REPORT OF TECHNICAL COMMITTEE ON DATA EXCHANGE

The meeting of the Technical Committee on Data Exchange (TCODE) was held from 13:30 – 18:00 hours on October 17, 2004. The Chairman, Dr. Igor I. Shevchenko, called the meeting to order and welcomed the participants (*TCODE Endnote 1*). The Committee reviewed the draft agenda and it was adopted as provided (*TCODE Endnote 2*).

Review progress on items in the 2003/2004 Workplan (Agenda Item 3)

TCODE/HAB-S work on HAE-DAT database

Mr. Robin Brown reported on the results of the joint PICES/IOC workshop on "Developing a North Pacific HAB data resource – II" at PICES XIII, and the trial with the IOC/ICES Harmful Algal Event Database (HAE-DAT; http://ioc. unesco.org/hab/data3.htm). Each country entered data from their nation for one year (selection of year varied by country) and reviewed their experiences using a "report card" approach. Several issues were identified:

- There were some variations in the definition of "event" used by each country;
- There were some difficulties in separating the coastline of each country into segments as recommended by IOC/ICES. Each country took a somewhat different approach, based on the length of their coastline and existing administrative area structures;
- Duration of a single event was often hard or impossible to determine;
- Each country reported that there were many blank fields in the database that they were unable to complete due to lack of data. The distribution of "blank fields" was highly variable amongst countries. For example, the only data that Canada could report was toxicity data from their shellfish monitoring program, and no data on taxonomic composition of harmful blooms was available. Russia, on the other hand, contributed taxonomic and abundance data on potentially harmful blooms, but has no toxicity data;
- Some countries reported that it was difficult to consolidate the data from within their nation, due to the high volume of data and/or administrative barriers.

There was an active discussion about the strengths and weaknesses of the HAE-DAT approach, but the workshop participants agreed that (within the limitations discussed), this was a useful approach and should be adopted by PICES member countries. The plan for 2004/2005 is for each country to enter data from a common year (2000) plus as many years after 2000 as is feasible. The participants also agreed that it is important that each country prepare a short, high-level description of the characteristics of the data entered for their country to allow users to interpret any data retrievals. This metadata could be provided on the PICES website and/or the HAE-DAT website.

GLOBEC/TCODE interactions on data management

Dr. Shevchenko reported that no replies were received from the GLOBEC International Project Office (IPO) on his queries about the data management activities. It appears that changes in personnel at the IPO have stalled these activities. After some discussion, the Committee agreed to drop this item from the TCODE Workplan for 2004/2005, with the suggestion that this could be re-introduced if there were signs of activity at the GLOBEC IPO. Members also noted that GLOBEC or GLOBEC-like programs in some countries were fully completed or approaching completion, and that the opportunity to efficiently inventory and archive the GLOBEC data holdings was rapidly disappearing.

North Pacific Ecosystem Metadatabase

Dr. Bernard A. Megrey and Mr. Allen Macklin reported on the status and ongoing activities with this database (http://www.pmel.noaa.gov/np/mdb/index.html). Specific activities include:

- Alaska Fisheries Science Center staff are continuing to actively search for and add records to the database;
- Automated processes are now in place to e-mail record contributors and solicit updates to the records they have submitted;
- Search queries have been modified and updated;
- Capabilities for "federated searching", where multiple metadatabases may be searched with a single query are being implemented. Dr. Megrey presented plans to implement this capability between the North Pacific Ecosystem Metadatabase and databases at the Korean Oceanographic Data Centre as a demonstration project across the PICES region. He submitted a proposal for this activity and a request for partial support from PICES (*TCODE Endnote 4*). The Committee endorsed this proposal;
- Dr. Megrey and Mr. Macklin worked with the Gulf of Alaska Ecosystem Monitoring (GEM) project to use a new standard (Ecological Metadatabase Language – EML; http://knb.ecoinformatics.org.software/eml/) which is better suited to *in situ* biological observational data;
- Mr. Macklin reminded the committee that there was an extensive article on the North Pacific Ecosystem Metadatabase in a recent issue of PICES Press (July 2004, Vol. 12, No. 2; http://pices.int/publications/pices_press/volume12/July04/pp_30_33.pdf.

