HYDROLOGY AND WATER RESOURCES IN ARIZONA AND THE SOUTHWEST, VOLUME 7, p. 1-6.
Proceedings of the 1977 meetings of the Arizona Section of the American Water Resources Association and the Hydrology Section of the Arizona Academy of Science, held in Las Vegas, Nevada, April 15-16.

A WATER SUPPLY DATA BASE

by

J. F. Nunamaker, David E. Pingry and Rex Riley

ABSTRACT

This paper describes a water supply data base being developed for the Colorado River Basin States by the University of Arizona under contract with the Electric Power Research Institute, Inc. This data base is a guide to existing natural, technical, economic, and legal water data and water data agencies in the states of Arizona, California, Colorado, Nevada, New Mexico, Utah and Wyoming.

INTRODUCTION

Water is one of the most studied and measured resources in the history of mankind. The amount of data available concerning water resources in the U.S. is enormous. Scores of private and public organizations collect and make available water resources information. In spite of the vast amount of time and money spent in accumulating water data it is still extremely difficult to quickly answer a specific question of the form:

> "Can water of a particular quantity and quality be delivered to a particular location for a specified time at an acceptable cost for the proposed use?"

The purpose the the data base design presented in this paper is to quickly and cheaply direct users to the sources of relevant information and data. The following sections describe the nature of existing water supply information and data, the design of the water supply data base structure and the implementation of that design.

WATER DATA STRUCTURE

The data necessary to provide an answer to the question posed above is extremely complex. An indication of this complexity can be illustrated by reviewing the process of supplying water. The process can be divided into five major sub-processes:

1. Remove water from natural water cycle
2. Alter water quality to meet damand process needs (including natural cycle for return water)
3. Store water
4. Transport water (including importing into and exporting from the region)
5. Return water to natural water cycle

For any single application of this process not all of these sub-processes need be used. For example, before the enactment of water quality laws, water was often removed from the natural cycle, used and returned with no treatment at all. Figure 1 illustrates the possible combination of sub-processes which might exist.

It is quickly evident that collecting some numbers on the quantities and qualities of water available at particular points in the water cycle alone does not answer the question posed above. To answer this question requires information not only on the location, quality and quantity of natural sources of supply, but also on factors which affect the delivery of the water to the point of use.

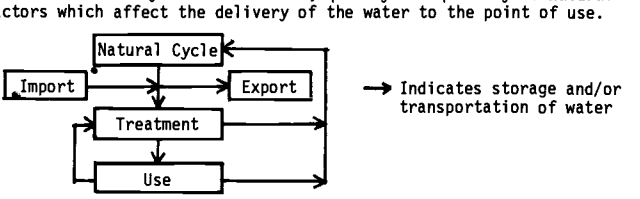

Figure 1 Regional Water Delivery Process

The authors are respectively Professor of Management Information Systems and Associate Professor of Economics at the University of Arizona, Tucson Arizona and Project Manager, Electric Power Research Institute, Inc., Palo Alto, California.

The factors which affect or constrain the supply of water can be divided into four major categories:

1. Natural The natural supply limitation for a region is all the known quantities of water in the natural cycle. Three broad sub-categories would include surface water, groundwater and rainfall.
2. Technical The technical limitations on supply for a region impinge on the subprocess of water delivery. Certain material supplies may be unusable because of lack of implemented technologies in the areas of extraction from the natural cycle, treatment, storage and transportation. In addition, these constraints may limit importation and recycling possibilities.
3. Economic The economic limitations further restrict the delivery of water. Water must be delivered at a cost low enough to make its proposed use economically feasible. Many technically possible alternatives may be ruled out by economic considerations.
4. Legal There are a large number of constraints on water delivery which are not directly associated with the level of technology and the cost of delivery. The constraints are imposed through the political process and include interstate compacts, laws (federal, state and local), agency regulations (federal, state and local) and judicial rulings. In addition to these formal restructions there are a number of supply possibilities which are politically unfeasible. That is, the public knowledge that one of these supply alternatives would be used, in all probability, will stimulate a legal response, either a court challenge or legislative action.

In light of the restrictions outlined above, two important observations can be made about the type of data base needed to answer water supply questions.

