Pauly, D. 1993. Data-rich books. BioScience 43(3): 167-168. [Reprinted as Essay no. 26, p. 185-188 In: D. Pauly. 1994. On the sex of fish and the gender of scientists: essays in fisheries science. Chapman & Hall, London]

Data-rich books

Reproducibility of results, a basic tenet of the scientific method extolled in principle by every practicing scientist (Merton 1973), is in reality, and especially in biology, subtly subverted by the mode in which results are presented (Collins 1985). Subversion may come about by editorial preference for short papers, the general disregard of careful description (Gould 1989), and the preference for quick tests of fashionable hypotheses (Janzen 1986). This problem is particularly acute in those disciplines that have a historical component and where published results are based on unique computer programs.

In this article, the problem is addressed from the perspective of fishery science, a strongly historical discipline, however much its practitioners would like to reduce it to time- and locale-invariant principles. The historic dimension of fishery science is illustrated by the common use of narratives, historians' main tool for presenting their material (see, for example, contributions in Cushing 1982 and Glantz and Thompson 1981).

S. J. Gould (1989) demonstrated rather convincingly that phylogeny is contingent—strongly structured by the consequences of chance events. It is thus not surprising that the evolution of exploited fish stocks and of the societal complexes that crystallize around fisheries should be contingent as well—hence the justification for narratives as a major form for the description and understanding of the evolution of fisheries.

I believe there are two interrelated approaches that can be used to overcome the problem posed by contingency and the simultaneous requirement for reproducibility. One is to emphasize the comparative method, which Mayr (1982) suggested was the most appropriate for evolutionary biology, an idea forcefully argued for fishery science as well by Bakun (1985). The other approach is to ensure that narratives and the analyses supporting them can be reproduced (that is, duplicated) not only in principle, but also in practice, by making the relevant data available.

by Daniel Pauly

Making data (as opposed to bibliographic references) available to a wide range of users is difficult, and established traditions of data exchanges are lacking in most disciplines. The glorious exception is oceanography, whose practitioners were convinced in the last century by Commodore M. F. Maury, US Navy, of the mutual advantages of shared data sets.

The most important international and widely available fisheries data set is the volumes-of-catch statistics published annually by the Food and Agriculture Organization of the United Nations. They are invaluable for studies of global or regional trends but too aggregated for the analysis of any specific fishery or stock. For those more specific analyses, I propose the concept of data-rich books, a variant of the format currently used to present fisheries and other biological narratives in monographic form.

Data-rich books consist of the standard elements of monographs, with the following added:

- As many tables as possible with detailed headings and footnotes, documenting as much of the data used in the analyses as possible. The tables should be in addition to graphs showing trends and interactions, which tables cannot do well.
- Short articles (as many as necessary), written by the technicians, junior scientists, or others who compiled the datasets used elsewhere in the book. These articles should describe, via reference to standard texts if appropriate, the method(s) used to collect the data in question.
- An appendix describing the entire set of data compiled by the authors of the short articles and the approach used for assembling these data into computer files.
- An address from which to order the data files described in the appendix, a price (covering handling and mailing costs, but not much more), and an explicit invitation to use the data for further analyses, as long as the authors are cited and acknowledged.

Two recent books on an international study of the Peruvian upwelling ecosystem (Pauly et al. 1989, Pauly and Tsukayama 1987) document the concept of data-rich books. The book, edited by Pauly and Tsukayama (1987), though lavishly illustrated, includes 88 pages of tables; therefore, 25% of the book consists of data that others can use directly.

Numerous scientists followed up on our invitation to use these data; for example, the time series of anchoveta recruitment (Table 1) was used by Restrepo and Hoenig (1988) to test one of their recruitment models and by Cury and Roy (1989). The latter's use of this series, in a comparative study of the relationship between upwelling intensity and clupeoid (herringlike fish) recruitment, won them a major scientific prize.

The second of these books, by Pauly et al., includes a similar number of tables and, like the first book, a number of articles by participants in the research. It also includes an appendix with details on the contents and potential use of the 15-diskette data set documenting the two books (Palomares et al. 1989; see Table 1).

With these publications, we have taken care that the authors of time series of data can get credit for their work (via citation of their articles) and ensured that others can test the replicability of our analysis. Thus, we enable future authors to go beyond

Table 1. Monthly time series on the Peruvian upwelling system, covering the years 1953 to mid-1980s (adapted from Palomares et al. 1989).

