

**MEETING OF THE AD HOC COMMITTEE FOR ADVICE
ON THE IMPLEMENTATION OF THE ICCAT RELATIONAL
DATABASE MANAGEMENT SYSTEM (ICCAT-RDB)**

(Madrid, May 13-14, 2001)

1 Opening, adoption of agenda and election of rapporteur

The meeting took place at the ICCAT Secretariat. The meeting was opened by Dr. Joseph E. Powers, SCRS Chairman. He welcomed the participants (see **Appendix 1**), who introduced themselves.

The Tentative Agenda (**Appendix 2**), which had been circulated previously, was adopted. Dr. Powers noted that the Chairman of the Sub-Committee on Statistics, Dr. Steve Turner was not able to attend and that he sent his regrets. At the suggestion of the group, Dr. Powers agreed to act as the Chairman of the meeting and (with the assistance of Dr. Carlos Palma) to be rapporteur. A List of Documents presented at the meeting is provided in **Appendix 3**.

2 Overview of current development status of ICCAT-RDB

Document SCRS/01/040 was presented which summarized the status of the development of the database.

2.a ICCAT-RDB characteristics

The ICCAT relational database management system adopted a Client/Server architecture with a two-tier design. The database server adopted was the Microsoft SQL server 2000 (MS-SQL). The client interface framework in development uses ActiveX Data Objects (ADO) as the data access model. The client interface will have distinct client programs (in development), each one with distinct goals:

Microsoft Access Projects: Programs planned to work as the “Data-Center” of each database. They will provide us with the standard input and output functionality. In other words, they will centralize in one single client interface, the data entry (manual and automatic) and validation processes, the edit/update processes, and the systematic querying processes (output with various levels of aggregated data, visual queries in the form of standard charts, etc.).

N-dimensional Cubes: OLAP (On-line Analytical Processes) cubes, developed with Microsoft Analysis Manager and published in Excel worksheets. These cubes structures are extremely powerful analytical tools. They allow us to dynamically arrange (transform, aggregate, join, etc.) the data in virtually any possible combination of dimensions (fields) available in the main database.

Specific executable programs: Programs developed in compiled languages (Delphi, MS-Visual C++, MS-Visual Basic, etc.) that will perform specific tasks, normally involving complex calculations (e.g., estimations of catch-at-size composition by species).

Apart from the above programs, the Microsoft SQL Query Analyzer program is a client program, an interface that provides the most powerful (however the most difficult) querying tool.

In terms of database structures, the ICCAT-RDB is made-up of several databases, which according to their priorities were grouped as shown below:

S Statistical (higher priority)

Task 1:	Best ICCAT estimates of catches
Task 2a:	Catch & effort data
Task 2b:	Size composition of samples
Tag:	Tag-release data
Trade:	Tuna import/export (BTSD, etc.)
Sharks:	Sharks landings data

S Secretariat Support (lower priority)

Contacts:	Management of a list of contacts
Meetings:	Management of ICCAT meetings
Publications:	Management of ICCAT publications

S Other (long-term priority)

Environment:	Geo-referenced physic-chemical parameters (temperature, salinity, rainfall, etc.)
GIS:	Geo-features data (sampling areas, EEZ areas, bathymetric lines, etc.)

2.b Work performed

The work performed during the last five months (Dec 2000 to April 2001) covered several tasks as originally scheduled in Doc. SCRS/00/043 (Palma 2001). They are summarized in Doc. SCRS/01/40, which also describes schematically the current status of each database by using a qualitative maturity scale of development stages: (i) data collection and database prototypes; (ii) data manipulation and normalization; (iii) data migration and validation of migration process; (iv) tuning final structures and development of client interfaces; and (v) full client/server functionality.

The Task 1 database is currently the database in the most advanced development stage.

At present, the last update of “Old T1” (March 2001) is fully contained in an MS-SQL database called “T1”, containing 25608 current records and 21846 historical updates. The migration process was validated and the client interface for T1 (T1- Data Center) is achieving its final development stage. The output results of some modules are already checked. The input processes are in the implementation phase. Given the current advanced development stage of Task-1 database, the old database version shall be discontinued next June.

2.c Changes from original plan

The major change made to the original plan (SCRS/00/43) was the adoption of a Client/Server platform from the beginning of Phase-1, without passing through an intermediate Client database (a recommendation of the *Ad Hoc* Advisory Committee on the Design and Implementation of the ICCAT-RDB); see DET/00/STAT (ICCAT 2001).

As a consequence, Phase 1 and Phase 2 were joined into one development Phase (Phase-1). Nevertheless, the overall time schedule is the same.

2.d Problems encountered

The ICCAT codes classification for Statistical Areas, Gears and Fleets do have some weaknesses. In

summary, they do not have a normalized (overlapped attributes) hierarchical classification, and present some structural incoherency among particular codes.

3 Measures and guidelines for ICCAT-RDB fine-tuning

Document SCRS/01/41 was also discussed, which presented in some detail, the work performed so far with the new Task-1 database (T1), and also several proposals envisaging the improvement of its overall potential.

3.a Structural and functional models

Several approaches were adopted during T1 implementation aiming to increase its flexibility and functionality. The relational database structure of T1 (SCRS/01/41) reflects these approaches. It is functionally a final relational structural (new tables, relationships, rules, hierarchical codes for Statistical Areas, Gears, Species, Fleets, etc.). It also contains some temporary fields responsible for preserving the connection and structure of “Old T1” (an indirect process of reverting every adoption made to the old structure). Adding other relevant record based characteristics (fields) is a simple and straightforward process.