1. A useful data base must account for all of the restrictions. Data on the natural supply without information on the other three categories is useless.
2. A useful data base must include more than numbers. It is not possible to quantitatively represent many of the restrictions outlined above.

The implication of these observations is that a useful water data base, that is one which assists the researcher in answering the proposed question, can not be constructed by simply extracting the "100 most important water numbers" from existing sources and organizing them into a new format. Information at that level is of little use. On the other hand, existing data sources are fragmented, requiring a user interested in the water supply situtation in a particular region to seek out sources which have information on all of the natural, technical, economic and legal constraints which are relevant. At the federal level, for example, the user would have to be familiar with 52 computerized data bases.

WATER DATA BASE DESIGN

HOST/SUPPORT DATA BASE DESIGN

The logical question which arises, in light of the discussion above, is "What does a useful data base contain?" As inticated before, the amount of information accumulated on water is enormous. It is in the form of computer accessible data bases, books, articles, research reports, agency reports, files, company equipment descriptions, legal documents, etc. Collectively this data represents the "state of knowledge" in the water area. The problem is that is never "collected." The reasons that this has never happened are obvious. The cost of assembling and organizing all of this data in some coherent manner would be extremely high. In addition, the cost of maintaining such a data base would be extreme. The Benefit-Cost ratio on such a project would be highly unfavorable.

Of course, the solution lies somewhere between the extreme of "100 important numbers" and all of the relevant data. Our approach is to create, in effect, a combination of data and a data dictionary. The strategy is to build a host data base with a structure of the form outlined in Figure 2. The host data base includes pointers to support data bases maintained by other organizations. The host water supply data base includes a central data description of all elements contained in the host data base as well as description of all data elements from the support data bases.

The necessity of the Host/Support approach is obvious when one examines some of the existing (support) data bases. For example, the WATSTORE Data Base maintained by the U.S. Geological Survey consists of 43,000 measurement sites with a number of values being recorded daily. The task of updating and maintaining such a large data base is tremendous. It would be costly to duplicate the updating and maintainance of support data bases of such magnitude. Instead, descriptions of the support data bases are maintained in the host data base. Certain data fields from a support data base which are used frequently or determined to be very valuable are included in the host data base.

The entries in the host data base consist of the following:

1. Number (data values) from other data bases
2. Descriptions of data elements in other data bases
3. Appropriate key works from other data bases
4. Abstracts of articles and research reports
5. References to abstracts of articles and research reports
6. References to laws, regulations and judicial rulings
7. Summaries of laws, regulations and judicial rulings
8. Copies of laws, abstracts and judicial rulings
9. Lists of federal, state and local agencies

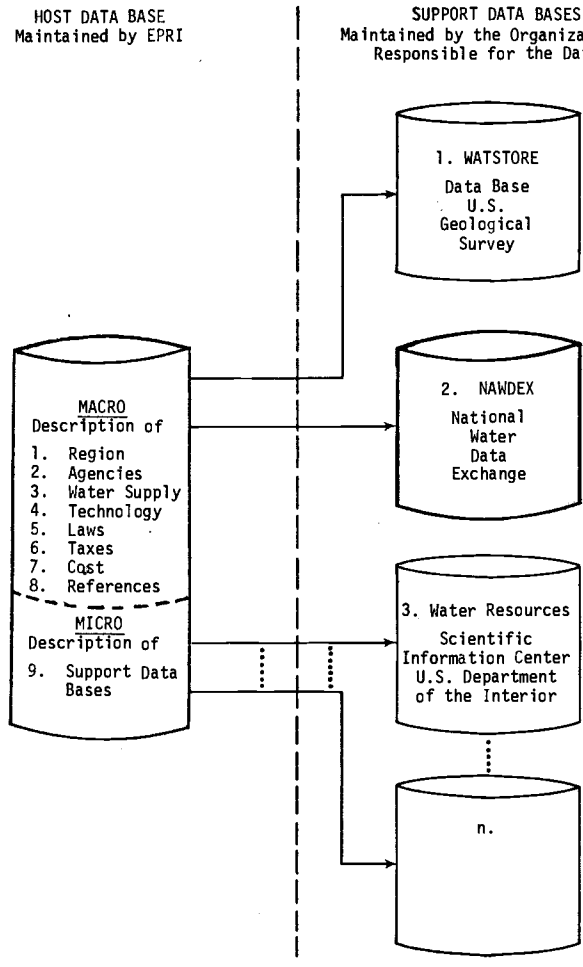

Figure 2 Methodology for the Development of a Water Supply Data Base System

MACRO/MICRO DATA BASE DESIGN

The host water supply data base is a logical data base structure consisting of two physical data bases: Macro data base and Micro data base. The Macro data base is a bibliographic data base with references and sources of information. The micro data base contains detailed descriptions of support data bases such as NAWDEX and WATSTORE. In addition, detailed structures of water data agencies are maintained.

