Short communication

French phytoplankton monitoring: an exploration of optimum data presentation

Catherine Belin

Different ways are explored for presenting 12 years of data on phytoplankton species and toxins in shellfish collected along the French coast and within the framework of the French Phytoplankton Monitoring Network. The advantages and difficulties of aggregating data in different dimensions are discussed. For example, graphs appear suitable only for restricted areas, while mapping is necessary for results on regional or larger scales. Within each of these broad categories, considerable care is required when identifying the appropriate type of information and presentation.

Catherine Belin: IFREMER, rue de l'Île d'Yeu, BP 21105, 44311 Nantes Cedex 03, France: tel: +33 (0)2 40 37 41 10; fax: +33 (0)2 40 37 40 73; e-mail: cbelin@ifremer.fr

Introduction

Monitoring programmes are often characterized by the inclusion of many sampling stations and many parameters measured simultaneously at each station, leading eventually to time-series data for each parameter. An important question is how the vast amount of data collected can best be synthesized and presented to show variations in both time and space. Based on information obtained within the framework of the French Phytoplankton Monitoring Network (REPHY), different methods of plotting and mapping the data are presented and discussed.

REPHY was set up in 1984 with three objectives: to enhance knowledge of phytoplankton populations, to safeguard public health, and to protect the marine environment. Water samples are taken regularly and, depending on the course of events, episodically. When a toxic species is detected at a regular station, both number of sampling stations in the area and sampling frequency are increased. Shellfish are also sampled for toxicity tests (Belin, 1993; Belin *et al.*, 1995).

Sampling and analyses are carried out at 12 coastal laboratories and the results entered into a central database (QUADRIGE) from terminals connected to a main-frame computer. This database also receives data from two other coastal monitoring networks (chemistry and microbiology). To facilitate data analysis, the French coast is divided into a hierarchy of sites and subsites common to the three networks. Each of the 43 sites is generally divided into two or three subsites comprising sampling stations belonging to one or more monitoring network.

Data presentation

Two types of data presentation, graphic plotting and mapping, are considered in relation to the appropriate space scales (Fig. 1). The examples primarily concern plots of phytoplankton species distribution. For each example, the presentation giving the best global and synthetic view is searched and the three dimensions of space, time, and data are examined.

Within the scale of the individual sampling station, time-series data can be plotted accurately by species without information being lost. If the number of species is limited, they may be combined within a single plot, but for the entire phytoplankton population several plots will be needed. To combine results in a single plot, data have to be aggregated, for instance by defining species groups (e.g., coastal vs. benthic species; Fig. 2). In this case, precision in space (single station) and precision in time (per sampling day) are maintained.

DIFFERENT SPACE SCALES	• sampling station • site = *a group of stations* • region = *a group of sites* • country • continent of world

	Detailed results	Aggregated results
Graphic Plotting	station	site region country
Mapping	station site	region country continent, world

Figure 1. A schematic representation of optimum methods of presentation of data in different space scales in relation to aggregation level.
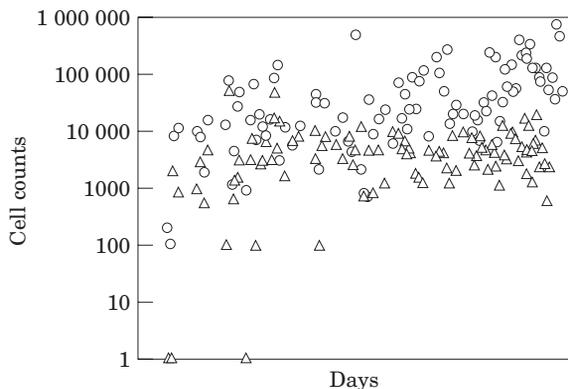
Figure 2. Time-series data of cell counts for two groups of species at a single station as an example of maintaining precision in space and time while aggregating data over species. Coastal species (○); benthic species (△).

