Crossover analysis toolbox for matlab

This set of m-files were created and developed during the secondary quality control of the CARINA data set. The scripts can (hopefully) be useful for consistency control of other collections of hydrographic data. The scripts are relatively easy to follow (….) and can easily be modified. However, these are simple scripts and you will likely have to modify some of them at any rate.

The scripts are not particularly fast, and can most likely be made much faster. But at least they work.

Toste Tanhua 2009.01.07

Content in crossover_toolbox

- read_ carina.m
- load_carina.m
- xovermap.m
- xover_RC.m
- xover_loop.m
- xover_find.m
- intprofile.m
- meanprofile.m
- matching.m
- posfig.m
- set_weights.m
- solvecarina.m
- offsetplot.m

In addition, you will need m-map (and a not too old version of matlab) to draw maps. There is an excellent documentation in the file "map.html". Most of you probably have this directory already; otherwise, you will be able to download the work package from here: http://www.eos.ubc.ca/~rich/map.html

It seems that you will need matlab version 7.0 or later to run these scripts.

The toolbox automatically performs crossover analysis and inversions of 8 parameters: salinity, tco2, alkalinity, phsws25 (these four are additative), nitrate, phosphate, silicate and oxygen (these 4 parameters are multiplicative).

Introduction:

This package allow users to perform consistency control (or secondary quality control) of hydrographic data using crossover analysis. Crossover refers to cruise tracks that crosses each other, or at least comes close to each other. For each crossover, comparisons of the parameter concentrations are made in the deep part of the water column, normally > 1500 meters depth. The result of each cross over is an offset and a standard deviation of the offset. These offsets (and its uncertainty) are used to generate a set of corrections for each cruise with a set of least square models (inversions). Using the results from the inversion, the analyst can come up with a set of adjustments that should (could) be applied to the data in order to make the dataset internally consistent. For more information, see Tanhua et al. (and references therein) in the CARINA special issue in Earth Systems Science Data, 2009 (in preparation).

Step-by-step

This example is for the Atlantic part of CARINA, but can easily be modified to whatever you need.

First, download the merged data file and the file with EXPOCODE information from CDIAC. Run read_carina to produce a matfile that is fast to read. The easiest way to start run "xover_find", which will find the cruise-pairs that actually cross (i.e. a crossover), and will save a file with information on the crossovers. The next step will be to run the automated crossover routine, xover_RC.m, This is best done with help of xover_loop.m, which will generate a set of files for viewing and as input to the inversions. The inversion is run with the script solvecarina.m. But you will need to create a file with weights first (by set-weights.m for instance).

I like to have one directory for the CARINA data, and another one for the calculations, remember that the crossover routines will create a few sub-directories for figures.

Here you will find a short description of the m-files in this package.

read_carina.m:

This script reads the merged data supplied in the csv format and stores the data in as a binary "mat-file". This format is fast to read by matlab, and the file is smaller. Use this routine at once to read the merged data-file, and also when there has been an update in the merged data-file. Make sure that you edit this routine so that it reads from the right directory, and that it stores the mat-file in the right directory. This script will also store a table of expocodes and cruisenumbers.

You will need:

- The csv file with the merged data as downloaded from CDIAC.
- A csv file with (at least) these two columns 1) cruisenum 2) expocodes.

load_carina.m

Use this routine for loading the desired data quickly. Make sure that you edit the routine with the right path so that matlab finds your data. Missing data will be assigned "NaN".

- You will need to specify "cruises" and "stations" to load exactly these data, no more and no less. If "stations = -999", you will load all stations for the specified cruise.

xover_RC.m

This function will perform crossover analysis of hydrographic data. This version does the "running cluster": This function calculates the difference of each pair of stations and the average of those differences is used to compute the Xover difference between the two cruises. This way it should not be necessary to do clustering of stations.

INPUT: CRUISE1: The ID of the first cruise in the crossover
CRUISE2: The ID of the second cruise in the crossover
PARAMETER: 1:tco2; 2:alk; 3:oxygen; 4: nitrate; 5:silicate;
6:phosphate; 7:salinity

OUTPUT: DIFF The weighted difference between the cruises
STDW The weighted standard deviation of the comparison
A figure will be generated in a new directory with the name of the parameter.

The profiles are interpolated using Piecewise Cubic Hermite Interpolating, and then averaged to an average profile for each cruise. Note that for less than 3 data for a depth/density no further analysis will be performed. You will thus need at least 3 stations per cruise to do a crossover analysis. This function calculates the difference of each pair of stations and the average of those differences is used to compute the crossover offset between the two cruises.

xover_find.m

This script will loop through the cruises in the merged data file. You need to run this script in order to identify cruise-pairs that forming crossovers.

If you run the script to the end (might take a while if you have lots of cruises) you will generate:

- A file with information (cruisnumber1 cruisnumber2 #stations1 #stations2 timedifference(years) latitude longitude)
- A file with the expocodes of the xovers.

This information is necessary for building a crossover table.

xover_loop.m

This routine will loop through the crossovers generated by the routine xover_find.m.

The routine will also prepare a txt file and two mat-files with the results from the crossover analysis that can be used for the inversion.

xovermap.m

Function to make a map of two cruises and mark stations that are selected for crossover analysis
INPUT: cruise1 = The first cruise
cruise2 = The second cruise
latlim = The maximum distance between two cruises to define a X-over (in deg lat)

OUTPUT: STAT1 = xover stations for the first cruise
STAT2 = xover stations for the second cruise

intprofile.m

Function to interpolate the value of a given parameter through the water column so that two profiles can be compared to each other without being sampled at the exact same depth.

meanprofile.m

Script to calculate an average profile from several profiles in an area. The profiles are first interpolated to standard depths (or densities). These are then averaged and the standard deviation is calculated.

Posfig.m

This is just a routine to make nicer plots.

Offsetplot.m

This script summarizes the offsets by individual cruises from a crossover analyses and creates one plot per cruise with all the determined offsets.

Remember to adjust your path to the where you store the output from your crossover analysis.

Set_weights.m

This routine will generate a table with weights that will be used for the WDLSQ inversions. If you want to change the weights for individual cruises you should manually do so before the next step (or better, add a little script that lowers the weight on the core cruises, for instance).