9

TECHNOLOGICAL INNOVATIONS

Convenors: Dr. Maurice Lynch and Dr. Robert Jaske

Toward a Sustainable Coastal Watershed:
The Chesapeake Experiment. Proceedings of a Conference
1-3 June 1994. Norfolk, VA
Chesapeake Research Consortium Publication No. 149

RESTRUCTURING OF THE CHESAPEAKE BAY PROGRAM'S DATABASE

Claudia P. Walters
U.S. Environmental Protection Agency

Abstract: After a decade of progress, the Chesapeake Bay Program (CBP) has generated a great deal of information about the ecological system and has changed the CBP's management information needs. The cornerstone of the CBP, the Water Quality Monitoring Program, has monitored Chesapeake Bay for nutrients over the past 8 years. Many other monitoring programs have been initiated during the past 10 years, including toxics and living resources. In addition, various historical data sets existed prior to the formal structure. This "measurement" type of data was used to create modeling systems of the environment, used to generate trends, or used in management activities. All of these subsequent activities generate an additional date type called "derived." The CBP is now on the verge of linking environmental information to people and land-based activities that affect the environment, such as population, forestry, and zoning. As a result, the CBP has generated a considerable amount of information and created new management information needs that must be reflected in the data base.

The solution to this data management dilemma is a well-organized collection of data and a standardized, interactive retrieval system for use in responding to the needs of scientists and policymakers around the Bay community. The data management system should permit integrated, cross-media data queries that are not platform-dependent. Queries should not require the services of skilled programmers but should be easily accessible to users. The best way to achieve this is to design and build a true data base using relational data structures. Oracle, the EPA's standard for scientific relational data base management, would be the best tool for implementing a true data base at the CBP. The overall design of the data management system will have Oracle as the relational data base, in which data will be loaded from our data entry and validation program, CAMS (Chesapeake Bay Program Automated Monitoring Data Transfer System) and accessed via COMPASS, a front-end, menu-driven, Windows-type application written by the National Oceanic and Atmospheric Administration .

Conversion of the current VAX VMS system to the Oracle RDBMS system will consist of a planning phase to identify all of the data base needs and involve the participation of various data base design experts. The planning has commenced. The process to design and implement the new data base is explained in this paper. In addition, the status of where we are in that process is identified.

Toward a Sustainable Coastal Watershed:
The Chesapeake Experiment. Proceedings of a Conference
1-3 June 1994. Norfolk, VA
Chesapeake Research Consortium Publication No. 149

USE OF TECHNOLOGICAL INNOVATIONS TO PROVIDE PUBLIC ACCESS TO CHESAPEAKE BAY PROGRAM DATA AND INFORMATION

Joseph Macknis
U.S. Environmental Protection Agency

Brand L. Niemann
National Biological Survey

Abstract: The purpose of this paper is to report progress in the use of technological innovations to provide public access to Chesapeake Bay Program data and information. The technological innovations are: (1) hypertext-hypermedia (Folio Views and Mosaic); (2) exploratory data analysis and visualizations (S-PLUS); and (3) writable CD technology (CD-Rs). They can be used with U.S. Environmental Protection Agencey standard personal computer hardware and software tools. The paper includes sample results of each of these innovations and a list of next steps. The results show how exploratory data analysis techniques can improve visualizations of complex data bases, support decisions on monitoring frequency, and assess achievement of environmental goals. In addition, this paper shows how exploratory data analysis with the Chesapeake Bay Program data set furthers the goals of data integration and multivariate analyses, both of which are next logical applications with the excellent data bases collected by the Chesapeake Bay Monitoring Program. In addition, the presentation includes a description of the CHESYROM CD-ROM, which includes Chesapeake Bay water quality data sets and key guides, reports and documentation, and information on theCD-ROM's availability and use.

INTRODUCTION

In his 4 November 1993 progress report, “Technology for Economic Growth,” President Clinton stated that information technology will be used to dramatically improve the way the federal government serves the people. In particular, government intends to use technology to improve the quality and timeliness of service, to provide new ways for the public to communicate with their government, and to make government information available to the public in an equitable manner (White House, 1993).

The U.S. Environmental Protection Agency(EPA) has been a leader in the federal government in providing public access to environmental data and information. The *1992 Toxics Release Inventory* (U.S. Environmental Protection Agency1994a) available on CD-ROM and in other forms. A Guide to Selected National Environmental Statistics (U.S. Environmental Protection Agency 1993) (available on diskette and the Internet) and the use of writable CD-ROMs (Computer World 1994) are some examples of the EPA’s leadership role.

The Chesapeake Bay Monitoring Subcommittee (MSC) has long recognized the need for public access to its monitoring data to promote broad multidisciplinary interpretation and reporting of results. The MSC has encouraged the use of technological innovations to make the monitoring data and metadata (“data about the data”) more accessible to everyone involved and interested in the Chesapeake Bay Program with access to personal computers (PCs). During the past 6 months the authors have undertaken an effort to make an initial selection of Chesapeake Bay Program data and information accessible in a PC format compatible with EPA standard PC hardware and software tools. The technological innovations described in this paper include hypertexthypermedia (Folio Views and World Wide Web Mosaic), exploratory data analysis and visualizations (S-PLUS), and writable CD technology. The technological innovations and prototype CHESYROM CD-ROM have been explained and demonstrated to the MSC and the agencies and institutions involved in the Chesapeake Bay Monitoring Program in a series of on-site mini-classes (Niemann 1994a).

The purpose of this paper is to report progress in the use of technological innovations to provide public access to Chesapeake Bay Program data and information. The paper includes sample results and a list of next steps.

Some Technological Innovations for Public Access

EPA Standard PC Hardware and Software Tools

The EPA provides its employees with IBM-compatible PCs with 386 and 486 processors, 4-8 MB of RAM, VGA monitors, standard 3.5- inch and/or 5.25 -inch disk drives, and hard drives with storage capacity ranging from about 100 MB to one GB. The EPA's standard PC contract also makes internal or external CD-ROM drives available at low cost. The EPA supports a suite of standard PC software applications that includes word processing (WordPerfect), spreadsheet (LOTUS 1-2-3), database management (dBASE III+ and IV), graphics (various), and communications (various) operating under the DOS and, increasingly, the Windows operating systems.

EPA employees have access to the Internet via the EPA All-In-One Email system for exchange of mail messages and transfer of small files. Some EPA employees also have direct access to the Internet that permits full use of Internet information resource tools such as FTP (File Transfer Protocol), Gopher, World Wide Web, WAIS (wide area information servers), and X.500-Directory Service for information distribution and retrieval worldwide. It is expected that more EPA employees will have the kind of Internet access that permits full use of Internet information resource tools such as Windows Mosaic.

Hypertext-Hypermedia (Folio Views and Mosaic)

With an ordinary book, you typically start reading with page one, followed by page two, then three, etc. The information is ordered sequentially, one sheet after the other. Hypertext breaks out of this linear arrangement. Different chunks of information can be arbitrarily linked together, providing multiple paths through the same information.

When you start up a World Wide Web server Mosaic document on the Internet, you'll notice that certain words and images are highlighted (in blue and underlined, unless you've changed the application's defaults). These are links to other documents. As an example, suppose you are looking at document A. The words "some text" are linked to document B, while the image of the tree is linked to document C. documents B and C could then in turn be linked to other documents, etc. The difference between hypertext and a book is that you can easily ignore information you are not interested in. If you don't care about document B, you would never have to look at it. Furthermore, documents A, B, and C could be on the same or different servers distributed around the world on the Internet. The "Web" is growing explosively. Use of the National Center for Supercomputing Applications Web more than doubles every 4 months. Recent estimates put the Internet Web traffic growth rate at 300,000% per year, which is attributable primarily to the release of the Windows Mosaic client/server in late 1993.

In the PC world, a Folio Views infobase is similar to an Internet Mosaic document in that an infobase is a dynamic, single-file repository of free-format data that can contain text, graphics, and multimedia objects (such as video, audio, and sound files). Infobases are (1) a place to turn for facts you need, (2) a tool to keep those facts current, (3) a utility to publish your information inexpensively, and (4) an effective method to consume information. Infobases can be distributed on diskettes, CD-ROMs, and the Internet.

Actually hypertext-hypermedia was conceived a long time ago as pointed out by Tufte (1990) as follows: "In 1613, when Galileo published the first telescopic observations of Saturn, word and drawing were as one. The stunning images, never seen before, were just another sentence element. Saturn, a drawing, a word, a noun. The wonderful becomes familiar and the familiar wonderful."