Figure 3. Time-series data of cell counts for one group of species at three stations belonging to one site as an example of maintaining precision in space and time while aggregating data over species. Station 1 (—●—); station 2 (—◆—); station 3 (—▲—).
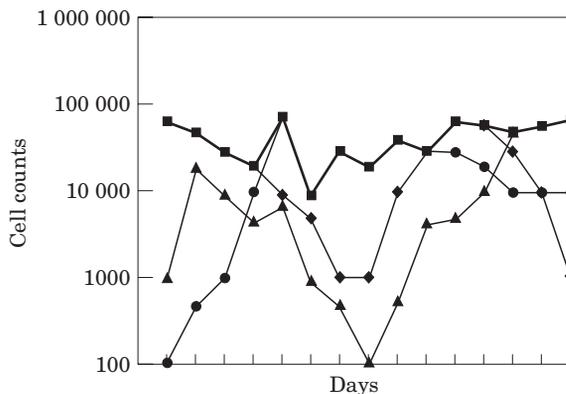
Figure 4. As Fig. 3, but highlighting the maximum cell counts observed at the site on each date as an example of maintaining precision in time while aggregating over space and species. Station 1 (—●—); station 2 (—◆—); station 3 (—▲—); maximum (—■—).

Within the scale of a site, which includes several stations, it is impossible to maintain precision in all three dimensions when presenting results for the entire phytoplankton population; aggregation in at least one dimension is required. A first case might be by aggregating data, while maintaining precision in space and time. A comparison of the results obtained at the different stations of the site for a selected group of species is shown in Fig. 3.

A second case might be aggregation over space. This requires the choice of an appropriate function allowing a meaningful interpretation, e.g., the mean, median, maximum, etc., value found for one species at the set of stations at the same time. If all stations have always been sampled simultaneously (Fig. 4), there is no problem and precision in time and of data can be maintained. The situation becomes more complex when stations are not

sampled simultaneously, because an appropriate time interval must be chosen (e.g. week, month, etc.) for calculating the function chosen. In this case, precision in space (single station) and precision of data (one species or species group) may be maintained, but there is no longer precision in time.

The last case is aggregation over all dimensions. As an example, Fig. 5 plots the maximum value observed for a group of species (aggregation of data) over the set of stations (aggregation over space) by month (aggregation over time).

Aggregation over space may be sensible in the scale of a site when individual stations are likely to give comparable results. However, aggregation over larger scales (e.g., region, country) may not be informative and, for the French monitoring data, site should always be the largest entity for spatial aggregation. Fig. 6 is an
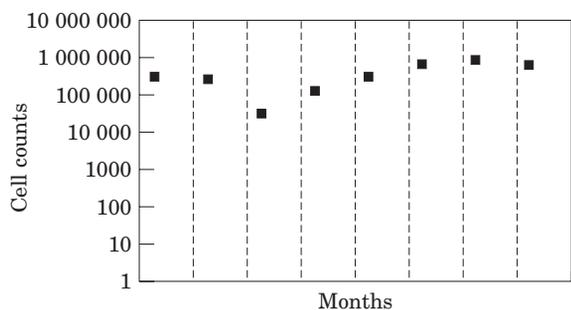
Figure 5. Maximum cell counts for one group of species observed at a particular site within a particular month as an example of aggregating in three dimensions.

example for toxic phytoplankton species at the Douarnenez Bay site (western Brittany), which contains four phytoplankton stations and is divided into three subsites, one of them containing two stations. Cell counts of all *Dinophysis* species were summed for each station, and the maximum value observed for the group of *Dinophysis* spp. in each subsite, within a month, was calculated. A similar procedure was applied for the toxic *Alexandrium minutum* (Joanny *et al.*, 1993).