Macro data base. The initial inquiries will be made using the macro data base. Although the macro data base is bibliographic in structure, its "entries" include regions, organizations, laws and data bases as well as the usual bibliographic entries. The queries in this system are made with keywords in the usual manner.

The logical structure of the Macro Water Supply Data Base is illustrated in Figure 3. The proposed scheme of the data base provides maximum flexibility in terms of taking alternative views of the data. A large number of specific queries can be formulated based on the suggested flexible record structure. Examples of the type of queries users will be able to ask are:

1. Which government agencies are responsible in a specific region?
2. Which laws are relevant in a specific region for a given technology?
3. Which technology and costs are relevant for specific regions?

4. Which papers and references are relevant to a specific technology or region?
5. Which supporting data bases are relevant to a specific region?
6. What are the relevant taxes, laws and costs for a specific technology in a given region?
7. Are there any data bases on cost of delivery of water by specific technology?

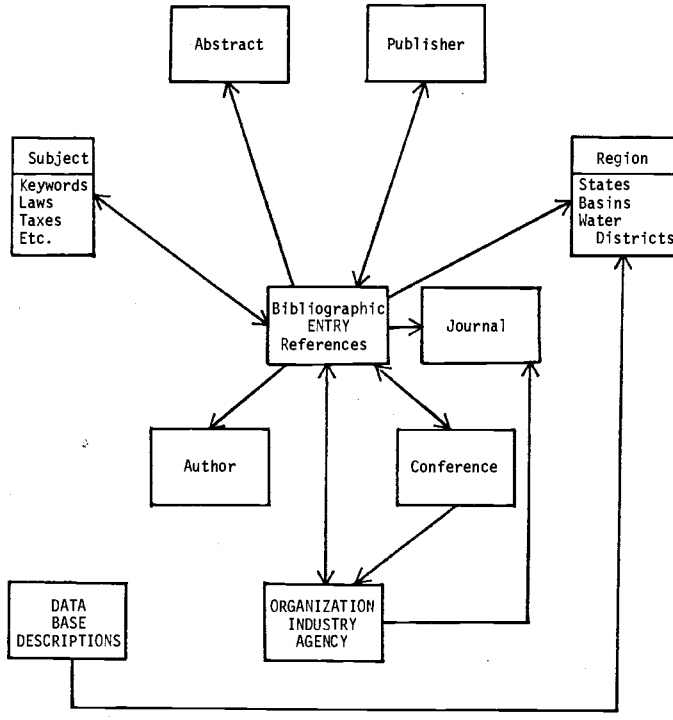

Figure 3 Water Supply Data Base: MACRO STRUCTURE

Micro data base. The other major portion of the host data base is the micro data base. This is probably the most unique feature of the data base design. The micro data base includes detailed information on water data agencies and water data bases. Use is made of a computer language called PSL (Problem Statement Language) and an associated software package called PSA (Problem Statement Analyzer). The combination of these packages (PSA/PSA) facilitates the common description of water agencies and water data bases, and allows for the analysis of their structures and interactions. The PSA system produces a number of "reports" which give the user different views of the interactions.

Users are directed from the macro data base to the micro data base. Data bases and agencies are briefly described in the macro data base. Pointers in the macro data base allow the user to obtain the more detailed information from the micro data base.

WATER DATA BASE IMPLEMENTATION

It was clear early in the design of this data base that anyone interested in the water supply situation in a specific region must have a good understanding of how the various governmental agencies interact to produce water policy. It was felt that one of the most useful services that could be provided would be to produce a guide through the maze of governmental agencies at all levels. To this end, a great deal of effort was expended in collecting information on the organization responsibilities and jurisdictions of the various federal, regional and state agencies. This information is contained in the micro data base. The PSL/PSA software allows the user to trace the links between the various organizations.