The migration process (made by a specific program that imported the database to the new structure) had a set of basic validation procedures. More advanced validation procedures (ex: species based range limits, etc.) shall be applied in a later phase (historical review & quality control).

The method used to validate the Task-1 migration process from the old to the new database, applied “cross checking” tests between equivalent outputs of both databases (element-wise subtraction of similar output matrices obtained from the old database and the new database).

At present, the largest portion of the development time is associated with the development of the T1-Data Center. However, the largest fraction of the modules (integrated routines) developed needs to be tested. Complete and test input and output routines of T1-Data Center will be a task for the next 2 months.

Several technical approaches were introduced with the aim of improving the T1 database development process (outlined in SCRS/01/41). These were discussed and approved at the meeting.

3.b Input processes and submission rules

The most efficient approach to manage data input (new records and data updates preserving an historical relation between updated records) in a database is to group them into “Input Processes”. This approach has many advantages: (i) data tracking system integrated with no need to manage data tracking outside the database; (ii) simplifies all the automation processes; (iii) several data layers can be used in the process; and (iv) it centralizes validation before passing data to Current or Historical tables. The drawback is that rigid input rules have to be implemented. Methods were discussed which would result in an efficient way of rebuilding the history of data submission processes. The details are in SCRS/01/41.

Statistical correspondents should be encouraged to submit data in a compatible and standard electronic format. Through the data base management system development, the Secretariat will supply software to allow compatible electronic formats to be used.

3.c Output procedures

Demonstrations of standardized real time outputs were given at the meeting, including Task 1 catch figures typically given in the Executive Summaries of the SCRS Report. The output system has the flexibility for additional standard graphical outputs to be generated. Species Working Groups should provide the Secretariat some guidance on the kind and extent of standardized outputs that are likely to be needed.

4 Review and improvement of current ICCAT look-up tables and codes

In the development of the data base management system, several difficulties have arisen in relation with the ICCAT codes and look-up tables that are used to classify data. In relational databases, the hierarchical classification of attributes (e.g., gears) is one of the most important techniques used in systematization of output processes and to increase flexibility (allow multiple levels of attributes in only one field). The current ICCAT codes used for Gears, Statistical Areas, Fleets lack a well structured classification. To increase the levels of flexibility of the T1 database, the ICCAT codes for the above cases should be changed to hierarchical codes. Hierarchical codes are needed for gears, species and statistical areas/stock boundaries. Hierarchical coding systems will make the use of the system in the future more efficient and will maintain the quality of the data. However, there will be difficulties in recreating historical data using new coding systems. Possible solutions were outlined in SCRS/01/41. The group agreed that the Secretariat should have the flexibility to change the coding system, as appropriate, recognizing that it will be a complex process. But the Secretariat should report on progress on this aspect at the next meeting and particularly any problems that arise. In any case, a document should be maintained showing the relationship between new and old codes.

5 Possible topics/areas for discussion at next meeting, future recommendations and other matters

The next meeting should include discussion of further progress on the data base management system development. In particular, it is expected that the development will be nearing completion for Task I (Task 1 database) and then the work will move toward Task II (Task 2a and Task 2b databases). Necessarily, converting of Task II databases will be more complex than that for Task I. The catch-at-size sampling data, and effort data (appropriately stratified) will be the emphasis of future work.

When addressing the Task II databases at the next meeting, some guidelines should be outlined in order to define standard substitution and raising rules (for each species) used in catch-at-size estimations.

Additionally, there is a need for databases that are specialized for individual species working groups. In particular, these might include biological parameter data (e.g. growth, weight-length and sex data). Processes should be developed to collect these data over the short term with the goal of fully integrating them into the database at a future date. Also, questions of how to handle NEI data, compliance data and confidential data will have to be addressed. The Secretariat will need guidance on the priorities for integrating these databases in the future.

6 Adoption of the report and closure of the meeting

The group adopted the report after reviewing it by e-mail. The group asked the Secretariat that the report be translated prior to the SCRS meeting. The meeting Chairman thanked the attendants and the Secretariat for their participation. The meeting was adjourned.

REFERENCES

PALMA, Carlos. 2001. Strategy Towards the ICCAT Relational Database Management System (ICCAT-RDB). Col. Vol. Sci. Pap. ICCAT, 52: 1875-1894.

ICCAT 2001. Report of the First Meeting of the *Ad Hoc* Advisory Committee on the Design and Implementation of the Relational Data Base System, in the Report of the Sub-Committee on Statistics. Col. Vol. Sci. Pap. ICCAT, 52: 1793-1831.

RÉUNION DU COMITÉ CONSULTATIF AD HOC SUR LA MISE EN PLACE DE LA BASE DE DONNÉES RELATIONNELLE DE L'ICCAT (ICCAT-RDB)

(Madrid, Espagne, 13-14 mai 2001)

1 Ouverture, adoption de l'ordre du jour et désignation du rapporteur

Les sessions se sont déroulées dans les locaux du Secrétariat de l'ICCAT. Les débats ont été ouverts par le D^r Joseph E. Powers, Président du SCRS, qui a souhaité la bienvenue aux participants (voir l' **Appendice 1**), qui se sont présentés.