The Committee noted that the North Pacific Ecosystem Metadatabase is supported through a variety of "soft money" sources and this makes the system somewhat vulnerable.

Updating TCODE web pages

Dr. Shevchenko reported on the updates to TCODE web pages (http://tcode.tinro.ru). All national reports and supporting documents for TCODE activities will be place on this web.

Electronic Poster Session at PICES XIII

Dr. Thomas C. Royer reported that nine submissions were received for the TCODE Electronic Poster session on "Data visualization of open ocean processes in the North Pacific) at PICES XIII. A summary of the session is included elsewhere in this Annual Report and posted on the TCODE web page.

Interactions with ICES-IOC Study Group on the Development of marine data exchange systems using Extensible Markup Language (XML)

Dr. Shevchenko reported on activities of the ICES-IOC Study Group on XML (this language is considered as a standard for data exchange). Dr. Georgiy Moiseenko attended a meeting of this group (May 2004, in Oostende, Belgium) as a PICES TCODE representative. Mr. Brown and Mr. Macklin offered to gather further information on the progress made by the ICES-IOC Study Group and provide this information to the TCODE Chairman for circulation. Dr. Shevchenko will post all these reports and links to supporting materials on the TCODE web page.

Annual reports and updates on data management activities in PICES member countries (Agenda Item 4)

Mr. Brown reported on key contacts and sources for data, software and educational/outreach material in Canada. In addition, he provided copies of the DFO Policy for management of scientific data to the Committee members. In terms of new data management activities in Canada, he noted that the VENUS and NEPTUNE cabled ocean observatories will require an advanced data management and archival system, and may present some interesting new opportunities.

Mr. Ling Tong reported that data management and exchange (including standards) is an acknowledged challenge in China, and they are working on a plan to address this issue. He noted that biological data are a particular challenge. Policies on charging for data (internally and externally) are further complicating the issue.

Dr. Kenji Asano reviewed the key data management contacts in Japan (JODC for oceanographic data and the Fisheries Research Agency for biological data). He also introduced a new Japan Fisheries Oceanography database implemented by the National Research Institute of Fisheries Science that will soon be publicly available.

Mr. Hae-Seok Kang reviewed the principal data management agencies for Korea (KORDI, NFRDI and NORI). He further reported that KORDI will assume responsibilities for real-time data, while NFRDI will focus on delayed-mode data.

Dr. Shevchenko reported on a workshop on "Mathematical modelling and information technologies in ocean studies" held in 2004 in Vladivostok (Russia). He also provided information on the recently published "Atlas of quantitative distribution of nekton species in the Okhotsk Sea".

Mr. Macklin reported on the Alaska Ocean Observing System (AOOS) and the central role that data management will play in this system. A pilot project will be undertaken in 2005 to integrate the Prince William Sound data sources. Key components of this data management system will be: metadata standards (FGDC), open data transport (OpenDAP), and Live Access Server (LAS) to browse data.

TCODE noted that all new mentioned projects require an advanced data management.

Second inter-sessional Science Board Meeting (Agenda Item 5)

Dr. Shevchenko provided a short report from this meeting.

Discussion of TCODE activity in relation to the PICES Strategic Plan and development of a draft Action Plan for current and future activity (Agenda Item 6)

Dr. Shevchenko and Mr. Brown agreed to prepare a draft TCODE Strategic/Action Plan for circulation and review by TCODE members, prior to the interim Science Board Meeting in April 2005.

Discussion of capacity building opportunities (Agenda Item 7)

TCODE noted that the PICES Capacity Building includes several items which have direct connections to TCODE's terms of reference: sharing data, methodologies, computing power, *etc*.

TCODE nominated Dr. Franz J. Mueter (Joint Institute for the Study of the Atmosphere and the Oceans, University of Washington, U.S.A.) as a potential member of an Organizing Committee of the proposed PICES/ICES Young Scientists Conference (2007).

PICES website (Agenda Item 8)

Ms. Julia Yazvenko (PICES Secretariat) reported on the new, much expanded PICES website and solicited feedback and advice from TCODE members. The Committee congratulated the Secretariat on the revised website, which is much more comprehensive and up-to-date than formerly. Dr. Megrey suggested that web-based collaboration software (SharePoint) might be a good investment for the PICES website, to more fully support the activities of various PICES groups. However, this software has a disadvantage – it works only with the MS Internet Explorer browser. It was proposed to provide services at the PICES website for the news subscription and website monitoring.