By placing drawings of Saturn in mid-sentence, Galileo had begun the tight integration of text and graphics. A modern implementation of Galileo's idea is hypertext and hypermedia where the reader is informed that a link to supplemental information exists and if that link is explored, the reader can then return to the departure point in the text.

Exploratory Data Analysis and Visualizations (S-PLUS)

All statistical analysis has, at its heart, a model that attempts to describe the structure or relationships in some objects or phenomena on which measurements (the data) are taken. The process of developing a statistical model varies depending on whether one follows a classical, hypothesis-driven approach (confirmatory data analysis) or a more modern, data-driven approach (exploratory data analysis). In many data analysis projects, both approaches are frequently used (StaSci 1994).

A new book by William Cleveland of the AT&T Bell Laboratories, *Visualizing Data* (1994), is about visualization tools and a new philosophy of data analysis that stresses a penetrating look at the structure of data. The book is organized around applications of the tools to data sets from scientific studies. Many of his visualization methods are widely used throughout the scientific, engineering, and business communities. Cleveland has developed a new module for S-PLUS, called Trellis Display, that makes it easier to create the new graphics in his book.

S-PLUS is a state-of-the-art, interactive computing environment that provides both a full-featured graphical data analysis system and an object-oriented language. The flexible S-PLUS system can be used for exploratory data analysis, graphics, statistics, and mathematical computing. S-PLUS is the commercial version and a superset of the original S language from AT&T Bell Laboratories available from StatSci. S and S-PLUS are at the leading edge of statistical research, and new developments usually appear sooner than in other statistical software packages. S-PLUS is a large system, with over 1,600 built-in functions and dozens of additional functions stored in libraries. S-PLUS runs under the DOS, Windows 3.1, and UNIX operating systems.

The S+ Interface for developing menu-style interfaces is available, S-PLUS for Arc/Info has just been released, and S+ Wavelets, a wavelet toolkit, and DataViewer, a new multivariate data visualization tool, are in development. StaSci recommends the product Data Junction to convert your data files to and from dozens of popular data bases, spreadsheets, and other applications for use with S-PLUS. S-PLUS requires a 386- or 486-based PC with a math co-processor, Windows 3.1, DOS 3,0 or higher, 8 MB of RAM, and 40 MB of hard disk space. The new S-PLUS 3.2 for Windows has just been released, and it allows users to embed S-PLUS high-resolution graphics in other Windows applications such as Folio Views.

The late W. Edwards Deming said that “visualization retains the information in the data.” Cleveland has observed that no matter how clever the information and no matter how technologically impressive the encoding (e.g., color and depth), a visualization fails if the decoding fails. Many excellent examples of the visualization of natural phenomena are found in the recent book by Wolff and Yaegar (1993) with accompanying animations on CD-ROM. The Virginia Institute of Marine Science(1994) has recently produced a *CD-ROM Digital Atlas of Chesapeake Bay Water Quality* using a volumetric interpolation to produce color-contoured maps and vertical sections at monthly intervals held in 8-bit-deep raster images. MacEachren (1994) has done some similar displays and suggests the use of saturation hues and focus (fuzziness) to display uncertainty and data quality.

Writable CD Technology

Compact discs (CDs) seem to be everywhere these days, especially for listening to music. They have found their way into a wide variety of computer applications because of their capacity, durability, portability, and interchangeability. One might ask, Why isn’t everyone using CDs for data storage and distribution? The answer is that until recently your desktop computer could not write to a CD the same way it can write to a floppy disk or other magnetic media. Instead, CDs have to be mastered and pressed an expensive operation requiring special production equipment. This made CDs expensive to produce and distribute in small quantities, but affordable in larger volumes.

Writable CD products (Eastman Kodak Company 1993) have solved that problem. Once computer files have been prepared, the writable CD process involves only one step taking about 30 minutes while pressed CDs require many steps and usually several days to manufacture. The advantages of writable CDs make them ideal for the following applications: low-volume publishing (large volumes of information to relatively small audiences); record-on-demand; desktop publishing; archiving; and system backup.

The break-even points for writable CDs at 1992 costs was about 100 or fewer disks (Pahwa 1994). The break-even point is expected to rise, as the cost of the writable media (about $20) drops, to about 250 disks in 1995. The cost per MB of writable CDs is estimated to be about $0.04 compared to about $4.54 for paper.

With the number of CD-ROM drives projected to reach 13 million by 1995, this means that more people can read the information you send to them on writable CDs. CD-ROMs in Print (1994) estimates that there are already 3,502 titles, an increase of nearly 20% over 1992! Standards ensure that what you write to a writable CD disk can be read by anyone with standard equipment and software in any of the standard formats.

Writable CDs (CD-Rs) are also called write-once or one-offs. However, write-once is a little misleading. For example with a Kodak writable CD disk you can append information dozens of times in fact each disk can hold up to 99 tracks or updates. Write-once simply means that information on the disk cannot be erased or altered physically. The writable CD system consists of three parts: the media, the writers, and the software.

RESULTS

Infobases

The *Guide to Using Chesapeake Bay Program Water Quality Monitoring Data* (Chesapeake Bay Program 1993b) and *Environmental Indicators: Measuring Our Progress* (Chesapeake Bay Program 1993b) provided the basis for the first two electronic info-bases of Chesapeake Bay Program information. The contents of these first two infobases and the remaining two are presented in figure 1. The user's guides for these four infobases (Niemann 1993a, 1993b, 1993c, and 1994b) are available from the author. These four infobases are included in the CHESYROM CD-ROM subdirectory structure (see figure 1).

The user can open the CBPOGUDE.NFO and search for how a chemical parameter is measured and hyperlink to a map of the mainstem monitoring locations and zoom in on their locations, among other things. The user can open the CBPOMAHA.NFO and select an indicator, read the description, and then hyperlink to the associated graphic, among other things. The user can open the VIZCHAL.NFO and select a monitoring location (e.g., CB1.1) and then hyperlink to time series of various nitrogen water quality parameters at that location. The user can open the CBPODO.NFO and review key documents on dissolved oxygen goals and then select a monitoring location with continuous dissolved oxygen data and display the times series of data at that location by means of hyperlinks on a map. Finally, the user can open the EMAPRIMR.NFO and search for a particular EMAP project or term in the glossary and then hyperlink to selected graphics.

Toxics Release Inventory Trends Analysis

A special Chesapeake Bay toxics release inventory (TRI) data base was developed to document trends in industrial loadings and releases from 1987 to 1991 and was included in the *Chesapeake Bay Basin Toxics Loading and Release Inventory* (Chesapeake Bay Program 1994). The steps in creating the special TRI inventory were as follows:

1. The TRI databases for the six states (DE, MD, NY, PA, VA, and WV) for the years 1987-91 were obtained as dBASE III+ or LOTUS .WKS files from the EPA Office of Pollution Prevention and Toxics on diskettes (U.S. Environmental Protection Agency 1994b).
2. The TRI databases for selected media categories were queried in LOTUS 1-2-3 R3.1+ for the subset of counties within the Chesapeake Bay watershed, for all toxic substances, for just the 126 priority pollutants, and for just the 14 “Toxics of Concern”.
3. The results of the queries were summarized and graphed in LOTUS 1-2-3 R3.1+.

Estimates of releases and discharges of sodium sulfate and aluminum oxide were not included in the analysis because those substances have low risks and are no longer reported to the TRI. In addition, the extremely large loadings of these substances would mask more important trends. The complete documentation of the methodology is included with the data bases and spreadsheets on the CHESYROM CD-ROM.

The Chesapeake Bay basin industry-reported releases and transfers of Chesapeake Bay Toxics of Concern shown in figure 2 account for 6% of all toxic substances reported to the TRI by Chesapeake Bay industries in 1991. Total reported releases and transfers of Chesapeake Bay Toxics of Concern increased from 1988 to 1990 and then declined sharply in 1991. Air and water release have remained fairly constant in quantity since 1987 and accounted for 7% and 3%, respectively, of the 1991 total reported releases and transfers of Chesapeake Bay Toxics of Concern.

The level of uncertainty associated with TRI-reported estimates is unknown. The individual reporting facility develops estimates of releases and transfers of toxic substances. Comparisons of these industry-reported surface water discharges with NPDES-calculated loads for common toxic substances show good agreement, strongly suggesting that TRI-reported loads to surface-waters have similar confidence levels.

CBP Interpolator Output and Continuous Dissolved Oxygen Data

The interpolator output (Bergstrom and Olson 1994,Chesapeake Bay Program 1993b) in the form of SAS files on the Chesapeake Bay Program Office VAX computer were converted to ASCII format and

Figure 1. Folio views infobases.

