

## **An Environmental Information Synthesizer for Expert Systems: A Framework for Use in Near Real-Time Detection of Harmful Algal Blooms**

James C. Hendee, Ph.D.  
(e-mail: jim.hendee@noaa.gov)  
Atlantic Oceanographic and Meteorological Laboratories  
National Oceanic and Atmospheric Administration  
4301 Rickenbacker Causeway  
Miami, FL 33149-1026

### **Abstract**

As an enhancement to the SEAKEYS environmental monitoring network in the Florida Keys National Marine Sanctuary, software called the Environmental Information Synthesizer for Expert Systems (EISES) has been utilized together with a specially developed expert system to model and report the near real-time sensing of environmental conditions conducive to the onset of a harmful algal bloom (HAB, e.g., “red tide”). Actual near real-time *in situ* fluorometry data was matched with wind speeds and Photosynthetically Active Radiation at the Long Key SEAKEYS station in Florida Bay to simulate the onset of an HAB. These incidences were e-mailed to the knowledge engineer as they occurred, and could in the future be e-mailed to regulatory agencies, or posted to a Web site, as is done with a similarly developed expert system for coral bleaching. This approach shows promise with the future remote detection of HAB pigment data via *in situ* or satellite sensors.

### **Introduction and Background**

Those of us who try to manage and make sense of large amounts of data streaming from satellites, continuously operating instrumented processes, or *in situ* instrumented arrays, have felt the stress of data overload. Having the data stored automatically in a database helps to organize the data, and having automated outputs of user-specified graphs and tables helps us to depict the course of events being monitored. However, these data products require interpretation, and for critically monitored processes, they require interpretation every day, or even more frequently. If we had nothing else to do in our lives, this might not be so overwhelming, but we do, and it is.

Many industries and research concerns have solved the problem of critical data overload using artificial intelligence techniques, such as expert systems. Artificial intelligence involves the capability of a device such as a computer to perform tasks that would be considered intelligent if they were performed by a human (Mockler & Dologite 1992). An expert system is a computer program that attempts to replicate the reasoning processes of experts and can make decisions and recommendations, or perform tasks, based on user input. Knowledge engineers construct expert systems in cooperation with problem domain experts so that the expert’s knowledge is available when the expert might not be, and so that the knowledge can be available at all times and in many places, as necessary. Expert systems derive their

input for decision making from prompts at the user interface, or from data files stored on the computer. The knowledge base upon which the input is matched is generally represented by a series of IF/THEN statements, called production rules, which are written to approximate the expert's reasoning. The degree of belief the expert has in a conclusion may be represented as a confidence factor (e.g., 0% to 100% confidence), or as a subjective term (e.g., "possibly," "probably," or "almost certainly") in the expert system. Real-time monitoring and control is mature technology, especially for industrial systems, and the use of expert systems in this arena continues to be a fast developing technology. On the other hand, monitoring and modeling of the environment in near real-time has come late in the evolution of expert systems, and only recently (Hendee 1998; Hendee 2000) for the marine environment.

Real-time expert systems were originally created to provide knowledge-based control of industrial processes and mechanical functions through feedback from sensors monitoring those processes. Real-time expert systems that monitor the environment can of course only hope to control the environment very indirectly. Through data gathering, inferencing, alerting and chronicling of monitored events, environmental expert systems can continuously deal with large amounts of data and save scientists and policy makers tremendous amounts of time. Recommendations can then be made to governments, which can then regulate human endeavors that directly or indirectly affect the environment. Under another scenario, however, the near real-time monitoring and reporting of an environmental event such as a harmful algal bloom (HAB, e.g., "red tide") could save lives: an alert to the predicted event would be sent to the appropriate authorities, who would in turn warn the public. The monitoring and alerting process discussed in the present report could eventually be adapted to just such a purpose.