Discussion of steps towards next major PICES scientific program(s) (Agenda Item 9)

The Committee felt that implementation of the Global Ocean Observing System (GOOS) in the open ocean and the continental seas of the North Pacific would provide a major focus for PICES activities over the next five years. This would require close coordination between MONITOR and TCODE to ensure that data management challenges are fully addressed.

Topic session proposals for PICES XIV (Agenda Item 10)

TCODE recommends a ½-day Topic Session on "Data management and data delivery systems to support ecosystem monitoring" (*TCODE Endnote 3*).

PICES XV theme (Agenda Item 11)

The Committee had no recommendation for the PICES XV theme.

Relations with other international programs and organizations (Agenda Item 12)

The Committee recommended adding IODE, CoML Ocean Biological Information System (http://www.obis.org/) and the IOC Marine XML project (http://marinexml.net) to the PICES Standing List.

Membership and Chairmanship of TCODE (Agenda Item 13)

TCODE has several ex-officio members who provide tremendous support to all Committee activities. The Committee supported the recommendation that a Vice-Chairman of TCODE be appointed to represent the Committee at selected meetings and help the Chairman with the duties of planning and running the committee business on an on-going basis throughout the year. The Vice-Chairman could potentially be eligible for election to Committee Chairman once the Chairman's term has been completed.

The Committee unanimously supported the extension of Dr. Shevchenko's term as TCODE Chairman for one additional year.

Summary of items with financial implications (Agenda Item 14)

TCODE requests:
- travel support for one invited speaker (an ICES XML expert) for the TCODE Topic Session on "Data management and data delivery systems to support ecosystem monitoring" at PICES XIV (*TCODE Endnote 3*);
- travel support for a TCODE representative to attend the International Conference on "Marine Biodiversity Data Management" (November 29 - December 1, 2004, Hamburg, Germany) to get a better understanding and linkage of such initiatives as the Generic Information Retrieval (DiGIR), OBIS, MarineXML, Global Biodiversity Information Facility (GBIF), etc., with TCODE activities;
- support for the pilot project to "Federate Metadata on North Pacific Ecosystems", and for publishing findings from this project in 2006 (*TCODE Endnote 4*).

TCODE Workplan for 2004/2005 (Agenda Item 16)

TCODE Workplan for 2004/2005:
- Convene a ½-day Topic Session on "Data management and data delivery systems to support ecosystem monitoring" at PICES XIV;
- Carry out a pilot project to "Federate Metadata on North Pacific Ecosystems";
- Develop TCODE Action Plan to complement the PICES Strategic Plan and Capacity Building initiative;
- Update TCODE web pages, including information on annual national reports, metadata languages, etc.;
- Provide assistance to the Section on *Ecology of harmful algal blooms in the North Pacific* with database activities and publicize HAB database activity;
- Coordinate TCODE activities with the new MONITOR Technical Committee;
- Develop improved understanding of the Ocean Biological Information Systems and other related projects.

TCODE Endnote 1

Participation List

<u>Members</u>

Robin Brown (Canada)
Hee-Dong Jeong (Korea)
Hae Seok Kang (Korea)
Bernard A. Megrey (U.S.A)
Thomas C. Royer (U.S.A)
Igor I. Shevchenko (Russia, Chairman))
Ling Tong (China)

<u>Observers</u>:

Kenji Asano (Japan)
Tatsu Kishida (Japan)
S. Allen Macklin (U.S.A)
Thomas Malone (U.S.A)
Phillip Mundy (U.S.A., MONITOR Chairman)