Air = release through smokestack or as fugitive emission; Water = surface water discharge; Land = landfill disposal; POWT = transfer off-site to publicly owned water treatment works; Others = off-site for treatment, storage, and/or disposal.(source: U.S. Environmental Protection Agency 1994a).

Figure 2. TRI trends analysis.

downloaded to PC diskettes. The ASCII files were imported into LOTUS 1-2-3 release 3.1+ for formatting for input to S-PLUS. Of the thirteen SAS files available, only six were used in this initial analysis, namely dissolved inorganic nitrogen (DIN), dissolved inorganic phosphorous (DIP), dissolved oxygen (DO), total nitrogen (TN), total phosphorous (TP), and Secchi depth (SD). The formatting in LOTUS created four different versions of the three data bases as follows:

1. Reported cruises for the entire period.
2. Just the first of the month cruises for the entire period;
3. Twice monthly data for the entire period with the unsampled mid-month periods interpolated from the monthly data; and
4. Twice monthly data for the entire period with the unsampled mid-month periods designated with NA for “not available”.

Recommended methods for statistical analysis of water quality data (see Helsel 1992, Helsel and Hirsch 1993) can be performed with S-PLUS (see Niemann submitted, 1994c). Figure 3 shows an example of the S-PLUS scripts, graphics, and interpretations accompanying each result in the earlier reference (Niemann 1994c) for figure 4.

The QQplots in figure 5 compare the rank order distributions of the DO concentrations for monthly-only cruises (12 cruises) with the complete data base (20 cruises per year). The QQplot shows how the monthly cruise data would underestimate the “true” DO concentrations from more complete data (8 more cruises per year).There is real power (benefit) to the 8 additional cruises to resolving the frequency of DO concentrations in the interpolator data base; however, the additional power is not in the range of DO values of interest to the DO goals.

The scatter plot matrix of interpolator DO concentrations for eight of the Bay segments in figure 6 shows, as expected, that the spatial correlation decreases as the separation increases between segments as one moves from the diagonal to the upper right-hand and lower left-hand sides of the matrix. Exploratory data analysis techniques like the scatter plot matrix can provide spatial analyses and statistics without the use of maps or at least complement and supplement the information on GIS maps.

The notched box plots of interpolator DO concentrations in figure 7 show that Bay segment CB4 has the lowest DO concentration overall, and it is not significantly different from the two adjacent segments CB3 and CB5, but is

Figure 3. Example of S-PLUS scripts, graphics, and interpretation in Niemann (1994c).

Figure 4. CBP interpolator output: Times series 12 versus 20 cruises.

Figure 5. CBP interpolator output: QQ plot for dissolved oxygen.

Figure 6. CBP interpolator output: Scatter plot matrix for dissolved oxygen. A box plot is a simple graphical representation showing the center and spread of distribution, along with a display of unusually deviant data points, called outliers. Box plots also indicate skewness as well. The most striking visual feature is the box that shows the limits of the middle half of the horizontial line in the interior of the box os located at the median of the data. This estimates the center of the distribution for the data. The height of the box is equal to the interquartile distance, IQD, which is the difference between the third quartile of the data and the first quartile. The IQD indicates the spread or width of the distribution of the data. The whiskers (the dotted line extending from the top and bottom of the box) extend to the extreme values of the data or a distance 1.5 X IQD from the center, whichever is less. For data having a Gaussian distribution, approximately 99.3% of the data falls inside the whiskers . Data points that fall outside the whiskers may be outliers, and are so indicated by horizontal lines The option varwidth = T means the box widths will be proportional to the square root of the number of observations for the box.The option notch = T means notched boxes are drawn. If the notches on two boxes do not overlap, this indicates a difference in the location (median) at a rough 5% significance level. Normal QQplots should also be used to check for departures from normality in the data.

Figure 7. CBP interpolator: Notched box plots for dissolved oxygen.

significantly different from the other seven segments in the mainstem of the Bay because the “notches” for the latter do not overlap.

It was found previously that various fits or smoothing methods give different trends for individual Chesapeake Bay stations (Niemann 1993c). In contrast, a very robust method (least square mean regression), which essentially ignores all the outliers and oscillations, shows essentially a straight line (no trend). In essence, trends are very sensitive to the smoothing or fitting method, period of record, and natural variability. A similar results was obtained for the interpolator output (see figure 8).

Figure 8. CBP interpolator output: Trends analysis for nitrogen.

The scatter plot matrix may also be used for modeling relationships between parameters as shown in figure 9, where water quality parameters have been combined with a living resource parameter, namely hectares of submerged aquatic vegetation by Bay segment. The results give rise to speculation about relationship or the lack thereof between the parameters based on spatial variability. Finally, the scatter plot matrix was also used to visualize the continuous dissolved oxygen data (see figure 10,) which shows considerable overplotting and generally good agreement between temperature and DO saturation as expected, in that one is calculated from the other.

CD-R Subdirectory Structure

Information on a CD-ROM is structured the same way as on a PC hard disk, namely as a hierarchy of subdirectories as shown in figure 11 for the CHESYROM CD-ROM. Each subdirectory contains a README file explaining the files. The subdirectories with infobases can be run by typing the name of the infobase (e.g., CBPOGUIDE) and ENTER. The rest of the subdirectories require applications software on the PC to retrieve and use (e.g., LOTUS 1-2-3, dBASE III+, and S-PLUS). The largest subdirectory by far is \MAINSTEM, which contains over 200 files totaling about 100 MB in size of the historical mainstem monitoring data. As explained previously, the writable CD-R can accommodate additional subdirectories in the future and updates to the information.

CD-R Review and Distribution

The plan is to distribute the initial copies of the CHESYROM CD-Rs to the participating agencies and institutions for review and verification of their particular portion of the mainstem monitoring data base. Upon incorporation of any changes or additions to the CHESYROM CD-R, it can then be distributed to a larger group to satisfy requests for PC access to the data and information. Any changes and updates to the mainstem or other data bases can be handled by writing them to the same individual CD-Rs in that multisession CD-Rs can be written to up to 99 times. The cost of the CD-R media (about $20) and their storage capacity (large) and mailing costs (low) are such that it is best to have the CD-Rs returned for writing updates and even custom requests for other data bases.

Additional Statistical Analyses and Advanced Visualizations

Some logical next steps for exploratory data analysis and visualization to consider are as follows:

1. QQ plots can be produced for the special continuous DO data in segment CB4 and the “composite” of twice monthly samples for all cruise sites in CB4 to produce a meaningful QQ plot for the DO range of interest to the DO goals.
2. The VIMS interpolator output could be analyzed using the same basic S-PlUS EDA used for the interpolator data base this would serve as a check on the consistency sensitivity in both results.
3. Selected sites in the mainstem data base could be analyzed using the same basic S-PLUS EDA used for the interpolator data base this would serve as a check on the consistency/sensitivity in both results.
4. Selected outputs from the CBP 3D model could be analyzed for comparisons with the interpolator and mainstem data S-PLUS outputs this would serve as a check for consistency/sensitivity in both results and guidance for the refinement effort to supply model inputs and evaluation outputs for the 3D model.

5. Additional multivariate analyses (S-PLUS scatter plot matrices, coplots, etc.) could be made on available data bases (e.g., Citizen's Monitoring CBP/TRS 71 92) and on data bases built by merging available water quality, flow, loading, living resources, climate, EMAP, and other data bases.

With the mainstem monitoring data base organized in PC format on CD-ROM, it is now possible to use advanced visualization tools such as PV-WAVE and AVS (advanced visualization system).

Figure 9. CBP interpolator output: Scatter plot matrix with fitting.

Figure 10. CBP continuous dissolved oxygen: Scatter plot matrix.

Figure 11. CHESYROM CD-ROM subdirectory structure.

World Wide Web Server

The Chesapeake Bay Program data and information could eventually be served to the world using a World Wide Web server on the Internet. Locke (1994) in *Information Week* stresses how the Web can enhance global communications and make a virtual corporation possible. Infobases and the CHESYROM CD-ROM are logical first steps for organizing the Chesapeake Bay Program data and information and adding value through improved access, navigation, and retrieval leading to the serving of that information on the Web.