L'ordre du jour provisoire qui avait été diffusé avant la réunion a été adopté (**Appendice 2**). Le D^r Powers a commenté que le Président du Sous-comité des Statistiques, le D^r Steve C. Turner, n'était pas en mesure d'assister à la rencontre, et tenait à transmettre ses regrets à cet égard. A la suggestion des participants, le D^r Powers a accepté de diriger les sessions, et (avec l'aide de M. Carlos Palma) d'en être le rapporteur. La liste des travaux présentés à la réunion figure en **Appendice 3**.

2 Vu d'ensemble des progrès réalisés dans la mise en place de l'ICCAT-RDB

Le document SCRS/01/040, qui récapitulait la situation de création de la base de données, a été présenté.

2.a Caractéristiques de l'ICCAT-RDB

Le système de gestion d'une base de données relationnelle adopté par l'ICCAT suivait une architecture client/serveur structurée sur trois plans. La base adoptée était la Microsoft SQL Server 2000 (MS-SQL). La structure interface du client qui est en cours d'élaboration utilise ActiveX Data Objects (ADO) comme modèle d'accès. L'interface du client comprendra plusieurs programmes client (en préparation) ayant chacun des objectifs distincts:

Applications sous Microsoft Access. Programmes conçus pour fonctionner en tant que "Data-Center" de chaque base de données. Ils nous assureront une fonctionnalité standard d'entrée et de sortie. Autrement dit, ils centraliseront sur une interface client unique les processus d'entrée des données (manuelles et automatiques) et de validation, les processus de correction/actualisation et les processus systématiques de consultation (données de sortie à divers niveau d'agrégation, consultations visuelles sous forme de cartes standard, etc.).

Cubes N-dimensionnels. Cubes OLAP (On-line Analytical Processes), élaborés au moyen du Microsoft Analysis Manager et publié sous forme de fiches de travail. Ces structures en cubes constituent un outil analytique extrêmement puissant. Elles nous permettent de disposer (transformer, agréger, unir, etc.) les données de façon dynamique selon pratiquement toute combinaison possible de dimensions (champs) disponibles dans la base principale de données.

Programmes spécifiques exécutables. Programmes élaborés au moyen de langages compilés (Delphi, MS-Visual C++, Ms-Visual Basic, etc.) qui exécuteront des tâches précises, mettant normalement en jeu des calculs complexes (par exemple, estimations de la prise par taille, composition spécifique).

A part les programmes ci-dessus, le programme Microsoft SQL Query Analyzer est un programme client, une interface qui constitue l'outil de consultation le plus puissant (bien que le plus difficile).

En termes de structure de bases de données, l'ICCAT-RDB se compose de plusieurs bases, qui sont regroupées ci-après par ordre de priorité:

- Statistique (priorité absolue)
 - Task 1: Meilleure estimation ICCAT des captures
 - Task 2a: Données de prise et effort
 - Task 2b: Composition en tailles des échantillons
 - Tag: Données de marquage-recapture
 - Trade: Importation/exportation de thon (BTSD, etc.)
 - Sharks: Données sur les débarquements de requins
- Appui du Secrétariat (moindre priorité)
 - Contacts: Gestion d'une liste de contacts
 - Meetings: Gestion des réunions ICCAT
 - Publications: Gestion des publications ICCAT
- Autres (priorité à long terme)
 - Environment: Paramètres physico-chimiques géo-associés (température, salinité, précipitations, etc.)
 - GIS: Données géo-caractéristiques (zones d'échantillonnage, ZEE, bathymétrie, etc.)

2.b Travaux réalisés

Le travail réalisé pendant cinq mois, décembre 2000-avril 2001, couvre plusieurs des tâches prévues au départ dans le document SCRS/00/043 (Palma 2001). Elles sont récapitulées dans le document SCRS/01/040, qui décrit aussi de façon schématique l'état actuel de chaque base de données au moyen d'une échelle qualitative du degré relatif de maturité des stades de développement: i) collecte de données et prototypes de base de données; ii) manipulation et normalisation des données; iii) migration de données et validation du processus de migration; iv) calibrage des structures définitives et élaboration des interfaces client; et v) fonctionnement intégral client/serveur.

La base de données Tâche 1 est actuellement celle qui en est au stade le plus avancé de développement.

A l'heure actuelle, la dernière actualisation de « Old T1 » (mars 2001) est entièrement reprise dans une base de données MS-SOL dénommée "T1", qui contient 25608 registres actuels et 21846 actualisations historiques. Le processus de migration a été validé, et l'interface client de T1 (T1-Data Center) en est au stade final d'élaboration. Les résultats de sortie de quelques modules ont déjà été vérifiés. Les processus d'entrée en sont au stade d'implantation. Vu l'état avancé de l'élaboration de la base de données Task-1, l'ancienne version de la base sera interrompue au mois de juin prochain.

2.c Modifications apportées au plan de travail original

La principale modification apportée au plan original (document SCRS/00/43) a été l'adoption d'une plate-forme client/serveur dès le début de la Phase 1, sans passer par une base intermédiaire client (il s'agit d'une recommandation du Comité consultatif *ad hoc* sur le concept et la réalisation de l'ICCAT-RDB, cf. DET/00/STAT, ICAT 2001).

Par conséquent, la Phase 1 et la Phase 2 ont été condensées en une seule phase d'élaboration (la Phase 1). Le calendrier global d'exécution est toutefois inchangé.