At the National Oceanic and Atmospheric Administration's (NOAA) Atlantic Oceanographic and Meteorological Laboratory (AOML), in Miami, Florida, environmental data are acquired from remote sites on lighthouses and navigational aids situated at coral reefs and other strategic locations within and near the Florida Keys National Marine Sanctuary via a satellite data archival site at Wallups Island, Virginia. The data are collected at the sites continuously then transmitted hourly. Oceanographic instruments (measuring sea temperature and salinity, minimally, but also other instruments at different sites) are maintained by the Florida Institute of Oceanography, and meteorological instruments for measuring wind speed, wind gust, wind direction, air temperature, dew point, and barometric pressure are maintained by the National Data Buoy Center of NOAA. There are currently seven sites, which have been termed the SEAKEYS (Sustained Ecological Research Related to Management of the Florida Keys Seascape) network: Fowey Rocks (in Biscayne National Park), Molasses Reef (near Key Largo), southern Florida Bay (near Long Key), Sombrero Reef (near Marathon), Sand Key (near Key West), Dry Tortugas (at the very end of the Florida Keys), and northwestern Florida Bay (near Cape Sable) (Ogden et al 1994). The data received are reformatted using software developed at AOML, called the Environmental Information Synthesizer for Expert Systems (EISES; Hendee 1999).

To effect a realistic test of a hypothetical HAB expert system, near real-time data from an *in situ* fluorometer in a marginally eutrophic region of Florida Bay, the Long Key SEAKEYS station, was

used to illustrate its utility. Using an approach very similar to that used to predict coral bleaching in the Florida Keys National Marine Sanctuary (Hendee et al 2000a) and the Great Barrier Reef (Hendee et al 2000b), a production rule was developed to show how an HAB could be monitored in near real-time, assuming the proper phytoplankton pigment sensor is eventually designed and refined, or that dinoflagellate pigment sensed data from satellites become available.

## **Methods**

### *Field Maintenance of the Instruments*

The instruments themselves were visited once a week by the SEAKEYS technicians so corrective maintenance of the oceanographic sensors could be attended to as necessary. In any case, the oceanographic instruments would become fouled and had to be cleaned periodically. Sea temperature and salinity sensors were “sea-truthed” during the station visits. That is, calibrated instruments were taken into the field and the parameters were measured at the same time the *in situ* instruments made their automated measurements, to see if the *in situ* instruments needed to be replaced, or their data needed to be corrected. Water samples were taken for fluorometry calibration for use in another study; however, the instrument was cleaned of biofouling at the same time as the sea temperature and salinity sensors. Meteorological instruments were maintained by NDBC. If any of the instruments malfunctioned, the expert system code was easily adjusted so that those values were not accounted for in the process (see Hendee 2000, for specifics).

### *The Expert System*

The C-Language Integrated Production System (CLIPS), developed by the National Aeronautics and Space Administration, was used for the expert system shell, and is an essential part of EISES. The development of the prototype HAB expert system actually proceeded under three different stages. The first stage was for parsing the raw data file and producing a columnar data report, which was easier to work with, and could also be used in daily postings of the raw data to the Web. ASCII columnar data files produced in Stage 1 were screened against production rules in Stage 2 to determine whether the values for the instruments were within realistic ranges, or whether the instrument appeared to be malfunctioning or off-line (garbled or no data). A comprehensive description of how Stage 1 and Stage 2 function can be seen in Hendee (2000); however, an overview of how Stage 2 operates is herewith described.

### *Stage 2*

To aid in the analysis of data, which might vary widely depending upon the time of day and the season of the year, values were averaged for eight three-hour periods per day, termed midnight (2200 to 0100 hrs local time), pre-dawn (0100 to 0400 hrs), dawn (0400 to 0700 hrs), morning (0700 to 1000 hrs), mid-day (1000 to 1300 hrs), pre-sunset (1300 to 1600 hrs), sunset (1600 to 1900) and evening (1900 to 2200 hrs). These groupings were convenient because meteorological, oceanographic and biological phenomena quite often show predictable fluctuations during these periods of the day, for instance the change of wind direction and wind speed, and the onset of crepuscular animal behavior and

