            Bytes  Class

GA02_var73               24x55            10560  double
GA02_var73_date           1x55              461  datetime
GA02_var73_depth         24x55            10560  double
GA02_var73_interp        33x55            14520  double
GA02_var73_latitude       1x55              440  double
GA02_var73_longitude      1x55              440  double
GA02_var73_vgrid          1x33              264  double
```

where the interpolated section (on the standard 33 WOA01 levels, ending with suffix _vgrid), is found in variable GA02_var73_interp.

Example 2

From the GA03 section, extract dissolved oxygen (var19) and interpolate onto the custom vertical grid created with FeMIP_makegrid.

```
GEOTRACES_section('GA03','var19','T')
```

This will produce the .mat file GA03_var19.mat.

4.1.4 model_section.m

`model_section` is intended to be used once FeMIP_regrid (optional FeMIP_makegrid) and GEOTRACES_section have been run (Fig. 2). The script uses the .mat file produced by GEOTRACES_section and extracts, for the same variable, the same section from the processed model output. The output of model_section is a .mat with the name of the model file appended with the model variable that was processed.

```
model_section(section,file,var,modelname)
```

`model_section` has four inputs of which three are mandatory:

(1) section file: this corresponds to the .mat file produced by `GEOTRACES_section`, such as 'GA02_var73'
(2) file: full path to the processed model file in the OUTDIR
(3) var: the model's corresponding variable to that extracted from the IDP2017. Use `ncdump -h [filename].nc` to inspect the model attribute headers. The code does not check if the variables correspond, thus the user should make sure that the comparison is done with the right variable.
(4) modelname: (optional) The user can save the created .mat file with a distinct name or if left blank, the .mat file will be prefixed with 'FeMIP'. This prefix will also be used for the generated variables; hence you cannot use special Matlab characters or it will generate an error (e.g. use underscore instead of dash)

Example 1

After running GEOTRACE_section, a file name GA02_var73 is present, containing dissolved iron data from the GA02 section. In the OUTDIR, the processed model file FeMIP_01_BFM.nc is chosen. The corresponding attribute name for dissolved iron in the processed model file is 'N7f' and the output .mat file is prefixed with the name 'BFM', producing the .mat file BFM_N7f.mat.

```
model_section('GA02_var73','../OUTPUTS/FeMIP_01_BFM.nc','N7f','BFM')
```

The file contains the following variables:

```
BFM_GA02_dfe
BFM_GA02_dfe_depth
BFM_GA02_dfe_latitude
BFM_GA02_dfe_longitude
```

Example 2

A similar example to the first except that the model-name variable is not specified and thus the output will be a file named FeMIP_FER.mat.

```
model_section('GA02_var73','FeMIP_PISCES.nc','FER')
```

4.1.5 modelplot.m

In comparing model-observational data, qualitative and quantitative tests can complement in each other to give a thorough skill evaluation. modelplot uses the output of GEOTRACES_section (Sec. 4.1.3) and model_section (Sec. 4.1.4) to create a composite plot (Fig 3) consisting of:

(a) Section plot of the observational data
(b) World map showing the location of cruise stations
(c) Section plot of the model outputs
(d) Goodness-of-fit plot
(e) Distribution plot
(f) Table of univariate statistics (refer to Sec. 2)

