Geophysical Research Abstracts, Vol. 7, 08311, 2005
SRef-ID: 1607-7962/gra/EGU05-A-08311

Retentive structures detection: an evolutionary approach

M. Segond (1), C. Fonlupt (1), D. Robilliard (1), B. Planque (2), P. Lazure (3)

(1) Laboratoire d'Informatique du Littoral, Calais, France, (2) IFREMER Nantes, France, (3) IFREMER Brest, France

The presence and density of animal species in the ocean and coastal waters are often conditioned by the presence of physical structures, such as upwellings, temperature fronts, or vortices. In the case of the anchovy in the Gulf of Biscay, biologists from the Ifremer (French Institute for Exploitation of the Sea) institute want to investigate the relationship between the presence of such structures and fish demography. Focus is put on so-called retentive structures, that could retain eggs and larvae in specific environmental conditions. Meso-scale vortices-like structures, whose size ranges from 10km to 300km, are supposed to be retentive structures of main importance. To verify this hypothesis, one difficulty is efficiently identifying such patterns in massive datasets, in order to match their presence against biologists observations and fisheries statistics.

There are several definitions of vortices such as Robinson's or Fielder's. As will be shown later, such "instantaneous" vortices, even if usefull for stream analysis and modelisation, do not always qualify as retentive structures for eggs and larvae. Here we relie on an example set of 30 maps tagged by oceanographers.

To perform such a detection, we have to work on large data sets. For that purpose, we use NetCDF files that are produced by the MARS3D model from the IFREMER institute. It generates series of 3D maps containing datas such as stream direction and velocity, but may also provide salt concentration or wind direction and velocity. These maps are stored in grids with cells representing 5 km in the horizontal plan, and 10 m in the vertical plan. From these 3D maps, we extract 2D maps at a fixed depth.

In this work, we have chosen several approaches to solve the problem in order to have the opportunity to fairly compare them. We can distinguish two types of methods: deterministic methods and evolutionary stochastic methods. Deterministic methods are based on mathematical principles and give always the same results, as the way to calculate them never changes. In this deterministic class of method, we tried two subclasses: local information based ones and global information based ones. On the other end, we tried an evolutionnary methods: ant based algorithms. This paper will give a summary of these methods and compare their performances for the problem we are looking at. For this purpose, we used ROC curves. These curves allow the representation of performances of several solutions by displaying the true positives rate on the y-axis and the false positives rate on the x-axis.

This problem is clearly difficult, since it partly relies on experts advice on the significance of structures and thus it has no complete formal characterization available. Only local properties could be based on strong physical background, and these properties appeared too weak to correctly classify retentive structures: sensibility (true positive rate) is excellent, but specificity (false positive rate) is a lag behind. Moreover, due to the same cause, local based methods seems unable to accurately retrieve the global shape of structures. Although it is difficult to quantify and thus it does not appear explicitely in the results section plots, this lack of accuracy was a real problem to oceanographers.

On the opposite, schemes based on the fusion of information at a more global level should prove more efficiency for outlining structure enveloppes. The streamline method is clean and efficient in this regard, but it behaves non-linearly when changing its parameters and also suffer from a not so good trade-off between sensibility and specifity.

Solving this problem with Genetic Programming could pretend as a difficult benchmark test. The approach we described here uses program components (such as matrix operators) that were deemed either strictly necessary or at least very useful to identify structures at a global level.

Perhaps this language was too general to allow for an efficient search of program space, thus explaining its deceitful results. However, from a methodology-oriented point of view, this means that one cannot easily rely on the power of mass computation to evolve successful answers to this problem, but needs to spend time to design involved, specialized program components.

At last, the ant algorithm was easily designed because it allowed an intuitive approach, as is often the case with artificial ants when the problem can be expressed in term of "optimal movements over a discrete structure". The directional bias was the only tricky point to set up. The performance is very good, reaching 100% true positives for the lowest number of false positives. It should be noted that it is obtained by combining ten detections results, in a multi-start fashion that is familiar to practitioneers of stochastic search. This is interesting since stochastic search was not favored at first by oceanographers (for fear of the variability of results) and also because the problem appeared at first as oriented either towards machine learning or specialized physical method, whereas ant algorithms are seldom used besides the pure combinatorial optimization domain.