STATUS OF THE TOWAN DATA REPOSITORY

Susan M. Starke
Naval Research Laboratory

Mark R. Fernandez
Planning Systems, Inc.

KEYWORDS
Environment, Data Repository

ABSTRACT

The Tactical Oceanography Wide Area Network (TOWAN) is a data repository for ocean environmental databases. Developed by the Naval Research Laboratory, TOWAN provides access to and rapid subsetting of large environmental descriptions in support of diverse modeling and simulation requirements. Database ingestion routines (manual and automatic), database management, subsetting of data, as well as efficiency of data delivery were design criteria imposed to ensure modelers, of all disciplines, could access and extract environmental data in an efficient and consistent manner.

This paper provides a description and status of TOWAN capabilities. Contained within the status report are descriptions of databases available, available extraction routines, typical access times for data via access through the commercial Internet and other networks. Also contained within the paper are sections on (1) future plans for TOWAN, (2) TOWAN status as a Master Environmental Library (MEL) node and (3) methods and requirements for accessing TOWAN.

INTRODUCTION

The Tactical Oceanography Wide Area Network (TOWAN) is an environmental data repository and network accessible server which allows Department of Defense personnel and their contractors to search for and retrieve environmental information for modeling, simulation and other analyses.

TOWAN was developed under the Naval Research Laboratory's 6.2 Science and Technology (S&T) Program. The objectives of this program included the ability to demonstrate the efficiency of exact computer recreations and simulations applied to sensor weapons and related environmental acoustic issues, and to apply this capability using real and realistic environments. To accomplish this objective, it was decided that the acoustic modeling would be decoupled from the data storage and management, and the modeling would require only that the data be delivered to it in a standard format. TOWAN was developed to provide this data storage, management and delivery.

The TOWAN server provides for the rapid subsetting of large environmental databases and delivery to the user of either an area, point or track of data in a standard data format. The term "rapid" is used purposefully. TOWAN is designed to be the fastest on-line network accessible repository of oceanographic data available. The design goal is to address the data needs of the researcher quickly enough for the researcher to do his job. The actual access times will be discussed later in the Access Times section of this paper.

TOWAN was designed as a tool for the research community and as such, was not designed to do data assimilation, manipulation or interpolation. Instead, the TOWAN design allows the user to easily extract data from the TOWAN databases, merge or manipulate the data in whatever manner is appropriate for the research being done, and then easily re-ingest the manipulated data onto TOWAN for storage and easy access. This allows the researcher to choose the assimilation routine best suited for the specific research being done, while still allowing the use of standard access routines for extracting data.

HARDWARE AND NETWORK CONNECTIONS

TOWAN uses the Silicon Graphics (SGI) Challenge S as a server with its built-in high speed bus (267 MBytes/sec) and dual Fast/Wide/Differential (F/W/D) SCCI-2 I/O capabilities. Although the Challenge S can be configured like a typical UNIX workstation, the system used by TOWAN is configured strictly for the purposes of rapid data delivery. For this reason, TOWAN has no keyboard, mouse, monitor, etc. Furthermore, there are no software development packages or libraries on the system, and there are login accounts only for the TOWAN DBA (DataBase Administrator) and server developer.

Attached to the dual F/W/D SCCI-2 controllers are two disk towers, each containing 9 Gbyte Seagate Barracuda disk drives with their 8-9 millisecond access times and 1 MByte on-board cache. All TOWAN data is on-line at all times. There is no secondary optical or tape storage. The commonly used databases are placed on separate disk drives evenly distributed over the SCCI-2 controllers to enhance system performance.

There are two identically configured TOWAN systems, one for classified data and one for unclassified data. The only difference in these systems is in network connectivity. Presently, the network connection is via the standard 10BaseT connection on the Challenge S to the local Naval Research Laboratory (NRL) - Stennis Space Center LAN (Local Area Network). The unclassified system is connected to the local TCP/IP Ethernet based LAN via a switched Bay Networks hub, which in turn is connected to the site-wide Cisco router. This affords the TOWAN system the full 10 Mbits/sec bandwidth to the Cisco router, which is connected to the local FDDI ring. Leaving the FDDI ring, and thus leaving the site, is done via dual DREN T-3 lines. The classified system is similarly connected, but a Network Encryption System (NES) is placed between the classified TOWAN system and the local LAN. Additionally, the TOWAN Challenge S systems have been equipped with FDDI boards, and a direct FDDI connection to the local FDDI ring is expected during the last quarter of FY97.

