MARE2DEM File Format Reference

This document describes the input and output file formats used by MARE2DEM. For instructions on how to use the code, refer to the website http://mare2dem.ucsd.edu

*** This is document is a work in progress and many features are not yet described fully ***

Version: December 8, 2013

Kerry Key
Scripps Institution of Oceanography
kkey@ucsd.edu
http://marineemlab.ucsd.edu/kkey

A product of the Seafloor Electromagnetic Methods Consortium
http://marineemlab.ucsd.edu/semc.html

Contents

1 Files Required by MARE2DEM

There are five files required to run MARE2DEM:

1. Resistivity iteration file
2. Polygon file
3. Penalty file
4. Data file
5. Settings file

The first two files describe the model geometry, resistivity and fixed and free inversion parameters (as well as other inversion settings). The penalty file stores the penalty matrix. These three files are all created automatically by Mamba2D.m, so the material below isn't required reading. The data file obviously contains the data, and the settings file contains settings for the parallel data decomposition and adaptive mesh refinement used by MARE2DEM. In addition to the Fortran subroutines used to read and write these files, there are also MATLAB versions in the folder "a_utils" in the MATLAB code distribution for MARE2DEM.

2 Input Model File Formats

The model consists of a series nodes connected by segments. A sequence of segments connected in a loop define arbitrary polygonal model parameter regions.

2.1 Resistivity Iteration File (*.resistivity)

The resistivity iteration file contains pointers to all the other files (data, polygon, penalty, and settings files), a few options for the inversion algorithm, and then lists the resistivity parameters for each region. Each inversion iteration writes out a new resistivity file with that iteration's results. The name convention is <BaseName>.<iteration# >.resistivity. The starting model should be <BaseName>.0.resistivity. See the example below for a description of the details.

2.2 Polygon File (*.poly)

The polygon file defines the model parameters using nodes and segments. It is assumed that the model has been built sensibly so that no two segments intersect (except where they share a common node), that the model is free of slivers (small angle intersections between segments, less than 20°), and that each segment bound region has an identifier in the regional attributes listing.

The *.poly file uses the format defined for J. Shewchuk's mesh generation code Triangle.c. The file format is:

- First line: <# of vertices><dimension (must be 2)><# of attributes><# of boundary markers (0 or 1)>
- Following lines: <vertex #><y><z> [attributes] [boundary marker]
- One line: <# of segments><# of boundary markers (0 or 1)>
- Following lines: <segment #><endpoint><endpoint> [boundary marker]
- One line: <# of holes>
- Following lines: <hole #><y><z>
- Optional line: <# of regional attributes and/or area constraints>
- Optional following lines: <region #><y><z><attribute><maximum area>

MARE2DEM uses the vertex locations, the segment endpoints, segment boundary markers, and the region locations and numbers. Holes are ignored. For the regional attributes, the attribute value and maximum area constraints are ignored by MARE2DEM, but the (y,z) location of each regional attribute is used to identify the region number of segment bound regions, and should be located at some point interior to the region. The segment boundary markers are used to store which segments have penalty cuts.

More detailed information on the Poly file format can be found at the Triangle.c website: http://www.cs.cmu.edu/~quake/triangle.html.

2.3 Example Files

Here's an example model and resistivity file for a layer of air over two quarter spaces (so three regions). This is only a toy example since it has only three parameters and uses a few of the special add-on features of the resistivity file (bounds and prejudices).

The file Test.0.resistivity specifies the resistivity of each region and if it is a free parameter. Free parameters can also have inversion bounds and prejudices. In this example regions 1 and 3 are free parameters so they have non-zero parameter numbers. Each has a staring value of 1 ohm-m and is bound between 0.1 to 1000 ohm-m, with a prejudiced value of 10 (and weight 1). Bounds that are 0 are ignored. Prejudices are ignored when the weights are 0. Region 2 is the overlying air layer, hence it is not a free parameter and has a large resistivity.

```
Format:              mare2dem_1.0         ! input
Model File:          Test.poly            ! input
Data File:           Test.emdata          ! input
Settings File:       mare2dem.params      ! input
Penalty File:        Test.penalty         ! input (required for inversion)
Maximum Iterations:  100                  ! input (optional for inversion)
Bounds Transform:    bandpass             ! input (optional for inversion)
Global Bounds:       0.1, 10000           ! input (optional for inversion)
Debug Level:         1                    ! input (optional for inversion)
Target Misfit:       1.0                  ! input (required for inversion)
```

```
Lagrange Value:     -0.274608798018983   ! input/output  (required for inversion)
Model Roughness:    155.668697550449     ! output from inversion
Model Misfit:       1.16009504231572     ! output from inversion
Date/Time:          01/26/2012 01:05:48  ! output from inversion
Anisotropy:         isotropic
Number of regions:  3
!#      Rho  Param # Lower/Upper BoundsPrejudices & Weights
1       1        1        0.1      1000     10       1
2       1e+12    0        0        0        0        0
3       1        2        0.1      1000     10       1
```