ACKNOWLEDGMENTS

The authors acknowledge the guidance and support of Dan Carr, Art Koines, Douglas Martin, Jerry McFaul, Phil Ross, John Williams, and the Chesapeake Bay Program Office and EPA Region III. The authors also express appreciation for assistance with data bases from John Posey, Peter Bergstom, Marcia Olson, Jim Roberts, Kent Mountford, Cynthia Stenger, Mike Mullen, Tom DeMoss, Carol Finch, Eric Hyatt, Ed Martinko, Bruce Potter, Jeff Rosen, Bob Shepanek, and Dan Vallero. One of the authors (Niemann) acknowledges the assistance of the Eastman Kodak Company for the loan of its PCD Writer 200 Plus and technical assistance.

DISCLAIMER

This paper has been produced as part of the author's work for the U.S. Environmental Protection Agency, but has not received formal agency review so no formal endorsement of its contents or of the commercial software mentioned should be inferred.

REFERENCES

Bergstrom, P., and M. Olson, 1994. Trends in hosphorus, nitrogen, secchi depth, and dissolved oxygen in the Chesapeake Bay, 1984 1992. CBP/TRS 90/93, Prepared by Computer Science Corp. for the Tennessee Center for Research and Development, Cooperative Agreement No. TCRD-93-08-01-000, draft for review.

CD-ROMs in Print. 1993, An international guide to CD-ROM, CD-I, CDTV, multimedia & electronic book products., London: Meckler, 1994.

Chesapeake Bay Citizen Monitoring Program. Report on Conestoga River (October 1986 - June 1990). CBP/TRS 71/92.

Chesapeake Bay Program. 1993a: Environmental indicators: Measuring our progress. Version 13, 46 pp.

Chesapeake Bay Program. 1993b: Guide to using Chesapeake Bay Program water quality monitoring data. Annapolis, MD.

Chesapeake Bay Program. 1993C: Volume tric interpolator: Users Guide. Annapolis, MD. Version 2.

Chesapeake Bay Program. 1994. Chesapeake Bay Basin toxics loading and release inventory, Basinwide toxics reduction strategy commitment Rep. CBP/TRS 102/94.

Cleveland, W.S. 1994. Visualizing data, Summit, N J: Hobart Press.Computer World. 1994 17, January. EPA gives public free access to its data. ComputerWorld, January 17, 1994: EPA gives public free access to its data.

Eastman Kodak Company, 1993. Writable CD Products.

Helsel, D.R., and R.M. Hirsch, 1992. Statistical methods in water resources. Elsevier Publishers, Amsterdam: 522 pp.

Helsel, D.H. 1993. Statistical analysis of water-quality data, In: Hydrology and stream water Quality: National water summary 1990-91. U.S. Geological Survey, Water Supply Paper No. 2400. 93 - 100.

Locke, C. 1994. Untangling the Web: The World Wide Web can enhance global business communications and make the virtual corporation possible. Information Week, 9 May p. 92.

MacEachren, A.M. 1994. Some truth with maps: A primer on symbolization & design. DC: Association of American Geographers, 1710

Niemann, B.L. 1993a. Chesapeake Bay and Mid-Atlantic Highlands regions: Environmental problem area fact sheet, indicator bulletin and infobase user's guide. Washington, D.C: U.S. EPA, 20 pp. (available from the author).

Niemann, B.L. 1993b Integration of metadata and space-time-depth variabilities and uncertainties in Chesapeake Bay monitoring data for improved visualization and spatial analysis. paper presented at the GIS/LIS'93 Annual Conference and Exposition in Minneapolis, MN: November 1993.

Niemann, B.L. 1993c Key CBPO Dissolved oxygen reports, databases, and analyses infobase. Washington, DC: U.S. EPA, 14 pp. (available from the author).

Niemann, B.L. submitted. Display of environmental data for EPA management: Putting it all together – but nicely.

Niemann, B.L. 1994a. Visualizing environmental data bases with S-PLUS, mini-class outline and notes.

Niemann, B.L. 1994b. The EMAP Primer Infobase. Washington, DC: U.S. EPA, (available from the author).

Niemann, B. 1994c. Refinement of the Chesapeake Bay Monitoring Program: Exploratory data analysis of the interpolator output. Unpublished paper and mini-class syllabus. Pahwa, Ash. 1994: The CD-recordable Bible, Wilton, C: Eight Bit Books.

StatSci, A Division of MathSoft, Inc. 1994 S-PLUS for Windows version 3.2 supplement.

Tufte, E. R. 1990. Envisioning information. Cheshire, CT: Graphics Press.

U.S. Environmental Protection Agency. 1993.A guide to selected national environmental statistics in the U.S. government, EPA Publ. EPA 230-R-93-003.

U.S. Environmental Protection Agency. 1994a. 1992 Toxics release inventory: Public Data release. EPA Publ. No. 745-R-94-001.

U.S. Environmental Protection Agency. 1994b. Toxic chemical release inventory state diskette documentation.

Virginia Institute of Marine Science. CD-ROM Digital Atlas of Chesapeake Bay Water Quality.

White House. 1993 The development of a communications network program between the federal government and the American public: Americans communicating electronically (ACE), 6 pp.

Wolff, R.S. , and Yaeger, L. 1993. Visualization of natural phenomena. Telos the electronic library of science. New York: Springer-Verlag.

Toward a Sustainable Coastal Watershed:
The Chesapeake Experiment. Proceedings of a Conference
1-3 June 1994. Norfolk, VA
Chesapeake Research Consortium Publication No. 149

GPS/GIS APPLICATIONS—AN INTEGRATED APPROACH TO SHORELINE STUDIES

Harry E. Berquist
Virginia Institute of Marine Science

Abstract: In cooperation with the National Park Service, a unique opportunity was presented to integrate the technologies of GIS, GPS, and remote sensing to map and inventory shoreline position and vegetation assemblages along the Cape Hatteras National Seashore. Two GPS campaigns were required. First, rectification data for high-altitude color infrared photography had to be obtained for the vegetation mapping component. Second, a survey of the high water shoreline was desired to compare the present-day shoreline position to previous surveys conducted. These data are to be imported into a GIS data base to develop a coastal risk assessment model that addresses the potential impacts of sea level rise on natural and cultural resources. As well, the data are to be integrated into the various GIS-based inventory and monitoring programs being implemented by the park. This paper reports on the planning and implementation of the GPS fieldwork and the application of the results.

Modifications of standard GPS procedures were required. In particular, the resolution of the scanned images (1 -2 meters) precluded using code-phase derived coordinates for georeference points. Without access to the latest dual-frequency, p-code, “swift static” receivers, an alternative was tested in which very short observations using L1 phase-measurement receivers were conducted, The results were surprising. Submeter and often decimeter positions were readily obtained. The geography of the region was conducive to using stop-and-go kinematic surveying for many points. The position of the high-water line was measured in situ on a falling tide using continuous kinematic techniques with the receiver mounted on a 4x4 vehicle.

Site selection and cataloguing of georeference points for image rectification required supporting large scale photography (1:32,000) of the area to augment the small-scale imagery (1:58,000) being rectified. Survey results from the georeference points were imported into the ERDAS image processing software to complete the rectifications. Results from the continuous kinematic shoreline survey were imported into an Arc/Info GIS data base to conduct a comparative analysis with existing digital shoreline surveys.

Toward a Sustainable Coastal Watershed:
The Chesapeake Experiment. Proceedings of a Conference
1-3 June 1994. Norfolk, VA
Chesapeake Research Consortium Publication No. 149

RESUSPENSION BEHAVIOR IN THE LOWER CHESAPEAKE BAY

J. P. -Y Maa and C. -H. Lee
Virginia Institute of Marine Science

Abstract: Using the VIMS Sea Carousel to conduct in-situ experiments in lower Chesapeake Bay, we found significant spatial and temporal difference of the critical bed shear stress for sediment resuspension, t_{cr}. At the Wolftrap site, t_{cr} varied from 1.1 Pa in the summer to 1.9 Pa in the winter. At the Burwell Bay site, although the bed was too soft for anchoring our research vessel against the changing tide, we identified that t_{cr} was 0.045 Pa. At the Old Plantation site, there was a small amount of fluffy material on top of the bed. At a winter deployment at the Cherrystone site, we found a layer of consolidating fluffy sediment at the surface. Within this layer, there was a transition from fluid mud (without erosion resistance) to bed (with erosion resistance).

A typical "type I" behavior of sediment resuspension has been identified at all sites. For this type of behavior, the resuspension rate decreases with time for a given constant bed shear stress larger than t_{cr}. This implies that the erosion resistance increases with depth. The natural sediments behave like cohesive sediments because of the biochemical processes.