Finally, TOWAN is designed to be up and available at all times. TOWAN is protected by Uninterruptable Power Supplies (UPS's) with 1 hour standby time. There is no regularly scheduled downtime. Ingestion and other database administration activities take place while on-line.

SOFTWARE DESCRIPTION

TOWAN was designed to support the needs of researchers, and as such is limited in its design. Of primary concern was rapid subsetting of existing oceanographic and atmospheric environmental databases which, for the most part, are already designed for use by researchers. Consequently, TOWAN has targeted query and extraction capabilities.

TOWAN supports only direct simple queries. It does not support compound, relational, or complex queries. The only subsetting queries supported deliver "all depths/heights, all parameters". In other words, the user may select data by type, geographic region, and time period. Delivered is a subsetted point, area, or track from the specified database.

To reduce costs, decrease start up time for a new user, and avoid a wide variety of problems associated with earlier systems, TOWAN contains no commercial Database Management System (DBMS), no licenses or licensed software, and no other proprietary software.

Software Interfaces

TOWAN has several layers of client software interfaces designed for use by the diverse community it must serve. Each interface layer is used to compose the next interface layer.

The first and "lowest" level is a library of C/C++ function calls which encapsulate the network transport mechanisms and the TOWAN queries. These were developed in-house and designed to be a minimalist set upon which to build command line routines. Compiled with g++ for portability and delivered statically linked, the user needs no additional supporting software, libraries, or licenses.

The command line routines stand alone as independent executables, are more user friendly, and have version and help options. Although as cryptic as most UNIX command line routines, they are a complete set upon which users may usefully access TOWAN. Again, this level of software is compiled with g++ for portability and delivered statically linked. The user needs no additional supporting software, libraries, or licenses.

Using the command line routines, an X/Motif Graphical User Interface (GUI) was constructed (See Figure 1). This minimal GUI serves the purpose of graphically depicting the small amount of information required by the user to make use of TOWAN. The heirarchy of the required information is illustrated, and the multiple query capability can be demonstrated. Additionally, the GUI echoes the command line routine call it is making on behalf of the user. These can be collected for use in other routines. X/Motif is required to use this level of interface. However, the routines use only Xlib calls and are compiled with

g++. Thus, portability is expected to be high. At present it is available on SGI with Sun and Hewlett Packard versions scheduled for release in February 1997. (See Table 1.)

Using the TOWAN GUI to browse for available data, the user traverses a short and simple hierarchy based loosely upon the Commander, Naval Meteorology and Oceanography Command (CNMOC) Reference Publication RP-2, Meteorological and Oceanographic Data for Modeling and Simulation [1]. The user must first select a "category" or "CAT", presently either oceanography or meteorology. Given a CAT, the user then selects a "CORE" area of research, such as acoustics, magnetics, or optics. The databases within TOWAN have been assigned a CAT and CORE and indexed accordingly to improve access times. Additionally, the user is presented with a "short list" of data "TYPES" which are commonly used for that particular core area of research. The TYPES presented are designed to be words and phrases describing types of data with which the researcher is familiar, such as bathymetric, and geoacoustic. Upon selecting a TYPE, the user is presented with a list of available "BRANDS" of data matching the requested CAT, CORE, and TYPE. The BRANDS are generally the names the databases have been given by the originators, such as DBDB5, GDEM, and HITS. Note that the BRAND names are generally not intuitive. Help at the BRAND level will present to the user in an HTML window all the metadata available on that particular BRAND.