The file Test.poly:

```
8 2 0 0
1 -100000 -100000
2 -100000 100000
3 100000 100000
4 100000 -100000
5 -100000 0
6 100000 0
7 0 0
8 0 100000
10 0
1 4 1
2 5 1
3 5 2
4 6 3
5 6 4
6 7 5
7 7 6
8 8 2
9 8 3
10 8 7
0
3
1 -70710.7 29289.3 1 -1
2 -70710.7 -29289.3 2 -1
3 -29289.3 70710.7 3 -1
```

There are several other features of the resistivity file that are not described here, and the format expands to handle various types of anisotropy MARE2DEM is capable of modeling.

2.4 Penalty File (*.penalty)

The Penalty file lists the parameters to difference for the model roughness penalty. Usually this file is generated by Mamba2D. The current format consists of a flexible compressed-sparse-row (CSR) storage that allows for any general penalty matrix. This represents the p by q penalty matrix $\mathbf{R}$, where p is the number of penalty rows and q is the number of model parameters, that is applied to the model vector with the matrix-vector product $\mathbf{Rm}$. This is automatically generated in Mamba2D.m and written using this format:

- Format: CSR_penalty_1.0
- <# non-zeros><p>
- # non-zeros following lines: <column_index><value>
- p following lines: <row_pointer>
- last line: <# non-zeros+1>

This is the usual CSR format where $value(k) = R(i, j)$ when $k = row\_pointer(i) : row\_pointer(i+1) - 1$ and $column\_index(k) = j$. The *value* parameters are used to store weights applied to the penalties, but the specific values depend on the structure of the matrix and the kind of penalty being applied. For example, a second difference penalty can be applied for parameter i by setting, for a given row, the value in the ith column to 1 and then using the value $1/n$ for the columns of all n surrounding model parameters, thereby making that row's contribution to the penalty be the difference of the ith parameter from the average of all its neighbors (which is equivalent to the numerical 2nd difference). However, the penalty output by Mamba2D approximates the integral of the model spatial gradient over each parameter using the formula

$$\|\mathbf{Rm}\|^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{N(i)} \left[w_j \left(m_i - m_j \right) \right]^2 \right) \tag{1}$$

where $N(i)$ is the number of neighboring parameters around parameter $\mathbf{m}_i$ and w_j are weights set so that the sum of differences approximates the integral of the local gradient.

For anisotropic models, the roughness is augmented by splitting the model vector into anisotropic subsets

$$\mathbf{m} = \begin{bmatrix} \mathbf{m}_x \\ \mathbf{m}_y \\ \mathbf{m}_z \end{bmatrix} \tag{2}$$

so that

$$\|\mathbf{Rm}\|^2 \equiv \|\mathbf{Rm}_x\|^2 + \|\mathbf{Rm}_y\|^2 + \|\mathbf{Rm}_z\|^2 + \alpha \|\mathbf{m} - \mathbf{m}'\|^2 \tag{3}$$

where the last term on the right is used to penalize anisotropy and can be dialed up or down with the scalar parameter α, and

$$\mathbf{m}' = \begin{bmatrix} \mathbf{m}_y \\ \mathbf{m}_z \\ \mathbf{m}_x \end{bmatrix} . \tag{4}$$

For the case of transversely isotropic models, there are only two anisotropic components instead of the three components shown above.

Note that the parameter indices correspond to the parameter numbers in the Resistivity File, not the region numbers (since each region can have fixed or free parameters, and because each region has more than one parameter for anisotropic models).

3 Settings File

Contains parameters that control the adaptive refinement computations and the parallel decompositions. See the Example files for further details.

4 Data File Format

Note (December 8, 2013): The current data format is EMData_2.2 but I don't have time to modify the description below before the code release this week, so you will have to look at the Example files to see what the new features are, for example finite length dipoles, MT static shifts options (use with caution!).