INTRODUCTION

When studying the sediment resuspension process, one needs to know the following parameters: the bed shear stress, t_b, the critical bed shear stress for sediment resuspension, t_{cr}, the resuspension behavior, and the resuspension rate, E. Recent advances in the study of hydrodynamics and computing resources provide a better estimation of t_b. However, no such progress has been made toward a better understanding of the other parameters. This is because of the complexity of natural sediments and associated biochemical processes in natural environments.

Natural estuarine sediments are usually a mixture of many different grain sizes and involve some biochemical processes. Their resuspension behavior changes with time as well as location because of the different physical, chemical, and biological activities. Collecting sediment samples and conducting laboratory experiments are usually not sufficient because of the difficulty in duplicating natural beds (Maa and Lee in review) and controlling biochemical processes. At the Virginia Institute of Marine Science(VIMS), we envisioned these difficulties, and thus developed the VIMS Sea Carousel (Maa et al. 1993) for scrutinizing the sediment resuspension process. Although this study is still in its early stages, the preliminary results show a quite different critical bed shear stress for sediment resuspension in lower Chesapeake Bay.

The VIMS Sea Carousel is an annular flume operating on the same principle as other annular flumes used in many laboratories (Fukuda 1978 Wainright 1988, Burt and Game 1985). The most significant difference is that the flume is designed specifically for field applications. Excepting some details, the VIMS Sea Carousel is similar to the Sea Carousel developed in Canada (Amos et al. 1992). Figure 1 shows the flume at a deployment. Two cylinders of 0.2 m height with diameters of 2.0 m and 2.3 m form the inner and outer walls of the flume. The two cylinders have a sharp edge so they can penetrate the seafloor when lowered from a boat. Two bearing plates stop the penetration at a selected elevation to give a 0.1 m channel depth. A ring, driven by a DC motor and controlled by a shipboard motor controller, rotates at the top of this channel at selected constant speeds to generate flows for sediment resuspension. Thus, an enclosed annular channel is formed with the natural substrate as the floor. Reasonably good seals between the rotating ring and the sidewalls permit the use of an OBS (Downing et al. 1981) to measure the concentration of suspended matter within this flume, and the total eroded sediment mass. Other details of the VIMS Sea Carousel can be found in Maa et al. (1993).

The rotating ring provides the driving shear force at the top boundary. This shear force induces a tangential flow as well as a secondary flow. This secondary circulation is beneficial for studying resuspension processes because it maintains a reasonably uniform suspension of eroded sediment in the flume. A numerical model study (Maa 1993) and laboratory experiments (Maa et al. in review) showed that reasonably uniform bed shear stresses could be achieved throughout the flume. The relationship between the spatially averaged bed shear stress and ring speed for a 0.1 m channel depth is given as $t = 0.0114\ \acute{y}^{1.693}$, where t is the spatially averaged bed shear stress given in Pa (N/m^2) and ý is the ring speed in rpm.

RESULTS OF FIELD EXPERIMENTS

To date, we have conducted several experiments at the following sites: Burwell Bay, Wolftrap, Old Plantation, Cherrystone, and Duck experimental site, N.C. Except for the last location, all are in lower Chesapeake Bay. Figure 2 shows these locations.

Wolftrap Site

We have conducted three experiments at the Wolftrap site (lat. 37½16'07", long. 76½09'52"). Details of the first two experiments have been reported elsewhere (Maa et al. 1993). The mean water depth at this site was 11.6 m. The tidal range was about 1 m. At this site, tidal currents were the major force of erosion except during storms. The measured maximum tidal current at 1 m above the sea floor was about 0.2 m/sec (Wright et al. 1992). The top 5 mm sediment sample collected during the last experiment (14 May 1992) at this site indicated that the sediment was composed of fine sand (2%), very fine sand (53%), silt (40%), and clay (5%). The median grain size was about 0.07 mm.

The first experiment was conducted on 19 June 1991. The bed shear stress was increased incrementally from a minimum of 0.025 Pa to a maxi-

Figure 1.VIMS Sea Carousel (after Maa et al. 1993).

mum of 0.34 Pa over a period of 140 minutes (Figure 3a). For the first six bed shear stresses, the concentration of suspended matter in the flume was practically unchanged. These values may reflect the original ambient concentration of suspended matter. It is clear that the resuspension began when t_b reached 0.19 Pa. When t_b further increased, more sediment was suspended.

Figure 2. The experiment sites.

Figure 3. Results from the Wolftrap experiments. (a) First experiment. (b) Second experiment.

The second experiment was conducted on 8 October 1991. Figure 3b shows the results. The concentration data show that the sediment was resuspended when $t_b = 0.11$ Pa (or more liberally, at $t_b = 0.09$ Pa). Thus we may say that $t_{cr} = 0.11$ Pa. This critical bed shear stress was considerably lower than that obtained from the first experiment, $t_{cr} = 0.19$ Pa.

The last two experiments on this site were conducted on 14 May 1992. The results are given in Figure 4. The concentration data shown in Figure 4a indicate that t_{cr} is about 0.14 Pa. (figure 4b.)shows the results of a resuspension test. We see a gradually decreasing concentration of suspended matter for each given t_b, with the rate of decrease increasing with t_b. This may be due to the leaking rate increasing with ring speed (i.e., t_b). If the water leakage is considered as a constant for a given ring speed, we can envision that the resuspension rate would decrease with time. We will discuss this issue later.

Figure 4. Results from the third wolftrap experiments. (a) For identifying t_{cr}; (b) A resuspension experiment.

Notice that the t_{cr}'s are quite different for the three experiments. The reason of these differences may come from bioturbation, the seasonal differences in biochemical process (Brekhovskikh et al. 1991), or spatial sediment heterogeneity. Further investigations are under way to address this issue.

Figure 5. Photo images of the seafloor at the Wolftrap Site. (a) A relatively smooth surface before experiment with a burrowing sea anemone in the flume. (b) After experiment, the surface is rough because of the exposure of polychaete tubes, amphipod tubes, and phoronid worms. (After Maa et al. 1993).

Figure 5 shows two photo images obtained from the third deployment. They clearly show that the original, relatively smooth surface Figure 5a was considerably roughened by erosion Figure 5b. The abundance of bioactivities shown in Figure 5b. further demonstrate the complications of natural sediments and the difficulty in duplicating the bed in laboratories to emulate the associated geobiochemical processes.

Burwell Bay Site

We only have a partially successful experiment at this site (lat. 37½2' 34", long. 76⁄36’4"). The analysis of Burwell Bay sediment samples reveals that the top 2 cm of sediment consists of 8% sand, 61% silt, and 31% clay. Organic content is about 6%. This site has a moderate sediment accumulation rate of 1 - 3 cm/ yr (Nichols 1972). The faunal characteristics at this site have been studied by Schaffner et al. (1987). For this soft muddy sediment, we selected a small increment of bed shear stress because we expected the critical bed shear stress to be small.

Figure 6 shows the experimental results. The initial spikes might have been caused by a rapid dispersion of a small pocket of fluid mud. Because of mixing within the flume, the suspended sediment concentration decreased sharply over the first 3 minutes. It then decreased at a much slower rate until the bed shear stress was increased to 0.027 Pa. At that time, more sediment was suspended and the concentration records showed a small “bump.” The constant flow shear stress that followed however, was insufficient to maintain the same concentration because of small leakage. When t_b = 0.048 Pa, a clear, sustained resuspension began at this bed shear stress. Thus, we can estimate that t_{cr} is about 0.045 Pa. This is a much lower value when compared with data obtained from the other sites. Possible reasons may include the large content of silt and clay and strong bioturbation.

Figure 6. Experiment for identifying t_{cr} at Burwell Bay.

Old Plantation Site

We conducted two experiments at the Old Plantation site (lat. 37½12’34", Long. 76⁄06’19") on 15 May 1992. At this site, the water depth was 10.7 m, and the tidal current was about 2 m/sec at the surface. Grain size analysis of the top 5 mm sediment samples showed that the sediment was mainly a mixture of very fine sand (77%) and silt (18%). The mean grain size was 0.067 mm. Figure 7a shows the results of the first experiment. At this site, we noticed that even at the first t_b, 0.06 Pa, sediment was stirred up. This may represent the dispersion of surficial fluffy particles. This response was further enhanced when t_b = 0.07 Pa. After reaching its maximum at t ³ 20 minutes, however, the concentration decreased with time and did not show any increase when we further increased t_b to 0.09 Pa and 0.12 Pa. The concentration raised again when t_b = 0.14 Pa. This behavior may be interpreted as follows. There was small amount of fluff on top of the bed. It was easily dispersed by the given small t_b, and because of leakage, it was completely dispersed in 30 min. The newly exposed sediment could resist some shear stress, and thus it did not response to the following t_b (i.e., 0.09 Pa and 0.12 Pa). The critical bed shear stress, t_{cr}, for this bed could be recognized as about 0.14 Pa.