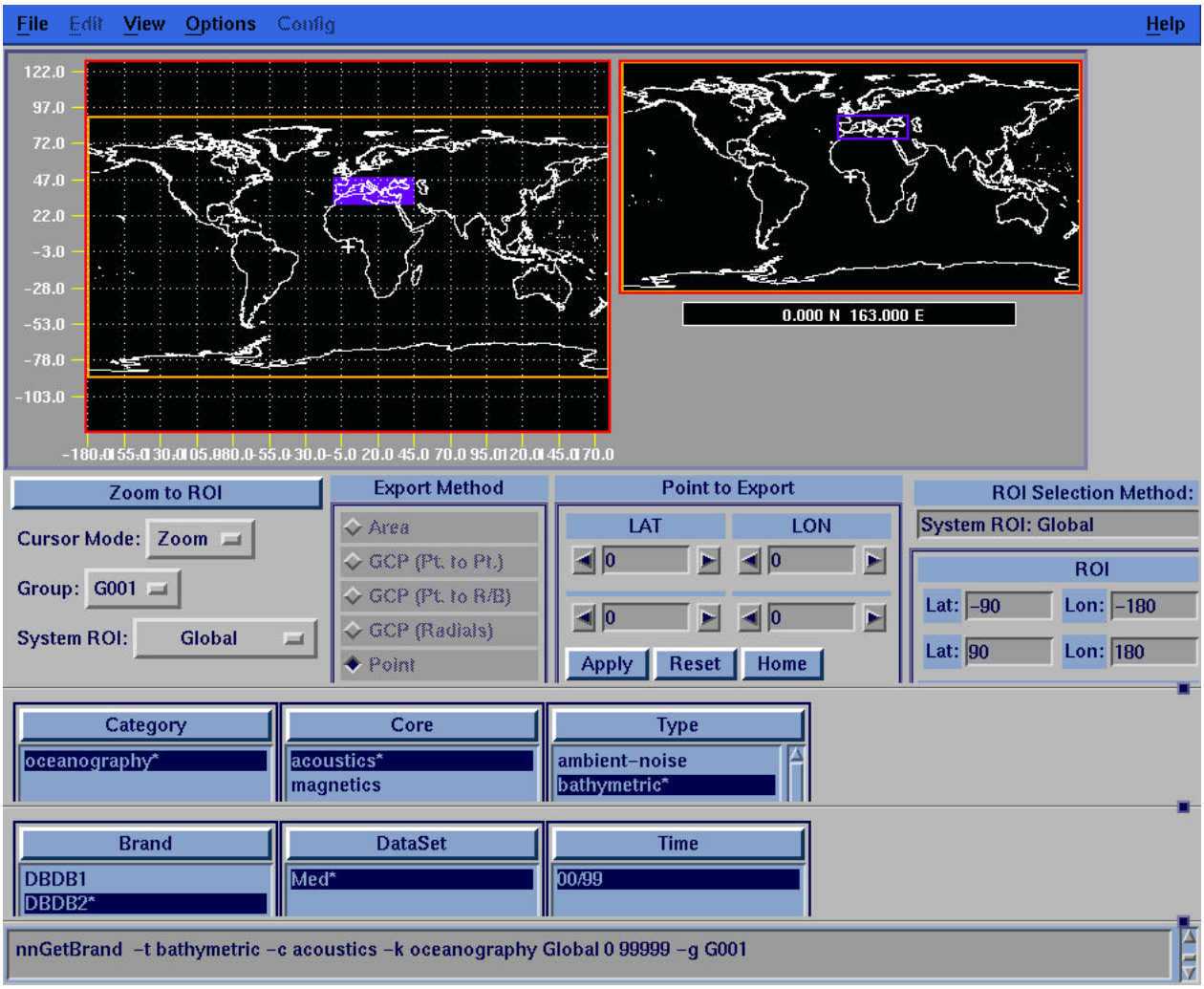

Figure 1. TOWAN GUI

Thus, given a geographic region of interest (ROI), a time of interest (TOI), a CAT, CORE, TYPE, and BRAND, TOWAN is designed to quickly retrieve the requested data. To do so, TOWAN uses flat binary direct access files. Furthermore, by centralizing the standard databases and the mechanisms for extraction, the user community is assured of consistency. Additionally, all the data within TOWAN has been test extracted prior to being made accessible. As part of the ingestion process, databases are automatically subjected to a test extraction process which performs a series of point, area, and track test extractions with automated comparisons to the original "raw" ingested data source. The process tests data on each edge/boundary of the original datasets as well as at interior points.