```
modelplot(section,model_section,scale_factor,units,axis)
```

modelplot allows for five inputs, for which two are required:

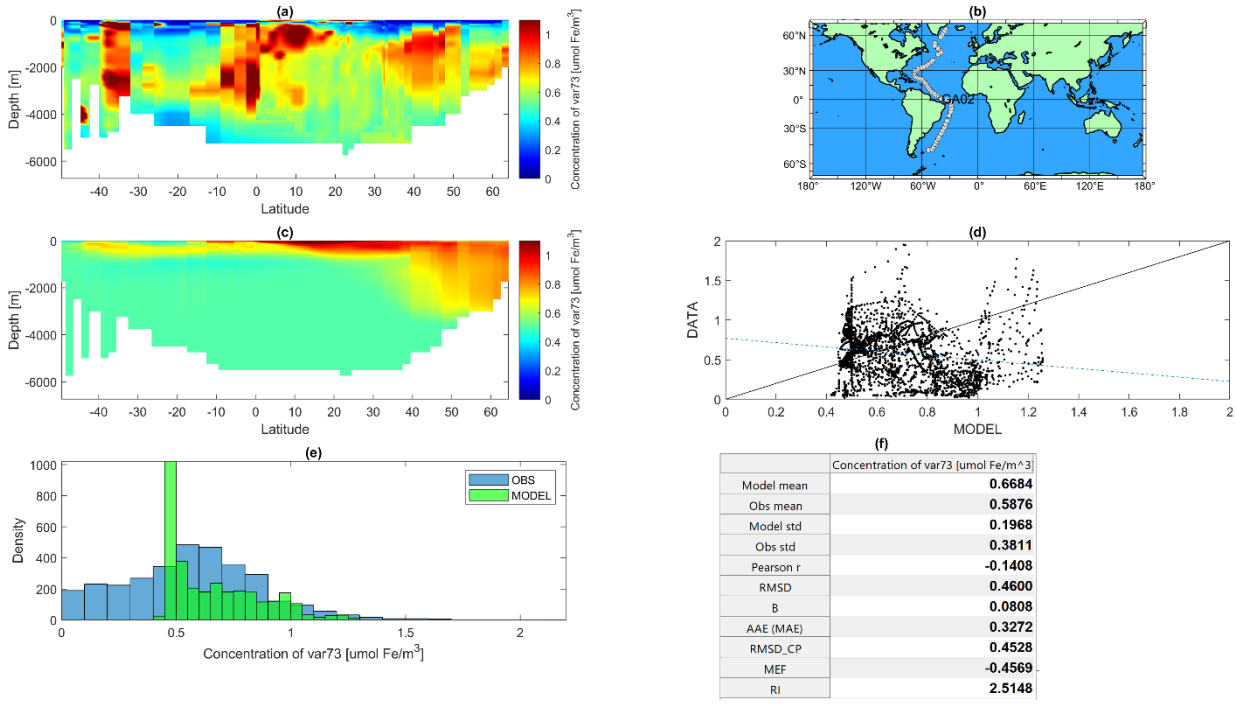
(1) section: the .mat file created with GEOTRACES_section
(2) model_section: the .mat file created by model_section
(3) scale_factor: (optional) Scale factor to apply to the model outputs to match the observational data. By default, scale is set to 1, that is same units of the IDP2017 variables (refer to Schlitzer *et al.* 2018 for the units of respective IDP2017 variables). If the model data is not scaled appropriately, this will obscure the results for the distribution plot, Goodness-of-Fit test, and the univariate statistical tests. For instance, the default dissolved iron units for CMIP data are

mol m^{-3}. The scale factor should thus be set to 1.e6, to compare with the IDP units of umol m^{-3}.

(4) units: (optional) Set the units that will be displayed on the model plots.

(5) axis: (optional) When plotting a section, modelplot by default plots along the latitudinal coordinate. Leave empty or set axis to 'lat' to plot along latitude or set to 'lon' for longitude. Note that sections in the IDP2017 can run equatorward or meridionally.

`modelplot` requires the mapping package M_Map (Sec. 3.2). Use the command `addpath` to add M_Map to your MATLAB pathways.



(f)

	Concentration of var73 [umol Fe/m^3]
Model mean	0.6684
Obs mean	0.5876
Model std	0.1968
Obs std	0.3811
Pearson r	-0.1408
RMSD	0.4600
B	0.0808
AAE (MAE)	0.3272
RMSD_CP	0.4528
MEF	-0.4569
RI	2.5148

Figure 3: (a) Section plot along GA02 from the GEOTRACES IDP2017 for var73 (dissolved iron) (b) World map showing the section. Grey dots represent locations of full-depth CTD profiles (c) Section plot of UKSEM model output for GA02 (d) Goodness-of-Fit plot for model and observational data. Solid black line is the 1:1 line while the blue dotted is the linear regression (e) Distribution plot (f) Table of statistics scores.

Example 1

Input the observational .mat file created by GEOTRACES_section and the model .mat from model_section.