2.d Problèmes rencontrés

La classification des codes ICCAT pour les Zones statistiques, les Engins et les Flottes présente quelques faiblesses. En résumé, elle ne comporte pas de classification hiérarchique normalisée (chevauchement d'attributions), et elle présente quelques incohérences structurelles entre des codes particuliers.

3 Etalons et directives pour le calibrage fin de l'ICCAT-RDB

Le document SCRS/01/041, qui contenait un exposé assez détaillé du travail réalisé jusqu'à maintenant en ce qui concerne la nouvelle base de données Tâche 1 (T1), ainsi que plusieurs propositions pour l'amélioration de son potentiel global, a également été discuté.

3.a Modèles structurels et fonctionnels

La mise en place de la T1 a été abordée de plusieurs manières, de façon à accroître sa flexibilité et sa fonctionnalité. La structure de base de données relationnelle de la T1 (SCRS/01/041) reflète ces approches. Il s'agit, du point de vue fonctionnel, d'une structure relationnelle définitive (nouvelles tables, relations, normes, codes hiérarchiques pour les Zones statistiques, les Engins, les Espèces, les Flottes, etc.). Elle contient également quelques champs temporaires responsables de maintenir la connexion et la structure de « Old T1 » (un processus indirect d'inversion de toute adoption faite dans l'ancienne structure). L'adjonction d'autres caractéristiques (champs) aux registres est un processus simple et direct.

Le processus de migration (effectué par un programme spécifique qui a importé la base de données dans la nouvelle structure) a un jeu de processus de base de validation. Des processus de validation plus poussés (par exemple, des limites de gammes basées sur les espèces, etc.) seront appliqués lors d'une phase ultérieure (révision historique et contrôle de qualité).

La méthode utilisée pour valider le processus de migration de la Tâche 1 de l'ancienne à la nouvelle base de données appliquait des tests de contre-vérification entre des sorties équivalentes des deux bases (soustraction par élément de matrices similaires de sortie obtenues de l'ancienne base et de la nouvelle).

A l'heure actuelle, la plus grande partie du temps consacré au développement est liée à l'élaboration du T1-Data Center. Toutefois, la majeure partie des modules (routines intégrées) créées doivent être testés. Compléter et tester les routines d'entrée et de sortie du T1-Data Center sera la tâche des deux mois à venir.

Plusieurs détails techniques ont été introduits dans le but d'améliorer le processus d'élaboration de la base de données T1 (description générale dans le document SCRS/01/041). Ceci a été discuté et approuvé à la réunion.

3.b Normes pour les processus d'entrée et la présentation

La façon la plus efficace de gérer l'entrée des données (nouveaux registres et mises à jour conservant une relation historique entre les registres actualisés) dans une base est de les regrouper dans des « processus d'entrée ». Cette méthode présente de nombreux avantages : i) elle intègre un système de suivi des données sans qu'il soit nécessaire de gérer le suivi de données hors de la base ; ii) elle simplifie tous les processus d'automatisation ; iii) elle permet d'utiliser plusieurs niveaux de données dans le processus ; et iv) elle centralise la validation avant de passer les données à des tables actuelles ou historiques. L'inconvénient en est qu'il faut imposer des normes rigides d'entrée. Les

débats ont porté sur les méthodes pouvant reconstituer de façon efficace l'évolution de la transmission des données. Elles sont décrites en détail dans le document SCRS/01/041.

Il faut encourager les correspondants pour les statistiques à soumettre les données sous un format compatible et standard. Dans le cadre du développement du système de gestion de la base de données, le Secrétariat fournira le logiciel nécessaire pour assurer l'utilisation de formats électroniques compatibles.

3.c Processus de sortie

Des démonstrations de sorties standardisées en temps réel ont été faites à la réunion, dont les figures de capture Tâche 1 qui accompagnent normalement les résumés exécutifs du rapport du SCRS. Le système de sortie est assez flexible pour créer des sorties graphiques standards additionnelles. Les Groupes d'espèce devraient donner au Secrétariat quelque orientation sur la nature et l'ampleur des sorties standards dont ils auront probablement besoin.

4 Examen et amélioration des tableaux "look-up" et des codes ICCAT

En élaborant la structure de gestion de la base de données, plusieurs difficultés ont surgi en ce qui concerne les codes ICCAT et les tables « look-up » qui sont utilisées pour classifier les données. Dans les bases relationnelles, le classement hiérarchique des attributs (par exemple, les engins) est l'une des plus importantes techniques de systématisation des processus de sortie dont il accroît la flexibilité (en permettant de multiples niveaux d'attribution dans un même champ). Les codes ICCAT qui sont utilisés actuellement pour les Engins, les Zones statistiques et les Flottes n'ont pas une classification bien structurée. Pour accroître la flexibilité de la base de données T1, les codes ICCAT des cas susmentionnés devraient être remplacés par des codes hiérarchiques. Les codes hiérarchiques sont nécessaires pour les engins, les espèces et les délimitations entre zones statistiques/stocks. Un codage hiérarchique permettra une utilisation plus efficace du système à l'avenir, et maintiendra la qualité des données. Toutefois, il y aura des difficultés au moment de recréer les données historiques selon le nouveau codage. Des solutions éventuelles sont ébauchées dans le document SCRS/01/041.