Hardware	Operating System	Available
SGI	IRIX 5.3	Jan-97
SGI	IRIX 6.2	Jan-97
HP	HP-UX 10.10	Feb-97
SUN	Solaris 2.0	Feb-97
SUN	SunOS 4.1.3	Feb-97

Table 1. User System Configurations

To deliver the data to the user over the network, TOWAN uses an object-oriented network-independent transport mechanism compliant with the philosophy of the HLA/RTI (High Level Architecture/Run Time Infrastructure). The InterLanguage Unification (ILU) implementation of CORBA (Common Object Request Broker Architecture) is used as the underlying basis for the network transport. The transport model and mechanism were developed in conjunction with Physitron and the University of Alabama - Huntsville (UAH) under a Small Business Innovative Research (SBIR) project. The requested data is delivered from server memory to client memory in object form. The transport is abstracted from the "final form", that is, the data is delivered to the client over the network and placed in the client's memory in object form. The client can then "use" the object via any of the object's methods. At present the only method implemented is one which will write the data in ASCII in the TOWAN Data Format (TDF) to the client's disk. The list of methods is expected to grow in FY97 to include outputting other formats such as HDF (Hierarchical Data Format), and direct use by applications without the need to write the data to disk. We expect to receive requests for other methods.

DOCUMENTATION

All documentation for TOWAN is on-line in HTML format. Information is available by accessing the Naval Research Laboratory Code 7180's homepage on the WWW at the URL address:

http://www7180.nrlssc.navy.mil

and clicking on the TOWAN hotlink, or going directly to the TOWAN homepage at the URL address:

http://www7180.nrlssc.navy.mil/TOWAN/TOWAN.html.

This site includes a description of TOWAN, documentation of the TOWAN Data Format (TDF) and some sample output files, along with access information. It also contains a listing of the databases available on TOWAN.

The TOWAN GUI is a user friendly browser for accessing the TOWAN databases. This browser provides three types of help. The first is a help menu located in the upper right hand corner of the GUI (see Figure 1) which has three options: (1) On Window: gives the user information on the way the GUI works, (2) On Data: brings up a listing of all the data available on TOWAN, and (3) On Version: gives the version number for the tool, along with the current development environment. The second type of help is available by clicking the Help button once a Brand of data is highlighted. This brings up general information on the specified Brand. The third type of help is more detailed information on a specific database and is available by clicking the More button once a DataSet and Time have been selected.

DATA EXTRACTIONS

Three types of extractions are available on TOWAN, point, area and track. Also available is the radial extraction which is a extension of track. Each is described below:

Point: This extraction returns the Latitude/Longitude of the point requested, along with the closest data point within the database. This type of extraction, for example, can be used for range independent acoustic model calculations, where only one environment is input to the model.

Area: This extraction returns the Latitude/Longitude and data associated with each data point within the area requested. This type of extraction can be used for display purposes.

Track: This extraction returns the Range along a track, the Latitude/Longitude associated with entering a new grid square, and the data associated with that grid square. This routine performs a great circle path calculation, and assumes that the actual data point is at the center of the grid square. This type of extraction for example, may be used for range dependent acoustic model calculations where a track of environments is required as input to a model.

Radials: This extraction returns a set of Tracks as described above, one for each azimuthal radial requested. The file name is appended with the three digit bearing associated with the respective azimuthal radial.

ACCESS TIMES

TOWAN was designed for the rapid subsetting of oceanographic and atmospheric environmental databases. As described in the introduction, the term "rapid" is used purposefully to refer to "researcher time". Researcher time is defined to be much faster than the snail-mail, e-mail, manual ftp, etc. that the researcher has endured in the past. Specifically, the design goal and test case identified was the delivery onto the network of all the data needed to perform a tactically significant underwater acoustic model run within five (5) seconds of receipt of the request. Network speed and traffic load will, of course, effect total time to respond to a request, but that is beyond the control of TOWAN. Current on-line systems (including the previous version of TOWAN) generally measure response times in minutes. The phrase "all the data needed to perform a tactically significant acoustic model run" was further defined to mean one track each of bathymetry, oceanography, and geoacoustics.

