DATA HARVESTING AND HOARDING IN OPERATIONAL OCEANOGRAPHY

Erik MOLLO-CHRISTENSEN, Jeffrey N. SHAUMEYER, and John M. BORDEN

Wavix, Incorporated

1. GLOBAL-SCALE CLIMATOLOGY

Remote sensing has significantly extended climatological research capabilities and global models are showing success. Pioneering global climatology, however, has had little technological support for global-scale research. Ocean observing was limited to buoys with data storage that limited the scope of investigation, particularly when data retrieval was delayed by months or years. Available (or practical) technology necessarily affected the questions that could be considered, and large-scale models have had to wait for large-scale observing systems to be effective.

The oceans play an important role in climate, responsible for a significant part of the global poleward heat transport. Part of the transport is by seasonally changing fluxes, exemplified by warm and cold core rings, modons, and eddies. The role of the ocean as a sink of carbon is also recognized, and local air-sea interactions, although reasonably well modeled now, will need significant amounts of data to verify steps and to improve approximations. Numerical models, previously crude because of computer limitations, can now be much more detailed in temporal and spatial resolution, provided that enhanced computations are constrained by increasingly comprehensive data. Improved models, will need continual updating of data for verification and assimilation, for example, to suppress chaotic tendencies and model the bifurcation dynamics present in the global system.

We therefore need routine observations of the oceans just as we now have of the atmosphere. Ocean color remote sensing and active and passive microwave satellites provide some of the information needed, but several crucial variables cannot currently be determined by remote sensing. One such variable is the poleward heat flux above the main thermocline at mid-latitudes, which vary seasonally and on time scales of weeks and months. We need frequent measurements of the current profile above the main thermocline, and the temperature profile as well.

2. OPERATIONAL OCEANOGRAPHY AND ENHANCED DATA HARVESTING

Changes in technology and the scale of ocean research have spawned a new concept in data collection: operational oceanography, the idea that large-scale programs of data collection operate independently from the users of the data. Until now, data users developed their own buoy programs at great expense and use of scarce internal resources. In operational oceanography, resources are shared and used more efficiently: much more data can be collected and analyzed by more researchers. Expertise in observational technology then reposes not only in the scientific community, but with commercial and government agencies as well, all with a role to play in operational oceanography.

Responsive, global-scale climate models and forecasting require comprehensive and timely data retrieval from observational networks, both remote sensing and *in situ*. While satellite-based remote sensing has revolutionized the practice of oceanography, the kinds of data that can be collected via satellite alone is

Paper for *Proceedings of the International Conference on The Ocean Observing System for Climate,* Saint-Raphael, France, 18-22 October 1999.

limited. There is a growing recognition of the need for much more *in situ* data to support modeling and forecasting efforts.

Programs such as CLIVAR, PIRATA, TOGA, WOCE, GOOS, and GCOS hope to expand observational networks with hundreds of additional buoys and enhance existing systems with additional sensors collecting vast data sets, but how is all that data to be retrieved? Most measurement buoys, using currently available satellite communications, can return only a few thousand bytes of data per day at most, and recovering complete observational data sets must wait for physical retrieval of the buoy. The timely harvest of data from *in situ* platforms must be increased by orders of magnitude if scientists are ever to understand the complexities of the Earth's climatic systems.

A major impediment has been the lack of an affordable communication infrastructure. Current data transmission services are not optimized for oceanographic data use and do not meet current and future needs. A relatively high-volume, low-cost communications system with two-way capability is required. Although some applications demand real-time transmissions, the majority of the market only requires what we refer to as “real-enough time” data: data received on the same day in which it was collected, which is appropriately timely for most purposes.

Improved satellite communications for buoys means the potential exists to integrate comprehensive *in situ* data with satellite data into a truly global, ocean-observing system. Wavix has developed a data transmission system for our own moorings, which we plan to make available to the oceanographic community. The WavSat satellite-communication system offers the potential of returning at least 500 kilobytes of data per day from each buoy directly to the researcher. In addition, the communications link is two way, allowing the researcher to send commands to the remote system to alter, say, its data-handling protocols or observing schedules in response to events observed in the harvested data stream.

3. DATA SHARING VS. DATA HOARDING

We suggest that at the global scale of research, hoarding data serves little useful purpose and may actually be a hazard to scientific progress. Indeed, the timely sharing of data in global research can offer unexpected positive bonuses: at this scale of inquiry, a single data set has little significance on its own, but in combination with others it can help support a variety of research programs. Such systems are partly in place from the combined efforts of scientists and government, to serve as a foundation for future growth. Experience from the public distribution of TOPEX/Poseidon data reveals some of the benefits; many papers have been written by those who would have been thought of as outsiders in the past, and the project has been exceptionally productive.

Scientists themselves may not wish to provide data handling and distribution services; this is best left to government and commercial agencies. Doing so can ease the impulse to hoard data. Assuredly, individual scientists will still deploy special-purpose equipment for attacking specific research questions. However, new paradigms must evolve to support a global network of data gathering and modeling and the community of scientists investigating questions that require global resources.

The ability to retrieve large amounts of data in real-enough time, coupled with interactive data platforms, will change the way remote oceanographic data collection is done, and enhance the utility of ocean-observing systems. Until now, global-scale research programs have been constrained by the limits of one-way, low-volume data returns. Eliminating the data-transmission barrier will lead to a revolution in operational oceanography, allowing more dense data harvests in real-enough time such that scientists can examine ocean processes in unprecedented detail.