Les participants ont convenu que le Secrétariat devrait disposer de flexibilité pour changer le système de codage, selon les besoins, tout en constatant qu'il s'agissait d'un processus complexe. Le Secrétariat devrait faire part de l'avancement des travaux à la prochaine réunion, en mentionnant notamment tout problème rencontré. De toutes façons, il faut dresser un document établissant les correspondances entre les codes nouveaux et anciens.

5 Thèmes/domaines à débattre éventuellement à la prochaine réunion, recommandations pour l'avenir et autres questions

La prochaine réunion devrait comprendre des débats sur l'élaboration du système de gestion de la base de données. Il est prévu, notamment, que le travail soit pratiquement achevé en ce qui concerne la Tâche I (base de données Task 1), et que le travail va pouvoir s'orienter vers la Tâche II (bases de données Task 2a et Task 2b). La conversion des bases Tâche II sera forcément plus complexe que pour la Tâche I. Le travail futur se centrera sur les données d'échantillonnage de prise par taille et sur les données d'effort (stratifiées de façon adéquate).

Au moment d'aborder les bases de données Tâche II à la prochaine réunion, quelques directives devront être dressées (pour chaque espèce) pour définir les normes standard de substitution et d'extrapolation qui sont utilisées dans les estimations de la prise par taille.

En outre, il faut des bases de données spécialisées pour chaque groupe de travail. Ces bases pourraient comprendre, en particulier, des données sur les paramètres biologiques (par exemple, la

croissance, la relation longueur-poids ou les données sur le sexe). Il faudrait mettre sur pied des processus pour collecter ces données à court terme dans le but de les intégrer entièrement dans la base de donnée à une date ultérieure. Il faudra également aborder la façon de manipuler les données NEI, les données sur l'application et les données confidentielles. Le Secrétariat devra recevoir à l'avenir une orientation sur l'ordre de priorité de l'intégration de ces bases de données.

6 Adoption du rapport et clôture

Le groupe a adopté le rapport après l'avoir révisé par e-mail. Les participants ont demandé au Secrétariat que ce rapport soit traduit avant la réunion du SCRS. Le président a remercié les participants et le Secrétariat de leur participation. Les débats ont été levés.

RÉFÉRENCES

PALMA, Carlos. 2001. Strategy towards the ICCAT Relational Database Management System (ICCAT-RDB).

Rec. Doc. Scient. ICCAT, 52 : 1875-1894.

ICCAT 2001. Rapport de la 1^e Réunion du Comité consultatif *ad hoc* sur l'Elaboration et la mise en place de la Base de Données relationnelle, Rapport du Sous-comité des Statistiques. Rec. Doc. Scient. ICCAT 52 : 1793-1831.

REUNIÓN DEL COMITÉ AD HOC DE ASESORAMIENTO SOBRE LA IMPLEMENTACIÓN DEL SISTEMA DE GESTIÓN DE LA BASE DE DATOS RELACIONAL DE ICCAT (ICCAT-RDB)

(Madrid, España, 13-14 de mayo, 2001)

1 Apertura, adopción de la agenda y elección del relator

La reunión tuvo lugar en la Secretaría de ICCAT. La reunión fue inaugurada por el Dr. Joseph E. Powers, Presidente del SCRS, que dio la bienvenida a los participantes, (**Apéndice 1**), quienes se presentaron ellos mismos.

La Agenda provisional (**Apéndice 2**), que había sido distribuida previamente, fue adoptada por el Dr. Powers, dado que el Presidente del Subcomité de Estadísticas, Dr. Steve Turner, no pudo asistir a la reunión, por lo que envió sus disculpas. A sugerencia del grupo, el Dr. Powers aceptó actuar como Presidente de la reunión y (con ayuda del Dr. Carlos Palma) actuar como relator. En el **Apéndice 3** se facilita una lista de los documentos presentados en la reunión.

2 Perspectiva del actual estado de desarrollo de la ICCAT-RDB

Se presentó el documento SCRS/01/040, que resumía el estado del desarrollo de la base de datos.

2.a Características de ICCAT-RDB

El sistema de gestión de la base de datos relacional de ICCAT adoptó una construcción Cliente/Servidor con un diseño de dos capas. El servidor de base de datos adoptado fue Microsoft SQL server 2000 (MS-SQL). El marco de interfaz de cliente en desarrollo utiliza objetos de datos ActiveX (ADO) como modelo de acceso de datos. La interfaz de cliente tendrá distintos programas cliente (en desarrollo), cada uno con diferentes objetivos:

Proyectos de Microsoft Access: Programas pensados para trabajar como “Centro de Datos” de cada base de datos. Proporcionan la funcionalidad estándar de entrada y salida. En otras palabras, centralizarán en una única interfaz de cliente, la entrada de datos (manual y automática) y los procesos de validación, los procesos de edición / actualización, y los procesos sistemáticos de consulta (salida con varios niveles de datos agregados, consultas visuales en forma de gráficos estándar, etc).

Cubos N-dimensionales: Cubos OLAP (On-line Analytical Processes), desarrollados con Microsoft Analysis Manager y publicados en hojas de cálculo de Excel. Estas estructuras de cubos son herramientas analíticas muy potentes que nos permiten disponer dinámicamente (transformar, agregar, unir, etc.) los datos en prácticamente cualquier posible combinación de dimensiones (campos) disponibles en la base de datos principal.

Programas ejecutables específicos: Programas desarrollados en lenguajes compilados (Delphi, MS-Visual C++, Ms-Visual Basic, etc.) que realizarán tareas específicas que normalmente implican cálculos complejos (por ejemplo, estimación de la composición de captura por tallas por especies).