TCODE Endnote 2

TCODE Meeting Agenda

1. Welcome and introduction of members
2. Adoption of agenda
3. Review progress on items in the 2003/2004 Workplan
   a. TCODE/HAB-S work on HAE-DAT database
   b. GLOBEC IPO/TCODE interactions on data management
   c. North Pacific Ecosystem Metadatabase
   d. Updating TCODE web pages
   e. Electronic Poster Session at PICES XIII
   f. Interactions with ICES-IOC Study Group on the Development of marine data exchange systems using Extensible Markup Language (XML)
4. Annual reports and updates on data management activities in PICES member countries
5. Second inter-sessional Science Board Meeting
6. Discussion of TCODE activity in relation to the PICES Strategic Plan and development of a draft Action Plan for current and future activity
7. Discussion of capacity building opportunities
8. PICES website
9. Discussion of steps towards next major PICES scientific program(s)
10. Topic session proposals for PICES XIV
11. PICES XV theme
12. Relations with other international programs and organizations
13. Membership and Chairmanship of TCODE (including possibility of having a Vice-Chairman position)
14. Summary of items with financial implications
15. New business
16. TCODE Work plan for 2004/2005

TCODE Endnote 3

**Proposal for a ½-day TCODE Topic Session at PICES XIV on
"Data management and delivery systems to support ecosystem monitoring"**

A stated objective of PICES is to provide data in exchangeable formats to better enable the evaluation of North Pacific ecosystems status and trends, and to support other strategic pursuits. PICES scientists face challenges in managing and delivering data in a shareable way. Furthermore, a growing number of ocean observing systems require data management and communication methodologies that conform to rigid standards and protocols. For the most part, traditional science education of the past century offered little training in data management. Today's typical scientist, although supportive of data exchange, lacks the background to understand techniques to facilitate it.

This session will acquaint PICES scientists with state-of-the-art information about metadata description, data delivery and data browsing techniques, with emphasis on existing standards and web services recommended for ocean observing systems. Basic to advanced methods will be presented in a graduated format. Presenters and participants will describe and explore existing, successful systems. Participants will learn ways to map existing data structures into conformant, exchangeable formats using no-cost, open-source applications.

Presenters and participants will describe and explore existing successful systems. Participants will learn ways to map existing data structures into conformant, exchangeable formats using no-cost, open-source applications.

Recommended convenors: S. Allen Macklin (U.S.A.), Bernard A. Megrey (U.S.A.) and Igor I. Shevchenko (Russia).

Suggested resources: 5-10 work stations with Internet connection.

TCODE Endnote 4
PICES TCODE data-sharing pilot project to "Federate Metadata on North Pacific Ecosystems"

Informal discussions between the Korean Oceanographic Data Center (KODC) and the North Pacific Ecosystem Metadatabase (NPEM) have taken place over the past year. The goal of these discussions is to ultimately connect PICES member nations' metadatabase systems into one integrated resource. In this way, a user of any one metadata inventory will have the ability to search for data catalogued by any and all other participating system with a single search request. Using modern data management techniques to cross-search separate metadatabases provides the advantages of shared metadata without compromising national ownership, data integrity, or security of national metadata products.

Our informal discussions indicated that this project is feasible and worthy of immediate pursuit. TCODE has adopted a pilot KODC-NPEM federation as part of its 2005 work plan. We plan to initiate a series of meetings, open to all PICES member nations, to discuss the required technical details. During the process of outlining and discovering the technical hurdles and the means to address and solve problems with KODC and other participants, we intend to document our findings in a report, to be published during 2006, that can be used by other PICES countries wishing to join the federation. In this way, success with this project should easily scale up to similar efforts with other PICES countries with a nominal investment of time and planning.

We believe this project is compatible with the PICES Strategic Plan and PICES scientific interests. It accomplishes capacity sharing because all subscribing systems will possess the ability to virtually expand their holdings to those of the complete network. Data sharing between PICES countries helps to advance PICES science.

We ask PICES to support this project on a shared-cost basis. Korea and U.S.A. will contribute ~37% of the total cost for two meetings of four persons from each organization. TCODE requests that PICES pay 25%. Other participants are welcome to join the meetings, as possible.

Projected budget:
    Total cost: US\$ 16,000
        Two 2-day meetings with 4 people from each country (2 managers, 2 technical staff) at US\$ 2,000/person/trip.
    Contribution from U.S.A.: US\$ 6,000
    Contribution from Korea: US\$ 6,000
    Contribution from PICES: US\$ 4,000

Proponents of the proposal:
    Hee-Dong Jeong (KODC/NFRDI, Korea)
    Hae-Seok Kang (KORDI, Korea)
    S. Allen Macklin (NOAA PMEL, U.S.A.)
    Bernard A. Megrey (NOAA AFSC, U.S.A.)