The test databases selected were all unclassified and included DBDB5 (bathymetry), GDEM-Winter (seasonal oceanography), and GEO (a research-grade geoacoustics database commonly used at NRL). An area 2 degrees in longitude by 2 degrees in latitude in the Straits of Sicily was chosen as the Area Of Interest (AOI). This constitutes subsetting of the worldwide DBDB5 and GDEM databases and subsetting of the Mediterranean GEO database. Two (2) test scenarios were constructed. One involved a network transport to a local UNIX workstation and another was to a UNIX workstation located across the country at our sister lab, NRL-Monterey (California). The traceroute routine indicates that NRL-Monterey is 7 hops on the DREN.

The local client was an older SGI Indigo R3000 based machine running IRIX 5.3 with 32 MBytes of memory, and a 1 Gbyte disk drive connected to the LAN via 10BaseT. The distant client was a SGI Indigo-2 Extreme R4000 based machine running IRIX 5.3 with 160 MBytes of memory, and a 1 Gbyte disk drive also connected to the local Monterey LAN via 10BaseT.

Selected were four (4) specific test cases: (1) a single point extraction out of all 3 databases, (2) an area extraction of the 2° x 2° selected area for all 3 databases, (3) a track extending from the southwest corner to the northeast corner of the 2° x 2° selected area for all 3 databases, a distance of about 80 nautical miles, and (4) a radial extraction at a 30 degree step increment extending 100 nautical miles from the center of the selected area, a total of 12 radials. A description of the type and amount of data moved is shown in the Tables 2-5.

	Lat/Lon Points	Depths	Bytes
DBDB5	1	1	152
GDEM	1	35	1777
GEO	1	43	6384
Total			8313

Table 2. Point Extraction Description

	Lat/Lon Points	Depths	Bytes
DBDB5	169	1	8499
GDEM	9	35	5113
GEO	169	43	704111
Total			717723

Table 3. Area Extraction Description

	Lat/Lon Points	Depths	Bytes
DBDB5	28	1	1874
GDEM	6	35	79591
GEO	28	43	1600
Total			83065

Table 4. Track Extraction Description

	Lat/Lon Points	Depths	Bytes
DBDB5	392	1	26159
GDEM	86	35	90380
GEO	392	43	891454
Total			1007993

Table 5. Radials (12) Extraction Description

The timings shown in Table 6 below represent an average of the individual timings of extractions performed once per hour every hour for a period of 7 days. The differences in timings are due primarily to the additional network distance traveled to get to Monterey. The design goal of 5 seconds has been met for local users, and is not too far off for distant users. Since the current implementation of the network transport does not include any encoding or compression, we fully expect to meet the design goals by the end of FY97.

	Average	Average	Minimum	Minimum	Maximum	Maximum
	Local	Distant	Local	Distant	Local	Distant
Point	1.6	3.9	1	3	4	4
Area	6.4	16.6	6	16	10	19
Track	3.1	6.1	3	6	4	7
Radials	41.7	86.5	41	81	45	104

Table 6. Preliminary timing data in seconds

TOWAN ACCESS

Access to TOWAN is available, free of charge, to all Department of Defense (DOD) personnel and their contractors. For access to the unclassified TOWAN system, the user is required to (1) download the TOWAN GUI and (2) register with the TOWAN System Administrator (shown below) as a user. These two steps will give the user access to all public release data on TOWAN.

To gain access to the Oceanographic and Atmospheric Master Library (OAML) products on TOWAN, the user must go through the normal process of requesting access from the Commander, Naval Meteorology and Oceanography Command. Once this access has been approved, the TOWAN System Administrator can give the user access to the specific OAML products requested. Access to any other restricted data on the unclassified system should be requested through the TOWAN System Administrator.

Access to the classified TOWAN system is via a Network Encryption System. Future plans include a direct link to the SIPRNET (Secret Internet Protocol

Routing Network). Please contact the TOWAN System Administrator for more information.

The TOWAN GUI is available via anonymous ftp from ftp7180.nrlssc.navy.mil (IP Address: 128.160.55.103) in the pub/towan directory. Included in the .tar.Z files are a READ.ME file and a man page for the executable. The GUI will be available in SUN, SGI and HP versions.