```
modelplot('GA02_var73','FeMIP_GA02_FER')
```

This will create a composite plot, with the section's plotted along latitude with no scaling of the model data and no units in the plot. This option is ideally used for quick visualization.

Example 2

Similar to the first example, input the two .mat files.

```
modelplot('GA02_var73','FeMIP_GA02_FER',1000,'umol Fe/m^3','lat')
```

The code will create a composite plot (similar to Fig 3), with the section's plotted along the latitudinal axis and the model data will be scaled by a factor of 10^3 with units of 'umol Fe/m^3' shown on the axis.

4.2 EXAMPLE MATLAB SCRIPT FOR RUNNING FEMIPEVAL

This short section showcases the MATLAB script used to create Fig. 3. Using the UKSEM model, makegrid was run to remap the vertical coordinates to that of ReCOM, after which, FeMIP_regrid was executed. Once FeMIP_regrid (and FeMIP_makegrid) had been successfully run, the processed model outputs were stored in the OUTDIR. Navigating to the OUTDIR, MATLAB was launched and assuming the directory structure of Sec. 3.3, the following short script was run.

```
%% Add the required paths
addpath '../../FeMIPeval'
addpath '../../DATA/m_map'
addpath '../../DATA/GEOTRACES_IDP2017_v2/discrete_sample_data/netcdf'

%% Extract observational data for iron along GA02
GEOTRACES_section('GA02','var73','T');

%% Process the model file
model_section('GA02_var73','FeMIP_UKSEM.nc','fer');

%% Make the plot (applying the scale factor to convert from mmol to umol)
modelplot('GA02_var73','FeMIP_GA02_fer',1000,'umol Fe/m^3','lat')
```

ACKNOWLEDGEMENTS

The work of WG 151 presented in this document results, in part, from funding provided by national committees of the Scientific Committee on Oceanic Research (SCOR) and from a grant to SCOR from the U.S. National Science Foundation (OCE-1840868).

REFERENCES

Conkright, M.E., Locarnini, R.A., Garcia, H.E., O'Brien, T.D., Boyer, T.P., Stephens, C. & Antonov, J.I., 2002, *World Ocean Atlas 2001: Objective analyses, data statistics, and figures CD-ROM documentation,* National Oceanographic Data Center, Silver Spring, MD.

Nash, J.E. & Sutcliffe, J.V., 1970, 'River flow forecasting through conceptual models part I—A discussion of principles', *Journal of hydrology* 10(3), 282-290.

Schlitzer, R., Anderson, R.F., Dodas, E.M., Lohan, M., Geibert, W., Tagliabue, A., Bowie, A., Jeandel, C., Maldonado, M.T., Landing, W.M. & Cockwell, D., 2018, 'The GEOTRACES intermediate data product 2017', *Chemical Geology* 493, 210-223.

Schulzweida, U., 2020, *CDO User's Guide: Climate Data Operators Version 1.9.9*, viewed 3 January 2021, from https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf.

Stow, C.A., Jolliff, J., McGillicuddy Jr, D.J., Doney, S.C., Allen, J.I., Friedrichs, M.A., Rose, K.A. & Wallhead, P., 2009, 'Skill assessment for coupled biological/physical models of marine systems', *Journal of Marine Systems* 76(1-2), 4-15.

Tagliabue, A., Aumont, O., DeAth, R., Dunne, J.P., Dutkiewicz, S., Galbraith, E., Misumi, K., Moore, J.K., Ridgwell, A., Sherman, E. & Stock, C., 2016, 'How well do global ocean biogeochemistry models simulate dissolved iron distributions?', *Global Biogeochemical Cycles* 30(2), 149-174.

Taylor, K.E., 2001, 'Summarizing multiple aspects of model performance in a single diagram', *Journal of Geophysical Research: Atmospheres* 106(D7), 7183-7192.

Vichi, M. & Masina, S., 2009, 'Skill assessment of the PELAGOS global ocean biogeochemistry model over the period 1980–2000', *Biogeosciences*.

Zender, C. S., 2021, *netCDF Operator (NCO) User Guide*, viewed 3 January 2021, from http://nco.sf.net/nco.pdf