Figure 7b shows the results of a resuspension test. The spikes occurring around t ³ 20 min may have been caused by a fish trapped in the flume. Later, divers found an escape trench dug out from the flume, presumably by the trapped fish. If we remove the spikes, we see a gradually decreasing concentration of suspended matter. For the first 10 min at the next t_b, 0.56 Pa, the rate of concentration decrease was higher, and possibly caused by a high leaking rate at the high ring speed. (We discuss this issue later.) After changing tidal direction at t ³45 min, the strong tidal current dragged the R/V *Bay Eagle* away from her anchoring location and tilted the flume. Thus, we aborted the experiment.

Cherrystone Site

We conducted two experiments at this site (lat. 37½14’04", long. 76⁄05’22") on 14 February 1994.

This site was not far from the Old Plantation site. The water depth was 13 m, and the maximum tidal current was about 0.3 m/sec at 1 m above the seafloor. Grain size analysis of the top 5 mm sediment samples showed that the sediment was mainly a mixture of very fine sand (50%), silt (33%), and clay (17%). The mean grain size was 0.063 mm. Figure 8a shows the results of the first experiment to identify t_{cr}. Even the smallest bed shear stress, 0.02 Pa, was able to stir up sediment. Although there were some fluctuations, the concentration of suspended matter always increased with the bed shear stress. We believe this response was caused by a consolidating layer of fluff on top of the consolidated bed. This fluffy layer may have developed some erosion resistance, and thus showed a gradual dispersion in 72 mi. There is no clearcut way of identifying t_{cr}. If we count this fluffy layer, then the t_{cr} is too small to be measured. If we exclude this layer, our best estimate for t_{cr} is about 0.12 - 0.13 Pa. This experiment is unique because it is the first time we have observed a layer of consolidating sediment in the lower Chesapeake Bay.

Figure 8b shows the results of a resuspension test. When t_b was 0.24 Pa, the concentration of suspended matter increased to 0.53 g/l and maintained that level until we increased t_b to 0.46 Pa. At the new t_b, the concentration reached a new level, 1.75 g/l, but decreased right after it reached the maximum. At the next t_b, 0.67 Pa, the same behavior was observed. At the last bed shear stress, t_b = 0.93 Pa, however, the concentration maintained at 3.6 g/l until we reduced t_b to 0.67 Pa. The two different behaviors may represent two different scenarios for sediment resuspension and leakage. We will now discuss this further.

Figure 7. Results from the Old Plantation site. (a) For identifying t_{cr}. (b) A resuspension experiment.

Figure 8. Results from the Cherrystone site. (a) For identifying t_{cr}. (b) A resuspension experiment.

CONCLUSION

Conducting field experiments for studying sediment transport is probably the most realistic approach because it has a minimum disturbance of the complex natural sediments and associated biochemical processes. To provide a controllable erosion force for field experiments with minimum resources, however, is the most difficult part in the design of a field instrument. At the VIMS, great teamwork led by the director for research and advisor service, Dr. R. Byrne, successfully developed the VIMS Sea Carousel. With this field instrument, we can bring the laboratory to fields and provide a wellstudied, controllable force to study sediment transport.

Our field experiments conducted in lower Chesapeake Bay demonstrated a great range of t_{cr}, which varied from 0.045 Pa to 0.19 Pa for different types of bed. Except at the Cherrystone site, we could clearly identify t_{cr}. At the Cherrystone site, a layer of fluffy material existed, preventing the observation of t_{cr}.

Notice that t_{cr} is not the same bed shear stress responsible for the incipient motion of sediment, t_{ci}. This is because our OBS sensor was mounted 5 cm above the bed. It would not sense the incipient motion,(e.g., rolling) of sediment particles at the bed surface. Only when the bed shear stress is large enough to entrain these particles at least 2-3 cm above the bed, is the OBS is able to sense them. For this reason, we know that t_{cr} must be larger than t_{ci}. Our laboratory experiments also found that the difference increases with grain size. For example, $t_{cr}{}^{3} 2 t_{ci}$ for very fine sand. For clay and silt, the difference is not significant (Maa and Lee in review).

Fluffy material can be dispersed into water columns even when t_b is small. After the fluffy material is dispersed, the true bed material is exposed. At the Burwell Bay site, we might have had a pocket of fluid mud or fluff. At the Old Plantation site, we probably had a small amount of fluff. At the Cherrystone site, we had a consolidating layer of fluff. This layer may have been changing from fluid mud to bed. At the Wolftrap site, however, we saw neither fluid mud nor fluffy sediment.

In our resuspension experiments, the concentration of suspended sediment always increased with the applied t_b during the first 7-9 min (if the duration for ring speed acceleration is 5 min). This is referred as the raise period. After this raise period, most experiments showed the concentration decrease with time. Only a few experiments showed a nearly constant concentration for suspended matter. Possible reasons for this behavior may include:

- Leakage from the flum.
- Initial oscillation of t_b caused by changing ring speed in a relatively short period.
- Decrease of the skin friction shear stress be cause of the increase of bed roughness by sediment resuspension.
- Deposition of large sediment particles within low bed shear stress areas in the flum.

Our early experiments had a relatively large fluid acceleration at the beginning of each constant t_b. This is because the duration for changing ring speeds from a low speed to the next higher speed was short: one min This short duration may cause t_b to oscillate and resuspend more bed material (Fukuda 1978, Maa et al. 1993). After the oscillatory t_b dissipated, the concentration of suspended materials also decreased. We have revised the operation procedures for our later experiments by giving 5 min for the ring to gradually change from one speed to the next. Under this new operational condition, we may assume that there is no more oscillation of bed shear stress.

Our auxiliary laboratory experiments conducted using a laboratory version of the VIMS Sea Carousel and sediment samples collected at the Wolftrap site indicated that the concentration of suspended matter does not decrease with time at all (Maa and Lee in review). This rules out the contribution from the above-stated third and fourth possible reasons. Thus, the observed variation of concentration for suspended matter should represent the net of sediment resuspension and leakage. When the resuspension rate is larger than the leakage, as happened in the raise period of a new t_b, the concentration increases with time. If these two rates are equal, then the concentration remains constant. In cases when the resuspension rate is smaller than the leakage, the concentration decreases with time.

Leakage is hard to prevent completely, especially for fine sediments, simply because of the large flume dimensions and the rotating ring. For coarse material (e.g., fine sand), the effects of leakage is probably negligible because leakage occurred at the top of the flume, where sediment concentration is low. For experiments conducted in this kind of bed, we see that the concentration of suspended matter can be reasonably maintained as a constant for a constant t_b (Maa et al. 1993). For sediments containing fine grain sediments, the concentration of suspended matter in the flume usually increased and reached a maximum during the raise period and then decreased.

We may choose to correlate the erosion rate as a function of (1) the excess bed shear stress, $t_{ex} = t_b - t_{cr}$, where t_{cr} may be a function of depth (Parchure and Mehta 1985), or (2) time (Lavelle et al. 1984). For the first approach, a detailed vertical profile of the bulk density for the top 0.5 cm is needed. This is difficult because a reliable method to measure this profile is not available yet. We are trying an acoustic approach for resolving this problem; however, it is still too early to make any comments. To correlate the erosion rate as a function of time, our findings will suffer from the drawback that the correlation would be site dependent, and could not be applied elsewhere. For the immediate future, however, the second approach seems to be the only feasible one, and we are working on this issue. At this time, we will concentrate on the interpretation of resuspension behavior.

The resuspension behavior observed can be explained if the resuspension rate decreases with time for a given constant t_b that is larger than t_{cr}.

For cohesive sediments, this behavior has been noticed a long time ago as "type I" response. Because of the biochemical processes, the sediments at our experiment sites all behave like cohesive sediments. When applying a new constant t_b, the erosion rate is high at the beginning because of the large t_{ex}. The erosion rate decreases with time because t_{ex} is decreasing for an increasing t_{cr} when erosion proceeds. In other words, the bed shows increasing erosion resistance with depth.

For all of the experiments, the resuspension rate is larger than the leakage at the raise period because of the large t_{ex}. Thus, the concentration always increases during this period. In most of the experiments, because the resuspension rate decreases with time and is smaller than the leakage after the initial raise period, the concentration decreases. In some cases, the concentration remained constant because the resuspension and leakage are balanced. This may be because t_{cr} is a constant and t_{ex} does not decrease while erosion proceeds.