The data file lists the transmitters, receivers and frequencies, and has a table of data parameters, values and standard errors. The current data format is named EMData_2.1 (but the older format EMData_2.0 is still supported). This format is specific to the MARE2DEM inversion code. The Data file can include arbitrary comment lines (using ! or %) or blank lines. Comments can also be placed at the end of a given line using ! or % symbols before the comment text. All positions are in units of meters and angles are in degrees. Here's an example (note that "..." indicates where some lines have been omitted for brevity):

```
Format:  EMData_2.1
UTM of x,y origin (UTM zone, N, E, 2D strike):  11N  3636717.0 476297.0   20.0
Phase Convention: lag
# Transmitters:   15
!            X            Y             Z       Azimuth         Dip       Type     Name
      0.000000  1498.210000  2168.120000    90.000000    0.000000    edipole    Tx01
      0.000000  4310.020000  2094.500000    90.000000    0.000000    edipole    Tx02
      0.000000  7121.830000  2020.870000    90.000000    0.000000    edipole    Tx03
      0.000000  9933.640000  1947.250000    90.000000    0.000000    edipole    Tx04
...
# CSEM Frequencies:    3
         0.1
         0.3
         0.5
...
# CSEM Receivers:      100
!            X            Y             Z    Theta    Alpha    Beta   Name
```

```
      0.000000 -1288.660000  2184.610000      0         0         0    Rx01
      0.000000  -862.010000  2182.370000      0         0         0    Rx02
      0.000000  -435.350000  2180.130000      0         0         0    Rx03
      0.000000    -8.700000  2177.900000      0         0         0    Rx04
      0.000000   417.950000  2166.910000      0         0         0    Rx05

...
# MT Frequencies:    21
      0.0001
    0.000158
    0.000251
...
# MT Receivers:      15
!            X            Y            Z    Theta    Alpha    Beta  SolveStatic  Name
      0.000000 -1288.660000  2184.610000      0        0        0        0        MT01
      0.000000  -862.010000  2182.370000      0        0        0        0        MT02
      0.000000  -435.350000  2180.130000      0        0        0        1        MT03
...
# Data:       5596
!         Type         Freq#        Tx#          Rx#        Data    Std_Error
           3             1            1            1  6.42506e-13   7.5608e-14
           4             1            1            1  1.20096e-12   7.5608e-14
           3             2            1            1  4.86059e-14  5.08595e-14
           4             2            1            1  8.69775e-13  5.08595e-14
           3             3            1            1 -4.24273e-13  3.82199e-14
...
         103             1            1            1      29.4792      1.67651
         104             1            1            1      26.1554      3.43775
         105             1            1            1      32.9502      2.12072
         106             1            1            1      24.4939      3.43775
...
```

The file consists of *token : value(s)* blocks where the *token* is a keyword or keywords. These are followed by a single *value* or multiple lines of values. Many of the tokens can appear in any order.

4.1 UTM of x,y origin

This block is not used by the MARE2DEM code, but is there so that plotting routines can convert the local 2D coordinate system used for the data and 2D model into geographical UTM or lat/lon coordinates. You can set these values to 0 if you don't need this. In this example:

```
UTM of x,y origin (UTM zone, N, E, 2D strike):  11 N  3636717.0 476297.0   20.0
```

the UTM origin is set to Scripps Institution of Oceanography and the 2D strike is at 20 degrees (clockwise from north). This means that the local 2D modeling coordinates corresponds to geographic coordinates where x is aligned along 20 degrees and y points along 110 degrees (so the 2D model conductivity strike is at 20 degrees, the 2D model profile runs along the angle 110 degrees from geographic North). Similarly, a receiver with a given theta angle of 0 degrees in the 2D coordinate system then corresponds to an angle of 20 degrees from geographic North.

4.2 Phase Convention

The default convention for MARE2DEM is phase lag (i.e., phases become increasingly positive with source-receiver offset), but you can instead specify that the data use a phase lead convention (i.e., phases become increasingly negative with source-receiver offset):

```
Phase Convention: lag         ! Optional, use lag (default) or lead
```

Note that the phase convention is ignored by MT data, and TM mode MT data are expected to have phases wrapped to the first quadrant. If in doubt, run a forward model of a half space to see what MARE2DEM outputs for a given data type.