Within the scale of a region or country, with data for many stations and many years, it is necessary to calculate statistics for all sites over longer periods. Either the entire period of observation is divided into years to describe the statistics (e.g., mean, maximum, etc., of all values measured at all stations), or statistics for the entire period may be described (e.g., the number of occurrences of a particular event). However, in these spatial scales, mapping provides more informative pictures of the results, because these highlight spatial variation in important events. Two mapping methods are considered, in both of which the coast is partitioned into sites: discontinuous presentation with symbols by site and continuous presentation using the coast line.

An example of a discontinuous presentation (Fig. 7) with symbol size proportional to the maximum concentration of *Dinophysis* spp. found on each site shows the hotspots during 1995. Similar results could be given for each individual year. In order to draw one representative map for the entire period, the maximum concentrations observed might be chosen, but this would result in the loss of important information on the irregularities in the annual occurrence of the species. In this case, it would be more informative to plot the frequency of particular
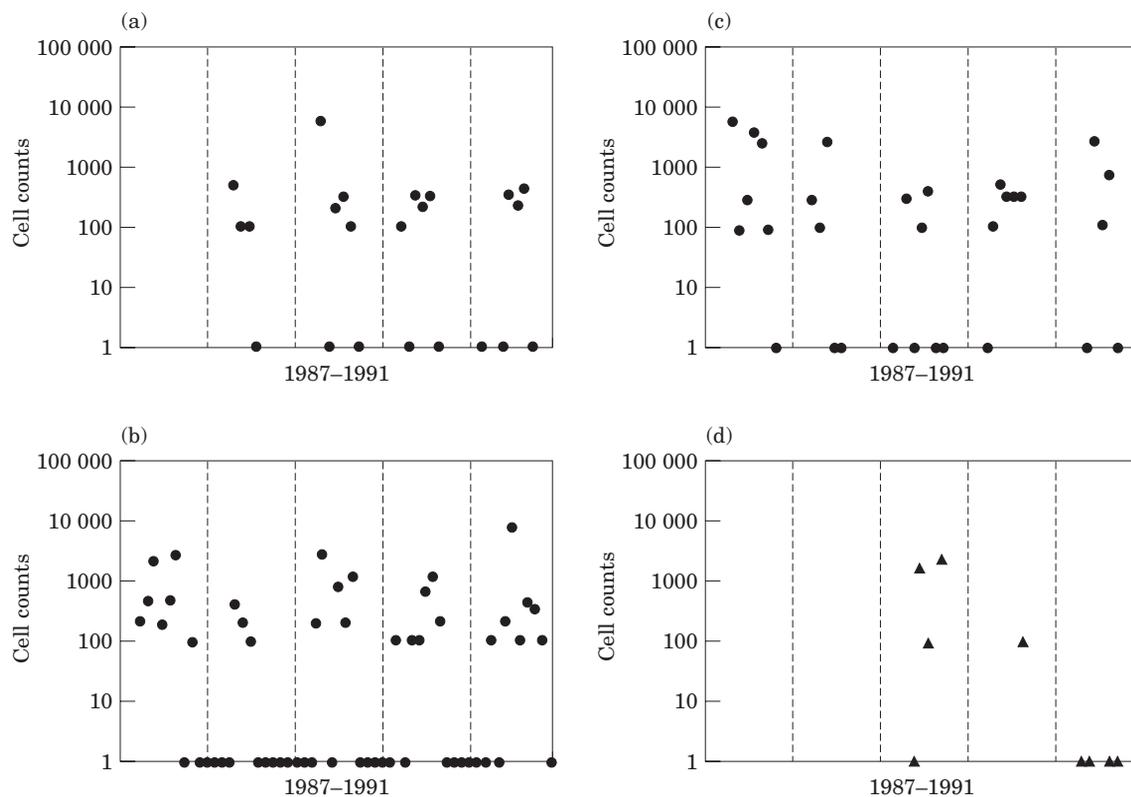
Figure 6. Maximum cell counts for two toxic phytoplankton species by subsite and month over 5 years as an example of aggregating over time and space while maintaining genus or species information. (a) *Dinophysis* sp.—sub-site 1; (b) *Dinophysis* sp.—sub-site 2; (c) *Dinophysis* sp.—sub-site 3; (d) *Alexandrium minutem*—sub-site 2.
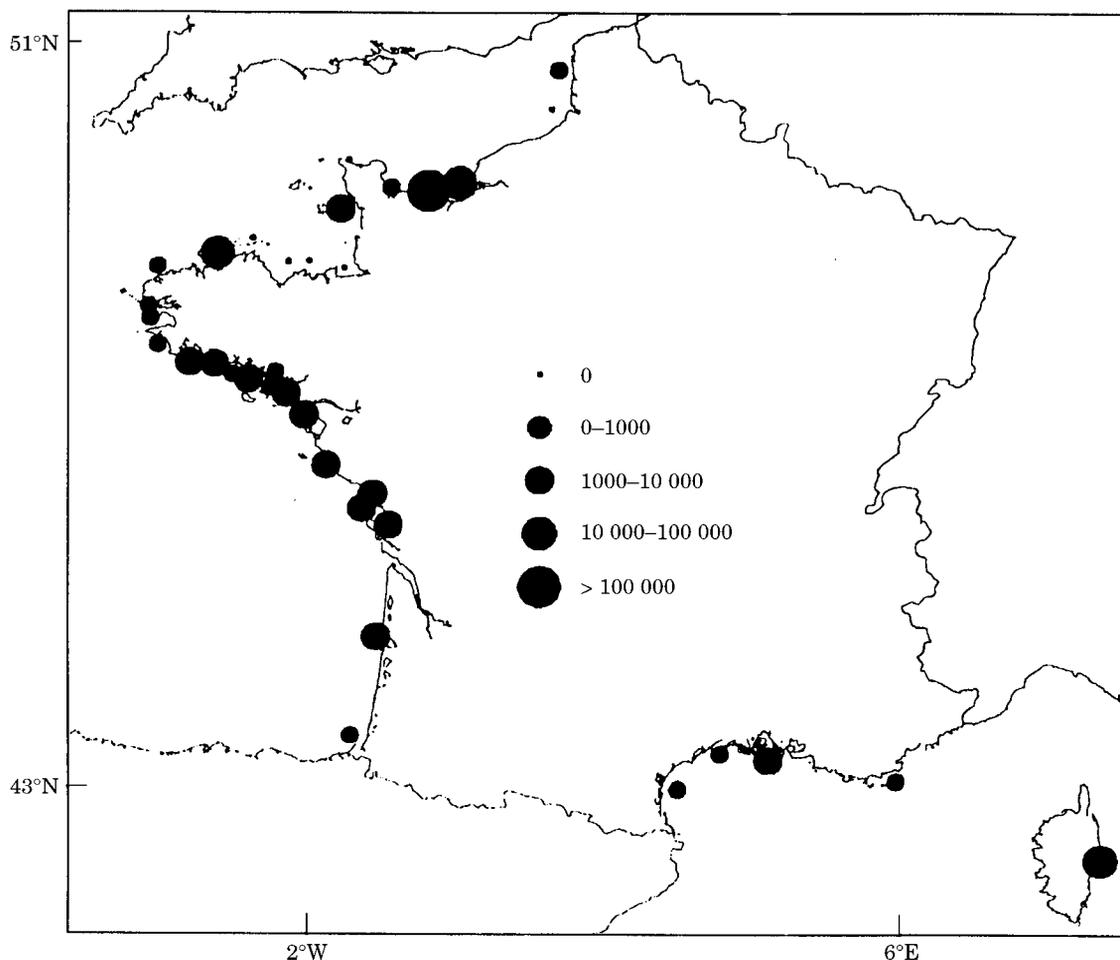
Figure 7. Map showing maximum cell counts for *Dinophysis* sp. (in classes) per site for 1995.

events during the whole period. As an example, Fig. 8 gives information on the number of red tides caused by *Noctiluca scintillans* during 1984–1992.