phytoplankton migration patterns, with sunrise and sunset. The averaged values within each of these categories were then subjectively assigned to fall within one of eleven categories: unbelievably low, drastically low, very low, low, somewhat low, average, somewhat high, high, very high, drastically high, and unbelievably high. These groupings were arbitrary terms, and parameters such as wind direction were further translated to different regions of the compass (e.g., NE-ENE). This approach was similar to that of Uhrmacher, Cellier, & Frye (1997) who argued that discretization of quantitative into qualitative values enabled them to represent a “gap-free history of the variables.” As they also pointed out, an approach such as this is of value if the knowledge of the system of interest is imprecise or incomplete, which is the usual case with ecological systems. The assignment of values to these categories depends upon what the season of year is (spring, summer, fall or winter), so, for instance, what might be considered “somewhat high” in winter might otherwise be considered to be “average” during summer. Values which are determined to be “unbelievably high” or “unbelievably low” represent values which are considered to be totally unrealistic for the parameter in question. However, should it happen that these values might begin to represent real-life values, the ranges may be easily reset to encompass the newer values. The expert system thus also serves as an environmental model utilizing subjective terms in its descriptions.

The status of the parameters, that is, where they were on the continuum from unbelievably low to unbelievably high, and when the values occurred (i.e., the period of the day), were saved as CLIPS-loadable “facts” represented in a text file, and thus available for Stage 3 processing. These facts thus represented subjective interpretations of the measured data, and therefore represented *information*, not just columns of numbers. An abbreviated example of Stage 2 CLIPS-loadable facts from Long Key can be seen in Figure 1.

### *Stage 3*

The third stage was the expert system developed for the monitoring of conditions hypothetically conducive to the onset of an HAB. The expert system looks for high fluorometry readings, low wind speed, and high photosynthetically active radiation (PAR). Wind direction is measured and reported, but not used in the inferencing system. Following is a pseudocode representation of the production rule used in the operational expert system.

IF fluorometry readings are high during late morning, mid-day and/or early afternoon  
AND wind speed is low during late morning, mid-day and/or early afternoon,  
AND PAR is high during late morning, mid-day and/or early afternoon  
THEN conditions are apparently consistent with, and conducive to, an HAB event.

The expert system was used to screen hourly data over the past 72 hours worth of data, once a day, and reported output via a file and an e-mail message to the knowledge engineer, as long as there was anything to report. The reports were cumulative over seven days, so that as long as any report was generated during that time, a cumulative report was generated and sent. Once conditions had not been met with seven days, reports were no longer sent. Thus, if conditions were met on March 1 only, you

would get the report until March 8; if conditions were met on March 1 and March 6 only, you would continue to get a daily report showing the results of those two days until March 8, but only for March 6 until March 13.

## Results

The expert system operated continuously and without intervention from the knowledge engineer throughout a trial period of four months. An example of the output of the report is represented in Figure 2. Wind speed was measured in knots, and PAR in microeinsteins/m<sup>2</sup>. The fluorometry output voltage was a function of the amplifier on the fluorometer. Actual chlorophyll values were not derived in this particular trial, but are a part of the usual field sampling and validation component of the SEAKEYS program. In a test of an actual HAB expert system, field validation of the phytoplankton pigment levels would be a necessary part of the knowledge engineering process.

## Discussion

Although wind direction and other parameters were not used in the operational expert system inferencing, they easily could be. In fact, any measured parameter could be accounted for in the matching pattern represented by a production rule similar to the above. For instance, following is a hypothetical example production rule (as pseudocode) that could be represented in an HAB expert system at mythical Deadman's Bay:

```
IF fluorometry was high during noon,  
    AND wind speed was high to very high,  
    AND tide was very low during noon,  
    AND wind direction was N-NE all day,  
    AND salinity was low during daylight hours,  
    AND sea temperature was high all day,  
THEN output a report that says northerly winds pushing high temperature,  
    low salinity water during the day has possibly resulted  
    in bloom conditions during a noon low tide at Deadman's Bay.
```

Thus, the expert system approach provides a powerful means for pattern matching of various measured parameters, something that is much more difficult to do with a standard procedural (vs CLIPS symbolic) programming approach.