Although all the beds show a typical type I behavior of sediment resuspension, the variation of t_{cr} with depth can be clearly identified by the different time series of suspended sediment concentrations. We are working on identifying the resuspension rate.

ACKNOWLEDGEMENTS

Sincere appreciation goes to Dr. R.J. Byrne for his initiation of this project as well as his generous support. We also thank Drs. L.D. Wright, A. Y. Kuo, and A.J. Mehta for their encouragement on the development of the Carousel. The great diving work from Mr. T. Shannon and R. Gammisch is gratefully acknowledged. Thanks also go to the captain of VIMS R/V *Bay Eagle,* Mr. D. Ward, and his crew for their beautiful three point anchorage. We also thank Mrs. S. Alexander, Mr. W. Matthews, and Mr. T. Nelson for their help on the electric circuits. Mr. M. Maa edited the manuscript. Support for the development of theVIMS Sea Carousel from the Environmental Protection Agency, R-817182-01-0, is sincerely acknowledged. Support for the field deployments at the Cherrystone site from the Office of Naval Research, Contract No. N00014-93-1-0986, is also acknowledged. This paper is a contribution of the Virginia Institute of Marine Science.

REFERENCES

Amos, C.L., J. Grant, G.R. Daborn, and K. Black. 1992. Sea carousel a benthic, annular flume. *Est., Coastal Shelf Sci.* 34:557-577.

Brekhovskikh, V.F., V.K. Debolsky, G.N. Vishnevskaya, and N.S. Zolotareva. 1991. Erosion of cohesive bottom sediments: The influence of benthos. *J. Hydraulic Res.,* 29(2):149-160.

Burt, T.N., and A.C. Game. 1985. The carousel: Commissioning of a circular flume for sediment transport research. *Report No. SR 33.* Wallingfold, England.Hydraulic Research.

Downing, J.P., R.W. Sternberg, and C.R.B. Lister. 1981. New instrumentation for the investigation of sediment suspension processes in the shallow marine environment. *Mar. Geol* 42:19-34.

Fukuda, M.K. 1978. *The entrainment of cohesive sediments in fresh water.* Ph.D. dissertation. Cleveland, Ohio: Case Western Reserve Univ.,

Lavelle, J.W., H.O., Mofjeld, and E.T. Baker. 1984. "An in-situ erosion rate for a fine-grained marine sediment." *J. of Geophysical Research,* 89(C4): 6543-6552.

Maa, J.P.-Y., L.D. Wright, C.-H. Lee, and T.W. Shannon.1993. "VIMS Sea Carousel: a field instrument for studying sediment transport," *Mar. Geol.,* 115 (3/4), 271-287.

Maa, J.P.-Y. 1993. VIMS Sea Carousel: Its hydrodynamic characteristics. In *Nearshore and estuarine cohesive sediment transport,* Mehta, A.J., ed. Washington,DC : American Geophysical Union, 265-280. Coastal and Estuarine Studies No. 42.

Maa, J.P.-Y., C.-H. Lee, and F.-J. Chen. in review. Bed shear stress measurements for VIMS Sea Carousel. Submitted to *Mar. Geol.*

Maa, J.P.-Y. and C.-H. Lee. In review. Resuspension behavior of natural sediments. Submitted to *Australian Journal of Marine and Freshwater Research.*

Nichols, M.M. 1972. Sediments of the James River estuary, Virginia. *Geol. Soc. Am. Mem.* 133:196-211.

Parchure, T.M. and A.J. Mehta. 1985. "Erosion of Soft Cohesive Sediment Deposits," *J. of Hydraulic Engrg.,*ASCE, 111(10), 1308-1326.

Schaffner, L.C., R.J. Diaz, C.R. Olsen, and I.L. Larsen. 1987. Faunal characteristics and sediment accumulation processes in the James River estuary, Virginia. *Est. Coastal Shelf Sci.,* 25:211-226.

Wainright, S.C. 1988. *The contribution of resus pended marine sediments to the planktonic food web.* Ph.D. dissertation, Athens: Univ. of Georgia.

Toward a Sustainable Coastal Watershed:
The Chesapeake Experiment. Proceedings of a Conference
1-3 June 1994. Norfolk, VA
Chesapeake Research Consortium Publication No. 149

LET'S MOVE OUT OF THE DARK AGES IN DATA COLLECTION AND PROCESSING
A COMMENTARY BY:

Robert T. Jaske
Public Safety Consultant

Francis Stetina
National Aeronautics and Space Administration

Abstract: In this presentation, the authors consider the progress in the state of information processing in the post -1970 period and make some predictions concerning how the research community will be doing its work in the near future. This near- term future will involve the processing of living resources data on an unprecedented scale, to positional accuracies never before thought feasible, and for acquisition costs less than those commonly accepted using current technology.

The near-term information system explosion is particularly pertinent to the Chesapeake Bay Restoration Project and its supporting research base because the depth of the Bay is quite nominal, and lasar spectroscopy can be expected be have an early foothold in operational analysis involving repetitive information sampling.

INTRODUCTION

Three major developments that are moving into common practice will catalyze a revolution in how basic data are collected, how time series data are related to modeling activities, and how all of this information is stored, exchanged, and prepared for analysis, publication and archiving. These developments are:

1. Lasar spectroscopy, already prominently in use for industrial control systems will be expanded in scope through the use of broad range of bandwidths from infrared to ultraviolet. Just as infrared spectroscopy in the chemical industry changed the world of analytical chemistry, the application of spectroscopy will move living resource research light-years ahead in its ability to sample in real time and move the information into models almost simultaneously.
2. The use of spatial data management systems that combine satellite geopositioning to as close as 6 cm on the North American Datum, object-oriented data management systems, and the standardization of spatial data in cellular format will open entirely new opportunities for the real-time coordination of diverse information sources. Techniques that exist today make possible the sectioning of data subsites from large fields, the electronic superpositioning of information from diverse sources, and the instant distribution for comment of conceptual ideas by means of high-speed networking.
3. Just as the Corliss engine served its purposes, the day of the large central processor is strictly limited. Astrophysics is now making routine use of parallel processing of arrays of relatively low cost workstations. Given a data management system that is both object oriented and spatially controlled by means of a binary compatible coordinate geometry, it is quite feasible to consider all of the operational research for the Chesapeake Bay Restoration Project being conducted in a client server network with linked workstations in real time. In such a networked array, each researcher could independently choose the level of detail for individual attention: (i.e., macro versus micro in a single data structure), leaving to the degree of significance the fineness of the grid and the computing horsepower. Central library functions could be carried out with mainframe technology such as Cray and array processors positioned as client/server in a workstation network. The transfer of data and data sets between and among client workstations would take place on either satellite or fiberglass networks, a direction now firmly established in

the commercial field. A major manufacturer has just proposed an "L" band system for public use with relatively small dish antennas on the order of 25 cm diameter.

BACKGROUND

As things stood in the mid 1970s, computing was dominated by large mainframes. Data sets and algorithms were closely tied into unified sets at the pleasure of individual practitioners in order to reduce cycle time, a severe limit at the time. Early efforts to produce array processors were only successful with narrowly standarized data fields such as telephone information systems. The processing of remotely sensed data, while feasible was slow and costly and exceeded storage limits by large margins. Without solidstate mass storage, data read rates were limited by tape transfer speeds and mechanical limits on tape positioning. All external transfer of data between remote users was by ASCII tape, a process that took large amounts of tape space. Large data transfer systems were limited to banking or government users such as National Aeronautics and Space Administration (NASA), and there were no modems or equivalent devices available to individuals. Early satellite imagery collected by the National Weather Service accumulated faster than it could be analyzed and simply piled up in tape storage. Early orthophoto concepts were seriously disadvantaged by storage costs.

Commencing in the 1980s, changes occurred rapidly. The development of virtual processing simplified programming, making interactive algorithm operating from common data bases feasible. Mass storage began to decrease in both size and price, and improved data bus systems reduced sequencing rates. The development of the first personal computers (PCs) had little or no effect on research type data analysis except for column and row work of simple nature. But it was not until the late 1980s that cheap mass storage and the 128-bit and larger chips along with faster modems began to open a new era. Even, however, at the end of the 1980s, mass storage in the multigigabyte size was limited to juke box arrays, although minicomputer manufacturers began to offer larger and larger storage capacity. Because of these storage limitations, however temporary, it became customary to continue programming practices that conserved space long after the need disappeared. Thus, the notion of storing data in polygons and the use of relational techniques both became established to meet needs already superseded by more comprehensive requirements of the future.