Aparte de los programas mencionados anteriormente, el Microsoft SQL Query Analyzer es un programa cliente, una interfaz que proporciona la herramienta de consulta más potente (aunque la más difícil).

En términos de estructuras de bases de datos, la ICCAT-RDB se compone de varias bases de datos, que según sus prioridades se agruparon de la siguiente forma:

- Estadística (alta prioridad)

Tarea 1:	Mejores estimaciones ICCAT de capturas
Tarea 2a:	Datos de captura y esfuerzo
Tarea 2b:	Composición por talla de las muestras
Marcas:	Datos de marcado-liberación
Comercial:	Importación / exportación de túnidos (DEAR, etc.)
Tiburones:	Datos de desembarques de tiburones

- Apoyo a la Secretaría (menor prioridad)

Contactos:	Gestión de una lista de contactos
Reuniones:	Gestión de las reuniones de ICCAT
Publicaciones:	Gestión de las publicaciones de ICCAT

- Otros (prioridad a largo plazo)

Medioambiente:	Parámetros físico-químicos geo-referenciados (temperatura, salinidad, pluviosidad, etc.,)
GIS:	Datos de objetos geográficos (zonas de muestreo, área de ZEE, líneas batimétricas, etc.,)

2.b Trabajo realizado

El trabajo realizado durante los 5 últimos meses (diciembre de 2000 hasta abril de 2001) cubrió varias tareas, tal y como se programó originalmente en el documento SCRS/00/043 (Palma 2001). Están resumidas en el documento SCRS/01/040, que describe también esquemáticamente el estado actual de cada base de datos, utilizando una escala de madurez cualitativa para las fases de desarrollo: i) recopilación de datos y prototipos de base de datos; ii) manipulación de datos y normalización; iii) migración de datos y validación de los procesos de migración; iv) puesta a punto de las estructuras finales y desarrollo de interfaces de cliente; y v) completa funcionalidad cliente/servidor.

Actualmente, la base de datos de la Tarea-1 es la que se encuentra en una fase de desarrollo más avanzada.

En la actualidad, la última actualización de la “Antigua T1” (marzo de 2001) se encuentra al completo en una base de datos MS-SQL llamada “T1”, y contiene 25608 registros actuales y 21846 actualizaciones históricas. El proceso de migración fue validado y la interfaz de cliente para T1 (T1-Centro de Datos) está alcanzando su fase final de desarrollo. Los resultados de salida de algunos módulos han sido ya comprobados. Los procesos de entrada se encuentran en fase de implementación. Dado el actual estado de desarrollo avanzado de la base de datos de la Tarea-1, la versión antigua de la base de datos se interrumpirá el próximo mes de junio.

2.c Cambios respecto al plan original

El principal cambio realizado respecto al plan original (SCRS/00/43) ha sido la adopción de una plataforma Cliente / Servidor desde el principio de la Fase-1, sin pasar a través de una base de datos Cliente intermedia (una recomendación del Comité *ad hoc* de Asesoramiento en el diseño e implementación de ICCAT-RDB); ver DET/00/STAT (ICCAT 2001).

Como consecuencia, la Fase 1 y la Fase 2 se unieron en una única fase de desarrollo (Fase-1) Sin embargo, el calendario global es el mismo.

2.d Problemas encontrados

La clasificación de los códigos ICCAT para Áreas Estadísticas, Artes y Flotas tiene algunos puntos débiles. En resumen, no tienen una clasificación jerárquica normalizada (atributos solapados) y presentan algunas incoherencias estructurales entre códigos especiales.

3 Medidas y directrices para optimizar la ICCAT-RDB

También se discutió el documento SCRS/01/041, que presenta más detalladamente el trabajo realizado hasta ahora con la nueva base de datos de la Tarea-1 (T1), además de nuevas propuestas con vistas a la mejora de su potencial global.

3.a Modelos estructurales y funcionales

Durante la implementación de T1 se adoptaron diversos enfoques con el fin de aumentar su flexibilidad y funcionalidad. La estructura de la base de datos relacional de T1 (SCRS/01/041) refleja estos enfoques. Funcionalmente, es una estructura final de base de datos relacional (nuevas tablas, relaciones, normas, códigos jerárquicos para Áreas Estadísticas, Artes, Especies, Flotas, etc). Además, contiene algunos campos temporales, que son los responsables de conservar la conexión y la estructura de “Antigua Tarea 1” (un proceso indirecto de revertir toda adopción realizada a la estructura antigua). Añadir otros registros pertinentes basados en características (campos) es un proceso simple y directo.

El proceso de migración (realizado por un programa específico que importó la base de datos a la nueva estructura) tenía un conjunto de procedimientos básicos de validación. En una fase posterior (revisión histórica y control de calidad), deberán aplicarse procedimientos de validación más avanzados (por ejemplo, límites de rango basados en las especies, etc.)

El método utilizado para validar el proceso de migración de la Tarea-1, de la antigua a la nueva base de datos, aplicó pruebas de “verificación cruzada” entre las salidas equivalentes de las dos bases de datos (sustracción elemento a elemento de matrices similares de salida obtenidas a partir de la antigua y la nueva base de datos).