## Outlook

Satellites have been used to remotely sense algal blooms in fresh water (e.g., Yacobi et al 1995; Jernakoff et al 1997) and salt water (for example, using Sea-viewing Wide Field-of-view Sensor, or, SeaWiFS satellites; e.g., Murtugudde, et al 1999, and many others), but there have been apparently few successful efforts at sensing HABs using satellites (but see Prasad and Haedrich 1993). Although there

has been some thought and effort at sensing HAB pigments using *in situ* and *in vivo* instrumentation (Gentien and Lunven 1993; Millie et al 1995), the goal remains elusive. However, even if such sensor development is far in the future, the fact that satellites can measure dinoflagellate pigments, and that the present expert system can be used with satellite data as well as *in situ* data, and finally, that artificial intelligence has been used in the study of determining chlorophyll concentrations from satellite data (Keiner & Brown, 1999), gives us cause for optimism in the development of an expert system for the early warning of the onset of HABs, such as is currently done for coral bleaching (Hendee et al 2000a).

## References

- Gentien, P; & Lunven, M. 1993. New perspectives in coastal marine environment management due to new development in instrumentation. International Council for the Exploration of the Sea, Copenhagen (Denmark). Marine Environmental Quality Comm., Council Meeting Papers, ICES, Copenhagen (Denmark), 11 pp.
- Hendee, J.C. 1998. An expert system for marine environmental monitoring in the Florida Keys National Marine Sanctuary and Florida Bay. *Proceedings, 2nd International Conference on the Coastal Environment, Cancun, Mexico, September 8-10, 1998.* Computational Mechanics Publications/WIT Press, Southampton, 57-66 (1998).
- Hendee, J.C. 1999. An environmental information synthesizer for expert systems. *Florida Bay and Adjacent Marine Systems Science Conference, Programs and Abstracts*, Key Largo, Florida, November 1-5, 1999. University of Florida Sea Grant, 217.
- Hendee, J.C. 2000 A data-driven soft real-time expert system for producing coral bleaching alerts. *Dissertation Abstracts International, Part B.* (in press) [Nova Southeastern University; Ft. Lauderdale, FL].
- Hendee, J.; Mueller, E.; Humphrey J.C. & Moore, T. 2000a. A data-driven expert system for producing coral bleaching alerts at Sombrero Reef. *Bulletin of Marine Science* (in press).
- Hendee, J.; Berkelmans, R., Sweatman, H., Coleman, G. & Gill, E. 2000b. Expert system generated coral bleaching alerts for Myrmidon Reef, Great Barrier Reef. *Proceedings of the Symposium on Integration in Environmental Information Systems. Zell Am See, Austria; May 30, 2000* (in press).
- Jernakoff, P; Hick, P; Ong, C; Hosja, W; & Grigo, S. 1997. Mapping algal blooms using airborne digital multi-spectral video and the importance of bloom dynamics in the collection of in-water data. *Marine Technology Society Journal* 30(4): 36-45.

