TUTORIAL to the FuGas model and software

Vasco Manuel Nobre de Carvalho da Silva Vieira

MARETEC, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001, Lisboa, Portugal.
vasco.vieira@tecnico.ulisboa.pt

Model overview and evolution

The FuGas 2.3 is an upgrade to the published by Vieira et al (2013; 2015a,b; 2016). The air-water gas flux is estimated in single processing running the GasFluxExe_field.m. This file was used for the model testing with the Baltic field data in Vieira et al (2015a,b; 2016). The parallel processing alternative simulating the gas fluxes in the Mediterranean and North Atlantic was done by the GasFluxExe_mp.m. Those are the parent scripts controlling the calculations. Each geophysical process is estimated by its specific child script, invoked by the parent script at the appropriate time. The variables, values for the constants, calculus settings and numerical options must all be defined in advance by the user in the beginning of the parent script.

The adaptation of the FuGas to parallel processing demanded changing its structure. The calculus was vectorized, which required replacing the for-loops by matrix algebra. This turned the computational speed 12 times faster. It also enabled running the gas flux model in parallel processing using the Single Program Multiple Data (spmd) solution. However, this spmd duplicates ram memory usage, turning it massive and thus demanding partitioning the spmd into blocks so that the data no longer required may be cleared in-between. Loading massive amounts of data into the RAM and duplicating it can be a slow process, particularly in older laptops. Furthermore, if the RAM is exhausted, the use of virtual memory turns the calculus extremely slow, revoking the advantages of parallel processing. So, the user should check the available RAM and its calculus demands. It may be useful to free RAM by turning off the internet, the anti-virus, Dropbox, Skype, and other non-essential software applications. This may be done directly on the respective software or on the task manager. My Windows 10 install with all these applications turned off uses ≈0.8 Gb RAM (games and other bloatware uninstalled). My Linux Mint install uses ≈0.4 Gb RAM.

The Matlab Parallel Processing Toolbox has a strict set of rules for managing variables. The 'transparency' rule disabled the hierarchical script Child.m run inside Parent.m management strategy. However, this problem was solved by estimating geophysical processes from their respective custom matlab @myfunction.m, and named after the property being modelled as 'estimate<property>.m'. Neither the calculus part of the GasFlux files nor the @myfunction files are to be edited by the user.

The numerical options

The model simulates several geophysical processes occurring at the air-water interface. Each of these can be simulated by alternative formulations taken from the literature. The user must define them in the respective block in the beginning of the script, where it is provided the tag for the respective formulation, bibliographic reference and other useful information.

The constants

These are divided in constants for the water, air and gas properties, automatically loaded into Matlab workspace by defining them in their respective blocks in the beginning of the script. The gas constants are specific for each gas. The script already has the values for the greenhouse gas CO_2, CH_4 and N_2O. These can be chosen un-commenting the values for the desired gas while commenting the values for the remaining gases. The values for other gases can be found in the works by Sander (1999, 2015) and Sarmiento and Gruber (2013).

The variables

These are automatically loaded into Matlab workspace by defining the file name, the directory and the respective data arrays. If it is simulated an area over a period, then 'location' should be along the lines and 'time' (t) along the columns. This way, the 2D ocean surface with the (m,n) matrix of geographical coordinates (longitude, latitude) must be squeezed into a (m×n,t) for each environmental variable, a (m×n,1) column vector with the longitude, and another (m×n,1) column vector with the latitude. If using single processing the variables are edited as (example with the Baltic data):

```
Tw= SST;
etc
```

If using multiple processing, Matlab requires the variables to be co-distributed among CPUs:

```
Tw=codistributed(SST);
etc
```

Coordination between variables and numerical options

Certain k_w formulations require u_*, the wind at the sea-surface. If this is not given, a preliminary guess is estimated (i) from the u' and w' components of Eddy-Covariance data, or (ii) the drag coefficient. Then, the final u_* is estimated from the wind log-linear profile, given the wind at height z, the sea-surface roughness and atmospheric stability of the surface boundary layer. However, in specific times and/or locations where some of the required variables are lacking, thus disabling this implementation, the model keeps the previous u_* estimate from the drag-coefficient (CD) or the eddy-covariance (E-C) formulation. Other k_w formulations require u_{10}, the wind at 10m, which sometimes is not available but instead the wind (u_z) at other heights. In these cases u_{10} is estimated from the wind log-linear profile.

Estimating transfer velocities from observed gas fluxes

The observed Eddy-Covariance gas fluxes and associated transfer velocities (k_w) are estimated by the 'getKfromEC.m' using the WPL correction for fluctuations in humidity and heat (Webb et al 1980). The file has two 'Settings' sections to be edited by the user and two 'Calculus' sections not to be edited by the user.