En la actualidad, la mayor parte del tiempo de desarrollo está asociado con el desarrollo del Centro de Datos-T1. Sin embargo, la mayoría de los módulos (rutinas integradas) desarrollados todavía deben ser comprobados. Completar y probar las rutinas de entrada y salida del Centro de Datos-T1 será una labor que durará los próximos dos meses.

Para mejorar el proceso de desarrollo de la base de datos T1 (descrito en el documento SCRS/01/041) se introdujeron diversos enfoques técnicos, que fueron discutidos y aprobados en la reunión.

3.b Procesos de entrada y normas de presentación

El enfoque más eficaz para gestionar la entrada de datos (nuevos registros y actualizaciones de datos manteniendo una relación histórica entre los registros actualizados) en una base de datos es agruparlos en “Procesos de Entrada”. Este enfoque presenta varias ventajas: i) una sistema integrado de rastreo de datos sin necesidad de gestionar el rastreo de datos fuera de la base de datos, ii) simplifica todos los procesos de automatización, iii) en el proceso pueden utilizarse varias capas de datos, y iv) centraliza la validación antes de enviar los datos a las tablas Actuales o Históricas. El inconveniente es que deben implementarse unas normas estrictas de entrada. Se discutieron varios

métodos que darían como resultado una manera eficaz de reconstruir el historial de los procesos de presentación de datos. Para más detalles, ver SCRS/01/041.

Debe solicitarse a los corresponsales estadísticos que presenten los datos en un formato electrónico estándar y compatible. Mediante el desarrollo del sistema de gestión de la base de datos, la Secretaría proporcionará programas informáticos que permitan la utilización de formatos electrónicos compatibles.

3.c Procedimientos de salida

Durante la reunión se realizaron demostraciones de salidas estandarizadas en tiempo real, incluyendo figuras de captura de la Tarea 1 que habitualmente aparecen en los resúmenes ejecutivos del Informe del SCRS. El sistema de salida tiene flexibilidad para generar salidas gráficas estándar adicionales. Los Grupos de Trabajo sobre Especies deberán facilitar alguna guía a la Secretaría sobre la clase y el alcance de las salidas estandarizadas que, probablemente, serán necesarias.

4 Revisión y mejora de las actuales tablas y códigos de consulta de ICCAT

En el desarrollo del sistema de gestión de la base de datos surgieron varias dificultades con respecto a los códigos y tablas de consulta de ICCAT que se utilizan para clasificar los datos. En las bases de datos relacionales, la clasificación jerárquica de los atributos (por ejemplo, artes) es una de las técnicas más importantes utilizadas en la sistematización de los procesos de salida y para aumentar la flexibilidad (permite diferentes niveles de atributos en un solo campo). Los actuales códigos ICCAT utilizados para Artes, Áreas Estadísticas y Flotas no tienen una clasificación bien estructurada. Para aumentar los niveles de flexibilidad de la base de datos T1, los códigos ICCAT para los casos anteriormente mencionados deben transformarse en códigos jerárquicos. Los códigos jerárquicos son necesarios para los artes, las especies y las áreas estadísticas/límites de stock. Los sistemas de codificación jerárquica lograrán, en el futuro, que el uso del sistema sea más eficaz y mantenga la calidad de los datos. Sin embargo, surgirán dificultades para recrear los datos históricos utilizando nuevos sistemas de codificación. Las posibles soluciones se describen en el documento SCRS/01/041. El grupo acordó que la Secretaría tendría la flexibilidad de cambiar el sistema de codificación, como juzgue oportuno, reconociendo que será un proceso complicado. Pero la Secretaría deberá informar de sus progresos a este respecto en la próxima reunión y especialmente deberá informar de los problemas que surjan. En cualquier caso, debe mantenerse un documento que muestre la relación entre los antiguos y nuevos códigos.

5 Posibles temas/áreas de discusión en la próxima reunión, recomendaciones futuras y otros asuntos

La próxima reunión deberá incluir un debate sobre el progreso realizado en el desarrollo del sistema de gestión de la base de datos. En especial, esperamos que la labor de desarrollar la Tarea I (base de datos de la Tarea I) esté casi terminada y se proceda a desarrollar la Tarea II (bases de datos de Tarea 2a y Tarea 2b). Convertir las bases de datos de la Tarea II será más difícil que convertir la Tarea I. Los datos de muestreo de captura por talla y los datos de esfuerzo (estratificados de manera adecuada) serán el centro del trabajo futuro.

En la próxima reunión, cuando se traten las bases de datos de la Tarea II, deben esbozarse algunas directrices con el fin de definir la sustitución estándar y las normas de extrapolación (para cada especie) utilizadas en las estimaciones de captura por talla.

Además, son necesarias bases de datos especializadas para los grupos de trabajo individuales sobre especies. En especial, estas bases de datos podrían incluir datos sobre parámetros biológicos (por ejemplo, datos de crecimiento, peso-talla, y sexo). Deben desarrollarse procesos para recopilar estos

datos a corto plazo, con el fin de integrarlos completamente en la base de datos en un futuro. Además, deben solucionarse temas como el manejo de los datos NEI, los datos de cumplimiento y los datos confidenciales. La Secretaría necesitará alguna guía sobre las prioridades para integrar estas bases de datos en el futuro.

6 Adopción del informe y clausura de la reunión

El grupo adoptó el informe tras revisarlo por e-mail. El grupo solicitó a la Secretaría que el informe fuera traducido antes de la reunión del SCRS. El Presidente agradeció a los participantes y a la Secretaría su participación. La reunión fue clausurada.

