

A Relational Database on the Abundance, Distribution, and Environmental Conditions of Pacific Salmon at Sea

Maxim V. Koval

Kamchatka Research Institute of Fisheries and Oceanography (KamchatNIRO)
18, Naberezhnaya Street, Petropavlovsk-Kamchatsky 683000, Russia



Keywords: RDBaseMIL, relational, database, tables, records, key fields, links

Since 1955, the Kamchatka Research Institute of Fisheries and Oceanography (KamchatNIRO) has investigated the marine life of Pacific salmon in various areas of the North Pacific Ocean. One of the objectives of this research was to get information about the abundance of salmon (and other species) collected with various fishing gears (gillnets, trawls, and seines). In addition, the scientists of the Laboratory of Marine Salmon Investigations (MIL) at KamchatNIRO collected biological information on Pacific salmon including biometric data (length, mass, etc.), biochemical composition of salmon and their prey, food habits (stomach content analysis), forage base data (zooplankton samples obtained using different plankton nets), tagging data, meteorological and hydrological observations in the study areas, etc.

The data collected in the field and in the laboratory during these surveys were presented in annual scientific reports in the 1950–1990s for one or more years of observation. Since then, the data have been kept in the laboratory in the form of gillnet and trawl cards, journals of biological analyses, cards of plankton processing and feeding, etc. Clearly, it is very problematic to retrieve the desired information by working with all the mediums at the same time.

The emergence of personal computers has allowed scientists to keep primary data in electronic form and access the necessary data much more effectively. Most of the KamchatNIRO laboratories, including the laboratory of MIL, have been using MS Excel for that purpose for almost 10 years, where primary data were kept in the form of electronic tables, and as a rule, on separate sheets and files. The format and contents of the data kept in these files varied among scientists in relation to their research needs and objectives. With the accumulation of primary data kept in MS Excel files, the problem of searching and selecting the necessary information occurred again, because MS Excel is not designed for keeping and processing massive databases and has limited connection between separate files.

MS Access is one of the most popular software among computer applications available for mass data processing. It was designed as a full-range Relational Database Management System (RDBMS) containing all the necessary means for data access and processing. MS Access has an advantage over similar products as it was created in Microsoft Office environment. Thus, data can be easily converted into MS Excel tables or MS Word files. In addition, MS Access allows to work out a run-time version of RDBMS, which can be realized with richer RDBMS. Thus, to manage the data reported in paper and electronic formats by the laboratory of MIL, we developed a relational database using Microsoft Access 2000/2002. Here, the author presents the structure of the database containing all the primary data collected by the laboratory of MIL.

For the convenience of describing the scheme of the Database, we saved the main tables used for keeping unique records and key fields connecting them (Fig. 1). For the demonstration of the RDBaseMIL's General scheme (Fig. 1), I have left only 20 basic tables (cruise data, station data, biometric data etc.) because the other 15 tables are auxiliary (i.e. vessel list, catching devices list, standard fishing areas list etc.). The complete version of the database RDBaseMIL contains 35 tables.

At the first stage, the main tasks for the author were to develop a scheme of the data, which will be included into the shareable database, and to adopt the scheme for a multi-user environment. While developing a relational database, it is quite difficult to foresee all the main tasks that will be performed with the application, and especially those tasks that are not currently used but that may appear in the future. That is why it is very important to create at the first stage a flexible database scheme that allows the addition of new components without changing the basic structure at the next stages.

The primary data kept in the laboratory were analyzed. Then some versions of the scheme were tested. As a result, the author has chosen the scheme given below. In our opinion, the scheme allows to solve the main tasks of the database and to insert all the necessary additions and updates (to create new fields and tables) (Fig. 1).

To adopt the database in a multi-users environment the following procedures were made: 1) A new MS Access file with the MDB extension was created (named RDBaseMILtbl). Then, 15 main tables of the database were imported into it. They contained the primary data collected during the scientific surveys of the laboratory of MIL;

2) Links with the RDBaseMILtbl tables of KamchatNIRO net file were connected with the main file RDBaseMIL.mdb, which contains all the “reference” tables.