To register as a user, you should contact the TOWAN System Administrator with a user name and IP address for the machine you intend to access TOWAN from.

TOWAN System Administrator:
Ms. Millie Ling
Naval Research Laboratory Detachment
Code 7183
Stennis Space Center, MS 39529-5004
Phone: (601) 688-4194
Fax: (601) 688-5341
email: ling@nrlssc.navy.mil

User System Requirements

The TOWAN GUI is presently supported on two system configurations with plans to expand to the five systems as shown in Table 1. All systems require access to the Internet, and 10 MBytes of disc space available for the GUI software. Additional space is required for any data to be retrieved from TOWAN. User software requirements include Motif, Xlib and X11R5.

MASTER ENVIRONMENTAL LIBRARY NODE

The TOWAN data repository is planned as a Level 1 MEL Node. The metadata for each database is in FGDC compliant format and access to this metadata via MEL will be available in the second quarter of FY97.

STATUS

The databases available on TOWAN are shown in Tables 7 and 8. Updated versions of the OAML products have been requested from the Naval Oceanographic Office, and will be ingested as soon as they become available. The TOWAN WWW site will be updated when new products are on-line and tested.

All data ingested on TOWAN to date are either 2-D or 3-D gridded databases. Point data, better known as LLT (Latitude, Longitude, Time) data is currently being worked on, and should be available in the 3rd quarter of FY97.

Brand	Region	Distribution
BOTTOM-CONDUCTIVITY	camp-lejeune	Public Release
BOTTOM-TYPE	camp-lejeune	Public Release
DBDB1	Med	Restricted
	SCal	Restricted
DBDB2	Med	Restricted
DBDB36sec	Korea	Restricted
DBDB3sec	SCal	Public Release
DBDB5	Worldwide	Public Release
DBHR	Black_Sea	Restricted
GDEM-PR	Worldwide	Public Release
	Indian	Public Release
GDEM-RES	Arctic	Restricted
GDEM-MONTHLY	Sea of Japan	Restricted
	Yellow Sea	Restricted
GEO	Korea	Restricted
	Sicily-Straits	Restricted
HITS-ANNUAL	Worldwide	Restricted
HITS-MONTHLY	Worldwide	Restricted
HITS-SEASONAL	Worldwide	Restricted
HOPS-PR	HIRes	Public Release
HOPS-RES	Arctic	Restricted
HWS	Worldwide	Restricted
MAGNETICS	camp-lejeune	Public Release
	korea	Public Release
MODAS	korea	Restricted
PH	camp-lejeune	Public Release
	korea	Public Release
SPECIFIC-GRAVITY	camp-lejeune	Public Release
	korea	Public Release
SWWC	worldwide	Restricted
	indian	Restricted
	korea	Restricted
VOLUME-SCATTERING	camp-lejeune	Restricted
WATER-CONDUCTIVITY	korea	Public Release
	camp-lejeune	Public Release

Table 7. TOWAN Unclassified Databases

Brand	Region
CLFBL	Worldwide
DBDB30sec	Gulf of Arabia
DBDB6sec	Korea
DBDBC	Worldwide
HFBL	Worldwide
LFBL	Worldwide
SNHR	kings-bay
	juan-de-fuca
	arctic
	wpac
SNLR	NHemp

Table 8. TOWAN Classified Databases

FUTURE PLANS

Future plans for TOWAN include Point and Vector database storage, along with area and point extractions from these databases. The Point database storage is presently being completed, and the Vector database storage will follow that effort. Within the Point database storage effort is the ability to search within the database on time. This type of time search will be implemented along with this type of storage structure.

A more general Time Search which allows the user to search at the database level for all data available during a specified time period is also planned. Current capabilities allow the user to see what the time period information is, but not to search on that information.

Other improvements planned for TOWAN include an HDF version of the TOWAN Data Format (TDF) for output to the user, allowing user specified Regions of Interest (ROI's), and a PC/Windows NT version of the GUI. These upgrades will be implemented when resources become available.

REFERENCES

1. Commander, Naval Meteorology and Oceanography Command, “Meteorological and Oceanographic Data for Modeling and Simulation”, Reference Publication RP-2, November 1994.