REFERENCIAS

- PALMA, Carlos. 2001. Strategy Towards the ICCAT Relational Database Management System (ICCAT-RDB). Col. Vol. Sci. Pap. ICCAT, 52: 1875-1894.
- ICCAT. 2001. Report of the First Meeting of the Ad Hoc Advisory Committee on the Design and Implementation of the Relational Database System, in the Report of the Sub-Committee on Statistics. Col. Vol. Sci. Pap. ICCAT, 52: 1793-1831.

Participants/Participants/Participantes

EUROPEAN COMMUNITY

EC-SPAIN

Pallarés, Pilar

Instituto Español de Oceanografía, Corazón de María, 8, 28002 - Madrid

Tel: +34 91 347 3620; Fax: +34 91 413 5597; E-mail: pilar.pallares@md.ieo.es

EC-PORTUGAL

Pereira, Joao Gil

Universidade dos Açores, Departamento de Oceanografia e Pescas, 9900 - Horta, Açores,

Tel: +351 292 29 2945; Fax: +351 292 29 2659; E-mail: pereira@notes.horta.uac.pt

JAPAN

Uosaki, Koji

Western Pacific Tuna and Skipjack Resources Division, National Research Institute of Far Seas Fisheries; 7-1, Orido 5-Chome, Shimizu-shi, Shizouka 424-8633

Tel: +81 543 36 6033; Fax: +81 543 35 9642; E-mail: uosaki@enyo.affrc.go.jp

Uehara, Kazuyuki

Oceanography and Southern Resources Division, National Research Institute of Far Seas Fisheries, 7-1, Orido 5-Chome, Shimizu-shi, Shizouka 424-8633

Tel: +81 543 36 6067; Fax: +81 543 35 9642; E-mail: kuehara@enyo.affrc.go.jp

MAROC

Abid, Nouredine

Center Regional de L'INRH á Nador, B.P. 493, Nador

Tel: +212 56 60 08 69; Fax: +212 56 60 38 28; E-mail: abid@nadornet.net.ma

El Ktiri, I, Taoufik

Direction de Pêches maritimes et de l'Aquaculture, Ministère des Pêches Maritimes, Nouveau Quartier Administratif, Agdal - Rabat

Tel: +212 37 68 81 15; Fax: +212 37 68 82 13; E-mail: elktiri@mp3m.gov.ma

UNITED STATES

Powers, Joseph E.

SCRS Chairman - Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL 33149-1099

Tel: +1 305 361 4295; Fax: +1 305 361 4278; E-mail: joseph.powers@noaa.gov

ICCAT SECRETARIAT

Restrepo, Víctor

E-mail: victor.restrepo@iccat.es

Kebe, Papa

E-mail: papa.kebe@iccat.es

Palma, Carlos

E-mail: carlos.palma@iccat.es

Agenda

- 1 Opening of the meeting
- 2 Adoption of the Agenda and arrangements for the meeting
- 3 Election of the Rapporteur
- 4 Overview of current development status of ICCAT-RDB
- 5 General presentation of prototypes
 - S Changes from original work plan
 - S Problems encountered
- 6 Measures and guidelines for ICCAT-RDB fine-tuning
 - S Structural and functional models
 - S Input processes and submission rules
 - S Output procedures
- 7 Review and improvement of current ICCAT lookup tables and codes
- 8 Possible topics/areas for discussion at next meeting
- 9 Future recommendations
- 10 Any other matters
- 11 Adoption of the Report

Ordre du jour

- 1 Ouverture de la réunion
- 2 Adoption de l'ordre du jour et organisation des sessions
- 3 Désignation du rapporteur
- 4 Vu d'ensemble des progrès réalisés dans la mise en place de l'ICCAT-RDB
 - Exposé général sur les prototypes
 - Modifications apportées au plan de travail original
 - Problèmes rencontrés
- 5 Etalons et directives pour le calibrage fin de l'ICCAT-RDB
 - Modèles structurels et fonctionnels
 - Normes pour les processus d'entrée et la présentation
 - Processus de sortie
- 6 Examen et amélioration des tableaux "look-up" et des codes ICCAT
- 7 Thèmes/domaines à débattre éventuellement à la prochaine réunion
- 8 Recommandations pour l'avenir
- 9 Autres questions
- 10 Adoption du rapport et clôture

Apéndice 2

Orden del día

1. Apertura de la reunión
2. Adopción de la Agenda y disposiciones para la reunión
3. Elección del relator
4. Perspectiva del estado actual de desarrollo de la ICCAT-RDB
 - Presentación general de prototipos
 - Cambios respecto al plan de trabajo original
 - Problemas encontrados
5. Medidas y directrices para ajustar la ICCAT-RDB
 - Modelos estructurales y funcionales
 - Procesos de entrada y normas de presentación
 - Procedimientos de salida
6. Revisión y mejora de los actuales códigos y tablas de consulta ICATT
7. Posibles temas / áreas para discutir en la próxima reunión
8. Recomendaciones futuras
9. Otros asuntos
10. Adopción del informe

Appendix 3

List of documents/Liste de documents/Lista de documentos

- | | |
|-------------|---|
| SCRS/01/040 | Current status of ICCAT relational database management system (ICCAT_RDB)-
Palma, C. |
| SCRS/01/041 | ICCAT-RDB management system: T1 database (Task 1) status – Palma, C. |