- Keiner, L.E. & C.W. Brown. 1999. Estimating oceanic chlorophyll concentrations with neural networks. *International Journal of Remote Sensing* 20: 189-194.
- Millie, D.F., Kirkpatrick, G.J., & Vinyard, B.T. 1995. Relating photosynthetic pigments and *in vivo* optical density spectra to irradiance for the Florida red-tide dinoflagellate *Gymnodinium breve*. *Marine Ecology Progress Series* 120(1-3): 65-75.
- Mockler, R.J. & Dologite, D.G. (1992). *Knowledge based systems. An introduction to expert systems*. New York: Macmillan Publishing.
- Murtugudde, R.G., Signorini, S.R., Christian, J.R., Busalacchi, A.J., McClain, C.R., & Picaut, J. 1999. Ocean color variability of the tropical Indo-Pacific basin observed by SeaWiFS during 1997-1998. *Journal of Geophysical Research. C. Oceans* 104 (C8) 18: 351-18.
- Ogden, J., Porter, J., Smith, N., Szmant, A., Jaap, W., & Forcucci, D. (1994). A long-term interdisciplinary study of the Florida Keys seascape. *Bulletin of Marine Science* 54(3), 1059-1071.
- Prasad, K.S., & Haedrich, R.L. 1993. Satellite observations of phytoplankton variability on the Grand Banks of Newfoundland during a spring bloom. *International Journal of Remote Sensing* 4(2): 241-252.
- Uhrmacher, A., Cellier, F., & Frye, R. (1997). Applying fuzzy-based inductive reasoning to analyze qualitatively the dynamic behavior of an ecological system. *AI Applications* 11(2), 1-10.
- Yacobi, Y.Z., Gitelson, A., & Mayo, M. 1995. Remote sensing of chlorophyll in Lake Kinneret using high-spectral-resolution radiometer and Landsat. *Journal of Plankton Research* 17(11): 2155-2173.

```

(lonf1 airT 23.1 average all-day of day 75)
(lonf1 airT 24.5 average all-day of day 77)
(lonf1 airT 25.1 average night-hours of day 81)
(lonf1 barom 1015 average night-hours of day 81)
(lonf1 barom 1017 average all-day of day 76)
(lonf1 fluoro 0.118 drastic-high morning of day 77)
(lonf1 fluoro 0.124 drastic-high evening of day 75)
(lonf1 fluoro 0.126 average all-day of day 75)
(lonf1 fluoro 0.126 drastic-high midnight of day 75)
(lonf1 parSurf 483 somewhat-low mid-day of day 75)
(lonf1 parSurf 53 very-low midnight of day 75)
(lonf1 parSurf 587 average mid-day of day 78)
(lonf1 parSurf 82 very-low pre-sunset of day 76)
(lonf1 salin1m 35.8 average night-hours of day 78)
(lonf1 salin1m 36.0 average all-day of day 76)
(lonf1 salin1m 36.1 somewhat-high mid-day of day 75)
(lonf1 salin1m 36.4 somewhat-high pre-sunset of day 75)
(lonf1 sealm 25.1 average afternoon of day 76)
(lonf1 sealm 26.0 average all-day of day 78)
(lonf1 sealm 27.0 average night-hours of day 81)
(lonf1 tidelm -9.4 low night-hours of day 81)
(lonf1 tidelm -9.4 low pre-sunset of day 75)
(lonf1 tidelm -9.6 very-low sunset of day 75)
(lonf1 transmiss 2.851 average all-day of day 76)
(lonf1 transmiss 2.852 average all-day of day 78)
(lonf1 transmiss 2.878 average all-day of day 77)
(lonf1 volts 14.6 very-high mid-day of day 75)
(lonf1 volts 14.6 very-high pre-sunset of day 75)
(lonf1 windDir 89 ENE-ESE pre-dawn of day 75)
(lonf1 windDir 94 ENE-ESE mid-day of day 78)
(lonf1 windDir 97 ENE-ESE afternoon of day 78)
(lonf1 windGu 9.5 somewhat-low sunset of day 78)
(lonf1 windGu 9.9 somewhat-low dawn-morning of day 77)
(lonf1 windSp 9.4 somewhat-low morning of day 80)
(lonf1 windSp 9.9 somewhat-low afternoon of day 78)
(lonf1 windSp 9.9 somewhat-low pre-dawn of day 81)

```

**Figure 1.** Example of CLIPS-loadable facts in the HAB expert system. Abbreviations: lonf1=identifier for Long Key, airT=air temperature, barom=barometric pressure, fluoro=fluorometer, parSurf=PAR (see text), salin1m=salinity at 1m, sea1m=sea temperature at 1m, tide1m=tide at 1m, transmiss=transmissometer, volts=voltage of instrumented array, windDir=wind Direction, windSp=wind speed.

```
Rule Flu-Wind-PAR-1 (high fluoro, low wind, high PAR)
(Bloom?)
Day 148
~~~~~
Fluorometry was about 0.073 (high mid-day)
Wind Speed was about 2.9 (very-low mid-day)
Wind Direction was from NE-ENE direction (mid-day)
PAR was about 868 (high mid-day)

Rule Flu-Wind-PAR-1 (high fluoro, low wind, high PAR)
(Bloom?)
Day 147
~~~~~
Fluorometry was about 0.066 (somewhat-high mid-day)
Wind Speed was about 1.4 (drastic-low mid-day)
Wind Direction was from WSW-W direction (mid-day)
PAR was about 862 (high mid-day)

Rule Flu-Wind-PAR-1 (high fluoro, low wind, high PAR)
(Bloom?)
Day 146
~~~~~
Fluorometry was about 0.082 (very-high mid-day)
Wind Speed was about 4.4 (low mid-day)
Wind Direction was from NE-ENE direction (mid-day)
PAR was about 969 (high mid-day)
```

**Figure 2.** File and e-mail output of the HAB expert system.