Parameter groups	Remarks
Temperature-related	
Sea surface temperature	Twelve localities and/or sources
Thermocline depth	For three areas off Peruvian coast
Wind-related	
Wind velocity and turbulence (wind^3)	Five series, including Lasker events
Upwelling intensity, Rossby radius	Four series (two coastal, three oceanic)
Offshore transport	One series, adjusted for depth of mixed layer
Meteorological variables	
Southern oscillation index	One series
Solar radiation, cloud cover	One series each
Nutrients and primary production	
Nitrate concentration and flux	Eleven inshore and offshore series, various models
Primary production	Five series from nitrate flux and various models
Biology of Anchoveta (*Engraulis ringens*)	
Losses to predation	Nine series, one for each of various predators (mammals, birds, and fish), based on species-specific predation models
Biomass and recruitment	Two series, estimated by virtual population analysis calibrated through acoustic surveys
Spawning stock	Derived from biomass and a temperature/size-dependent maturation model
Fishery catches	One series by length group
Natural and fishing mortality	Two series, estimated via virtual population analysis
Miscellaneous features of individual Anchoveta	Stomach contents, visceral fat, and maturation stage of ~10^4 specimens, with sampling locations and sea surface temperature
Maps of egg distribution	89 maps, from March 1964 to October 1985
Biology of other system components	
Biomass of zooplankton	One series each for three areas of the Peruvian coast for the years 1964–1985
Biomass and catches of bonito	One series each
Numbers and biomass of cormorants	One series each, based on analysis of more than 10,000 distribution maps
Numbers and biomass of boobies	Same as for cormorant
Numbers and biomass of pelicans	Same as for cormorant
Population and food consumption of sea lion	One series each, based on calibrated population and energetics model
Population and food consumption of fur seal	Same as for sea lion
Biomass of predatory fish in distribution area of Anchoveta	One series each for mackerel, horse mackerel, and hake

these results, as Cury and Roy (1989) did and as Jacqueline McGlade illustrated when she showed at the World Fisheries Congress of May 1992 in Athens, Greece, that these data can be used to demonstrate the existence, in the Peruvian upwelling system, of a two-state strange attractor.

Several other data-rich books, notably on multivariate analyses of fish culture experiments and on trophic modeling of ecosystems, were in press at ICLARM in late 1992. We would be delighted to discuss these products and/or the concept behind them with interested colleagues.

Acknowledgment

This article is ICLARM Contribution No. 897.

References cited

Bakun, A. 1985. Comparative studies and the recruitment problem: searching for generalizations. *Calif. Coop. Oceanic Fish. Invest. Rep.* 26: 30–40.

Collins, H. M. 1985. *Changing Order: Replication and Induction in Scientific Practice.* Sage Publ., London, UK.

Cury, P., and C. Roy. 1989. Optimal environmental window and fish recruitment success in upwelling areas. *Can. J. Fish. Aquat. Sci.* 46: 670–680.

Cushing, D. H. 1982. *Climate and Fisheries.* Academic Press, New York.

Glantz, M. H., and J. D. Thompson, eds. 1981. *Resources and Environmental Uncertainty: Lessons from Coastal Upwelling Fisheries.* John Wiley & Sons, New York.

Gould, S. J. 1989. *Wonderful Life: The Burgess Shale and the Nature of History.* W.W. Norton, New York.

Janzen, D. 1986. Science is forever. *Oikos* 46: 281–283.

Mayr, E. 1982. *The Growth of Biological Thought: Diversity, Evolution and Inheritance.* Harvard Univ. Press, Cambridge, MA.

Merton, R. K. 1973. The normative structure of science. Pages 267–278 in N. W. Storer, ed. *The Sociology of Science.* University of Chicago Press, Chicago.

Palomares, M. L., A. Jarre, and V. Sambilay Jr. 1989. Documentation of available 5 1/4" MS-DOS data diskettes on the Peruvian upwelling system. Pages 408–416 in D. Pauly, P. Muck, J. Mendo, and I. Tsukayama. 1989. *The Peruvian Upwelling Ecosystem: Dynamics and Interactions.* ICLARM Conference Proceedings 18, Manila.

Pauly, D., P. Muck, J. Mendo, and I. Tsukayama. 1989. *The Peruvian Upwelling Ecosystem: Dynamics and Interactions.* ICLARM Conference Proceedings 18, Manila.

Pauly, D., and I. Tsukayama. 1987. *The Peruvian Anchoveta and Its Upwelling Ecosystem: Three Decades of Change.* ICLARM Studies and Reviews 15, Manila.

Restrepo, V. R., and J. H. Hoenig. 1988. An empirical investigation of Peruvian anchovy recruitment. *American Fisheries Society Marine Fisheries Section Newsletter* 6(2): 7.

Daniel Pauly is the director of the Capture Management Fisheries Program, International Center for Living Aquatic Resources Management (ICLARM), MC P.O. Box 2631, Makati, Metro Manila 0718, Philippines. He is interested in the dynamics and management of tropical fish.