4.3 # Transmitters:

The Transmitter block lists the number of transmitters and then the x, y, z (meters) locations of the transmitters, and their rotation angle (degrees clockwise from x), dip angle (degrees positive down), the type of transmitter and optionally the name assigned to each transmitter (used for plotting purposes only). MARE2DEM currently only supports point electric dipoles (edipole) or point magnetic dipoles (bdipole). It has been assumed that all data responses have been normalized by the transmitter dipole moment (i.e., divided by Am or Am^2 for electric and magnetic dipoles, respectively), so each transmitter is considered a unit dipole. Also MARE2DEM works best for 2D or weakly 2.5D data, so try keeping the x coordinates of the transmitters and receivers within a few 100 meters of the origin. If there are no CSEM data (e..g, and MT only inversion), all the CSEM blocks can be omitted from the data file.

4.4 # CSEM Frequencies:

This block lists the number of CSEM frequencies on the first lines and the following lines give the specific values (Hz).

4.5 # CSEM Receivers:

This block lists the number of CSEM receivers and then the x, y, z (meters) positions of the receivers, each receiver's rotation angles and optionally the name assigned to each CSEM receiver (used for plotting purposes only). Theta corresponds to the angle from the 2D modeling coordinate x to the receiver's x channel. Alpha and Beta are dip angles. Alpha is the angle of the y channel, positive down from horizontal. Beta is the angle of the receiver's z channel from the 2D model z axis.

4.6 # MT Frequencies:

This block lists the number of MT frequencies on the first lines and the following lines give the specific values (Hz). If there are no MT data, all the MT blocks can be omitted from the data file.

4.7 # MT Receivers:

This block lists the number of MT receivers and then the x, y, z (meters) positions of the receivers, each receiver's rotation angles, a flag for solving for static shifts, and optionally the name assigned to each MT receiver (used for plotting purposes only). The angles are the same convention as used for CSEM receivers. MT receivers rotations are ignored, except for the beta parameter, which specifies the tilt of the receiver in the y-z plane. If a non-zero beta is used, the MT responses and Tippers will be computed using this angle (this could be the local topographic tilt, for instance). The angle is computed clockwise from positive y towards z. So if the receiver is on a slope that is downhill to the right (increasing y), the beta angle will be positive. Conversely, if the receiver is on a slope that is uphill to the right, the beta angle should be negative.

The SolveStatic column for the MT receivers should be normally set to 0; if this value is set to 1, then MARE2DEM will solve for the static shift for the TE and TM modes (separately) by simply using the average of the model fit residuals to log10(apparent resistivity). This requires that the input data be formatted as either apparent resistivity or its log10 equivalent. This approach to estimating the static shift works best when only a few stations are suspected of having static offsets. The resulting static shifts are shown for each iteration in the Occam log file. Synthetic tests show that this method works well for estimating static shifts, however, those same tests show that for truly non-static shifted data it can estimate shifts of up to 10-50%, so us this option only when necessary.

4.8 # Data:

The Data block lists the number of data and then a table of data parameters, the data and standard errors. The first column of the data table specifies the data type. The currently supported data types for CSEM and MT data are given in Table 1. The data type can be specified using either the string codes or the numeric codes given in Table 1. The string values are not case sensitive. Note that for each receiver, x, y, z component are in the local receiver coordinate frame defined by its given theta, alpha and beta angles.

The second column of the data table is the frequency index for each datum. This index refers to either the CSEM or MT frequencies, depending on the data type specified.

The third column is the transmitter index and the fourth column gives the receiver index. The receiver index refers to either an MT or CSEM receiver, again depending on the data type specified.For MT data the transmitter index is ignored if it is less than or equal to 0, otherwise the transmitter index is used to specify which receiver should be used for the magnetic fields of the MT response, whereas the receiver index specifies which receiver to use for the electric fields. In this way, hybrid MT stations can be modeled (i.e., magnetic fields from one location and electric fields form another location).

The fifth and sixth columns are the data and standard errors. Electric fields should be given in units of V/Am^2 and magnetic fields in units of T/Am for CSEM data from electric dipoles. CSEM data from magnetic dipoles should be given in units V/Am^3 for electric field responses and units of T/Am^2 for magnetic field responses. For MT data, the impedances are in units of ohms, apparent resistivities in units of linear ohm-m and phases in degrees. Impedance phases (for Zyx) should be moved to the first quadrant by adding 180 degrees (so they are nominally 45 degrees for a half-space, not -135 degrees).

While CSEM receivers are allowed to have arbitrary orientation, MT receivers rotations are ignored, except for the beta parameter, which specifies the tilt of the receiver in the y-z plane. If a non-zero beta is used, the MT responses and Tippers will be computed using this angle (this could be the local topographic tilt, for instance).