In the 1990s, we have emerged into an almost limitless period in which custom and prejudice govern choices more than realistic needs. GIS technology has become popular, but tied to obsolescent concepts of data base management. We have also emerged into an age where automating old processes is no longer a proper goal. Today, the computer needs to be considered as an exploratory tool by which we can do work never before attempted. The key element in system technology in the future will be data storage, standarization, transmission and archiving. Along with communications, these elements of computing will occupy 90% or more of system developer's time, with equipment costs shrinking into insignficance. Today, in effect, hardware is moving toward being expendable.

Three important elements of the near future will revolutionize how in situ biological research is done. These are:

- laser spectroscopy,
- standardized coordinate geometry, and object oriented data base management, along with vastly improved communication networking,
- the use of parallel processors on an unprecedented scale for interactive analysis of real-time data.

Laser Spectroscopy

Born in the early space age as the by-product of holographic development work, feasibility of classifying oil slicks and other organic chemicals by analysis of return radiation was demonstrated in early 1972. Effective use was, however, limited to chemical process technology in the form of velometers and in situ process controllers because of the difficulty of indexing geographic position and standardizing a spatial data management system. Security classification also seriously impacted public use, and government has been slow to declassify the technology although it was already in use in Viet Nam. Today, none of that is a valid reason for not using the techniques. The positioning problem has been resolved by the use of GPS, which in its most accurate form can resolve position to 6 cm on the North American Datum from an aircraft at 5,000 feet using portable equipment. The laser technology has been extensively tested to deliver very accurate (better than 1 cm) vertical information, and is being tested on forest canopies and related three-dimensional geometries. Bathymetry to 30 meters is common, and lower depths are feasible with higher power. The equipment is esentially miniaturized, and can be flown from single-engine light aircraft using standard access hatches. All data are recorded on magnetic disk (optical if you please), and read out using shelf model 386 microcomputers is routinely done.

Accordingly, it is now feasible to plan and carry out in situ research on surface wet lands and submerged aquatic vegetation with cell sizes in the centimeters if justified, and with returns to the same position without ground monitoring to within 6 cm. Given some calibration research, essentially real time analysis of chemical transfer rates in watersheds is feasible or will be shortly. In addition, it is also feasible to index the study cells and share the information with any other interested researcher by networking. For those bound to raster, pixel technology, it is feasible to code the coordinate geometry of individual pixels.

Our conclusion is that future research to establish intricate relationships in a complex environment will be economically carried out from aircraft using laser/GPS technology. The tools are there, where are the users?

Data Base Management

Because all of the above will produce unprecedented quantities of raw and relational data, there must be vastly improved data base management. Research is now showing that object-oriented data management is feasible in which points, points forming lines, and lines forming polygons allows creation of cell-type data systems that have spatial characteristics and are not tied to layers as is the case with most present GIS systems. While the dust has not settled on the matter of attribute storage sites, the identification of attributes with points primarily and with point-related objects allows standardization at the binary level and complete freedom from machine and software constraints. Early research by the authors in 1984 identified object-oriented systems as the best approach to research application because the individual objects can be editied, copied, or transferred without a relational reorientation. Lest users of relational databases feel threatened, combination systems are under development and use of SQL will probably spread to object-oriented systems. Sampling of raster imagery data to provide cell-type storage arrays is already standard practice, and ratios as high as 20:1 are common in raster/vector combination systems.

Accordingly, researchers can plan on standardized cell arrays in which cell dimensions are some resonant frequency of some world standard such as 0.1 second of arc. Individuals can then create accurately positioned cell data as fractions or multiples of the standard, knowing that the spatial characteristics will all pass a binary filter among networked client machines. This means that all data will play on everyone's machines, the limit coming in the degree of detail the individual is able to process for his or her individual problem. Some unresolved problems involve version management in shared data base regimes. By tradition, data elements are locked out during an edit process, but GIS applications involve, relatively long lock times, so this must be resolved to keep track of the edit function when multiple users are on at the same time.

Such a coordinated system will place heavy demands on library functions, a relatively new concept in computing. Up to the present, users have usually created data bases of their own, carried large overheads in empty spaces, and generally had little regard for interchangability. Mainframes to the rescue! The filing of enormous amounts of data in coordinate geometry format, and the management of huge relational and object-oriented data bases with fixed arrays is duck soup for the big boys. The telephone companies are already offering such service, and public agencies can be expected to follow suite shortly. In the future, a researcher can order up an array to suit his or her budget and accuracy interval and than add the raw data to compute interactions. If the results are satisfactory, the new data can be sent back to the library (hopefully, with peer review.)

Networking? One large industrial firm has just announced an "L" band system for public application using small 15- inch antennas and miniaturized components costing under $1,000 and with data rates in multimegabytes/second. Not to be outdone, the telephone companies are set to offer competitive service on fiberglass systems.

All of the above fits into the future concept of high-definition, television which has been announced as the combined successor to both the PC and the television ste. Modular systems such as high-fidelity sets are soon to emerge, and researchers will all benefit from the new capabilities.

Extended Parallel Processing

The rapid increase in data and the latent possibilities of new research opportunities are being accompanied by an increase in sophistication of algorithms that allow us to convert the data into meaningful information. These operations have taxed the capabilities of traditional high-speed mainframes such as the Cray and seriously limit the level of detail at which researchers can address studies. The answer appears to lie in two dimensions, one evolving rapidly, the other more slowly. The former is an extension of array processing in which the dimensional arrays within a given core assignment are replaced with an entire 64- or 128-bit processor. The result of such a combination multiplies the processing speed such that the resulting speed is as much as 50 times the speed of each individual computer in the array when doing arithmetic, and floating-point operations, and about 20 times for attribute-loaded graphics operations. Parallel processing machines up to 512 units are currently in place and projections of 2,048 units are being forecast. Enormous amounts of repetitive computation can be made, but the degree of flexibility is limited by the data relationships similar to those of relational data base management systems.

Further on the horizon is the networked array of mixed processors working in unison on several related but separate problems at one time. In this feasible, but not yet ideal, world, the entire Chesapeake Bay could be modeled by a networked combination of several hundred individual processors, each contributing a portion of the final result, and each responding in a client/server mode to new secondary data being generated by the combination. In effect, such a system will run all the algorithms of individual researchers in unison on their own local processors, doing in a single operational step what the present system does in several steps, with much hassling over data and transformations.

CONCLUSION

These developments are not for tomorrow, they are here today. For example, the world meteorological community passes 30 megabytes per second of data in a worldwide satellite network by means of a binary compatible transfer code entitled BUFR (binary universal format regime). Any or all of this information is available to desk-type workstations by direct satellite interchange.

The real challenge is ourselves. Too often the research community finds little comfort in new approaches without being spoonfed by vendors or peers. Political correctness takes its toll here too. Each of the researchers in the Chesapeake Bay Restoration Project can also become a part of an evolution in information processing if he or she has the will to become a part. Just as the astrophysicists found they could do the work of several Crays with networked IBM 6000 RISC machines for a tenth of the cost, the Chesapeake Bay Program can decide to move out of the Dark Ages of central processing and data management into distributed data processing on networked work-stations. And the researchers can have fun doing it, each on his or her own.

BIBLIOGRAPHY

Buzbee, William, et al. 1993, Supercomputing: The view from NCAR. Boulder, CO, National Center for Atmospheric Research, 191 pp.

Dobson, Jerome E. 1994. GIS integration requires a new kind of parallel processing. In: GIS world. Fort Collins, CO: 28.

Fisher, Michael J. 1993, New chips and operating systems: What do they mean to GIS users? In: GIS World, Fort Collins, CO: GIS World, Inc., 39-42. pp.

Hill, Jack, and Stetina, Fran. 1991. The Airborne LIDAR Topographic Mapping System, (ALTMS) Houston: Houston Advanced Research Center, (available from the authors).

Jameson, Donald. 1994. Object-oriented approaches add value to GIS. In GIS world, Fort Collins, CO: 43-47.

Maillard, W.D., and W.D. Meitzler 1990. An integrated software environment for tactical decision making aids that require a geographic information system. Richland, WA: Battelle Pacific Northwest Laboratories, 9 pp.

Proc., Seventh International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, 1991. Boston, MA: American Meteorology Society, 504 pp

Proc. , International Symposium on Spectral Sensing Research, 1992. Maul: HI. Science and Technology Corp. Hampton, VA: 1366 pp

Walklet, Donn C. 1994. Commercial control of remote sensing predicted. In GIS World. Fort Collins, CO: GIS World, Inc. 18.