An example of a continuous presentation (Fig. 9) identifies the areas which were closed during 1995 because of the presence of toxins in shellfish. The choice of a thickened coast line gives a clear idea of the discontinuities between the affected areas.

Discussion

For an entire continent (or the world), the change in scale is much more crucial because the coherence in data collection within national monitoring programmes ceases to exist when information from different countries is merged. Many attempts at a synthetic presentation of toxic events related to phytoplankton blooms have been made on a world scale (Granéli *et al.*, 1990; Andersen, 1996), but difficulties have always been linked

to the fact that the criteria used differed between countries.

The most important choice is the kind of information to present on such maps. If the distribution of toxic species is aimed for, difficulties arise from the fact that some species may be toxic in some parts of the world and not in others. If observations of toxic effects (human intoxication or animal mortality) are mapped, this may not represent the actual size of the problem because monitoring programmes have been put in place to reduce the potential hazards. The presence of toxins, regardless of the level of toxicity or source, is probably the most objective criterion when producing a reliable picture of toxic hazards throughout the world.

Secondly, a choice must be made regarding the appropriate way of presenting the impact of toxic events. Apparently, an informative and synthetic picture can be obtained by indicating their frequency during a given
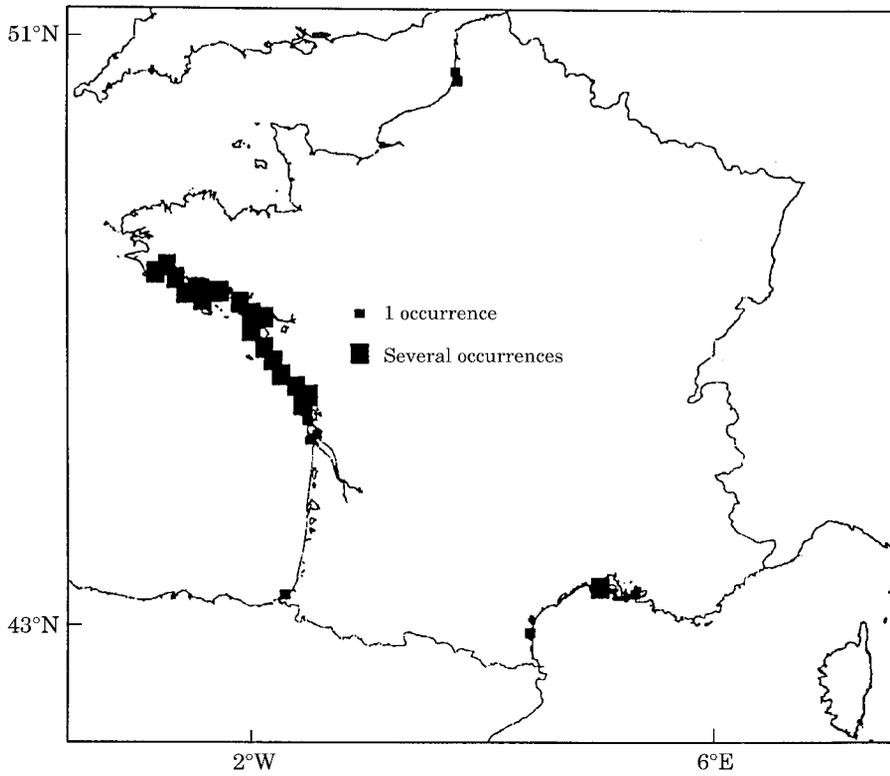
Figure 8. Frequency of occurrence of *Noctiluca scintillans* blooms per site for 1984–1992.
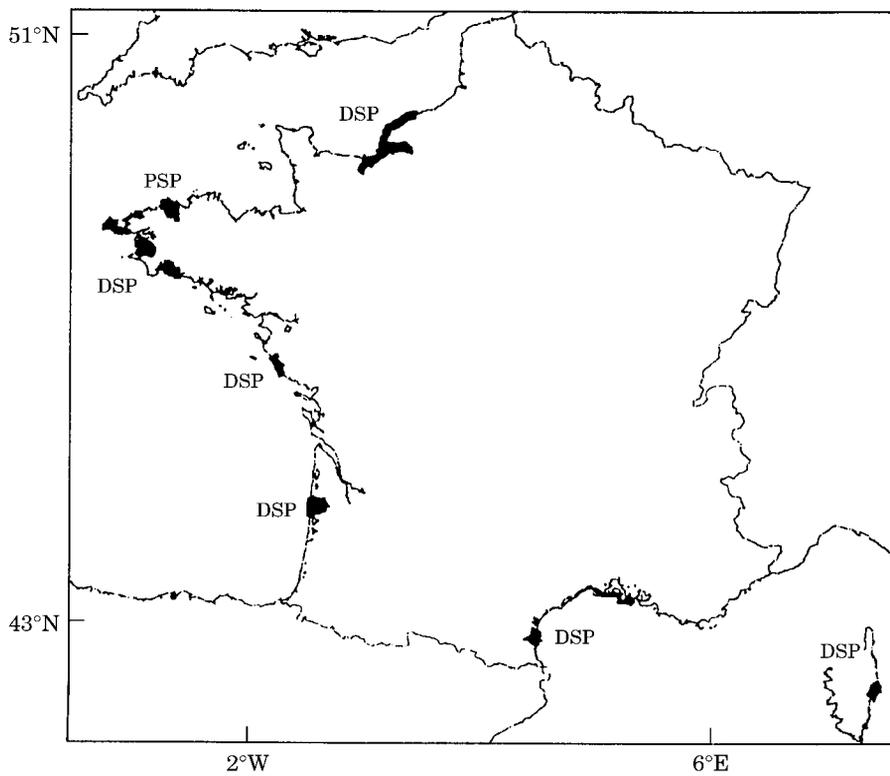
Figure 9. Areas temporarily closed to shellfish fisheries because of the presence of DSP or PSP toxins in 1995.