Finally, the use of different receiver and transmitter indices for MT responses allows for differential tilts between electric and magnetic sensors for land MT (where the magnetics are usually horizontal and the electric dipoles are slope parallel). Simply create a “receiver” for each sensor type and specify its tilts in the receiver section. The use the receiver and transmitter indices to specify which to use for electric and which for the magnetic magnetic components of the MT apparent resistivity and phase calculations.

Table 1: Data types used in format EMData_2.0

	Integer Code	String Code	Description
CSEM Data:	1	RealEx	Real Ex
	2	ImagEx	Imaginary Ex
	3	RealEy	Real Ey
	4	ImagEy	Imaginary Ey
	5	RealEz	Real Ez
	6	ImagEz	Imaginary Ez
	11	RealBx	Real Bx
	12	ImagBx	Imaginary Bx
	13	RealBy	Real By
	14	ImagBy	Imaginary By
	15	RealBz	Real Bz
	16	ImagBz	Imaginary Bz
	21	AmpEx	Amplitude Ex
	22	PhsEx	Phase Ex
	23	AmpEy	Amplitude Ey
	24	PhsEy	Phase Ey
	25	AmpEz	Amplitude Ez
	26	PhsEz	Phase Ez
	27	log10AmpEx	log10Amplitude Ex
	28	log10AmpEy	log10Amplitude Ey
	29	log10AmpEz	log10Amplitude Ez
	31	AmpBx	Amplitude Bx
	32	PhsBx	Phase Bx
	33	AmpBy	Amplitude By
	34	PhsBy	Phase By
	35	AmpBz	Amplitude Bz
	36	PhsBz	Phase Bz
	37	log10AmpBx	log10Amplitude Bx
	38	log10AmpBy	log10Amplitude By
	39	log10AmpBz	log10Amplitude Bz
	41	PEmax	Electric horizontal polarization ellipse maximum
	42	PEmin	Electric horizontal polarization ellipse minimum
	43	PBmax	Magnetic horizontal polarization ellipse maximum
	44	PBmin	Magnetic horizontal polarization ellipse minimum
MT Data:	101	RhoZxx	Apparent Resistivity Zxx, not used in 2D (reserved for future MARE3DEM)
	102	PhsZxx	Phase Zxx, not used in 2D (reserved for future MARE3DEM)
	103	RhoZxy	Apparent Resistivity Zxy , 2D TE mode
	104	PhsZxy	Phase Zxy, 2D TE mode
	105	RhoZyx	Apparent Resistivity Zyx, 2D TM mode
	106	PhsZyx	Phase Zyx, 2D TM mode
	107	RhoZyy	Apparent Resistivity Zyy, not used in 2D (reserved for future MARE3DEM)
	108	PhsZyy	Phase Zyy, not used in 2D (reserved for future MARE3DEM)
	111	RealZxx	Real Zxx, not used in 2D (reserved for future MARE3DEM)
	112	ImagZxx	Imaginary Zxx, not used in 2D (reserved for future MARE3DEM)
	113	RealZxy	Real Zxy, 2D TE mode
	114	ImagZxy	Imaginary Zxy, 2D TE mode
	115	RealZyx	Real Zyx, 2D TM mode
	116	ImagZyx	Imaginary Zyx, 2D TM mode
	117	RealZyy	Real Zyy, not used in 2D (reserved for future MARE3DEM)
	118	ImagZyy	Imaginary Zyy, not used in 2D (reserved for future MARE3DEM)
	121	log10RhoZxx	log10 Apparent Resistivity Zxx, not used in 2D (reserved for future MARE3DEM)
	123	log10RhoZxy	log10 Apparent Resistivity Zxy, 2D TE mode
	125	log10RhoZyx	log10 Apparent Resistivity Zyx, 2D TM mode

127	log10RhoZyy	log10 Apparent Resistivity Zyy, not used in 2D (reserved for future MARE3DEM)
MT Tipper:		
		$H_z = M_{zy} H_y$
133	RealMzy	Real tipper, TE mode, only for unrotated receivers
134	ImagMzy	Imaginary tipper, TE mode, only for unrotated receivers

5 Response File Format

The response file is identical to the data file, except that the data section has two more columns, one for the model response and another for the weighted residual. The format line is EMResp_2.0 instead of EMData_2.0. The responses files are written from MARE2DEM as filename.1.resp, filename.2.resp, ... for each inversion iteration.