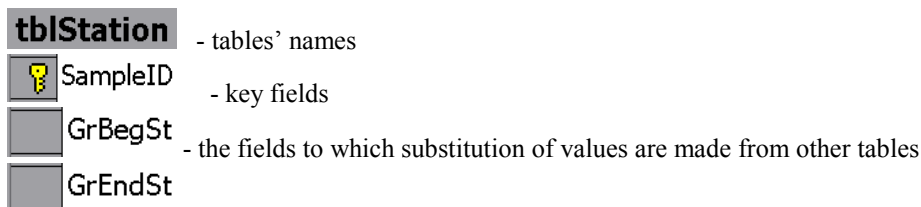
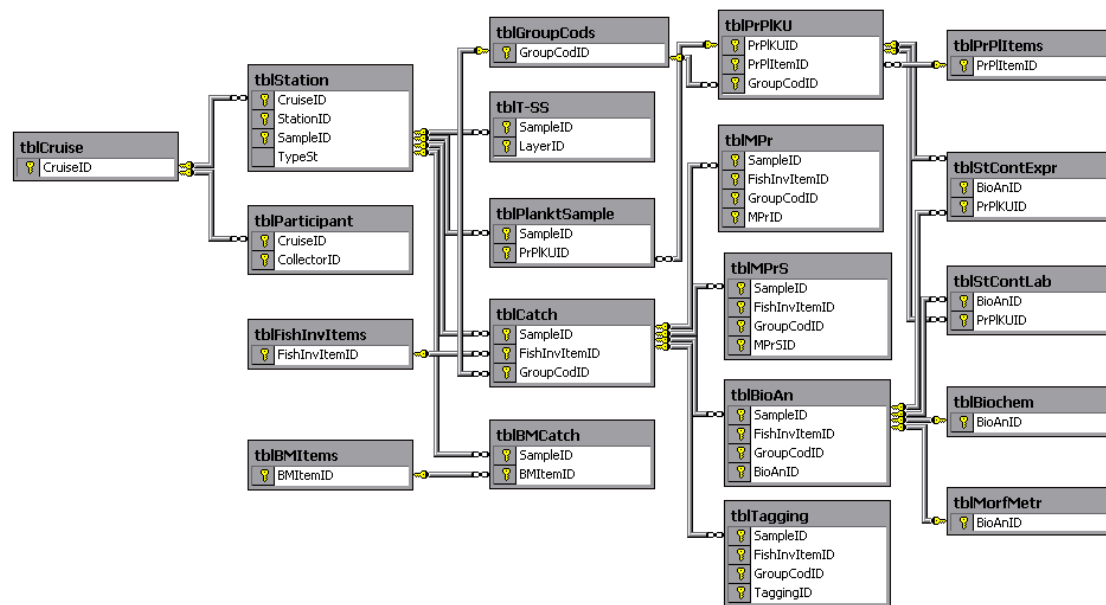
RDBaseMIL.mdb files are installed in the computers of the scientists having access to the database. It allows several users to access the database simultaneously. The users can view, add and review data of the database, depending on their accessibility to the database and their affiliation.

Primarily, the project was developed for the scientists from the laboratory of MIL for their work during cruises. The main step used to build this database was by manually entering the raw data using the keyboard. For this purpose, MS Access has special objects, called “forms”. This is why that when the database was designed a special attention was paid to these objects, as the information, obtained during cruises can be overwhelming. Maintainability of the “forms” greatly influences on the results of the work of scientists, because the data are entered manually.

In conclusion we would like to add that this version of the relational data base RDBaseMIL is not final, because we can encounter new inevitable tasks, not provided by the database. In our opinion, today’s scheme of the project is optimal and «normalized», because it allows to input the primary data of the laboratory and create new tables, without changing the base structure of the database RDBaseMIL. Now the database is being implemented for a more effective usage, but its base structure remains.

The detailed description of the database RDBaseMIL was submitted in the following document: Koval, M.V. A Relational Database on the Abundance, Distribution, and Environmental Conditions of Pacific Salmon at Sea (NPAFC Doc. 800). 30p. Kamchatka Research Institute of Fisheries & Oceanography (KamchatNIRO), Russia, Petropavlovsk-Kamchatsky.

Fig.1. General scheme of the relational database RDBaseMIL (the basic tables).



Description of links:

- link “one-to-many” (each record of the table A (from side 1) can have several records of the table B (from side ∞), but a record of the table B cannot have more than one record in the table A.

- link «one-to-one» (each record of the table A can have only one record of the table B and vice versa)