In addition, detailed information on water data bases was obtained. This includes a description of the data base contents down to the parameter level. The data bases are also linked to each other and the appropriate water agencies. This feature is particularly useful since a great deal of the data is duplicated in different data bases. The user, with the appropriate information, is in a position to select the data base with the cheapest or most convenient access.

THE REGION OF IMPLEMENTATION

The region defined by the states containing the Colorado River Basin was selected for implementation of the host data base concept. This region includes the states of Arizona, California, Colorado, Nevada, New Mexico, Utah and Wyoming and is unique in the sense that it contains the full range of water quantity and quality problems.

These states collectively have virtually every major energy-related water problem. The basic unifying feature of these states is the Colorado River Basin. The allocation of this river among the states and among the alternative uses within the states is a continuing and controversial process. The alternative uses are numerous and include domestic oncsumption, industrial consumption, irrigated agriculture, metal mining, coal mining, oil shale processing and power production. Many of these alternative uses are directly energy related. Others are indirectly energy-related. Copper, for example, palys a major role in the energy industry, although it is not an alternative energy source.

The trade-offs associated with water use in the Colorado River Basin are numerous and complex. The area is a major energy and water supplier. In addition, it is a major consumer of energy and water, particularly in the Los Angeles, San Francisco, Denver and Phoenix areas. The application of the Host Water Supply Data base for this region should test power of the data base structure as well as be directly useful for current issues.

To give an overview of data available in the Colorado River Basin states, it is useful to catagorize the data elements and give some summary statistics. In Table 1 the agencies described in the Water Supply Data Base are classified by level of geographic coverage. In addition, the associated computerized data bases are listed along with the data classification (Natural, Technical, Economic, Legal). The regional catagory is used for any organization which overlaps two or more states. This catagory includes the regional offices of federal agencies.

The preliminary summary statistics themselves are impressive. There are 31 U.S. agencies and 2 independent organizations with access to 72 computerized data bases which cover all four data catagories. The most completely covered catagory is natural data group. There are 32 regional and 118 state organizations which have data in the Colorado River Basin States. In Arizona, for example, there are 24 organizations and 10 computerized data bases. This does not include the U.S. data base with Arizona data. For similar information on the other states see Table 1.

Some comment should be made about the character of the data. The natural data tends for the most part to be numbers, that is, measurements of flow, quality parameters, depth to water, etc. The most comprehensive directory of this data is the new NAWDEX system operated by the Geological Survey. In fact, most of the water quality data available is developed through USGS programs. There are also various documents such as the Geological Survey Water Supply Papers which discuss in great detail the various groundwater aquifers. The technical and economic data is mainly held in bibliographic form. References to various studies which cover water supply technologies both from a technical and economic standpoint of view are included in a number of computerized data bases. Some numbers are directly available such as the economic projections included in the Water Resources Council Natural Assessment. The legal data available in the limit is all of the laws and court decisions relating ot those laws. However, there are legal data bases as well as various summary information available in report form and data bases which are useful to the legal layman.

Table 1 Summary Table (Preliminary)

Agency Classification	Data Catagory									
	Total		Natural		Technical		Economic		Legal	
	Data Bases	Agencies	Data Bases	Agencies	Data Bases	Agencies	Data Bases	Agencies	Data Bases	Agencies
Federal	52	31	50	9	22	5	5	4	7	6
Non-Federal U.S.	22	2	21	2	11	2	7	2	4	2
Regional	2	32	2	2	0	0	0	0	0	0
State Totals	41	118	36	30	0	0	1	1	5	5
Arizona	10	24	10	9	0	0	0	0	0	0
California	14	34	13	9	0	0	0	0	1	1
Colorado	2	16	2	2	0	0	0	0	0	0
Nevada	7	12	6	5	0	0	0	0	1	1
New Mexico	3	14	3	3	0	0	1	1	1	1
Utah	2	11	0	0	0	0	0	0	1	1
Wyoming	3	7	2	2	0	0	0	0	1	1

TOTAL AGENCIES 283
TOTAL DATA BASES 117

NOTE: The number of agencies and data bases in the four data categories does not always equal the total number of agencies or data bases because some agencies have more than one data base and some data bases contain several data catagories.