period. Furthermore, maps of this kind should indicate countries or regions which are regularly monitored, since the presence of toxins can only reliably be assessed in these regions (ICES, 1996).

This exploration of different methods of data presentation has focused on particular points. More importantly, however, is the description of a general quality index of the coastal environment based on the ongoing monitoring programmes.

References

Andersen, P. 1996. Design and implementation of some harmful algal monitoring systems. IOC Technical Series no. 44, UNESCO. 102 pp.

Belin, C. 1993. Distribution of *Dinophysis* spp. and *Alexandrium minutum* along French coasts since 1984 and their DSP and PSP toxicity levels. *In* Toxic phytoplankton blooms in the sea, pp. 469–474. Ed. by T. Smayda and Shimizu. Elsevier, Amsterdam. 952 pp.

Belin, C., Beliaeff, B., Raffin, B., Rabia, M., and Ibanez, F. 1995. Phytoplankton time-series data of the French phytoplankton monitoring network: toxic and dominant species. *In* Harmful marine algal blooms (Proliférations d'algues marines nuisibles), pp. 771–776. Ed. by Lassus, Arzul, Erard-Le Denn, Gentien and Marcaillou-Le Baut. Lavoisier. 878 pp.

Granéli, E., Sundström, B., Edler L., and Anderson, D. M. 1990. Toxic marine phytoplankton. Elsevier, Amsterdam. 554 pp.

ICES. 1996. Report of the ICES/IOC Working Group on Harmful Algal Bloom Dynamics. ICES CM 1996/L: 4.

Joanny, M., Belin, C., Claisse, D., Miossec, L., Berthomé, J. P., Grouhel, A., and Raffin, B. 1993. Qualité du milieu marin littoral. IFREMER, DEL. 241 pp.